



HAL
open science

L'Origine du langage et des langues

Béatrice Fracchiolla

► **To cite this version:**

Béatrice Fracchiolla (Dir.). L'Origine du langage et des langues. Marges Linguistiques. 11, 294 p., 2006, Les origines du langage et des langues, Michel Santacrose. hal-01492095

HAL Id: hal-01492095

<https://hal.univ-lorraine.fr/hal-01492095>

Submitted on 20 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

M

a r g e s

L**i n g u i s t i q u e s**

Numéro 11, Mai 2006

Langage – Communication – Représentations

REVUE ELECTRONIQUE EN SCIENCES DU LANGAGE

**L'Origine du langage et
des langues**

Béatrice Fracchiolla (ed.)

SommaireLes embarras de l'origine des langues
*Par Sylvain Auroux*Origine du langage et origine des langues :
réflexions sur la permanence et le renouvellement
d'un questionnement des Lumières
*Par Robert Nicolai*La question de l'origine du langage chez Saussure
*Par Marcio Alexandre Cruz*Du protolangage au langage : modèle d'une transition
*Par Jean-Louis Dessalles*La mosaïque du langage / The Mosaic of Language
*Par Stéphane Robert et Georges Chapouthier*Language Evolution : What Evolved ?
*By John H. Schumann & al.*From Calls to Words : How ethology can bridge the divide
*By Eric A. Salzen*Evolution of Language as a Gestural System
*By Michael C. Corballis*De la communication chimique à la parole
*Par Béatrice Fracchiolla*Language Evolution : the Population Genetics Way
*By Salikoko S. Mufwene*Genes and Languages
*By Luigi Luca Cavalli Sforza*La convergence pluridisciplinaire
dans la recherche de l'origine des langues :
l'exemple américain de la « Nouvelle Synthèse ».
Aspects historiques, théoriques et étude critique
Par Laurent MétozDe l'origine du langage à l'émergence
du milieu sémiotique
*Par François Rastier*La question de l'origine des langues :
vaine quête du Graal ?
Par Anne Szulmajster-Celnikier

Comité scientifique

Jean-Michel Adam (Université de Lausanne, Suisse) — Jean-Jacques Boutaud (Université de Bourgogne, France) — Josiane Boutet (Université de Paris VII, France) — Thierry Bulot (Université de Rouen, France) — Paul Cappeau (Université de Poitiers, France), Jean Caron (Université de Poitiers, France), Chantal Charnet (Université Paul Valéry — Montpellier III, France) — Joseph Courtés (Université de Toulouse II, France) — Béatrice Daille (IRIN — Université de Nantes, France) — Marcelo Dascal (Université de Tel Aviv, Israël) — Françoise Gadet (Université de Paris-X Nanterre, France) — Alain Giacomi (Université de Provence, France) — Benoit Habert (Laboratoire LIMSI, Université Paris X, France) — Monica Heller (Université de Toronto, Canada) — Thérèse Jeanneret (Université de Neuchâtel, Suisse) — Catherine Kerbrat-Orecchioni (GRIC (Groupe de Recherches sur les Interactions Communicatives, CNRS-Lyon2) Université Lumière Lyon II, France) — Norman Labrie (Université de Toronto, Canada) — Guy Lapalme (Université de Montréal, Québec, Canada) — Olivier Laügt (Université Michel de Montaigne - Bordeaux III, France) — Marinette Matthey (Université de Neuchâtel, Suisse) — Jacques Maurais (Conseil de la langue française, Québec, Canada) — Piet Mertens (Katholieke Universiteit Leuven, Département de Linguistique, Belgique) — Sophie Moirand (Université de la Sorbonne Nouvelle, France) — Claudine Moise (Université d'Avignon, France) — Lorenza Mondada (Université de Bâle, Suisse) — Marie-Louise Moreau (Université de Mons-Hainaut, Belgique) — Bernard Py (Université de Neuchâtel, Suisse) — François Rastier (CNRS, Paris, France) — Véronique Rey (Université de Provence, France) — Didier de Robillard (Université de Tours, France) — Eddy Roulet (Université de Genève, Suisse) — Daniel Véronique (Université de Paris III : Sorbonne nouvelle, France) — Jean Véronis (Université de Provence, France) — Evelyne Viegas (Natural Language Group, Microsoft Corporation, USA) — Diane Vincent (Université de Laval, Québec, Canada) — Robert Vion (Université de Provence, France).

Consultants associés

Michel Arrivé (Université de Paris X Nanterre, France) — Louis-Jean Calvet (Université de Provence, France) — Jacques Fontanille (Université de Limoges, Centre de Recherches Sémiotiques (FRE2208 CNRS), France) — Jacques Moeschler, Département de linguistique, Université de Genève, Suisse) — Geneviève Dominique de Salins, Faculté Arts, Lettres et Langues, CIREC (EA 3068), Université de Saint-Etienne, France) — Andréa Semprini (Université de Lille III, France).

Comité de rédaction

Michel Arrivé (Université de Paris X Nanterre, France) — Mireille Bastien (Université de Provence, France) — Thierry Bulot (Université de Rouen, France) — Stéphanie Clerc (Université d'Avignon, France) — Véronique Fillol (Université de Nouméa, Nouvelle Calédonie) — Béatrice Fracchiolla (Université de Toulon, France) — Alain Giacomi (Université de Provence, France) — Véronique Magaud (Université de Provence, France) — Marinette Matthey (Université de Neuchâtel, Suisse) — Michèle Monte (Université de Toulon, France) — Philippe Rapatel (Université de Franche Comté, France) — François Rastier (Cnrs, Paris, France) — Didier de Robillard (Université de Tours, France) — Michel Santacroce (Université de Provence, France) — Yvonne Touchard (IUFM de Marseille, France) — Daniel Véronique (Université de Paris III : Sorbonne nouvelle, France) — Jean Véronis (Université de Provence, France) — Chantal Wionet (Université d'Avignon, France).

Rédacteur en chef

Michel Santacroce (Université de Provence, France).

Présentation éditoriale

Par Béatrice Fracchiolla
Université de Toulon,
UFR de Lettres et Sciences Humaines, France

Mai 2006

« Le problème de l'origine concerne évidemment autre chose que l'origine du langage : *son essence, son rapport à la pensée et à la société, l'organisation de la pensée humaine (empirisme vs rationalisme) sa différence à la pensée animale et plus généralement la position de l'humanité dans l'ensemble du règne animal* ». (S. Auroux).

Comme le rappelle Sylvain Auroux dans le texte qui inaugure ce numéro, la question de l'origine des langues est irrémédiablement liée à celle de l'origine de l'homme, et a de tout temps suscité nombre de débats passionnés et tentatives d'explications diverses. La question de l'interdit renvoie en réalité à celle de l'impossibilité scientifique. Cette question de l'interdit, qui est posée de façon spécifique par rapport à l'essence même de l'objet de la linguistique chez Saussure, M.-A. Cruz l'examine à son tour plus spécifiquement sous l'aspect du rapport système/temps.

L'étendue de la problématique multiplie les questions : il s'agit de s'interroger à la fois sur les manières dont les langues modernes sont le reflet de ce qui a existé d'abord puis évolué, tout autant que de se demander comment le langage a évolué ? et qu'est-ce qui a évolué exactement ? Mais il s'agit aussi, de distinguer fondamentalement entre la problématique liée à l'émergence du langage même et celle de l'évolution des langues (R. Nicolai). Pour R. Nicolai, la recherche sur l'origine des langues est sans aucun doute un mythe qu'il convient d'aborder en termes de « constructions intersubjectives ».

Cette thématique du « mythe de l'origine des langues » est, on le verra, commune à plusieurs auteurs, dont A. Szulmajster-Celnikier qui la reprend à l'occasion d'une interrogation plus vaste sur les modalités et les raisons de la naissance et de la mort des langues. Quel est le rôle de la « transmission », du contact des langues et comment traiter de la « variation » dans cette problématique ? (S. Mufwene) ? Quelles sont les différentes approches de classification des langues (A. Szulmajster-Celnikier) et qu'en est-il d'une corrélation possible entre évolution génétique des populations et diversification des langues sur la planète (L.-L. Cavalli-Sforza) ? L. Métoz expose ici les problématiques de la « nouvelle synthèse » et comment un certain nombre de travaux (dont ceux, précisément de L.-L. Cavalli-Sforza, mais aussi de Greenberg, Ruhlen, Renfrew) recourent à la linguistique, l'archéologie et la biologie pour appréhender le problème dans sa complexe globalité.

La question centrale, toujours présente et rappelée par S. Auroux est : l'origine des langues est-elle inaccessible à l'histoire ? D'après Schumann & al., elle n'est plus aujourd'hui inabordable, en raison plus particulièrement des découvertes relativement récentes sur le cerveau et sur les processus impliqués dans l'évolution. L'exploration des fondements biologiques du langage et les questionnements afférents sont du moins devenus possibles et légitimes sous certains de leurs aspects, plus résolument grâce aux découvertes et réflexions venant d'autres disciplines que de la linguistique, comme la biologie, la paléanthropologie, l'éthologie, la neurolinguistique (Schumann & al. ; E. Salzen ; M. Corballis ; B. Fracchiolla). Les contributions de G. Chapouthier & S. Robert, mais aussi de J.-L. Dessales mettent en regard, à des degrés divers, les données observées en termes de « construction par juxtaposition » au niveau à la fois des langues et du vivant, et reprennent respectivement l'analogie entre mosaïque « génétique », constitutive du vivant et mosaïque « syntaxique » constitutive du sens.

Partant du principe qu'il est illégitime de discuter d'éléments qui ne peuvent être selon lui, ni confirmés ni infirmés, F. Rastier, reprend à son compte la question de l'interdit comme

impossibilité scientifique, et réfute pour sa part la question même de l'origine du langage pour développer ce qui fait selon lui la spécificité des langues. Or, sur l'origine du langage et des langues – et comme le laisse supposer la conclusion de Sylvain Auroux – c'est sans aucun doute dans le débat lui-même – qui participe de la volonté tout autant que du besoin humain de comprendre le monde – qu'il convient de chercher la légitimité, et non dans les thèses visant à expliquer l'un et l'autre de façon définitive.

Anne Szulmajster-Celnikier remarque en conclusion de son article : « l'attention portée à la pluralité linguistique a en outre l'avantage de mettre en relief, sur fond d'universalisme (l'unité écologique et psycho-physiologique du genre humain, et le fait que l'on puisse traduire d'une langue à l'autre) un relativisme linguistique fondamental. Faire voir ce que, selon la formule de Jakobson reprise par Hagège, les langues « obligent » à dire ou « empêchent de dire », c'est aussi prendre conscience que les structurations spécifiques des idiomes imposent aux sujets parlants des représentations diversifiées du monde qu'il importe d'analyser dans leur richesse, non de réduire à l'unique ».

Aussi est-ce par l'hétérogénéité des approches que nous avons voulu refléter, à l'échelle de ce numéro et des auteurs qui nous ont fait l'amitié d'y participer, la difficulté d'approcher le vivant et l'humain dans leurs grandes diversité et complexité, en tenant compte des apports récents de l'interdisciplinarité. Comme à la fin du XIX^e siècle les débats restent vifs, sans doute parce qu'ils engagent aussi des visions de l'homme différentes, et que des désaccords profonds existent. C'est pourquoi, sans doute, il convient d'envisager ce numéro plus sous l'angle des questions qui y sont posées, que sous celui de réponses qui y seraient données.

Sous la direction de *Béatrice Fracchiolla*, Université de Toulon, UFR de Lettres et Sciences Humaines, France.

01- Les embarras de l'origine des langues

Par **Sylvain Auroux**, Laboratoire d'histoire des théories linguistiques, UMR 7597 : Cnrs, Paris 7, École Normale Supérieure-LSH, France

Pages 058 — 092

02- Origine du langage et origine des langues : réflexions sur la permanence et le renouvellement d'un questionnement des Lumières

Par **Robert Nicolai**, Université de Nice et Institut universitaire de France, France

Pages 093 — 129

03- La question de l'origine du langage chez Saussure

Par **Marcio Alexandre Cruz**, Université Paris 3 : Sorbonne Nouvelle, École doctorale 268 UPRES SYLED, EA 2290, Paris, France

Pages 130 — 141

04- Du protolangage au langage : modèle d'une transition

Par **Jean-Louis Dessalles**, École Nationale Supérieure des Télécommunications, Paris, France

Pages 142 — 152

05- La mosaïque du langage / The Mosaic of Language

Par **Stéphane Robert**, CNRS-LLACAN & INALCO et **Georges Chapouthier**, Université de Paris 6, France

Pages 153 — 166

06- Language Evolution : What Evolved ?

By **John H. Schumann & al.**, University of California, Los Angeles, USA

Pages 167 — 199

07- From Calls to Words : How Ethology can Bridge the Divide

By **Eric A. Salzen**, Psychology Department, Aberdeen University, Scotland

Pages 200 — 217

08- Evolution of Language as a Gestural System

By **Michael C. Corballis**, Department of Psychology, University of Auckland, New Zealand

Pages 218 — 229

09- De la communication chimique à la parole

Par **Béatrice Fracchiolla**, UFR de Lettres et Sciences Humaines, Université de Toulon, France

Pages 230 — 242

10- Language Evolution : the Population Genetics Way

By **Salikoko S. Mufwene**, University of Chicago, Department of Linguistics, USA

Pages 243 — 260

11- Genes and Languages

By **Luigi Luca Cavalli Sforza**, Stanford University, USA

Pages 261 — 280

12- *La convergence pluridisciplinaire dans la recherche de l'origine des langues : l'exemple américain de la « Nouvelle Synthèse ». Aspects historiques, théoriques et étude critique*

Par **Laurent Métoz**, Laboratoire HTL, UMR CNRS 7597, ENS LSH, Lyon, France

Pages 281 — 296

13- *De l'origine du langage à l'émergence du milieu sémiotique*

Par **François Rastier**, Directeur de recherche CNRS, Paris, France

Pages 297 — 326

14- *La question de l'origine des langues : vaine quête du Graal ?*

Par **Anne Szulmajster-Celnikier**, Collège de France, Chaire de Théorie Linguistique, Paris, France

Pages 327 — 342

Mai 2006

La question de l'origine des langues, comme celle de l'origine de l'homme à laquelle elle est irrémédiablement liée, a été abordée de multiples façons tout au long de notre histoire intellectuelle : mythes archaïques et/ou religieux, théories philosophiques (cf. le *Cratyle* de Platon), spéculations plus ou moins oiseuses, récits d'expériences improbables colportées durant des siècles (voire des millénaires, lorsqu'on les attribue à un pharaon), travaux des linguistes, anthropologues, archéologues et biologistes. Peu de sujets ont bénéficié d'une préoccupation aussi constante et passionnée. À l'époque moderne¹, on peut ponctuer l'histoire de son traitement par cinq articulations :

- laïcisation² de la question par les philosophes des lumières (Condillac, Rousseau, Maupertuis, Herder) ;
- recherches multiples des linguistes du XIXe siècle qui pensent faire avancer la question grâce aux nouvelles méthodes de la *grammaire comparée* ;
- interdiction à ses membres d'étudier la question (ainsi que celle de la langue universelle) par la Société de linguistique de Paris (article 2 des statuts de 1866) ;
- discussions philosophiques dans les années soixante du XXe siècle dans la communauté francophone (notamment autour de Derrida ; voir Derrida 1967a, b) suscitée par la « redécouverte » de l'*Essai sur l'origine des langues* de Rousseau (voir Porset 1970), discussions qui vont dans le sens de l'*épigénéisme*³ et préservent ainsi l'interdit, justifié par la *circularité des origines* (numéro spécial de la *Revue Internationale de philosophie*, 82-4, 1967) ;
- conférence organisée en 1975 sous les auspices de l'Académie des sciences de New York sur les *Origins and Evolution of Language and Speech* (les actes sont publiés l'année suivante, Harnad *et al.* 1976). C'est de cette conférence – qui est encadrée par la querelle autour de Swadesh et sa glottochronologie dans les années soixante d'un côté et le projet d'une « nouvelle synthèse » à la fin des années quatre vingt de

¹ Pour des raisons de compétence linguistique personnelle, j'ai dû laisser de côté la littérature rédigée en russe ; dans ces conditions, j'ai choisi d'exclure la discussion du marrisme, à laquelle je n'ai accès que de seconde main (voir dans le t. 3 de Auroux (éd.) 1990-2000, la contribution de P. Sériot). La tradition russe semble relativement isolée jusqu'à la récente discussion sur la famille nostratique (A. Dolgoposky et V. Illich-Svytich dans les années 1960) que l'on rapproche de la famille eurasiatique de Greenberg (Ruhlen, 1994, pp. 72-74 ; à cette date, les principaux textes sont accessible en anglais).

² On peut percevoir l'étendue de cette laïcisation par la question que l'Académie de Berlin met au concours en 1769 : « En supposant que les hommes abandonnés à leurs facultés naturelles, sont-ils en état d'inventer le langage ? Et par quels moyens parviendront-ils à cette invention ? On demande une hypothèse qui explique la chose clairement et qui satisfait à toutes les difficultés ». C'est Herder qui sera primé en 1771. Cette laïcisation ne signifie pas la disparition des partisans de l'origine divine comme J.-P. Süssmilch, auquel Herder s'oppose ; au reste, Herder aura ultérieurement tendance à re-théologiser la question.

³ J'utilise ce terme de façon quelque peu néologique par rapport à son sens courant en embryologie pour désigner une théorie qui soutient qu'il faut distinguer entre le commencement (qui est une question empirique) et l'origine (qui concerne l'appréhension rationnelle) d'un phénomène. C'est Kant qui, dans la Critique de la raison pure (1781) a introduit cette distinction à propos de nos connaissances. Husserl la reprend lorsqu'il traite de la géométrie : « La question de l'origine de la géométrie (...) ne doit pas être la question philologico-historique, ni par conséquent l'enquête menée à propos des premiers géomètres (...). (...), notre préoccupation doit aller plutôt vers une question en retour sur le sens le plus originaire selon lequel la géométrie est née un jour <et,> dès lors, est restée présente comme tradition millénaire, le reste pour nous et se tient dans le vif d'une élaboration incessante » (Derrida, 1962, pp. 174-175). La longue introduction de Derrida (171 pages) a eu un grand retentissement lors des discussions sur Rousseau.

l'autre - que l'on peut dater symboliquement le développement considérable de la recherche contemporaine (colloques, publications, sociétés savantes¹, soutien des grands organismes de recherche à l'aide de crédits conséquents).

Il est arrivé que d'importantes questions qui ont fasciné des générations fassent l'objet d'interdits de recherche par des organismes officiels, les plus célèbres étant la *quadrature du cercle* et le *mouvement perpétuel*. Ils avaient de solides raisons scientifiques (appartenance de Pi au corps des réels ; second principe de la thermo-dynamique). Dans chaque cas, l'interdit a été considéré comme un progrès scientifique considérable et – bien entendu – ne fut jamais levé ; il désigne une *impossibilité scientifique*. Dans d'autres cas, c'est le fait de braver un interdit qui ne provenait pas de la communauté savante (par exemple, la dissection des cadavres) qui fut considéré comme un progrès scientifique ; on peut considérer qu'il s'agissait seulement d'un *interdit social*. L'origine des langues et la langue universelle constituent, à ma connaissance, la seule intervention d'interdits scientifiques dans le domaine des sciences sociales. Encore que la brève chronologie de l'époque moderne que nous venons de présenter montre clairement que la position de la Société de linguistique ne fit pas l'unanimité dans la communauté ; le moins que l'on puisse dire, à voir la vitalité de la recherche contemporaine, c'est que l'arrêt n'a pas été définitif. Devons-nous introduire une autre catégorie d'interdit scientifique, celle des *interdits conjoncturels* qui dépendent de l'état de la science et n'ont pas de valeur définitive ?

1. Naissance d'un interdit

La tradition linguistique qui passe par le *Cours de linguistique générale* de Saussure (1916) se caractérise par le refus de s'engager dans des recherches sur l'origine des langues. On peut emprunter aux *Antinomies linguistiques* de V. Henry une formulation précise et claire des termes de ce refus :

Que le linguiste doive s'interdire toute recherche sur l'origine du langage, c'est un point qui semble définitivement acquis, tout au moins parmi les linguistes, si paradoxale qu'en soit la première apparence : l'origine du langage n'est pas, *a priori*, un problème linguistique, puisque la linguistique ne se propose pour objets que des langues toutes formées, dans leur état actuel, historique ou préhistorique, et qu'il ne lui est donné que de constater l'évolution, jamais la naissance d'un langage (1896, pp. 25-26).

C'est vers 1870 que ce refus a pris une tournure institutionnelle. La Société de Linguistique de Paris, fondée en 1866, doit une partie de sa célébrité (cf. Jespersen, 1922 : chap. IV) à l'article II de ses premiers statuts :

art. 2 - La société n'admet aucune communication concernant soit l'origine du langage, soit la création d'une langue universelle².

Si le débat concernant le refus est très caractéristique du développement de la recherche française, le refus lui-même n'est nullement limité à ce pays. Dans son adresse présidentielle de 1873 à la Philological Society (*TPS*, 1873-1874, pp. 251-52), Alexander J. Ellis, déclare que pour lui les questions d'origines sont « out of the field of philology proper », et que les linguistes feront davantage en s'intéressant au développement historique d'un seul vernaculaire qu'en remplissant des corbeilles « with ream of paper covered with speculations on the origin of all languages » (voir Aarsleff, 1983, p. 230). Déjà F. Bopp, qu'on considérait à l'époque comme le fondateur de la linguistique, avait posé des limites à ses recherches :

Il n'y a que le mystère des racines ou, en d'autres termes, la cause pour laquelle telle conception primitive est marquée par tel son et non par tel autre, que nous nous abstenons de pénétrer (*Vergleichende Gram.*, trad. fr., M. Bréal, 1866, p. 1).

Cette situation peut paraître étonnante : le siècle le plus fécond en études sur l'origine du langage est indubitablement le dix-neuvième. La plupart des grands noms de la linguistique (W. von Humboldt, F. Schlegel, J. Grimm, A. Schleicher, A. Steinthal, H. Paul, E. Renan, M. Muller, D. G. Withney, M. Bréal), ont apporté leur contribution, parmi une foule de travaux d'auteurs moins connus dont notre bibliographie ne reprend qu'une partie. Le développement

¹ Cf. *Mother Tongue*, revue publiée à partir de 1986 à Boston par H. Fleming pour le compte de l'Association for the Study of Language in Prehistory.

² Cf. pour les statuts de 1866, *Mémoires* I, III-V et ceux de 1876, *BSL* n° 15, i-iii. Voir Auroux 1984, p. 317. On notera que la langue universelle a été l'objet explicite et unique de la Société internationale des linguistes (1850-1860) qui a eu pour organe la *Tribune des linguistes* (1858-1859).

de la linguistique – la claire conscience qu'il s'agit d'une « science nouvelle » reposant sur la connaissance des faits et sur une nouvelle méthode (la comparaison des langues) - n'a nullement tari le projet de résoudre la question de l'origine des langues. Bien au contraire, les progrès de la philologie comparée, ont d'abord incité les auteurs, à en appliquer les résultats aux recherches sur l'origine, c'est-à-dire « à étudier la possibilité de résoudre un tel problème d'une manière scientifique » (Renan, 1858, p. 5), selon l'exemple de Grimm lui-même (1852, p. 10 et pp. 54-55 ; cf. également Schlegel, 1808, p. 60). Ce qu'on reproche aux prédécesseurs du siècle précédent, c'est d'avoir voulu résoudre la question sans la méthode comparatiste (Renan, 1858, p. 76). En France, l'Académie des Sciences Morales et Politiques - vingt ans après l'interdit de la Société de linguistique - consacre en 1886 son Prix Bordin au problème. La rédaction du sujet mis au concours est très claire :

1. Exposer et apprécier les différents systèmes qui depuis l'antiquité jusqu'à nos jours ont eu pour but d'expliquer philosophiquement les origines et les lois du langage.
2. Recueillir dans les œuvres les plus importantes de la philologie contemporaine les principes et les faits qui pourraient servir à la formation d'une philosophie du langage.

Le lauréat¹ du prix (1887), sera P. Regnaud, professeur de sanskrit et de grammaire comparée à la Faculté des Lettres de Lyon ; Regnaud n'était pas membre de la Société de Linguistique et ne le sera jamais ; aucune de ses théories n'est passée à la postérité et l'histoire n'a pas retenu son nom. Son ouvrage s'attirera les foudres de V. Henry (1888), lequel critiquera l'initiative de l'Académie (1896, p. 26).

La découverte et la diffusion des théories évolutionnistes contribuèrent largement à alimenter le débat sur l'origine des langues. La première édition de *On the origin of Species* date de 1859 ; mais l'ouvrage de Darwin qui touche au plus près le sujet est *The Expression of the Emotions in Man and Animals* (1872). L'évolutionnisme peut interférer avec les études sur l'origine des langues de trois façons. D'abord en présentant l'origine animale de l'homme, il induit logiquement l'idée d'une émergence de la faculté du langage, tandis qu'il fournit un cadre de représentation très général qu'on pense pouvoir appliquer aux phénomènes culturels (cf. Tylor 1871, qui remet à propos du langage le président De Brogues au goût du jour). L'idée d'un évolutionnisme linguistique se développe rapidement avec A. Schleicher dont certaines thèses sont antérieures à sa connaissance de Darwin (Zaborowski, 1878, p. 35), ou avec Wedgwood (1866), le beau-frère du naturaliste. Le premier ouvrage de Regnaud était intitulé *Essais de linguistique évolutionniste* (1866). *Le darwinisme incite au mono-génétilisme en matière de langage*. A. Hovelacque, professeur de linguistique à l'École d'Anthropologie de Broca, résume ainsi les conséquences de l'idée d'évolution pour l'origine des langues :

[...] en présence de ce perpétuel spectacle d'évolution qui se déroule sous nos yeux dans la nature entière, nous ne pouvons pas ne pas admettre que la faculté du langage articulé ne se soit acquise petit à petit, grâce à un développement progressif des organes (1877, p. 38).

Ce sont des arguments que l'on retrouvera encore vingt-cinq ans plus tard dans la tradition américaine chez Sapir². La thèse du développement progressif du langage se retrouve toutefois chez des auteurs qui doivent moins au gradualisme³ du darwinisme qu'à l'uniformitarisme⁴ géologique de Lyell¹.

¹ Deux mémoires seulement ont été présentés au concours ; l'autre - qui a pour auteur Destrem - obtient une mention honorable. Dans la seconde édition de son livre Regnaud publie le rapport de Vacherot sur les deux mémoires et le discours de J. Zeller lors de la séance publique de l'Académie (17 déc. 1887). Le premier (*l.c.* pp. XXIV-LIX), est intéressant en ce qu'il présente les motivations du corps savant : le modèle est la paléontologie, il s'agit de trouver « des témoins dans les vieilles langues qui nous ont été conservées » et de faire que « la science des faits vienne au secours de la logique ».

² « Despite Max Müller, however, it seems to me that the path for future work on the prime problem, more especially the origin of language lies in the direction pointed out by evolutions » ([1905], 1984, p. 388). Il s'agit d'un mémoire de maîtrise. Sapir soutiendra par la suite une conception totalement historique et culturelle du langage, voir le Chapitre I de *Language* (1922).

³ L'évolution se fait régulièrement et de façon continue (existence de formes intermédiaires). Le gradualisme se heurte au problème posé par le fait qu'il arrive qu'il manque des formes intermédiaires. Il est aujourd'hui remplacé par la théorie des « équilibres ponctués » qui admet des variations de vitesse dans l'évolution et leur inégale distribution dans les différentes lignées évolutives.

⁴ Pour l'uniformitarisme les causes actives dans le passé (ou l'avenir) doivent être observables par leurs effets actuels et agir de la même façon, dans les mêmes circonstances et avec la même force

Le second point d'interférence est la possibilité d'utiliser les recherches sur l'origine des langues contre la théorie darwinienne (le langage prouverait l'existence d'un fossé infranchissable entre l'homme et l'animal²) ; tel est en grande partie l'enjeu qui oppose Müller et Whitney.

Le troisième point est l'utilisation directe de la théorie des émotions (Carrau, 1878 ; Girard de Rialle, 1875) ; cette voie correspond à la tentative traditionnelle de l'explication par l'interjection en particulier telle qu'elle est développée par Chavée (1878), Heyse (1856) ou Steinthal (1858).

La Société de Linguistique de Paris a résisté aux études d'origine ; par exemple, elle repousse en 1869, une proposition de Dufriche-Desgenettes pour abolir l'article II (Vendryes, 1955, p. 13). Si cet article disparaît des statuts adoptés en 1876, la nouvelle formulation peut (à la rigueur) être interprétée comme englobant encore l'ancien interdit :

art. 1 - La société a pour objet l'étude des langues et l'histoire du langage. Tout autre sujet d'études est rigoureusement interdit.

Dans un article de 1955 Vendryes n'hésite pas à soutenir que l'interdit ne fut jamais levé (1955, p. 13). Son appréciation traduit au moins la permanence d'un certain état d'esprit. La question toutefois n'est pas si simple. La Société publia dans ses *Mémoires*, au moins un article concernant le thème de l'origine, il s'agit des « Commencements du verbe » (1900, tome IX), dont l'auteur est Michel Bréal, qui fut sa vie durant la cheville ouvrière et le secrétaire de la Société³. Les termes dans lesquels il s'exprime pour présenter son travail, peuvent paraître étonnants par rapport à ses fonctions :

Depuis environ trente ans, on a cherché à jeter le discrédit sur les questions d'origine : on les a déclarées insolubles. Mais le jour où la linguistique laisserait rayer ces questions de son programme, elle me ferait l'effet d'une science découronnée [...]. Si la première génération de linguistes, à qui rien ne semblait inaccessible, a été remplacée par une génération plus prudente, suivie elle-même d'une génération disposée à s'interdire les grands problèmes, nous ne nous résignons point à une diminution qui dépouillerait ces études de leur principal attrait, et presque de leur raison d'être (1904, p. 334).

L'interdit serait conjoncturel. En 1922, Jespersen n'aura pas une autre position : « La linguistique ne peut s'abstenir éternellement de s'interroger sur l'origine (et le pourquoi) de l'évolution du langage ». (1922, chap. XXI).

Il y a donc, chez les linguistes professionnels mêmes, un malaise devant l'abstention institutionnelle vis-à-vis de grandes questions de la philosophie linguistique. Ce sont ces questions auxquelles s'affronte V. Henry en 1896, pour conclure que « les découvertes les plus importantes de la philologie contemporaine ne fournissent ni faits ni principes qui puissent jamais conduire directement à la formation d'une philosophie du langage ». Ces découvertes auront simplement montré les voies que la philosophie « devra désormais s'interdire » (l.c., 44). *L'interdit serait donc vraiment scientifique et indépassable.*

Le titre choisi par Henry est une allusion transparente à Kant, qui dans la « Dialectique Transcendentale » de la *Critique de la Raison Pure*, désignait sous le nom d'*antinomies* les quatre questions (commencement du monde ; décomposition du réel ; liberté ; cause du monde) pour lesquelles des argumentations tout aussi valables conduisent à des résultats contradictoires, que Kant expose sous forme d'une thèse et de son antithèse. Henry dégage également quatre questions : nature du langage, origine du langage, langage et pensée, légalité de l'évolution (ce dernier point prévu dans le projet du livre étant finalement relégué

qu'aujourd'hui. Voir Cristy (1983), qui défend la thèse, généralement acceptée, que l'uniformitarisme - géologique constitue le modèle des néogrammairiens.

¹ « Es existieren nur graduelle Unterschiede zwischen den ersten Anfängen der Sprache und den späteren Epochen » (H. Paul, [1880], 1975, p. 36 ; cf. déjà Grimm, 1852, p. 12).

² Darwin est contraint d'aborder la question parce que le langage est considéré comme le trait distinctif entre l'homme et l'animal ; il lui importe de montrer que « the faculty of articulate speech in itself « does not » offer any insuperable objection to the belief that man has been developed from some lower form » (*Descent of Man*, Chap. 2, 1871 ; reproduit dans Harris (ed.). 1996, p. 146). Les linguistes « naturalistes » (voir Desmet, 1996), tout comme Henry, raisonnent à l'inverse : en prenant la théorie de l'évolution pour acquise, il doit y avoir un commencement au langage articulé.

³ Voir Auroux 1984, pp. 305-308.

en note ; 1896, p. 3), et les présente sous forme de thèse et d'antithèse¹. Pour la question de l'origine cela donne :

Thèse : Le bon sens à lui seul, à défaut d'aucun document, indique que le langage, comme toute chose au monde, a dû avoir un commencement, et l'intérêt qui s'attache à cette haute caractéristique de l'humanité fut de tout temps un puissant stimulant à la recherche de l'origine.

Antithèse : L'origine du langage est un problème, non seulement inabordable à la science du langage, mais dont tous les documents qu'elle étale ou accumulera dans l'avenir ne sauraient jamais lui faire entrevoir même la plus lointaine solution.

Le recours du paradigme kantien d'un côté, la divergence des appréciations de l'autre, conduisent à penser que l'interdit en question n'est pas quelque chose qui va de soi. Si l'on suit, par exemple, l'histoire du Prix Volney de l'Institut², qui jusqu'au dernier tiers du XIXe siècle constitue en France la matrice où s'élaborent les critères de scientificité³ de la nouvelle linguistique, on ne peut manquer d'être frappé par l'extrême modération. De nombreux mémoires présentés pour le Prix traitent de l'origine des langues (la plupart relèvent de la *linguistique fantastique*⁴), ou de la langue universelle⁵. Aucun n'a jamais été primé ; mais si le rejet de la langue universelle est catégorique (« L'idée première en est chimérique », séance du 22.9.1853), celui de l'origine des langues (Renan sera membre du jury à partir de 1876) est beaucoup plus prudent (« n'appartient pas à l'objet du concours », séance de 1872). La raison en est - comme le notera Regnaud⁶ - que la question touche aux principes mêmes qui sont au fondement des programmes de recherche linguistique.

2. Le labyrinthe de l'origine

La question de l'origine des langues concerne des lieux, qui semblent s'organiser de façon dichotomique, et par conséquent reposer sur des thèses antinomiques. Ces couples sont bien répertoriés depuis l'antiquité. L'origine est :

01- divine	ou	02- humaine
03- conventionnelle	ou	04- naturelle
05- progressive	ou	06- momentanée
07- arbitraire	ou	08- motivée
09- contingente	ou	10- nécessaire

On peut envisager la structure de la discussion comme un puzzle logique : soit le choix d'un des termes d'une dichotomie, quelles sont les contraintes que ce choix impose pour celui des termes des autres dichotomies ? L'historien qui se penche sur la question ne peut manquer d'être étonné par la relative faiblesse des contraintes. Si l'origine divine répugne à la thèse conventionnaliste, la première n'est pas nécessairement contradictoire avec la thèse d'une origine humaine, et la thèse conventionnaliste peut s'accorder avec celle de la motivation.

¹ On a attribué à Henry « la manera de argumentacion antitética hegeliana » (Koerner, 1982, p. 313, *ibid.*, p. 290, p. 416, p. 496), ce renvoi est insoutenable et relève simplement de l'ignorance du kantisme de la part du commentateur (voir Auroux, 2004).

² Institué par un legs de Volney, le prix avait initialement pour but la constitution d'un alphabet permettant la transcription de toutes les langues ; à côté de la question originale, on introduit en 1838, la philologie comparée et il sera transformé en prix consacré à la philologie comparée. Voir registres conservés aux Archives de l'Institut sous les côtes 40 H*1 (séances 1821-1840), 40 H2 (*id.*, 1841-1878).

³ Je veux dire les critères d'acceptabilité au sein d'une communauté professionnelle.

⁴ On entend par là l'ensemble des pratiques marginalisées et explicitement refusés par les groupes de recherche officiels, et qui se caractérisent par des formes d'argumentation et/ou des objets déviants. Cf. Auroux & al. (eds.). 1985.

⁵ L'objet du prix induisait sans doute ce type de confusion.

⁶ cf. Regnaud, 1888a, p. 347, note 1 : « La Société de Linguistique de Paris a eu tort, à notre avis, de placer sur le même pied, en les excluant l'une et l'autre de son programme, la question de la langue universelle et celle de l'origine du langage. Il y a entre les deux toute la distance qui sépare l'utopie de la recherche des principes ; et les fondateurs de la Société peuvent être comparés à cet égard à des géomètres qui s'interdiraient tout à la fois de s'occuper de la quadrature du cercle et des axiomes qui servent de base à leur science ». Le philologue est encore plus explicite dans un article de la *Revue de Linguistique* (1888b, p. 175) : « ... non seulement ils [nos linguistes] limitent volontiers leur domaine aux racines à tiroirs et aux fameuses lettres sonnantes de MM. Brugmann et de Saussure, mais ils s'irritent contre ceux qui s'y sentent à l'étroit, ils les anathématisent et, pour un rien ils les mettraient hors de la science ».

Parmi les chemins de ce « labyrinthe », trois peuvent caractériser les solutions classiques du XVIII^e siècle: (1, 3, 6, 7, 9), (2, 5, 7, 10) et (1, 4, 5, 8, 10). On aurait là des doctrines clairement déterminées, la seconde étant par exemple soutenue par Beauzée dans l'article « langue » de l'Encyclopédie, la première étant assez caractéristique des empiristes, la troisième de Court de Gebelin (*Monde Primitif*, 1773-1780). Toutefois la seconde philosophie de Condillac correspond au chemin (1, 4, 3, 5, 8, 10), elle étonne parce qu'elle n'oppose ni le naturel, ni la motivation à la convention (Auroux 1983b). Au XIX^e siècle la motivation est généralement conçue comme un argument contre la convention. Dans les deux premiers tiers du siècle les thèses de type cratyléen (motivation des signes primitifs par imitation de l'objet) sont les plus répandues en France (Chervel, 1979) ; en Allemagne ce sont les thèses de type révélation ou innéisme (compatibles avec l'arbitraire) qui sont les plus fréquentes. La deuxième moitié du XIX^e siècle change notablement la problématique, en se faisant plus naturaliste (c'était la voie choisie par Court de Gébelin) :

à partir du jour où la science des langues fut devenue une des sciences de la vie, le problème des origines se trouva transporté sur son véritable terrain, sur le terrain de la conscience créatrice (Renan, 1858, p. 87).

L'étude des facultés humaines est censée apporter des éléments à la question ; l'innéisme étant le relais normal de la thèse de l'origine divine (la nature humaine est une création de Dieu). Jusqu'au développement de l'école française de linguistique, orchestré par l'activité autant institutionnelle que théorique de Bréal et qui se poursuivra dans le sociologisme de Meillet, les théories du XVIII^e siècle français (Condillac, Maupertuis, Rousseau, Condorcet, Turgot, Volney) et anglais (Smith, Reid, Steward) sont assez généralement repoussées :

ils commettaient une erreur en attribuant aux facultés réfléchies et à une combinaison, voulue de l'intelligence un produit spontané de cette force vive que recèlent les facultés humaines, qui n'est ni la convention, ni le calcul, qui produit son effet d'elle-même et par sa propre tension (Renan, 1858, p. 80).

Ce jugement est évidemment erroné, le seul auteur postulant une origine réfléchie du langage, est, à ma connaissance, Maupertuis. Turgot trouve grâce aux yeux de Renan pour avoir écrit dans sa critique de Maupertuis que « les langues ne sont pas l'ouvrage d'une raison présente à elle-même ». Cette incompréhension des thèses conventionnalistes (je veux dire du fait qu'elles n'impliquent pas nécessairement une intention), va de pair avec l'espoir de trouver une solution dans le développement de l'anthropologie physique.

Il est évident que la possibilité de choisir tant de chemins différents dans le labyrinthe de l'origine, correspond à la possibilité d'infléchir largement le contenu que l'on assigne à ses lieux. Par là le problème d'origine concerne évidemment autre chose que l'origine du langage : *son essence, son rapport à la pensée et à la société, l'organisation de la pensée humaine (empirisme vs rationalisme) sa différence à la pensée animale et plus généralement la position de l'humanité dans l'ensemble du règne animal*. Comme les auteurs font la plupart du temps un historique des solutions qui les ont précédées, la façon dont ils classent ces solutions indique assez clairement les points qui leur paraissent sensibles. Pour Regnaud¹ il importe encore de se préoccuper de la question de la révélation, de l'innéisme et de repousser le conventionnalisme ; mais la thèse de l'origine naturelle est celle qui requiert le plus d'attention. Elle est en quelque sorte la matrice de nombreuses théories modernes. En 1922, Jespersen présentera une classification qui ne concerne plus que cette seule thèse : *bow-wow theory* (onomatopées), *pooh-pooh theory* (interjection), *ding-dong theory* (résonance interne des impressions), *yo-he-ho theory* (émission à la suite d'efforts). La référence théologique a enfin disparu.

¹ En témoigne sa classification des auteurs dans le labyrinthe de l'origine :

I - Révélation divine : Süssmilch (1766), Hamann, Sicard (1799), de Bonald, de Maistre, Kruse (1827), Gengenigl (1846), Noiré (1879), Gerber (1884). *II Innéisme* : W. von Humboldt, Steintal (1851), Heyse (1856), Renan (1848, 1858), Kelle (1856), Wackernagel (1876). *III - Invention humaine (= Conventionnalisme)* : R. Simon (1678), Maupertuis (1748), Harris (1752), Condillac (1748-75), A. Smith (1759), Abbé Copineau (1774), Monboddo (1774-96), Hourwitz (1807), Charma (1846). *IV - Origine naturelle* : Lucrèce, Diodore de Sicile, Rousseau, Schleicher, Zaborowski (1879) ; *IVa - Interjections* : Chavée (1878), Müller, Gerber, Sayce ; *IVb - Onomatopée* : Platon, Liebniz, de Brosses (1765), Court de Gebelin (1774), Herder, Nodier (1808), Renan (1848), Farrar (1860), Burgraff (1863), Wedgwood (1866), Whitney (1870), Chavée (1878), Zaborowski (1879), Paul (1880), Egger (1881), Henry (1883).

3. Arguments et raison de l'interdit

À l'exception d'un souhait assumé d'autonomie de la linguistique par rapport à l'ethnologie et à l'anthropologie, nous n'avons pas de connaissances sur les motivations et les discussions qui ont amené la Société à adopter son fameux article II ; nous ne pouvons qu'en reconstruire le contexte. L'idée qu'on doive s'abstenir de recherches sur l'origine, parce qu'elles sont impossibles sous certaines conditions a une longue histoire, et n'est certainement pas une nouveauté de « la linguistique scientifique » du XIX^e siècle. Par ailleurs la pratique qui consiste à interdire certaines discussions dans une société dont les statuts officiels définissent l'objet, semble tout à fait courante dans le droit français. Déjà en 1840, la première Société de Linguistique de Paris, après avoir décrété « (...) s'occuper de la grammaire et de l'histoire des langues en général, et spécialement de la langue française », ajoutait que « (...) *toute lecture - et toute discussion étrangères à la science qui est [son but] (...) sont formellement interdites* [mes italiques, SA] » (cf. Auroux, 1983a, p. 247).

Afin de mieux cerner l'originalité des discussions contemporaines de l'article II et leur rôle essentiel dans le développement de la linguistique moderne, le plus simple est d'énumérer ces interdits et leurs raisons. En examinant les textes qui nous servent de corpus, on peut lister de la façon suivante les arguments utilisés pour rejeter la question de l'origine des langues :

Raisons de caractère social ou conjoncturel :

- **[A]** il faut éviter les discussions qui touchent à un dogme essentiel de la religion.
- **[B]** il faut s'abstenir des hypothèses générales qui touchent des questions obscures.
- **[C]** les recherches sur l'origine des langues n'ont aucune fécondité scientifique.
- Raisons de caractère « scientifique » :
- **[D]** le passage d'un état non-linguistique à un état linguistique de l'humanité est impossible.
- **[E]** Il n'y a aucun observable qui corresponde à l'origine des langues.
- **[Fa]** La première langue est irrémédiablement perdue.
- **[Fb]** La naissance d'une langue n'est pas un observable.
- **[Fc]** La naissance du langage n'est pas un fait.
- **[G]** Les racines correspondant aux différentes familles linguistiques, sont des faits primitifs au-delà desquels la science n'a pas le droit de remonter.

La religion judéo-chrétienne comme tous les mythes archaïques contient une cosmogonie. On peut rendre cette cosmogonie compatible avec plusieurs thèses sur l'origine des langues, en particulier le mono- et le poly-généisme¹. Il n'en demeure pas moins que les discussions sur l'origine ne sont pas bienvenues. Le type de restriction **A** n'offre aucun intérêt épistémologique ; il concerne également la biologie comme le notent certains auteurs (Renan, 1858, p. 6 ; La Calle, 1881, p. 343). Le débat est cependant crucial au XIX^e siècle. F. Schlegel avait publié ses fantasmagories, sur la perfection originelle du sanskrit (1808, p. 69), l'année de sa conversion au catholicisme. Il s'accroît dans la seconde moitié du siècle avec l'apparition du darwinisme. La lutte contre le créationnisme, qui va permettre l'un des progrès les plus décisifs de la pensée occidentale, a bien entendu des connotations politiques. La première encyclopédie française de linguistique (Jehan 1858) qui reprend les thèses de Bonald avec celles de Gêbelin est un tome de *l'Encyclopédie théologique* de l'Abbé Migne. Les animateurs² de la *Revue de linguistique et de philologie comparée* (1867) sont résolument - anti-créationnistes: « la transformation, qui bientôt sans doute passera à l'état de vérité indiscutable, livre seule une raison scientifique de l'origine du langage » (Hovelacque, 1869, p. 103). Même opinion chez Zaborowski, chargé du cours d'ethnographie à l'École d'Anthropologie de Broca et proche du naturalisme des disciples de Chavée :

l'école qui procède de Schleicher est seule capable d'aborder utilement le problème de l'origine du langage, seule en possession d'aboutir à une solution scientifique, à une explication positive, naturelle et compréhensive (1878, p. 41).

Les catholiques n'en ont pas moins un poids social important en France. En ouvrant le premier Congrès des Américanistes à Nancy (1875), L. Adam est obligé de préciser que la science (en admettant la séparation des espèces) est compatible avec la foi : il a en face de lui le Président de la Société *Foi et lumière*, qui en 1843 a consacré des séances à la question (cf.

¹ Dieu donne un langage à Adam ; mais après le déluge plusieurs naissent selon des processus naturels.

² Son fondateur le séminariste belge H. Chavée a perdu la foi en devenant linguiste.

P.-G. de Dumast, *Mémoire sur la question de l'Unité des Langues*, Paris, Maisonneuve). La Société de Linguistique évite des débats socialement embarrassants (Vendryes, 1955, p. 13) ; son attitude pouvait rassembler un large consensus.

La prescription **B**, typiquement positiviste, se retrouve à peu près à toutes les époques. Au XVIII^e siècle, on rencontre souvent l'idée qu'il faut se borner à la recherche des *étymologies proches*, c'est-à-dire éviter de les chercher dans des langues très anciennes (cf. Turgot, art. « étymologie » de l'*Encyclopédie*). Ce type de limitation n'est pas nécessairement très fécond. Il conduit E. Burnouf à refuser d'envisager la cause de la proximité morphologique des langues indo-européennes et donc d'adopter l'hypothèse génétique au moment où se développe la grammaire comparée :

d'où vient ce rapport ? Quels événements l'expliquent ? À quelle époque faut-il en placer la date ? Ce sont là des questions auxquelles la philologie seule n'a ni le choix ni le pouvoir de répondre (*Journal Asiatique* V, 124, 1825).

C, peut en être considérée comme une variante. C'est une constatation sur le développement de la recherche que faisait déjà, Cl. Fauchet, face à la multiplication des thèses sur le caractère originaire de telle ou telle langue nationale¹. F. Thurot retrouve l'argument en 1796, dans l'étude historique qui constitue sa préface à la traduction de l'*Hermès* de J. Harris (1752) : « tout ce qu'ont pu faire les savants les plus laborieux, après des recherches immenses, a été d'établir une série de conjectures plus ou moins probables » (1970, p. 67). On a vu J. Ellis le reprendre ; Withney (1870, p. 84) fera de même. Ce constat d'un échec n'a évidemment aucune valeur épistémologique s'il n'est pas complété par l'idée que ses causes sont insurmontables.

L'impossibilité d'un passage d'un état non-linguistique à un état linguistique (**D**) correspond, par exemple, à l'argumentation de Beauzée dans l'article « Langue » de l'*Encyclopédie* ; il se fonde sur une *aporie* des recherches sur l'origine : l'origine de la société suppose le langage, celle du langage suppose la société. L'*argument de la circularité* est emprunté à Rousseau (*Discours sur l'origine de l'inégalité*), mais on le trouvait déjà chez Platon (*Cratyle*, 438a-b). Beauzée se sert de l'argument pour soutenir que le langage a été donné d'un seul coup à l'homme par Dieu ; Rousseau pour avancer sa thèse de l'origine instinctive (la passion). Il sera décisif dans le contexte de l'épigénétisme français des années soixante du XX^e siècle.

Au XIX^e siècle les théories « momentanéistes », qui opposent irrémédiablement l'homme à l'animal, transforment l'argument. Elles ont eu de nombreux partisans L'idée de base semble provenir de Humboldt et de la conception de la langue comme une *totalité*². Steinthal (1851), qui est à bien des égards son disciple, la reprendra et, aussi, Renan³ :

C'est (...) un rêve d'imaginer un premier état où l'homme ne parla pas, suivi d'un autre état où il conquiert l'usage de la parole. L'homme est naturellement parlant (...) (1858, p. 90).

Toutefois le contexte est bien différent. Celui dans lequel travaille Humboldt peut s'accorder avec l'*épigénétisme* moderne de l'époque structuraliste (voir *infra*). La solution de Renan est que « chaque famille d'idiomes est (...) sortie du génie de chaque race, sans effort comme sans tâtonnement » (*ibid.*, p. 94). La constitution génétique de la race (que défendra Chavée en 1878), prend la place de l'origine divine. Le darwinisme rend difficile la manipulation de l'argument (et du coup occulte la force de l'idée humboldtienne de la langue comme *totalité*), comme ses solutions : puisque l'homme n'est pas dans la nature un phénomène primitif, puisque l'on n'est pas disposé à admettre qu'un quelconque animal en dehors de l'homme possède le langage, il y a nécessairement une naissance du langage à partir d'un état non-linguistique. En soi, l'origine du langage n'est pas la quadrature du cercle. Henry prend acte du darwinisme en posant la thèse de son antinomie.

L'argument de l'inobservabilité (**F**) est incontestablement le plus fréquent ; Zaborowski (1878, p. 16) note « la répugnance des naturalistes à se mêler d'une question soustraite dans l'opinion générale à la méthode d'observation positive, sans base matérielle et sans lien avec

¹ « Quant à vouloir rechercher quelle fut la langue de nos premiers pères, je pense que ce serait une trop pénible, et encore plus vaine curiosité » (*Recueil de l'origine de la langue et poésie française [...]*. Paris. 1581. p. 3).

² « Es kann auch die Sprache nicht anders als auf einmal entstehen (...), sie muss in jedem Augenblick ihres Daseyns dasjenige besitzen, was sie zu einem Ganzen macht » ([1820] 1963, p. 2).

³ « [...] le désaccord entre les vues de M. Steinthal et les miennes est fort subtil » (1858, p. 37).

la réalité des faits sensibles ». Le problème est toutefois beaucoup plus complexe, et correspond à des thèses très différentes.

Sous sa forme **Fa** l'argument est bien formulé par Thurot ([1796], 1970, p. 67) : il n'y a pas de « monuments authentiques », d'où sa conclusion « il est en quelque sorte impossible de répandre une lumière satisfaisante sur l'origine de l'art de la parole » (*ibid.*, p. 68). Les auteurs du XIX^e siècle l'admettent généralement, comme Renan: « Les langues primitives ont, il est vrai disparu pour la science avec l'état psychologique qu'elles représentaient, et personne n'est désormais tenté de se fatiguer à leur poursuite avec l'ancienne philologie » (1858, p. 69). La plupart des grammairiens philosophes du XVIII^e siècle ont refusé que nous puissions avoir connaissance du **fait** de la première langue. Pour eux, l'origine des langues est inaccessible à l'histoire ; d'où leur orientation génétique, qui dégage un modèle possible sans asserter qu'il est réel en tant que fait historique¹. L'argumentation toutefois a été refusée par Court de Gébelin, sur des fondements qui vont constituer l'une des bases de la méthode comparative : *les langues actuelles ne sont que la première langue devenue selon certaines lois*. Les véritables monuments de la langue primitive sont les langues actuelles, à partir desquelles on peut reconstruire selon des procédures codifiées (« scientifiques ») leur état primitif. La reconstruction qui se situe dans la dimension chronologique de l'histoire, est une démarche totalement différente de l'induction qui, à partir d'observables qui ne sont pas ordonnés chronologiquement, autorise des hypothèses génétiques.

L'uniformitarisme permet de justifier l'induction à partir des phénomènes actuels. H. Paul consacre le chap. 9 des *Principien der Sprachgeschichte*, à l'*Urschöpfung* (la création originaire), sans supposer d'autres processus que ceux qu'il décrit dans son livre, et que les néogrammairiens ont posé à la base du fonctionnement des langues². L'argument peut se retourner ; si les processus par lesquels le langage se crée ne diffèrent en rien de ceux par lesquels il fonctionne actuellement, aucun des deux types de processus ne nous apprendra rien sur l'autre (Saussure, Engler I, p. 30): pourquoi le langage serait-il apparu ? Renan était lui aussi partisan d'une induction à partir du présent :

l'expérimentation directe (...) [de l'âge primitif] nous est impossible ; mais l'induction, en s'exerçant sur le présent, peut nous faire remonter à l'état spontané, dont les époques réfléchies ne sont que l'épanouissement (1858, p. 66).

Ses conceptions étaient très en deçà de l'uniformitarisme. Il suppose que des langues « parlées par des peuples presque voués à l'immobilité », « subsistent encore comme des témoins non pas (...) de la langue primitive, ni même d'une langue primitive, mais des procédés primitifs » (*ibid.*, p. 70). En quelque sorte des *éléments fossiles dans des langues archaïques*. La stratégie, à première vue, n'avait rien d'absurde : c'est la vie des iroquois et des algonquins, qui a permis à Lafitau (*Mœurs des Sauvages Américains, comparées aux mœurs des anciens tems*, 1724), de faire des hypothèses sur le passé lointain des grecs, et c'est l'utilisation d'armes en pierre par des tribus indigènes qui mit fin au mythe des « pierres de feu »³ pour ouvrir la voie à l'étude de la taille des silex (Daniel, 1983, pp. 36-38). Elle suppose qu'il y a un sens à considérer que certaines langues sont plus anciennes que d'autres, ce qu'à première vue le bon sens accorde facilement.

Que la naissance d'une langue ne soit pas un observable (argument **Fb**) est l'une des thèses des grammairiens philosophes du XVIII^e siècle : « Si on pouvait observer une langue dans ses progrès successifs, on verrait les règles s'établir peu à peu. Cela est impossible » (Condillac, *Gram.*, 1775, I.IX). L'argument repose chez eux moins sur la continuité des processus de changement linguistique (voir l'argument suivant), que sur l'impossibilité pratique d'une observation : ne restent de l'histoire des langues que des *monuments*, le processus par lequel des structures linguistiques sont nées est irrémédiablement perdu.

Si la naissance d'une langue n'est pas un observable⁴, elle n'a jamais été observée. On peut transformer cette dernière proposition en constatation empirique et en tirer des

¹ Cf. Rousseau, *Disc. Or. Inég.* « Commençons donc par écarter tous les faits... » ; Condillac, *Grammaire*, 1775, « Quand je parle d'une première langue, je ne prétends pas établir que les hommes l'ont faite, je pense seulement qu'ils l'ont pu faire » ; cf. également de Brosses, *Mécanique des Langues*, 1765, § 65.

² « [...] selbst in den Sprachen der europäischen Kulturvölker die Schöpfung neuen Stoffes niemals ganz aufgehört hat » ([1880] 1975, p. 175).

³ Les outils paléolithiques étaient considérés comme des espèces de météorites tombées du ciel.

⁴ Saussure : « Le moment de la genèse n'est pas lui-même saisissable ; on ne le voit pas » ; note de Riedlinger sur le second cours ; Engler I, p. 30.

conclusions sur l'origine des langues, même si l'assertion qu'on n'a pas observé quelque chose ne prouve logiquement rien. C'est ce que fait Humboldt ([1820], 1963, p. 2), pour soutenir son momentanéisme ; on trouve la même démarche chez Renan (1858, p. 1030). Henry (1896, p. 10) et Saussure¹, avanceront également que la naissance d'une langue n'a jamais été observée, mais cela signifiera pour eux qu'elle n'est pas un observable, parce qu'elle n'est pas un fait ayant eu lieu (voir *infra*), autrement dit, *qu'une langue ne naît pas*. Cette thèse est également une conséquence de la thèse selon laquelle le passage du linguistique au non-linguistique est impossible. Elle prend dans le dernier tiers du XIXe siècle un tout autre sens. Voici comment elle est présentée par A. de la Calle :

[...] le langage humain n'a point d'origine. Il prend sa source communément, avec le langage animal, dans les étapes les plus infimes de l'évolution ; dans les organismes les plus élémentaires. Il n'est en définitive que le résultat d'un développement supérieur de l'expression (1881, p. 320).

On ne saurait scientifiquement assigner un point quelconque ou une limitation absolue et réelle, dans l'ordre chronologique à l'origine du langage humain. Donc [...] il est aussi inutile de spéculer sur l'origine du langage, que de spéculer sur l'origine de l'humanité (1881, p. 343).

À première vue le raisonnement ressemble aux sophismes de Zénon d'Elée sur le mouvement ou aux arguties sur la calvitie (à partir de quand devient-on chauve ?). Sa valeur s'éclaire dans le contexte de la querelle qui opposa P. Meyer à G. Ascoli (cf. Auroux 1979b, p. 168) ; nous y reviendrons. L'idée que l'origine du langage corresponde à un type de phénomène qui n'est pas assignable en un point spatio-temporel, est également utilisée par Saussure, contre les recherches d'origine : « La question de l'origine du langage n'existe même pas. Ce serait étudier où commence le Rhône, localement et temporellement, question - absolument puérile » (Engler 41, p. 30, notes de L. Gautier sur le second cours).

L'argument **G** est le seul élément véritablement nouveau qu'aient apporté les débuts du comparatisme ; avec la notion de **racine**, il concerne, en effet, l'un des éléments clés de cette nouvelle discipline. Il remonte au moins à Bopp (voir *supra*), et semble accepté par la majorité des linguistes, même ceux qui écrivent sur l'origine des langues ; il est à première vue compatible **soit** avec une étude génétique (voir *infra*), **soit** avec la thèse d'une détermination biologique des racines propres à chaque groupe de langues (hypothèse de type *racialiste*, cf. Renan, 1858 ; Chavée, 1878), **soit** avec une thèse créationniste. Renan n'hésite pas à proclamer le caractère *anti-scientifique* de toute tentative d'analyse des racines :

C'est là une recherche qui est comme interdite à la science, parce que l'objet qu'il s'agit d'analyser ne laisse aucune prise à nos moyens d'attaque. Les racines des langues se montrent à nous, non pas comme des unités absolues, mais comme des faits constitués, au-delà desquels la philologie ne doit pas songer à remonter, sans encourir les mêmes reproches que l'alchimie (1858, p. 111, note 1).

Bopp ne s'est pas justifié ; Renan nous ramène à l'inobservabilité. L'exclusion de l'origine des racines des familles linguistiques hors du champ scientifique est pourtant une démarche constante chez les comparatistes spécialistes de l'indo-européen. À lire Renan, il semble que cette exclusion ne soit pas incompatible avec l'étude de l'origine du langage.

4. La force des arguments

L'interdit posé par la Société de Linguistique ne manquait donc pas d'arguments, tant des bons que des mauvais ou des faibles. À première vue, l'argument de l'*inobservabilité* ne paraît pas très probant. Tout le comparatisme a été orienté vers *la reconstruction de la structure des proto-langues*, qui ne correspondent évidemment pas à des faits attestés, donc observables. Dans le cas du passage des formes historiquement attestées aux racines de chaque famille linguistique l'argument ne vaudrait pas, tandis qu'il s'appliquerait dans le cas du passage de ces racines à des formes qui leur seraient antérieures. Il faut donc justifier cette situation par une différence de statut.

Lorsque Saussure dans le fameux *Mémoire* de 1879 postule les coefficients sonantiques, c'est pour expliquer l'existence d'anomalies dans la réalisation de lois phonétiques (apophonie e/o avec vocalisation de la sonante² suivante, par exemple) bien connues des langues

¹ Engler IV, p. 7, « Jamais on n'a signalé sur le globe la naissance d'une langue nouvelle ».

² Une « sonante » peut se définir ainsi : une consonne (par exemple un *m*) qui se réalise de façon vocalique (échappement libre de l'air phonatoire).

européennes. L'entité reconstruite, permet d'expliquer que par exemple le /e/ du grec ne disparaisse pas (*tithemi/theto*). Autrement dit l'entité reconstruite a des *effets observables*, à la limite on pourrait dire qu'elle n'est rien d'autre que la somme de ses effets. Saussure ira jusqu'à soutenir que la reconstruction n'a d'autre but que d'enregistrer les progrès de la linguistique (Engler III, p. 492) ; *c'est une convention qui nous permet de classer les phénomènes* : on est loin de Schleicher qui, dans le second tiers du siècle, utilisait les entités qu'il reconstruisait pour écrire une fable en proto-indoeuropéen, leur donnant par là valeur de faits historiques réels (Schleicher, 1868). Il est clair que la reconstruction n'a de sens, que si l'entité est la cause *prochaine* de ses effets. Henry (1896, p. 26) peut avancer l'argument de l'allongement considérable de la période anté-historique que la préhistoire moderne¹ attribue à l'homme, contre les recherches sur l'origine. Mais il me semble que l'essentiel réside dans une analyse plus fine de l'argument **G**.

Le programme comparatiste concerne l'apparement par familles. Si on va au-delà des racines qui doivent être différentes pour chaque proto-langue², **ou** on postule leur pluralité d'origine (mais on n'a pas d'éléments pour poursuivre plus loin), **ou** on postule leur communauté d'origine. Dans ce dernier cas, il faut rechercher les *causes et les motivations des racines*, d'où la thèse de l'onomatopée, c'est-à-dire *la recherche d'éléments communs à toutes les familles*. On ne peut rechercher à la fois le commun et le différent. En 1799, lorsque Gyarmathi³ s'efforce de lister les analogies syntaxiques du hongrois, du lapon et du finnois pour prouver leur parenté, il rejette explicitement ce qui lui paraît relever de la grammaire universelle. Nous avons montré ailleurs comment la recherche de la langue primitive conduisait Court de Gebelin à obscurcir la reconnaissance de la famille austronésienne (Auroux & Boès, 1981, p. 47). L'apparement par famille suppose qu'on écarte d'*abord* ce qu'il peut y avoir d'universellement commun, comme quelque chose de parasite et d'inessentiel à la recherche en cours. Autrement dit, il faut postuler l'*arbitraire* des racines, écarter les faits qui témoigneraient du contraire⁴, parce qu'ils empêcheraient le programme comparatiste de fonctionner. M. Müller dans ses *Nouvelles leçons sur la science du langage* a parfaitement vu l'enjeu de la question. Soit l'allemand /*feind*/ ennemi ; si on recherche l'origine onomatopéique du mot, on cherche dans toutes les langues quelque chose qui signifie à peu près la même chose et commence par /*f*/, ou on se permet n'importe quel changement de la consonne, et on peut dire n'importe quoi. Le second cas est évidemment sans intérêt. Dans le premier, on viole la *loi de Grimm* (1822), principe de base de l'identification parmi les langues indo-européennes de la famille des langues germaniques⁵. « La théorie de l'onomatopée (...) serait la mort de toute étymologie scientifique » (t.f., 1867, I., p. 114). *Il faut considérer le rejet de la recherche de l'origine (celle qui figure au-delà des racines) comme un postulat qui figure parmi les conditions de possibilité du programme comparatiste*.

Pour défendre le programme, qui se justifie par ses résultats, on peut construire des hypothèses auxiliaires destinées à protéger l'arbitraire. La motivation actuelle qu'on croit reconnaître dans un groupe de mots ne préjuge en rien de la communauté de leur origine, et donc de la valeur causale de la prétendue onomatopée. La ressemblance peut venir d'une rencontre homonymique⁶. À l'inverse les reconstructions peuvent conduire à postuler

¹ La préhistoire a fait de considérables progrès au XIXe siècle, notamment par la présentation de la théorie des trois âges *Pierre, Bronze, Fer*, (cf. C.-J. Thomson, 1836. *Ledetraad till Nordisk Old Kindighed*. Copenhague). Ces progrès conduisent à reculer constamment l'apparition de l'homme, qui était encore fixée aux environs de 7 000 ans au début du siècle ; voir Daniel (G.). 1981. *A Short History of Archeology*. New York : Thames and Hudson (voir dans l'édition paperback, 1983, p. 48-55) ; Laming-Emperaire (A.). 1964. *Origines de l'archéologie préhistorique en France. Des superstitions médiévales à la découverte de l'homme fossile*. Paris : A. et J. Picard Cie.

² Dans le cas contraire, la proto-langue ne permettrait pas de définir une famille.

³ Gyarmathi (S.). 1799. *Affinitas Lingvae Hvngaricae cvm Lingvis Fnnicae Originis Grammaticae Demonstrata nec non Vocabularia Dialectorum Tataricarvm et Slavicarvm cvm Hvngarica comparata*. Göttingen [Trad. Angl., avec Introd. par V. Hanzeli, 1983. Amsterdam : John Benjamins].

⁴ On trouvera dans Auroux & Boès, l.c. un certain nombre de citations, tout à fait probantes, comme celle-ci de Lanjuinais (1816) : « De telles affinités ne serviraient ni à l'intelligence d'autres langues, ni à constater la filiation des peuples ».

⁵ Le *p* indo-européen y est remplacé par un *f* ; /*p*/ > /*f*/ : angl. /*fiend*/ ; got. /*fijant*/, mais skt. /*pîyant*/, participe de /*pîy*/, haïr.

⁶ L'onomatopée est un *fait naturel*, donc universel ; la rencontre homonymique un *fait historique*, donc contingent.

l'homonymie de certaines racines comme le fait Sayce pour /dhà/, « sucer » et /dhà/ « placer » (cf. Chervel, 1979, p. 17). *L'affirmation de l'arbitraire linguistique et l'abandon de toute recherche concernant les motivations des racines (donc l'origine du langage) sont un seul et même acte épistémologique qui rend possible le comparatisme.* Quand on veut contourner cet abandon on n'aboutit à rien. Regnaud s'efforce de montrer que toutes les racines peuvent dériver phonétiquement les unes des autres¹. En admettant que ses arguments soient valables, ils ne prouvent rien d'autre que la possibilité de la naissance des racines. Cela ne fait pas avancer d'un pas le comparatisme ou l'histoire des langues.

La thèse de l'arbitraire des racines semble bien être la seule raison du refus de considérer l'origine des langues dont on puisse reconstruire l'articulation logique. Elle est si profondément ancrée chez la plupart des linguistes que non seulement ils finissent par admettre l'arbitraire des onomatopées², mais qu'il arrive à Pictet de soutenir que s'il a pu exister des formes originaires et naturelles (/pa/ et /ma/) pour désigner le père et la mère, ces formes sont différentes des racines /pâ/ et /mâ/. Cette remarque, qui scandalise Zaborowsky (1878, p. 164), revient à reconnaître à la racine³ un statut de *pure entité théorique*. La reconstruction des formes cesse de présupposer leur réalité historique. Henry soutient qu'un /*⁴curratis/ de la première conjugaison latine est le même mot que fr. /courrez/ « bien qu'il n'ait pas été créé par les latins » (1896, p. 21). La notion de racine recouvre en fin de compte la façon dont on se représente abstraitement la parenté d'une famille de langues (Henry, 1896, p. 13 : « [les racines sont des] catégories abstraites qui servent [au linguiste] à classer les mots »). Telle n'est évidemment pas la conception que les premiers comparatistes avaient de la racine, souvent prise pour un véritable mot (Schleicher prétendait écrire en proto-indo-européen) et - objet de vives discussions. Elle correspond à peu près aux réflexions de Saussure signalées plus haut. Pourquoi ce cheminement vers l'abstraction des racines, qui ne semble pas propre au néo-kantisme de Henry ? Il faut y voir un déplacement épistémologique de la thèse de Bopp concernant leur arbitraire : comme elles sont des *conventions* de la science linguistique, on ôte à cet « arbitraire » le statut discutable de fait linguistique historique, on n'a pas à s'occuper de leur éventuelle motivation onomatopéique, leur seule motivation est celle de la démarche du linguiste⁵.

5. L'impasse génétique

Là où le réel du *fait historique* est en défaut, la voie génétique, que favorise l'hypothèse uniformitariste (cf. Paul [1880], 1975, pp. 174-188 ; 1907) peut sembler la seule possible. Soit un ensemble de faits, c'est-à-dire de références de propositions singulières vraies, dont la vérité peut s'argumenter à partir de données accessibles à tout chercheur. Soient les points

¹ Regnaud s'efforce de montrer qu'un grand nombre de racines sont à l'état de variantes phonétiques évidentes les unes par rapport aux autres, interprétant ainsi le concept de *Neben Formen* des comparatistes allemands. Dès lors « toutes les racines peuvent se rattacher phonétiquement les unes aux autres, ou en d'autres termes, (...) elles peuvent descendre par voie d'évolution phonétique d'un seul type primitif » (1888, p. 178). Cela conduit à rejeter le caractère obligatoire des lois phonétiques des néogrammairiens, au profit d'un modèle par bifurcation (*ibid.*, pp. 182-183) et syncrétisme des formes ainsi produites (*ibid.*, p. 185). La différenciation des formes (= *doublets phonétiques*) précède celle des sons que l'auteur explique par l'instinct qui répugne à l'existence de synonymes.

² Cf. Bréal (M.). 1904, p. 401 ; c'est une thèse que Regnaud soutient également (114 s.), en montrant en particulier que la série des « onomatopées » de même signification (sk. /kokilla/, gr. /kokkuk/, lat. /cuculus/, fr. /coucou/), ne correspond pas à une dérivation phonétique (présence du /i/ en sanskrit). Les onomatopées sont *a posteriori* (cf. Egger, 1881, p. 257 note 1 : « l'instinct populaire n'a-t-il pas souvent transformé des mots à signification arbitraire de façon à les rapprocher du son propre de la chose qu'ils signifiaient ? »).

³ La notion au reste perd de son importance avec l'uniformitarisme néogrammairien : le manuel de Paul (1880) n'y consacre pas même un chapitre.

⁴ L'étoile * pour noter les formes reconstruites (c'est-à-dire non attestées) apparaît vers le milieu du XIXe siècle, en 1843 selon Koerner (1978, p. 211).

⁵ Situation que Hjelmslev formulera parfaitement : « Lors de la comparaison des langues indo-européennes, nous avons désigné nos formules communes comme indo-européennes. Les romanistes enregistrent les leurs comme appartenant au roman commun, à l'ancien roman ou au latin vulgaire. Pour des raisons pratiques ils choisissent, pour représenter chaque formule, un signe qui rappelle l'élément d'expression latin correspondant, mais ces dénominations sont aussi arbitraires que pour les formules indo-européennes, et il y a de nombreux cas où il serait impossible de les accorder avec le éléments d'expression du latin connus dans les mêmes mots » (1965, pp. 44-45).

d'une ligne temporelle universelle, c'est-à-dire ordonnés de façon stricte. Il y a une seule projection des faits sur la ligne, elle constitue leur datation ; l'ensemble des couples ordonnés ainsi obtenus permet ce qu'on appelle de la façon la plus plate un *agencement chronologique*. Supposons maintenant un ensemble de propriétés assignables au langage (par exemple : avoir des verbes, être holophrasique, etc..). J'appellerai *genèse* toute structure d'ordre assignée à un tel ensemble de propriétés ; elle suppose par conséquent un état initial. Par définition une genèse n'est pas une histoire, non parce qu'elle n'a pas de valeur empirique, non parce qu'elle ne pourrait pas être connectée avec des faits, mais parce que les faits qu'on peut relier aux différents états n'obéissent pas nécessairement à l'ordre strict d'un agencement chronologique universel. Soit par exemple une séquence génétique. Je puis très bien l'argumenter à l'aide du développement du langage enfantin, ou de celui de plusieurs groupes de langues, par des faits que la chronologie universelle ordonnera d'une façon que la genèse ne respectera pas. Contrairement à l'histoire, la genèse ne se préoccupe pas de l'ordre unique des faits, mais cherche à soutenir que des faits d'un certain type dans une ligne chronologique donnée sont toujours précédés par des faits d'un autre type.

Si on remonte à la Renaissance c'est d'abord une approche historique qui a prévalu dans la question de l'origine des langues, par l'assignation (erronée) d'une langue mère universelle réelle (par exemple l'hébreu ; voir Dubois 1970). Le débat sur l'origine des langues possède au XVIIIe siècle une orientation essentiellement génétique (voir Schreyer, 1985). Il change toutefois de nature avec de Brosses (1765) et Court de Gébelin (1773-81). Ce dernier entend inférer la structure et le vocabulaire de la « langue primitive » à partir de l'histoire des langues ; il proposera un « dictionnaire des racines primitives ». Autrement dit, il affirme des faits non attestés, mais reconstruits ; une pratique que le comparatisme, avec d'autres techniques, raffinerait au siècle suivant, notamment dans le cas des *proto-langues* (*Ursprache*). Le développement du comparatisme aura pour conséquence de ramener le problème vers l'histoire ; on en a vu les impasses. Certaines tentatives de la fin du siècle (Bréal, 1900 ; Brugmann, 1897) seront assez clairement génétiques, comme celle de Jespersen (1922). Les distinctions entre les deux orientations ne sont toutefois pas absolues et ouvrent le double problème de savoir à quelles conditions la genèse permet des inférences historiques, et l'histoire des inférences génétiques. Il me semble que ces conditions sont respectivement les suivantes : les circonstances du développement linguistique chez l'enfant sont : (i) les mêmes dans tous les contextes linguistiques ; (ii) les mêmes que celles du développement du langage. Cela implique que les étapes du développement de toutes les langues sont les mêmes, et qu'elles correspondent à un ordre linéaire (non réversible), avec un *état initial*.

La première démarche envisageable est celle qui consiste à construire un modèle génétique sur des phénomènes observables, et à le projeter sur l'histoire des langues. L'apprentissage du langage par les enfants correspond par excellence aux phénomènes - permettant ce type d'inférence. Le dernier tiers du XIXe siècle y consacre de nombreuses études expérimentales, et les néogrammairiens en tirent de nombreux exemples. Mais le développement du langage enfantin est une procédure d'acquisition, qui est en elle-même un objet de recherche avec ses problèmes propres (langue maternelle, bilinguisme, aspect inconscient du processus, Henry 1896, pp. 48-78). Comme objet d'étude empirique l'acquisition suppose toujours déjà là « l'unique objet » de la linguistique, un « langage conventionnel, [...] que nul ne parle ni ne comprend sans qu'on le lui ait enseigné » (Henry, 1896, p. 48). On ne peut en tirer de conclusions sur l'origine des langues : l'une des stipulations que nous avons décrites plus haut (en résumé, l'ontogénèse répète la phylogénèse) n'est en effet pas réalisée.

On pourrait envisager d'inférer un modèle plausible d'évolution linguistique, à partir de faits historiques dispersés dans des langues différentes. L'organicisme¹ dominant du XIXe siècle établit ainsi des « stades » dans la « vie » du langage. Ces stades vont jusqu'à sept chez Curtius (1870 ; cf. Pott, 1884, p. 8), mais ils se présentent le plus souvent sous la forme de la triade schleicherienne : *monosyllabisme*, *agglutination*, *flexion*. Le génétisme typique du XVIIIe siècle, et la théorie des stades linguistiques (c'est-à-dire l'interprétation chronologique

¹ « Ce qui distingue la linguistique moderne des spéculations du passé sur l'origine des langues, c'est que cette science toute contemporaine, a reconnu et proclamé qu'il existait une *vie du langage* ; que chaque langue passait inévitablement par telles ou telles périodes biologiques ; en d'autres termes, qu'elle partageait le sort commun de tous les organismes » (Hovelacque & Vinson, 1878, p. 4).

d'une typologie) possèdent la même structure logique, dont nous avons énoncé plus haut les contraintes. Pour qu'un schéma génétique - disons (A > B > C > D) - soit historiquement valable, il faut que n'importe quel fait relevant de A, qui se situe sur la même ligne évolutive qu'un fait relevant de B, ait une date antérieure à ce dernier ; si ce n'est pas le cas, il faut qu'avant lui il y ait achèvement d'une première série. Dans ces conditions aucun fait observé ne peut à lui seul témoigner de l'antériorité historique d'un type de forme. C'est un argument que Henry avait repéré ; il s'efforce de montrer que le monosyllabisme peut suivre la flexion (naissance des articles), et préparer une nouvelle phase d'agglutination ; dès lors :

Loin que le monosyllabisme théorique du chinois actuel nous représente l'état primitif du langage, il recouvre peut-être vingt couches sous-jacentes et à jamais inaccessibles d'évolution linguistique à trois étages chacune (1896, p. 17).

Empiriquement l'argument¹ sur le chinois n'a peut-être pas des fondements très solides ; ce point ne nous intéresse pas ici. Abstraitement, il revient à soutenir que les types généraux qu'on pourrait attribuer à des classes de langues ne peuvent et réordonnés du plus « primitif » au plus « évolué » : *on ne dispose pas d'état initial*. Contrairement à ce que pouvait penser Renan, quelque chose d'analogue au concept de « fossile », n'a pas de place dans les sciences du langage ; une stratigraphie linguistique est impossible.

La véritable valeur de l'argument est d'ordre logique, il porte aussi plus loin que l'origine du langage. Toute genèse ou toute suite de stades évolutifs suppose un ordre nécessaire, puisque immuable. Pour Henry, les faits linguistiques sont *historiques* en ceci qu'ils sont contingents (ce qui n'exclut pas la régularité) : « le langage significatif est un heureux et sublime accident » (1896, p. 40). Ou comme le disait Saussure en 1891 dans sa première - conférence à Genève : « la langue est *histoire* (...) elle se compose de *faits*, et non de *lois*, (...) tout ce qui semble organique dans le langage est en réalité *contingent* et complètement accidentel » (Engler IV : 5). Dans le *Cours* (Engler III, p. 507, s.) l'argument lui servira pour disqualifier la thèse de l'immutabilité des « racines » ou des traits grammaticaux propres à certaines familles de langues (cf. *supra*, les thèses de Renan et Steinthal). Cette conception de l'historicité des langues, étrangère à Schlegel comme à Schleicher, bloque la possibilité de toute inférence génétique sur leur origine. **Ou** il y a des lois² d'évolution linguistique (biologiques ou autres), et certains faits permettent d'inférer quelque chose sur l'origine des langues ; **ou** les phénomènes linguistiques sont des entités historiquement contingentes, et l'absence de *faits* datant de leur origine est indépassable.

6. Les langues n'ont pas de date

Dans ce contexte le continuisme que l'on a vu à l'œuvre dans l'argument **Fc**, possède des conséquences paradoxales. On sait (Auroux, 1979b) qu'en rendant compte du tome III de *l'Archivio glottologico* (1874) de G. Ascoli, P. Meyer dans la revue *Romania* (1875), critiquait l'idée qu'il puisse y avoir discontinuité réelle entre les dialectes néo-latins : *on ne peut pas*

¹ De façon rhétorique, sa structure peut servir également à libérer certaines recherches du poids de l'interdit : puisqu'on ne s'occupe pas de l'état initial inaccessible, on peut avancer sur des cas plus récents. Lorsque dans un chapitre de ses *Essais de Sémantique*, M. Bréal se pose la question canonique du XVIIIe siècle de la formation des catégories grammaticales, il note : « La question ne se confond pas avec le problème de l'origine du langage, car il y a des langues qui, encore aujourd'hui, ne distinguent point de catégories grammaticales, et il se peut fort bien que nos idiomes aient passé par un état semblable [mes italiques, SA]. Il s'agit donc de faits relativement récents, pour lesquels l'observation ne doit pas être, *a priori* déclarée impossible » (1897, p. 185).

² Les « lois phonétiques » (par exemple, la loi de Grimm) postulent : i) que les langues divergent (hypothèse génétique d'un ancêtre commun) ; ii) que ces divergences affectent de façon discontinue la structure phonétique qui va définir la nouvelle langue ou la nouvelles famille de langues (par exemple, les langues germaniques dans le cas de la loi de Grimm). La formulation théorique de ce dernier point provient des néogrammairiens qui assertent que le changement d'une valeur phonétique en une autre affecte *nécessairement et simultanément toutes les occurrences* de cette valeur dans la langue considérée. La dialectologie (voir plus loin) et le développement des thèses continuistes remettent largement en question cette conception de la loi phonétique : le *modèle diffusionniste* (ou, en Allemagne, la « théorie des vagues ») ressurgit contre le *modèle génétique* et certains n'hésitent pas à penser qu'entre les deux éléments d'un changement phonétique peuvent prendre place quantité de variations historiques ou dialectales (c'est l'argumentation de Gilliéron, 1919). La théorie conventionaliste qui prévaut de Henry et Saussure jusqu'à Hjelmslev (voir plus loin) permet d'utiliser les formules correspondant aux « lois phonétiques » sans pâtir de ces arguments.

assigner de frontière dialectale, la variation est continue ; on peut seulement tracer des frontières pour certains traits (eux sont discontinus). *Langues et dialectes, sont des vues de l'esprit, des abstractions*. Le romaniste G. Paris projettera dans la chronologie, cette continuité d'abord affirmée dans la géographie : au sens propre il n'y a ni langue mère, ni langue fille, la prétendue fille n'est que la mère insensiblement devenue¹. C'est sur ces problèmes que Henry ouvre ses antinomies en refusant tout *réalisme* des langues.

Il en résulte *qu'une langue n'a point d'âge* (1896, p. 12) puisqu'il n'y a pas de point discontinu de datation, conséquence² que Saussure admettait dès 1891 (Engler IV, pp. 7-8) et qu'il reprendra dans son *Cours* (Engler III, pp. 483-485) :

Toutes les langues se parlant à la même époque sont de même âge en ce sens qu'elles remontent à un passé égal. Il n'est pas nécessaire de déterminer la longueur de ce passé. Si l'on veut c'est l'origine du langage, mais sans remonter aux périodes inaccessibles (Engler IV, p. 7).

On dispose de monuments *datés*, cela n'indique assurément rien sur l'âge de la langue qui s'y réalise. On peut parler de langues plus archaïques que d'autres en fonction de leur éloignement d'un modèle primitif, mais alors c'est « en dehors de toute question de date » (Engler III : 485). Avec la fin de l'organicisme, c'est le rapport de l'objet « langue » à la temporalité qui a changé : le temps comme l'espace est l'une des dimensions où se réalisent les phénomènes linguistiques, mais les langues ne sont pas des réalités que le changement affecterait dans un rapport défini et nécessaire à la temporalité, comme la naissance, la croissance et la mort affectent les vivants. Le véritable rapport à la temporalité est la contingence de l'histoire qui ne concerne que les faits. L'entité théorique construite comme objet de la linguistique avec les propriétés qui lui sont assignées dans cet acte de construction (Henry, 1896, p. 4, « entité abstraite » ; Saussure, Engler I, p. 26, « C'est le point de vue qui crée l'objet ») est incompatible avec une étude d'origine (Milner, 1978, pp. 26-27, p. 36). C'est à peu près ce que consignait Saussure dans une *Note* en vue de son cours :

ORIGINE DU LANGAGE : Inanité de la question pour qui prend une juste idée de ce qu'est un système sémiologique et de ses conditions de *vie*, avant de considérer ses conditions de *genèse* (Engler I, p. 30).

Dans ces conditions, on peut se demander en quoi ce qui va rester du projet comparatiste, l'apparemment génétique des langues, concerne encore l'histoire. C'est le structuralisme qui assumera le paradoxe : dans une langue où tout se tient (un système), les faits concernent la parole, pas la langue. Hjelmslev n'hésitera pas à donner une définition totalement abstraite de la parenté linguistique : « *la parenté génétique est une fonction reliant les langues : elle consiste dans le fait que chaque élément d'expression d'une langue est relié par une fonction à un élément d'expression d'une autre* » (1965, pp. 52-53). La seule façon d'expliquer la parenté est de concevoir que les langues en question ont une origine commune : « C'est la seule explication que l'on puisse imaginer et il faut donc l'admettre au même titre que la parenté génétique » (1965, p. 112). Mais tout ce que nous savons avec certitude sur la langue originelle, c'est ce que nous pouvons extraire des formules que nous avons établies. Du coup, nous n'avons aucune certitude concernant la date et la localisation de son emploi (elle a pu ne jamais être employée), ni la façon dont pouvait se présenter l'usage de cette langue. Sur ces questions on en est réduit aux hypothèses ; elles sont tentantes, « le champ est libre aux natures romantiques et riches d'imagination ». Mais :

[La linguistique] est une science avec des résultats sûrs et une méthode exacte ; [...] elle ne peut se soutenir en tant que telle que *si elle s'interdit le domaine hypothétique* [je souligne, SA] et s'en tient à ce que l'on sait avec certitude et ne prétend pas tirer de la méthode plus qu'elle ne peut donner (1966, p. 116).

7. La synthèse impossible

La linguistique comparée indo-européenne recueille les faits destinés à ses reconstructions dans des « monuments », ce sont des *langues écrites*. L'apparemment par famille disposait d'autres techniques développées entre le XVI^e et le XVIII^e siècles par les explorateurs et des

¹ Il est clair que cette conception affecte le concept même de « proto-langue » (*Ursprache*), comme langue originelle d'une famille. En 1908, Meillet en tirera les conséquences : la diversité dialectale existe dès le départ, on ne peut parler d'une langue *originelle*, mais d'une langue *commune*.

² L'idée que les langues ont le même âge se trouve déjà chez Whitney (1880, p. 340 : « Languages are all equally old »), mais dans un tout autre contexte, où elle sert à justifier que le fait que l'on n'ait pas observé le passage d'un état isolant à un état agglutinant ne prouve pas qu'il n'ait pas eu lieu.

érudits qui entendaient par là classer les *langues à tradition orale*, étudier les migrations¹ des peuples qui les parlent et leurs apparentements. Il s'agit de recueillir des vocabulaires² succincts qui peuvent rapidement se rassembler au cours d'une escale et se compiler. Le plus souvent on utilise une vingtaine de mots d'un vocabulaire de base canonique : comme le notait Reeland, pour découvrir la « *convenientia* » entre les langues³ le mieux que l'on puisse faire c'est de comparer les mots les plus communs comme sont ceux des nombres, du soleil, de la lune, de la terre, et d'autres semblables ; liste à laquelle on ajoute souvent les parties du corps et les pronoms personnels. De bons résultats ont été obtenus sur la famille austronésienne dont on parvient, dès le commencement du XVIII^e siècle, à identifier un certain nombre de langues, parallèlement à une aventure qui conduira à la découverte de l'Australie par les Anglais. De Brosses (*Histoire des navigations aux terres australes*, 1756), Court de Gébelin (voir Auroux & Boès, 1981) et quantité d'autres comme W. von Humboldt et J. C. E. Buschmann y apporteront leur contribution. L'apparentement des langues orales par la méthode du vocabulaire de base permet une classification génétique ramifiée, si l'on admet que la ressemblance décroît avec la distance temporelle de séparation (Dumont d'Urville⁴, 1834). Tout au long du siècle les recherches amérindiennes procéderont par un recours massif aux vocabulaires, celles de Brinton (1887, 1891), qui pourtant préfère les rapprochements morphologiques, comme celles du Bureau of American Ethnology autour de Powell (1880) et de ses collaborateurs. *Géographie, linguistique et ethnographie*⁵ concourent harmonieusement. Le navigateur Dumont d'Urville ajoute l'*anthropologie physique* à la géographie, lorsqu'il propose les nominations de familles polynésiennes, micronésiennes et *mélanésiennes* (ceux qui les parlent ont la peau noire). En 1878, Pictet en étudiant le vocabulaire essaiera de situer le peuplement des « aryas » primitifs et de faire contribuer la linguistique à la paléontologie. L'étude de la préhistoire de l'humanité, comme celle des langues orales, suppose une *synthèse entre différentes disciplines*.

Une part importante du crédit accordé à la linguistique au XIX^e siècle tient au rôle que l'on entend lui faire jouer dans l'ensemble des disciplines qui s'intéressent à la nature et à l'histoire de l'humanité. Les succès du comparatisme indo-européen paraissent témoigner en ce sens, puisqu'ils dessinent peu à peu les traits d'une famille complexe. Les développements de

¹ Cf., parmi d'autres, Grotius (H.). 1642. *De origine gentium americanarum*. Le célèbre spécialiste du droit naturel montrait (en permutant *ad libido* des lettres sur un vocabulaire restreint) que le peuplement de l'Amérique provenait d'un groupe d'immigrants norvégiens.

² Muller (1984) fait remonter la discussion sur la pratique du vocabulaire de base à une critique des thèses de Grotius par Jan de Laet (*Notae ad Dissertationem Hugonis Grotii*. Amsterdam, 1643) ; on doit à Laet l'ajout des termes de parenté. Les ressemblances peuvent s'interpréter par l'emprunt (modèle diffusionniste) ou par une origine commune. La thèse génétique, cependant, cache les apparentements locaux, si on passe directement à l'origine des langues. Monboddo (1773-1792) procède comme son contemporain Court de Gébelin : « it follows of necessary consequence that they must be all dialects of the same parent-language » (vol. I, 1773, p. 605).

³ « [convenientia] optime possit ostendi vocibus communioribus, quales sunt nomina numerorum, solis, lunae, aquae, terrae, boni, mali et similia », « Dissertatio de linguis insularum quarumdam orientalium », *Dissertationum Miscellanearum pars tertis et ultima*, Trajecti ad Rhenum, ex Officina Gulielmi Broedelet, 1709, p. 145.

⁴ Entre deux langues, il établit une échelle des valeurs des ressemblances : 1 (mots parfaitement semblables), 0.8, 0.6, 0.4, 0.2, 0 (disparate absolue entre les deux mots) ; pour une langue donnée, on somme les degrés de ressemblance de chaque mot ; ce nombre divisé par celui des mots considérés donne la valeur de la ressemblance entre les deux langues. Du coup on peut construire des tableaux à double entrée qui permettent de placer les langues les unes par rapport aux autres (ces tableaux sont équivalents à des arbres généalogiques). Il semblerait que ce type de lexico-statistique soit né de travaux au sein de la Société de Géographie (créée en 1821, par Malte-Brun) et de la discussion autour d'un travail de Rafinesque qui répondait à un concours proposé par la Société sur l'origine de populations noires localisées en Chine (Métoz, 2005, pp. 187 et s.).

⁵ Lorsqu'en 1826, le géographe A. Balbi publie son *Atlas ethnographique du globe ou classification des peuples anciens et modernes d'après leurs langues*, le rôle des langues apparaît clairement dans le projet ethnographique. Les *Nouvelles Annales des Voyages de la Géographie et de l'Histoire* annoncent ainsi l'ouvrage dans un texte anonyme (probablement rédigé par Malte-Brun) dans leur numéro de juillet 1826 : « Ce n'est pas un abrégé corrigé de *Mithridate* que M. Balbi donne, ce sont des éléments nouveaux et raisonnés d'une vaste science sœur de la géographie et de l'histoire. M. Balbi sera reconnu pour tous les savants comme le nouveau fondateur, comme le *Linnaeus* de l'ethnographie, et plus particulièrement de l'*idiomatographie* ou statistique des langues » (pp. 132-133).

l'anthropologie physique, la découverte des localisations cérébrales par Broca (1861, 1865) renforcent la possibilité d'une liaison entre les différents domaines, allant, au reste dans le sens proposé par Darwin¹. Certains vont plus loin, Chavée n'hésite pas à relier les langues et les races (1862a) : à chaque langue correspond la structure cérébrale spécifique d'une race. À l'École d'anthropologie fondée par Broca, le titulaire de la chaire de linguistique ne renoncera jamais à lier linguistique et ethnographie, s'associant parfois à J. Vinson (Hovelacque & Vinson, 1878), aux tendances raciales plus affirmées.

C'est de Broca que viendra la critique la plus sévère contre ce genre d'entreprise dans une Conférence présentée en 1862 à la Société d'Anthropologie (« La linguistique et l'anthropologie », 1863), ce qui peut sembler paradoxal, venant du découvreur des localisations cérébrales². Sa première attaque vise la théorie raciale de Chavée, présentée devant la Société (Chavée, 1862b) selon laquelle le langage n'est pas une création arbitraire et le type de langage qu'il parle ne dépend pas de la volonté de l'individu : « c'est un produit en quelque sorte fatal de l'organisation propre à une race ». Les arguments de Broca sont massifs : si la thèse de Chavée était correcte, les caractères linguistiques devraient présenter la même permanence que les autres caractères physiologiques ; l'homme ne pourrait changer de langage qu'en faisant violence à sa nature ; « une race dépossédée de sa langue naturelle devrait manifester une certaine tendance à revenir ultérieurement sinon à son ancien idiome effacé et oublié, du moins à des formes, à des procédés grammaticaux analogues à ceux de sa langue primitive, comme on voit reparaître peu à peu, par la suite des générations, les types organiques naguère altérés par des circonstances accidentelles » (1863, pp. 286-287). L'évolution linguistique peut concorder avec les caractéristiques physiques dans le cas des isolats ; lorsque les communautés linguistiques sont en contact, l'évolution est très rapide (emprunts, substitutions de langues) ; la substitution linguistique, enfin, peut ne pas correspondre à l'éventuel mélange des races. « La linguistique, par conséquent, fournit à l'anthropologie des renseignements et non des arrêts, et elle doit intervenir, non à titre de juge, mais à titre de témoin » (1863, pp. 318-319).

À elle seule, peut-elle aller plus loin dans l'approche préhistorique ? L'évolution d'une langue est-elle comparable à ces courbes algébriques que l'on peut construire dans toute leur étendue lorsqu'on connaît quelques points ? Broca utilise les travaux de Dumont et ceux de Hales pour construire une « glottochronologie » permettant d'évaluer les temps de séparation entre des langues apparentées³. *L'évolution spontanée des langues est de 1% pour 1 500 ans* : « [...] l'altération spontanée ne peut défigurer une langue qu'au bout d'un très grand nombre de siècles » (1863, p. 309). Par conséquent, l'approche linguistique ne mène à rien, on devrait aller bien au-delà de ce qui est observable ; tout ce qu'il reste ce sont des conjectures :

[...] les conjectures (...) ne semblent pas des impossibilités. Mais quel usage en pourrions-nous faire ? Aucun. Prouver qu'une chose n'est pas impossible, c'est aller au-devant d'une objection ; c'est déblayer la place pour donner toute carrière à son imagination ; ce n'est pas même faire le premier pas vers les démonstrations qu'exige la science (1863, p. 312).

Les conclusions de Broca ne sont guère éloignées de la position de la Société de linguistique, sur lesquelles elles pourraient avoir eu une influence⁴. Elles ont des conséquences importantes pour l'idée même d'une participation de plusieurs disciplines à la question de l'origine des langues. Dans l'ouvrage de Henry qui nous a souvent servi de fil directeur, l'essentiel est moins la présentation antinomique, que la poursuite jusqu'au bout de la démarche kantienne, par la production d'une *synthèse*, c'est-à-dire d'une solution qui

² Métoz (2005, pp. 223 s.) a formulé l'hypothèse plausible qu'il était motivé par une réfutation du monogénéisme linguistique inspiré par Darwin.

³ La diversité linguistique est fonction du temps. Les langues qui d'après Dumont d'Urville sont les plus différentes sont celles qui d'après Hales sont séparées depuis le plus longtemps (1863, p. 307). Hales s'appuie sur les généalogies retenues par les populations ; à 30 ans par génération, l'occupation d'Hawaï aurait à peu près quinze siècles. Prenons les langues de Tahiti et d'Hawaï et voyons ce qui subsiste après leur séparation. Considérons non seulement comme Dumont les mots très semblables (74 %), mais tous ceux qui ne sont pas essentiellement différents ; on en trouve 219 sur 221, soit 99 % de permanence au bout de quinze siècles.

⁴ Renan lui-même conviendra qu'on ne peut rien inférer des familles linguistiques pour les races humaines et réciproquement ([1878] 1958, p. 1224). Saussure reprendra les arguments de la discussion (Engler III, pp. 436-515).

considère la question du point de vue des fondements de la connaissance¹. Si synthèse il y a, elle se situe donc au niveau du problème épistémologique, pas dans la collaboration des disciplines. Pour l'essentiel Henry procède en désambiguïsant ce qu'on entend par « langage ». L'origine du « langage » en tant que *faculté* de la parole articulée appartient à l'anatomie comparée ; en tant qu'*exercice* de cette faculté, à la physiologie, en tant que *réflexe d'une excitation interne ou externe*, à la psycho-physiologie, en tant que *phénomène affectif et cognitif*, c'est l'histoire d'un processus psychologique. *La question de l'origine du langage n'est pas insoluble* (1896, p. 37), *mais ce n'est pas un problème linguistique*. Le psychologue E. B. Leroy, (cf. Nerlich, 1984, p. 79) trouvera quelques années plus tard le vocabulaire (en français ; l'anglais lui interdirait ce genre de solution) adéquat pour exprimer la différence de cet objet à tout autre : « En somme, les linguistes étudient la structure des *langues*, les psychologues étudient le *langage* [mes italiques, S.A.] » (1905, p. 2). Le rejet de l'origine des langues hors de la linguistique, est désormais la même chose que l'affirmation de l'autonomie² de cette discipline (Saussure, Engler I, p. 30). Il n'y a pas de synthèse inter-disciplinaire possible. Comme le notera Lévi-Strauss, pour les linguistes la méthode ne permet à aucun égard de conclure de la langue à la culture, de la langue au peuple ou de la langue à la « race » (1966, p. 116) ; la seule solution, c'est l'épigénétisme :

Quels qu'aient été le moment et les circonstances de son apparition dans l'échelle de la vie animale, le langage n'a pu naître que tout d'un coup. Les choses n'ont pas pu se mettre à signifier progressivement. À la suite d'une transformation dont l'étude ne relève pas des sciences sociales, mais de la biologie et de la psychologie, un passage s'est effectué, d'un stade où rien n'avait de sens à un autre où tout en possédait (« Introduction », *Œuvres* de M. Mauss, Paris, Minuit, 1968, t. 1, XLVII).

8. La théorie comme après-coup de l'acte institutionnel

Cassirer soutenait que c'est avec le thème de l'organicisme romantique que la philosophie du langage a renoncé « une fois pour toutes à tenter de découvrir derrière la diversité individuelle et la contingence historique des langues singulières la structure universelle d'une langue mère » ([1923], 1972, pp. 102-103). Il a raison sur deux points. D'un côté, ce type d'organicisme (c'est-à-dire la conception d'une langue comme totalité, cf. la position de Humboldt ou Renan) induit qu'on ne puisse fragmenter la grammaire d'une langue dans l'ensemble des éléments d'une apparition progressive, ni éparpiller une langue mère supposée dans des morceaux hétérogènes qui aboutiraient à la différence radicale où se trouvent les diverses familles linguistiques les unes par rapport aux autres. De l'autre, tout rejet d'une approche quelconque de l'origine des langues, n'est pas une simple constatation de l'absence de données empiriques, mais la contre partie des propriétés générales par lesquelles dans une théorie linguistique quelque chose peut être construit comme fait (cf. Milner, 1984), et auxquelles par certains côtés l'origine ne répond pas. Mais Cassirer est dans l'illusion lorsqu'il entend dater le rejet de l'origine des langues en Allemagne au début du XIX^e siècle. La stabilisation de l'interdit, a lieu, bien plus tardivement, en France dans une ambiance rationaliste, qui est contemporaine du néokantisme positiviste de Vaihinger³.

L'étude historique de la question de l'interdit sur l'origine des langues a de quoi laisser perplexe. Nous avons certes pu reconstituer un corps d'arguments, qui est devenu canonique

¹ « nach der Fundamente der Erkenntnis, worauf sie gegründet ist » (*Kri. Rein. Vern.*, Kant Werke, Suhrkamp, 1956, IV, p. 456).

² Il est certain que l'autonomie est dès le départ un projet conscient, même s'il n'a pas de fondement méthodologique explicite. Cf. *Bulletin de la Société de linguistique* n° 15, séance du 5 janvier 1879, adresse du président (Mowat) : « Notre compagnie a été assez près d'entrer dans une voie différente de celle qu'elle a définitivement suivie. Il s'agissait du nom qu'elle prendrait, et parmi les diverses dénominations proposées, celle de *Société d'ethnologie et de philologie* ne fut écartée après un débat prolongé que sur l'observation d'un membre qui nous fit sentir la nécessité de délimiter d'un seul mot le champ de nos futurs travaux, et d'accentuer avec netteté notre personnalité à côté des sociétés vouées aux études anthropologiques et ethnologiques ».

³ Hans Vaihinger publie son commentaire de la *Critique de la raison pure* entre 1881 et 1892 ; il fonde les *Kant-Studien* en 1896 ; c'est en 1911 seulement qu'il publiera son fameux *Die Philosophie des Als-Ob* (*La philosophie du « comme si »*), rédigé pour l'essentiel dès 1876. Il y défend l'idée du rôle scientifique des constructions fictives, correspondant souvent à des concepts contradictoires, généralement reconnus comme tels, mais indispensables à l'obtention de certains résultats. Il définit sa conception comme un « positivisme idéaliste ».

(cf. Vendryes, 1921, pp. 6-20 ; Perrot, p. 20 ; Mounin, 1967, pp. 17-32). Nous avons pu également déterminer le rapport de cet interdit à la constitution de la « langue » comme objet de la linguistique, constatation dont l'aboutissement est incontestablement le *Cours* de Saussure et sans laquelle la valeur théorique de l'interdit n'est ni probant, ni même intelligible. Il est clair par ailleurs que cet interdit structure les possibilités de compréhension théorique des linguistes de la tradition saussurienne en général, et de la tradition française en particulier. Pour le voir, il suffit de constater l'accueil fait, à la glottochronologie de Swadesh (voir plus loin) parmi les saussuriens et à la Société de Linguistique de Paris¹.

Les arguments que nous avons recensés contre l'origine l'ont été sur la *chronologie longue*, de part et d'autre de la décision de la société de linguistique. Nous n'avons pas pu donner à cette décision de véritable consistance théorique. Il faut en conclure que les fondateurs de la Société de linguistique n'ont pas de théorie apte à refuser la question de l'origine comme en ont les mathématiciens pour la quadrature du cercle. Il est probable que *le principal motif du refus est un acte de négociation sociale* (Vendryes, 1955, p. 13), dont on ne peut nier l'opportunité en considérant les remous de la querelle du darwinisme² ou la tournure que prend majoritairement l'anthropologie physique tout au long du siècle, dont la liaison avec la linguistique sert souvent à opposer les aryens aux sémites³. Les contraintes logiques qui régissent les relations entre les différents arguments du rejet semblent aussi floues que celles qui s'établissent entre les différents lieux du labyrinthe de l'origine. Le cas Renan est un exemple caricatural ; il rejette comme anti-scientifique l'étude de l'origine des racines ; mais il écrit l'ouvrage français le plus connu sur l'origine des langues, ce qui ne l'empêche pas d'être l'un des premiers présidents (1867) de la Société de Linguistique, au temps du fameux article II.

Nous sommes en présence de trois types d'éléments : des *théories* (par exemple celle qui s'exprime dans le *Cours* de Saussure, mais aussi l'organicisme de Schleicher et de ses disciples français, ou encore le corps de doctrine des néogrammairiens), des *arguments* ou *thèmes* théoriques, et des *institutions*. Il est assez vain de dater ponctuellement l'apparition des thèmes théoriques ; s'ils ont un intérêt pour l'historien, c'est par leur permanence sur la longue durée (l'argument de la circularité remonte à l'Antiquité), et par leur situation argumentative dans les structures théoriques souvent très différentes où on les rencontre de façon récurrente. Les institutions, par contre, donnent lieu à des *actes institutionnels*, qui constituent le point de contact obligé entre les entités abstraites que sont les théories et un réel qui est de nature autant événementielle que sociale. Sans institutions, les sciences, comme toutes les activités intellectuelles, n'auraient ni histoire ni même existence. Entre les théories et les actes institutionnels existe un réseau complexe de causalités, qui n'est pas la simple réalisation des premières dans la publicité des seconds. L'article II de la société de linguistique a évidemment des causes (ne serait-ce que stratégiques), mais comme acte institutionnel, il est *sur le coup* « théoriquement contingent » : *il n'est pas la mise en œuvre d'une théorie linguistique*.

Sa réalité institutionnelle est pourtant incontournable et efficace⁴. La majeure partie des discussions que nous avons évoquées ne datent pas de 1866, mais plutôt du dernier quart du

¹ A. Sauvageot (*BSL*, 48/3, pp. 25 s.) a donné un compte-rendu dévastateur (mettant en cause les compétences linguistiques de l'auteur) de la monographie, aujourd'hui classique de Swadesh (Swadesh (M.). 1951. « Diffusional Cumulation and Archaic Residue as Historical Explanations ». in : *Southwestern Journal of Anthropology*, 7, pp.1-21). Ce dernier a répondu l'année suivante (*BSL*, 49/3, pp. 1 s.) en se basant sur le fait qu'il s'adressait à des anthropologues.

² La Société de linguistique évita les remous de la Société d'Anthropologie (créée en 1859) à propos du darwinisme ainsi que de l'opposition entre poly-génétilisme et mono-génétilisme, le poly-génétilisme étant initialement opposé au darwinisme. Voir Harvey, 1984.

³ On se reportera au livre remarquable de L. Poliakov (1971), notamment au passage du chapitre V de la seconde partie, intitulé la « Responsabilité des linguistes ». Les citations (notamment celles de Pictet) sont édifiantes. Voir également Olender (2002, 2005).

⁴ Cette efficacité se poursuit jusqu'à une distorsion de la représentation de l'histoire des sciences du langage, dont on a vu un bon exemple avec Cassirer. On tenait encore récemment pour acquis i) que la question de l'origine des langues est l'affaire des « philosophes » du XVIIIe siècle ; ii) qu'elle a été repoussée en dehors de la science par l'avènement de la « linguistique scientifique » avec la grammaire comparée du siècle suivant. Une telle représentation fausse complètement les analyses de l'excellent livre de Droixhe (1978), qui reste utile pour sa documentation. À chaque fois que cet auteur relève une remarque contre l'étude de l'origine, il s'empresse d'en conclure qu'il s'agit d'un pressentiment de la « véritable » science, sans remarquer combien l'argument que nous avons noté **B**, a été un obstacle

siècle (elles vont même jusqu'à la seconde moitié du XX^e siècle), et elles sont animées par des membres de la Société de Linguistique. Sans l'article II, la théorie saussurienne qui par bien des aspects, est l'aboutissement de cette discussion, n'aurait sans doute pas été ce qu'elle est ; les considérations paradoxales sur l'absence de datation d'une langue ou la position raffinée des indo-européanistes sur le caractère épistémologiquement conventionnel des racines (qui est une incontestable innovation dans l'histoire des sciences du langage) n'auraient sans doute pas eu de raison d'être. Somme toute, la maturation des théories est parfois *l'après coup des actes institutionnels*. Faut-il en conclure qu'*après coup* l'interdit *conjecturel* est devenu un véritable interdit *scientifique*, comme le proclament de nombreux linguistes ?

9. Les limites de l'hégémonie

La différence entre une société savante et la « science », c'est que la première relève de la concurrence sociale et l'autre de l'universalité abstraite. Le développement de la seconde dépend largement de l'hégémonie de la première. Dans le dernier tiers du XIX^e siècle cette hégémonie n'est certainement pas totale. La Société de linguistique reste volontairement cantonnée dans les langues indo-européennes, son refus de la question de l'origine et le maintien du statut de totale indépendance de la discipline « linguistique », dont elle s'efforce d'assurer le développement dans un monde universitaire en pleine mutation depuis la fin du Second Empire (ministère V. Duruy) et, surtout, de l'établissement de la République (J. Ferry). En 1864, on voit naître la Société d'Ethnographie qui s'intéresse à l'Océanie et au Pacifique, le principal contributeur de ses *Bulletins* étant le japonologue Rosny (voir Rosny, 1901), professeur à l'École des langues orientales qui publiera sur l'Amérique pré-colombienne et l'origine des langues (Rosny, 1869, 1904) ; après y avoir donné un cours sur l'histoire de l'Amérique précolombienne, il sera vice-président de l'École Pratique des Hautes Études. En 1869, autour du Comte de Charencey est créée la Société Philologique, liée aux missionnaires et aux langues exotiques, qui penche vers le mono-génétilisme et n'hésite pas à accueillir favorablement l'idée d'une parenté entre langues sémitiques et indo-européennes. En 1876, avec l'appui de la Faculté de Médecine, l'aide de souscripteurs privés et malgré le déchaînement de la presse catholique, les membres de la Société d'Anthropologie ouvrent l'École d'Anthropologie où exerce A. Hovelacque, qui, en 1873 n'avait pas hésité à décrire les aptitudes linguistiques du « précurseur de l'homme », anticipé dans son aspect physique par Mortillet (Hovelacque, 1873 ; Mortillet, 1873). On voit clairement que cette concurrence sur le marché de la linguistique est déterminée par une demande sociale, dont les motivations sont très largement extérieures à toute idée d'une autonomie de la science. L'origine du langage et celle de l'homme sont l'objet d'un *intérêt social général*, alimenté tout au long du siècle par le développement de la théorie de l'évolution et les découvertes de la paléontologie. On ne remet pas en question tout le socle mythologique d'une culture sans que cela n'affecte profondément la société. Le succès populaire du premier « roman préhistorique », *La guerre du feu* publié en 1911 par Rosny l'Aîné¹, témoigne en ce sens.

À l'exception de la Société d'Anthropologie², qui était à l'origine de l'École, aucune de ces associations ne survivra à la mort de ses animateurs et à la coupure de la Première Guerre Mondiale (voir Auroux, 1984), faute d'avoir pu intégrer le nouveau système universitaire. La Société de Linguistique³ de Paris, fortement ancrée sur les chaires universitaires et l'École pratique des hautes études acquiert enfin l'hégémonie ; ses thématiques seront relayées et prolongées par le structuralisme européen. Toutefois, malgré l'influence de la linguistique européenne allemande sur Boas ou Bloomfield⁴, l'histoire originale des recherches linguistiques

considérable au développement du comparatisme, y compris tardivement chez Burnouf. Ce n'est qu'à la fin des années quatre-vingt qu'une étude empirique de la documentation a fait apparaître aux historiens que le XIX^e siècle était profondément engagé dans la recherche de l'origine des langues (voir, par exemple, ma contribution dans Gessinger & von Rahden, 1989 et l'anthologie de Harris, 1996).

¹ Il s'agit du pseudonyme du Belge Joseph Henri Boex, qui figure avec son frère parmi les premiers membres de l'Académie Goncourt.

² Plusieurs membres de cette société, partagée entre positivistes et matérialistes, participèrent à la Commune de Paris. Hovelacque sera député d'extrême gauche et s'opposera à J. Ferry.

³ Son secrétaire, M. Bréal, personnage officiel de la III^e République, Professeur au Collège de France, Inspecteur général de l'Instruction publique, a joué les bonnes cartes politiques.

⁴ Boas était un juif allemand émigré aux Etats-Unis en 1887 ; son élève Sapir est également né en Allemagne, mais a émigré à l'âge de 5 ans. Bloomfield est allé étudier en Allemagne auprès des

aux États Unis appartient à une autre problématique : la grande question est la classification des langues amérindiennes et, jusqu'au début de la seconde moitié du XX^e siècle, les chaires de linguistique seront souvent abritées dans des départements d'anthropologie. D. Hymes a rassemblé (Hymes, 1964) une anthologie des textes fondateurs de l'anthropologie linguistique¹. Tout un pan, actif et fécond, de la recherche linguistique est ainsi resté éloigné des préoccupations des indo-européanistes européens.

10. Le renouveau d'intérêt pour la question

Même si l'on peut remarquer une relative stagnation de la question de l'origine des langues durant la première moitié du XX^e siècle, on ne peut pas dire qu'elle ait jamais été abandonnée. Plutôt qu'une véritable solution de continuité, comme le proclament les organisateurs² du colloque de 1976 (concurrence oblige, il faut bien reprendre la main³), on doit voir dans ce dernier un acte institutionnel symétrique de celui de la Société de linguistique : les organismes officiels valident ce champ de recherche, des crédits vont être distribués, des chercheurs vont pouvoir faire carrière, une masse de publications va paraître⁴. On peut résumer à quatre lieux ce renouveau des études d'origine : *les avancées considérables de la paléontologie humaine, l'extension du « paradigme généalogique » à la préhistoire* (Swadesh, Greenberg, Ruhlen), *le rapprochement biologique* (Cavalli-Sforza, Bickerton, Pinker) et *l'idée d'une nouvelle synthèse disciplinaire* (Renfrew). Le contexte général peut être défini par la domination des chomskyens, l'explosion des études psycho-linguistiques et celle sur les aptitudes linguistiques des primates⁵, ainsi qu'une ignorance abyssale de l'histoire de la discipline.

Les découvertes de la paléontologie humaine, notamment en Afrique, à partir de celle du *Zinjanthropus boisei*, un Australopithéciné de taille humaine entouré d'un outillage par L. B. S. Leakey (15 juillet 1959), ont considérablement changé la problématique du développement de l'humanité⁶. D'un côté, l'apparition de l'« homme moderne » est considérablement reculée et elle ne provient pas de l'humanisation des primates. De l'autre, en considérant le développement de la boîte crânienne tant des primates que des hominiens, à l'aide de ce que l'on connaît des aires cérébrales, on parvient à deux conclusions que met bien en lumière

néogrammariens et a travaillé sur l'apophonie. Le plus célèbres des linguistes américains, Whitney, était un sanskritiste très orthodoxe. Notons, enfin, que la première application cohérente de la notion de loi phonétique à une famille amérindienne (en l'occurrence, la famille caribe) est due à l'explorateur allemand K. von den Steinen dès 1892.

¹ « Anthropological linguistics may briefly be defined as an area of linguistics research which is devoted in the main to studies, synchronic and diachronic, of the language of people who have no writing », H. Hoijer, 1961. Le terme est peu utilisé en France où son équivalent « ethnolinguistique » apparaît en 1970 chez les américanistes (n° 18 de la revue *Langages*, dirigée par B. Pottier).

² « Virtually all aspects of our relevant knowledge have changed radically since the nineteenth century. Our concepts of the nature of language is totally altered and has become more profound and more complex. The revolution in linguistics due to Noam Chomsky has provided a very different idea of what the nature of the « target » for the evolutionary process might actually be » (Harnad & al., 1976, p. 1).

³ Certains confondent le temps de leur formation et de leurs premières publications avec l'histoire de la linguistique : « Most we know about language has been learned in the last three decades » (Bickerton, 1990, p. 5)

⁴ On peut rester interloqué devant les assertions de certains auteurs : « With few exceptions, linguists have refrained from any consideration of the origins of language (...). Noam Chomsky, arguably the Newton of our field, has dismissed the origin of language as an issue of no more scientific interest than the origin of the heart » (Bickerton, 1990, p. 5).

⁵ À la fin des années soixante commence aux États Unis une série de programmes expérimentaux visant à faire acquérir à des primates supérieurs la capacité de manipulation des systèmes symboliques complexes permettant, notamment, la communication avec les humains. On signalera parmi les principaux : i) WASHOE : B.T. et A. Garner (à partir de 1966 ; en 1970, les études sur Washoe ont été poursuivies à l'Université d'Oklahoma, où des programmes ont été développés sous la conduite de R. Fouts) ; ii) SARAH : A. et D. Premack (à partir de 1970) ; iii) LANA (LAN-guage ANAlogue) : D. M. Rumbaugh (Atlanta : à partir de 1972) ; iv) KOKO : F. Patterson (Stanford ; à partir de 1972) ; v) NIM : H. S. Terrace (New York ; à partir de 1973) ; vi) CHANTEK : L. Miles (Université du Tennessee ; à partir de 1977) ; vii) KANZI : E. S. Savage-Rumbaugh (Atlanta, à partir de 1981). Les résultats sont très controversés ; de manière générale, il ne semble pas que les primates puissent maîtriser des éléments syntaxiques significatifs (ce qui renforce les partisans d'une discontinuité biologique à l'origine du langage).

⁶ « La découverte de Leakey marque un point final au mythe de l'homme singe, tout au moins dans les milieux scientifiques », Leroi-Gourhan, 1964, pp. 128 s.

Leroi-Gourhan : liaison de la faculté de langage à la motricité propre à l'outillage¹ et séparation nette de l'homme et des primates². Dès lors, le préhistorien dispose d'une « voie détournée » (*ibid.*, p. 161), pour aborder l'origine et le développement du langage :

[...] il y a possibilité de langage à partir du moment où la préhistoire livre des outils, puisque outil et langage sont liés neurologiquement et puisque l'un et l'autre sont indissociables dans la structure sociale de l'humanité. [...]. À partir de là, peut-être, une paléontologie du langage pourrait être tentée, paléontologie toute squelettique d'ailleurs, car il n'y a guère d'espoir de retrouver jamais la chair des langages fossiles (Leroi-Gourhan, 1964, p. 163).

Leroi-Gourhan reste dans les limites déterminées par la synthèse de V. Henry. Il ne déborde pas sur les caractéristiques proprement linguistiques³ du langage humain et donne des arguments empiriques en faveur d'un détachement progressif du développement de l'humanité par rapport à son contexte biologique⁴. Quoi qu'il en soit, la voie était ouverte pour lier la question de l'origine du langage aux avancées factuelles de la préhistoire humaine qui ont été considérables dans le dernier tiers du XX^e siècle (nouveaux sites avec outillage et autres objets, nouveaux fossiles d'hominiens, renforcement de la piste africaine). La liaison de la complexité des chaînes opératoires nécessaires aux façonnages d'outils (voire à d'autres activités plus proprement symboliques comme l'ornement individuel ou la peinture pariétale) est désormais une technique éprouvée pour aborder la datation du développement linguistique de l'humanité. Les nouvelles techniques de datation des objets matériels et de traçage génétique affinent l'approche préhistorique du peuplement de la planète et semblent actuellement valider l'hypothèse d'une origine africaine (Stringer & McKie, 1996). L'archéologie plaide en faveur d'une sorte de révolution⁵ technique et cognitive entre le paléolithique moyen et le paléolithique supérieur dans des populations « anatomiquement modernes ». Cette révolution serait introduite en Europe par des nouvelles populations « plus probablement d'une source africaine antérieure » (Mellars 1998 : 109). Le langage a évidemment quelque chose à voir avec cette révolution qui remonterait à quelque 100 000 ans ; on imagine bien la corrélation entre des sociétés complexes et un langage qui dispose de « traits » semblables au nôtre (syntaxe complexe, système temporel, conditionnel, subordination) ; mais comment orienter cette corrélation en évitant le fameux « cercle des origines »⁶ ? Ou l'on reste dans la prudence de Leroi-Gourhan s'en tenant au parallélisme permettant aux monuments de la préhistoire de dater grossièrement l'apparition du langage humain et quelques étapes probables et abstraites de son évolution ou l'on doit franchir un pas de plus et relier cette datation à l'apparition de déterminations biologiques qui contraignent mécaniquement telle ou telle forme de langage. C'est ainsi que pour Mellars, l'archaïque homme de néanderthal disposerait de ce que Bickerton (voir plus loin) reconnaît comme un « proto-langage » chez l'enfant (Mellars, 1998, p. 102), génétiquement déterminé. Ce type d'hypothèse ne relève plus de la paléontologie.

¹ La remarque initiale concerne la proximité entre les fibres de projection manuelle et les fibres faciales. « [...] à partir d'une formule identique à celle des Primates, l'homme fabrique des outils concrets et des symboles, les uns et les autres relevant du même processus ou plutôt recourant dans le cerveau au même équipement fondamental » (Leroi-Gourhan, 1964, p. 162).

² « non seulement le langage est aussi caractéristique de l'homme que l'outil, mais [...] ils ne sont que l'expression de la même propriété de l'homme, exactement comme les trente signes vocaux différents du chimpanzé sont l'exact correspondant vocal des bâtons emmanchés pour attirer la banane suspendue, c'est-à-dire aussi peu un langage que l'opération des bâtons n'est une technique au sens propre » (Leroi-Gourhan, 1964, pp. 162-163)

³ « Avant l'écriture toute saisie directe du langage est irréalisable » (Leroi-Gourhan, 1964, p. 161).

⁴ En prenant comme indice d'avancée technologique la courbe de développement du rapport tranchant/poids des outils et le volume crânien pour celui du développement biologique, on remarque qu'à partir de la fin du moustérien, la première devient quasiment verticale, tandis que la seconde stagne : « La technique n'est plus liée chez l'*homo-sapiens* au progrès cellulaire mais elle paraît s'extérioriser complètement et vivre en quelque sorte de sa vie propre » (1964, p. 197).

⁵ Non seulement changement quantitatif dans les outils, mais semble-t-il apparition de pratiques funéraires, instruments producteurs de sons, transports d'artefacts sur des distances de 400 à 600 km, ce qui suppose une complexification de l'organisation des groupes sociaux.

⁶ Dans le modèle présenté par Dessales (2000) le « récit » (référence à des événements passés, chronologie, système temporel) est présenté comme la naissance du langage qui rend possible le développement social ; on pourra ajouter « et réciproquement » et remarquer que ce type de trivialité ne fait pas avancer d'un pas la question de l'origine des langues.

11. De la glottochronologie à l'étymologie globale

Les linguistes de la tradition américaine ont tenté d'apporter une contribution proprement linguistique à la préhistoire du langage. On doit à Swadesh la méthode mathématique connue sous le nom de « glottochronologie¹ ». Cette méthode² naît et se développe autant sous l'effet des limitations de la linguistique historique concernant les langues sans écriture que des progrès considérables quant à la connaissance empirique de la diversité des langues du monde et des difficultés à les regrouper par familles (Amériques, Afrique, Asie). On retrouve systématisées et appuyées sur un nombre sans précédent de données la démarche de Dumont d'Urville et Broca³. Les hypothèses nécessaires sont très fortes : existence d'un vocabulaire de base universel (environ cent mots) plus résistant que le reste aux emprunts⁴, fiabilité des données et de la reconnaissance des ressemblances, constance historique du taux d'usure et universalité de ce taux sur toutes les familles linguistiques, notamment. En dépit de multiples exposés de la méthode⁵ et du fait que dans bien des cas (langues amérindiennes ou africaines) on ne dispose guère d'autres moyens pour proposer des datations, chacun de ces points a fait l'objet de violentes contestations. Le choix du vocabulaire et son exploitation ont été critiqués ; la fiabilité a été prise en défaut ; on a montré statistiquement qu'en l'absence d'autres critères, la technique d'apparement par le vocabulaire ne donnait pas de meilleurs résultats que des rapprochements aléatoires (Ringe, 1992). Compte tenu de la lenteur de la séparation éventuelle, la datation que l'on donne, forcément ponctuelle, est celle de quelque chose qui n'existe pas⁶. L'idée d'un taux constant gomme complètement les aléas de l'histoire. Enfin, dans tous les cas où l'on disposait d'autres méthodes de datation, notamment, parce qu'il s'agissait de langues écrites disposant de documentation historique (langues indo-européennes, langues romanes, langues slaves, etc.) les datations données par la glottochronologie ne sont pas validées, ce qui ressemble fort à une réfutation empirique. Même si l'idée de relier la similarité entre langues et leur proximité génétique paraît relativement évidente⁷, au mieux la fiabilité des méthodes proposées n'est pas satisfaisante sans qu'il soit possible de l'améliorer (Fox 1995 : 290), au pire elles relèvent d'un fantasme pseudo-scientifique totalement étranger à la nature du langage humain. Largement discréditée dès les années soixante-dix, la glottochronologie a pourtant connu des raffinements ultérieurs (Trask 1996). Nichols en a proposé une variante significative, qui admet dès le départ que la vitesse du taux de remplacement et ses variations n'est pas une question linguistique mais dépend de « facteurs géographiques et économiques » (Nichols, 1998, p. 136). Elle admet également que la piste du vocabulaire ne permet guère d'aller au-delà de la période historique⁸. C'est

¹La glottochronologie s'efforce de dater la séparation des idiomes à l'aide d'une formule mathématique : $t = \log C / 2 \log r$, où t a pour valeur le temps (en millénaire), C indique le pourcentage de termes apparentés subsistant dans les langues considérées, r est une constante qui a pour valeur le pourcentage de termes apparentés qui demeurent dans des langues de même origine après un millénaire de séparation. La vitesse d'évolution des langues devient une constante, ce qui, à première vue, contredit la contingence de l'histoire, sauf à supposer qu'il s'agit d'une moyenne globale (dans ce cas on peut se demander quel sens linguistique elle pourrait avoir).

² Dans l'ouvrage posthume de 1971, Swadesh reconstruit un schéma d'apparition du langage à partir de la lente évolution des cris instinctifs qui sont à peu près les mêmes pour toute l'humanité. Cette base biologique induit le mono-généisme.

³ Il ne semble pas que Swadesh ait eu connaissance des travaux antérieurs allant dans cette direction. Il faut attendre Hymes (1973) pour faire le rapprochement. On peut supposer que c'est la considération des vocabulaires comme moyen d'apparement qui, en l'absence de tout autre documentation, conduit à rechercher une solution par « calcul ».

⁴ Ce vocabulaire est conçu comme l'équivalent du Carbone 14 pour la datation des objets matériels.

⁵ Voir les textes repris dans Hymes (1964), notamment Gudschinsky (1956).

⁶ On retrouve l'un des arguments canoniques (continuité du processus) contre l'étude de l'origine des langues.

⁷ On pourrait admettre que sous une formulation aussi générale que celle que l'on vient de donner, il s'agit du principe de base de la grammaire comparée : « C'est la persistance de formes anciennes à l'intérieur du système nouveau qui a rendu possible la grammaire comparée » (Meillet, 1937, p. 450). Toutefois l'application en est inversée : la grammaire comparée constate les apparements et s'efforce de les expliquer (changements phonétiques), la glottochronologie, à l'inverse, s'intéresse à l'érosion et doit ajouter quantité d'hypothèses auxiliaires (ce sont elles qui sont indésirables) pour en tirer des conclusions.

⁸ La perte de 20 % du vocabulaire par millénaire impose que l'on ne puisse remonter à plus de 6 000 ans (Nichols, 1998, p. 128).

pourquoi elle propose une méthode de calcul basée sur le nombre de langues que l'on reconnaît descendre d'un ancêtre commun¹. On parvient ainsi à 132 000 ans pour « l'âge linguistique du monde » dans l'hypothèse mono-génétique (*ibid.*, p. 139) et 100 000 dans une hypothèse poly-génétique comportant 10 familles primitives (*ibid.*, p. 165). En admettant que la méthode puisse donner des ordres de grandeur, elle est largement arbitraire. Si une migration que nous ne connaissons pas a rajouté des langues apparentées cela rallonge la période ; dans le cas contraire de la disparition de langues inconnues, cela la raccourcit. Ces deux facteurs permettent d'expliquer les nombreux cas où les données archéologiques ne concordent pas avec le calcul qui, dès lors, devient d'autant plus infalsifiable que l'on a retiré de son domaine le cas bien documenté des langues indo-européennes.

Ce sont partiellement les critiques de la glottochronologie qui ont conduit Greenberg à renoncer à la datation et à se concentrer sur l'utilisation des vocabulaires pour améliorer les classifications linguistiques (méthode dite « multilatérale »). En 1987, il propose une réduction des langues amérindiennes en trois familles : *eskimo-aléoute*, *na-dene*, *amerind*. Les critiques (Campbell, 1988) ont été largement les mêmes que pour la glottochronologie (Métoz, 2004, pp. 245-381) ; elles portent d'autant plus qu'il n'a pas donné le matériel qui a servi de base aux regroupements et que ses carnets de notes sont inexploitablement, donnant l'impression qu'il a tenté de justifier une hypothèse *a priori*. Son élève M. Ruhlen a porté la méthode de comparaison multilatérale à son paroxysme en défendant à toute force le mono-génétilisme.

Pour comprendre l'enjeu de la méthodologie proposée par Ruhlen (« l'étymologie globale »), il faut revenir aux origines de la grammaire comparée et à l'innovation proposée par Grimm en 1822 dans la seconde édition de sa *Vergleichende Grammatik der deutsche Sprache*. L'idée de comparer les langues n'est pas une nouveauté ; elle survient mécaniquement dès que l'on dispose d'une connaissance suffisante de plusieurs langues (Auroux (ed.). 1990-2000). Mais pendant longtemps la base de comparaison reste le mot que l'on identifie par sa ressemblance phonétique et sémantique. Comme d'une langue à l'autre la constitution et la valeur des mots varient, leur rapprochement nécessite quelques aménagements pour « rendre » leur identité plus évidente. On dispose depuis l'Antiquité de techniques appropriées à cette fin, il s'agit des *figures*. Le *métaplasme* permet dans un mot de changer une lettre par une autre (par exemple, v pour b) ; les *tropes* expliquent pareillement les changements de sens (par exemple, un mot qui signifie « bouche » peut signifier « tête » par métonymie). Il est évident que cette technique donne une grande latitude. La phonétique articulatoire proposée au XVIII^e siècle par De Brosses et Court de Gébelin (Auroux, 1979c) pose quelques principes limitatifs : remplacement réciproque privilégié des voyelles ayant le même organe d'articulation (par exemple, une labiale pour une autre labiale), indifférence au changement vocalique². C'est à l'aide de ces principes et de la plausibilité du changement sémantique à l'aide d'une chaîne de tropes que sont construits le « Dictionnaire de la langue primitive » de Court ou le projet de « Nomenclateur universel » du Président de Brosses. Grimm apporte une vraie révolution³ épistémologique : *l'unité du changement ce n'est plus le*

¹ On prend les grandes familles connues et on compte le nombre de branchements initiaux, en excluant l'indo-européen qui est trop « particulier » par rapport à ce qui se passe ailleurs (*sic* !) ; on évalue le nombre moyen de branchements sur un ensemble de familles (Nichols a pris celles de l'hémisphère Nord, mieux documenté), ce qui donne un taux moyen de dispersion (en l'occurrence 1,5) sur la période historique de 6 000 ans. Dès lors, lorsque l'on se trouve en présence d'un nombre n de branchements, il suffit de diviser n par 1,5 et ainsi de suite sur chacun des résultats jusqu'à ce que l'on trouve un nombre inférieur à 2, puis de multiplier le nombre d'étapes par 6 000 pour avoir la datation de l'ancêtre commun (*ibid.*, p. 136).

² Cette dernière thèse, rendue possible par la théorie des voyelles conçues comme des sons différenciés par la hauteur de ton (on ne distingue pas encore les formants) est totalement erronée. Elle conduit à ne pas comprendre la valeur sémantique de l'apophonie (c'est Bréal qui introduira tardivement le mot en français) et interdit toute idée de la reconstruction du vocalisme indo-européen (voir Benware, 1974), reconstruction qui sera l'un des grands résultats de la grammaire comparée allemande (et du fameux *Mémoire* que Saussure publie en 1879).

³ Sur toutes ces questions voir dans Auroux (ed.). 1990-2000, t. 2, pp. 538-579 et t. 3, pp. 155-172. Dès que l'on dispose de suffisamment d'informations sur un groupe de langues, on propose les apparentements sur la base des ressemblances intuitives : XVII^e siècle, langues sémitiques et néo-latines ; XVIII^e siècle, familles caribe, algonkine, sioux, austronésienne, finno-ougrienne. On pourrait se demander pourquoi il faut attendre le début du XIX^e siècle pour la famille indo-européenne, alors que dès le XVI^e siècle on sait rapprocher sk. *asti* et gr. *esti* (3^e pers. Sing. du verbe « être »). La meilleure

mot mais le phonème (la « lettre »), il advient dans toutes les occurrences du phonème considéré, il est orienté et spécifique à chaque langue (ou, au plus, à un groupe de langues bien déterminées comme dans le cas des langues germaniques). Ce sont ces principes fondateurs du comparatisme que refuse l'étymologie globale. Elle choisit le mot plutôt que les correspondances phonétiques ; la langue n'est pas considérée comme un système. Ruhlen soutient que les correspondances phonétiques ne viennent qu'après l'apparement sur la base du vocabulaire de base (1994a : 205)¹ et que ce ne sont pas elles qui prouvent l'apparement. Il se gausse de correspondances contre-intuitives comme le fameux *dw>erk² (1994a, p ; 39) qui n'aurait jamais joué aucun rôle dans la classification des langues indo-européennes et invoque les « changements sporadiques » des néogrammairiens qui n'obéissent pas à la régularité des correspondances et donnent donc des raisons de préférer le mot. La première affirmation peut se soutenir, au moins partiellement, sur les bases de nos connaissances de l'histoire du comparatisme³. La seconde est extrêmement discutable. D'abord, les correspondances phonétiques ont l'immense avantage de nous débarrasser des acrobaties sémantiques autorisées par les reconstructions à la Court de Gébelin. Elles réduisent d'autant les effets du hasard ou des emprunts qui peuvent perturber nos apparements sur la seule base du vocabulaire. Elles ont, enfin, une puissance de généralisation incomparable, au point de permettre l'anticipation dans la classification des faits inconnus⁴. Dans ces conditions, le travail de Ruhlen fait souvent appel à des faits d'apparements isolés sur les langues du monde (« global cognates ») : la racine TIK (doigt, un), ou encore MALIQ'A (gosier) qu'on retrouverait en Afrique du Nord, Eurasie, Amériques du Nord et du Sud (Ruhlen, 1994b), etc. ; il s'appuie également sur des discussions anecdotiques et non probantes⁵. Bien qu'il s'agisse d'un livre de vulgarisation⁶, contrairement à Greenberg (1987), Ruhlen (1994a) fournit un matériel linguistique. Il s'agit d'une trentaine de mots (les 27 « racines mondiales ») pris dans une douzaine de familles linguistiques représentant les quelques 5 000 langues qui existent aujourd'hui (voir également Bengston & Ruhlen, 1994). Les transformations phonétiques autorisées sont nombreuses et les variations sémantiques excèdent la dizaine⁷. La réception chez les linguistes a été assez dure¹. Les chercheurs de

hypothèse me paraît la suivante : pour construire une famille, il faut accepter la thèse génétique (au XVIIIe siècle le modèle dominant est le *diffusionnisme* qui s'oppose à la généalogie biblique, en proposant l'explication laïque des emprunts dus au commerce ; voir Auroux, 1990) ; or, si on le fait pour les langues indo-européennes, *il faut admettre que l'hébreu ne figure pas parmi nos ancêtres linguistiques*. Ce verrou est largement dû à des facteurs idéologiques ; quand on le fera sauter d'autres facteurs interviendront (opposition sémites/ariens ; voir Olender, 2002).

¹ « [...] sound laws are discovered once one has identified a language family by means of the method of comparison of basic vocabulary ».

² Il s'agit de ce que l'on nomme la « loi de Meillet » pour l'arménien (par exemple, *dwoo en proto-indoeuropéen donne *erko*, deux, en arménien). Voir Lamberterie (1988) qui retrace l'histoire de cette loi et de sa réception, tout en donnant une argumentation solide en sa faveur.

³ Voir *supra* le début de notre section sur « La synthèse impossible ».

⁴ Lorsque Saussure en 1879 décrit le système vocalique proto-indoeuropéen il ne pouvait savoir que Kurylowicz en 1927 retrouverait dans un phonème du hittite, langue que l'on venait de déchiffrer, l'un des éléments qu'il avait reconstruits (voir dans le vol. 3 de Auroux (ed.). 1990-2000, la section intitulée « Des coefficients sonantiques à la théorie des laryngales », rédigée par M.-J. Reichler-Béguelin ; sur la théorie des laryngales, voir Lindeman, 1987). Bloomfield postule l'existence d'une consonne occlusive abstraite dans sa reconstruction du proto-algonquin (*Language I*, 1925, pp. 130-156), dont il confirme par la suite l'existence à partir d'un dialecte algonquin encore peu étudié (*Language IV*, 1928, pp. 99-100).

⁵ De Brosses, comme Monbodo, avait utilisé le couple papa/maman qu'on retrouve dans différentes langues pour justifier le mono-généisme à base organique (utilisation de bilabiales étayées sur le suçotement et de valeurs vocaliques proches de la voyelle générique *a*). L'origine de l'argument a disparu, mais il est toujours présent chez les linguistes (par exemple, Jakobson). Ruhlen tire parti de ce que le mot KAKA (oncle), auquel on n'a jamais donné d'origine organique, se retrouverait dans plusieurs langues pour soutenir que les deux autres termes de parenté sont des preuves en faveur d'une origine commune. Outre la discussion possible sur KAKA, on ne voit pas en quoi est exclue l'origine organique !

⁶ Avec son argumentation propre de recours à un public de non spécialistes contre la communauté des linguistes, qui, conservatrice, brimerait une innovation révolutionnaire. Sur tous ces aspects rhétoriques qui entourent le renouveau de la question de l'origine, voir Nicolai (2002) et Métoz (2005).

⁷ La moyenne serait de 24 (Métoz, 2005, p. 405) ; la racine TIK (doigt, un) se retrouve sous différentes formes phonétiques avec les significations suivantes : ongle, premier, cinq, pied, main, index, plusieurs, seulement, paume, patte, dix, désigner, dire, montrer, chose, orteil.

L'Institut de la Communication Parlée (Grenoble) ont monté un programme de simulation statistique de la méthode utilisée. Les résultats sont écrasants et définitifs : dès que l'on utilise plus de deux changements sémantiques, la probabilité que les rapprochements soient dus au hasard avoisine 1 (Boé & al., 2003). La façon dont Ruhlen utilise la phonétique ne le met pas en meilleure posture : inversion des correspondances de la loi de Grimm (Métoz, 2005, p. 417), autrement dit utilisation réversible des correspondances, étroitesse du matériel phonétique (moins de 3 000 formes différentes, K et ses équivalents représentent 80% de toutes les consonnes dans les référents principaux), tout conduit à rejoindre le rapprochement aléatoire. On est retourné aux vieux démons d'avant Bopp et Grimm. On comprend pourquoi Salmons, dès 1992, qualifiait l' « étymologie globale » de « linguistique pré-copernicienne » ; il s'agit d'un véritable retour, en deçà de la révolution comparatiste, aux méthodes contestées du XVIII^e siècle². Avec Ruhlen la « comparaison multilatérale » aboutit à un véritable fiasco, dont l'étude fournit en quelque sorte le meilleur plaidoyer contemporain pour l'attitude de la Société de linguistique.

12. La « naturalisation » du langage

Si la « méthode multilatérale » et l' « étymologie globale » continuent en quelque sorte une tradition bien ancrée chez les américanistes où elle a connu d'incontestables succès, le rapprochement avec la biologie possède d'autres sources. Comme en témoigne l'introduction de Harnad à la publication du fameux congrès de l'Académie de New-york (Harnad & al., 1976), la motivation provient de la grammaire générative. Dès le départ Chomsky classait la grammaire parmi les sciences biologiques et défendait l'innéisme, à l'aide d'arguments ressassés à l'infini par ses disciples, comme la fameuse sous-détermination de la créativité linguistique de l'enfant par les stimuli environnementaux³. Par rapport au racialisme du XIX^e siècle, les conséquences du recours à la biologie sont inversées : on admet l'évidence (à savoir qu'un enfant peut apprendre n'importe quelle langue), et l'ancrage biologique est requis pour témoigner de *l'universalité de la faculté linguistique humaine*, quitte à noyer la diversité des langues dans des hypothèses *ad hoc*⁴. En outre, ce type de naturalisme a l'avantage de résoudre la question de l'état initial, dont nous avons vu qu'elle a bloqué la démarche « génétique » au XIX^e siècle. Si le langage est inné et son développement chez l'enfant programmé, on peut utiliser les observations⁵ des psychologues spécialistes de l'apprentissage pour aborder le développement linguistique de l'humanité (Pinker, 1984). Chomsky et des disciples comme le psychologue Fodor (Fodor, 1983) se sont toujours contentés d'arguments abstraits en faveur de l'innéisme⁶. La problématique s'est précisée lors de discussions autour d'une forme d'aphasie et celles concernant la nature des créoles.

Au début des années quatre dix, M. Gopnick (Gopnick, 1990 ; Gopnick & Crago, 1991) fait état d'une aphasie spécifique dont la distribution coïncide si bien avec des paramètres familiaux qu'on peut conclure à son caractère héréditaire accentué chez les mâles de la lignée ; il s'agit d'une authentique découverte. Des études épidémiologiques (Tomblin, 1997) le confirmeront au point qu'ils incitent à explorer comment le sexe et la génétique interagissent avec les substrats biologiques du langage⁷. Pinker⁸, n'hésitera pas à lancer l'idée

¹ Dans les années 90, les linguistes se livrent *via* Internet et un forum de discussion (« Linguist List ») à une critique des méthodes utilisées par les protagonistes du retour à l'origine des langues.

² Ses démêlés avec les indo-européanistes sont rapportés comme l' « histoire d'une hystérie » (Ruhlen, 1994a, pp. 76 et s.). Il prend volontiers Meillet comme tête de turc (Ruhlen, 1994a, pp. 78-79) et se réfère à des sottises éculées sur le rôle de D. Jones comme « découvreur » du sanskrit. Le seul argument intéressant est utilisé de façon purement rhétorique, sans véritable argumentation, ni évaluation de l'effet potentiel sur l'argumentation scientifique: le refus d'aller au-delà de l'indo-européen tient à l'eurocentrisme et à la crainte de voir d'autres populations jouer un rôle prépondérant dans l'histoire de l'humanité (Ruhlen s'appuie sur une citation de Sweet (1901) sans référence).

³ Pour une critique, voir Auroux, 1998, pp. 80 et s. ; *passim*.

⁴ Notre équipement inné comporterait différents paramètres correspondant aux différents types de langues et qui seraient sélectivement activés en fonction des environnements d'apprentissage.

⁵ À la condition forte qu'elles soient universelles et indépendantes des langues observées.

⁶ D'autres comme Katz (1981) ont dérivé vers un idéalisme de type platonicien.

⁷ « We are now at the stage where biological and neurodevelopmental theories must be explored to consider ways in which maleness and genotypes interact to influence the neural substrates of language » (Tomblin, 1997, p. 107).

⁸ *The Language Instinct* (1994) est un livre étonnant, dont la structure argumentative dans son recours au public et le mode de promotion ressemblent à celui de Ruhlen (1994a) ; il dispose, en plus, d'un style

qu'il y a des gènes de la grammaire (1994, chap. 10, « Language Organs and grammar Genes »). On remarquera simplement qu'entre l'idée générale, acceptée par tous tant elle est banale et largement confirmée depuis Broca, selon laquelle il y a nécessairement un substrat biologique au langage humain et, l'hypothèse selon laquelle il y a des gènes spécifiques expliquant tel ou tel trait particulier d'une grammaire, il y a un gouffre scientifique¹.

Bickerton, quant à lui, travaille plus spécifiquement sur du matériel linguistique de terrain, les créoles². Selon lui l'apparition de tous les créoles est marquée par les mêmes traits. Par exemple, il remarque que les créoles n'ont pas d'articles et que l'existence d'articles ne figure pas dans les *requisites* de la grammaire universelle (Bickerton, 1988, p. 279). De là l'hypothèse que les créoles témoignent de la résurgence du « bio programme » linguistique de l'humanité ; ils sont la réalisation du programme biologique avec le minimum d'ajout culturel³. La condition est évidemment que l'on retrouve partout les mêmes traits, et dans tous les créoles et dans tous les processus d'apprentissage. Les créolistes ont largement refusé, sur des bases empiriques, ces hypothèses (voir un résumé de la discussion dans Romaine, 1988, pp. 256-310). L'hypothèse du « bio programme » assume que les créoles sont des systèmes plus simples que les autres langues, plus semblables entre eux qu'avec les autres langues et « mixtes » ; ce sont des caractéristiques qui ne vont pas sans poser de problèmes, notamment au vu de l'histoire des langues (voir Nicolai, 1990). Les créoles sont des langues comme les autres (Muysken, 1988, p. 300⁴) et le « bio programme » une théorie largement arbitraire.

13. Les chances de la « nouvelle synthèse »

À la fin des années quatre vingt paraissent un certain nombre de travaux dont le rapprochement (Métoz, 2005, p. 571) est significatif : Greenberg & al., 1986, Greenberg, 1987 ; Ruhlen, 1986 ; Renfrew, 1987 ; Cavalli-Sforza, 1988. Ils ont pour traits communs de recourir à des données provenant d'au moins deux, sinon toutes les disciplines suivantes : linguistique, archéologie, biologie. Il s'agit de « bringing together Genetic, Archeological and Linguistic Data » comme le précise le titre de Cavalli-Sforza (1988). C'est ce que l'on reconnaîtra comme « nouvelle synthèse en cours d'émergence » (Renfrew, 1992a, 1992b). Ce thème identifie un groupe de chercheurs assez distinct du précédent⁵, on voit bien comment leur rapprochement a pu naître de certaines difficultés. Par exemple, les trois macro familles linguistiques amérindiennes auxquelles aboutit Greenberg nécessitent pour produire la

aisé et familier qui n'hésite pas à recourir à de petits poèmes. Nous sommes dans la vulgarisation de théories qui n'ont pas reçues l'approbation de la communauté scientifique, mais dont les auteurs disposent d'une reconnaissance institutionnelle forte (Pinker dirige un Laboratoire du MIT).

¹ Prenons une analogie. Il y a un substrat physico-chimique à tous les phénomènes biologique. Cependant les lois de la biologie ou ses objets spécifiques (par exemple, les hormones ou les chromosomes) ne se déduisent pas mécaniquement des lois et des objets de la biologie.

² Les études créoles démarrent dans les années quatre vingt du XIXe siècle (voir Adam (1883) et les textes rassemblés dans Schuchardt 1922). On pourrait se demander pourquoi elles ne sont pas, dès cette époque, intervenues directement dans la question de l'origine des langues. D'abord, les créoles sont considérés comme des langues formées ; ensuite, ils invalident les hypothèses biologiques sous la forme racialisée (une langue = une race) spécifique de l'époque, puisqu'on parle à leur propos de langues hybrides (Adam) ou mixtes (Schuchardt) ; enfin, ils ne respectent pas le modèle arborescent (1 langue mère, n descendantes par divergence), en mettant en lumière le schéma inverse (n « langues mères », 1 descendante par convergence). Au mieux, ils posent des problèmes au modèle génétique du comparatisme (voir Baggioni (1988), sur la polémique Meillet/Schuchardt), tout en renforçant l'aspect historique et contingent de l'évolution linguistique (Schuchardt finira par admettre que toutes les langues sont des langues mixtes).

³ Les fondements biologiques universels du langage peuvent donc disparaître d'une langue en fonction de son évolution historique (*contra*, Chomsky pour qui les universaux sont présents dans toutes les langues). Cf. Romaine, 1988, p. 268.

⁴ « The very notion of a « creole » language from the linguistic point of view tends to disappear if one looks closely ; what we have is just a language ». Le bio programme est le seul cas que nous rencontrons où les partisans du biologisme ne font pas preuve d'un « politiquement correct » impeccable et rejoignent les préjugés raciaux de la fin du XIXe siècle (le créole, le « petit nègre », est un langage simplifié dans la mesure où ceux qui le parlent sont incapables d'apprendre les langues de culture plus compliquées). Voir Hagège, 1985, chap. 2, « Le laboratoire créole » : les créoles ne sont pas des langues simples aux traits rudimentaires, ce sont des langues riches en traits dominants.

⁵ Pinker participe souvent à des recueils collectifs qui concernent le groupe, mais il ne semble faire que de la figuration. Voir, par exemple, la généralité et la platitude consternantes de sa contribution à Jablonski & Aiello, 1998.

diversification observable actuellement un temps beaucoup plus long (au moins 50 000 ans) que celui que les données archéologiques reconnaît au peuplement de l'Amérique (environ 12 000 ans). D'où le recours à un modèle biologique fondé sur la dentition pour retrouver trois vagues de peuplement (Greenberg & al., 1987). Renfrew (1992b, p. 175) présente un concours harmonieux de disciplines: à la base, la génétique moléculaire qui donne des modèles de variabilité génétique, puis parallèlement les modèles du développement culturel et ceux du changement linguistique, qui aboutissent aux données de l'archéologie préhistorique et de la linguistique préhistorique. Toutefois la valeur d'une synthèse dépend de celle de ses éléments. Nous avons déjà analysé le fiasco linguistique.

Le projet de Cavalli-Sforza repose sur un champ dont il a été le pionnier, la recherche des lignées génétiques des populations humaines sur la base de l'ADN. On en voit la nouveauté en le comparant avec la « statistique linguistique » de Balbi (1826). Balbi tenait pour acquis qu'une classification linguistique (effectuée sur la base d'une comparaison restreinte de vocabulaires) était *ipso facto* une classification ethnographique. Hypothèse qui s'invalide avec les mêmes arguments que Broca employait contre le racialisme de Chavée. Cavalli-Sforza fournit une classification génétique de la population¹ mondiale **et** la compare avec une classification des langues. C'est la coïncidence des deux qui a fait le succès médiatique du travail. Or c'est justement dans cette coïncidence qu'il faut voir des traces d'inconséquences (voir Métoz, 2005). On a fait remarquer qu'une partie des populations considérées n'avaient d'autre existence que celle de la famille linguistique : la coïncidence étonnera d'autant moins. Sur un échantillon de trente-huit populations humaines, dix-neuf ne montrent aucune corrélation entre les gènes et les langues ; à 50% le rapprochement ne prouve plus grand-chose ! Enfin, certains cas sont franchement des contre exemples : les hongrois sont génétiquement apparentés à leurs voisins européens, leur langue (famille finno-ougrienne) ne l'est pas ; les pygmées n'ont rien à voir génétiquement avec leurs voisins bantous, mais ils parlent des langues bantou, etc. Aux arguments de Broca rappelant la contingence de l'histoire, on peut ajouter les effets de la convergence aréale : certains traits typologiques (ordre des mots) manifestent une relative stabilité géographique, mais pas généalogique (Nichols, 1992 ; Fox, 1995, p. 297).

Pour sa part, l'archéologue Renfrew retrouve le type de modèle proposé par Leroi-Gourhan quelque vingt ans auparavant, avec une grande fécondité comme on l'a vu: lier évolution linguistique, culturelle et technique, ces dernières bénéficiant de traces archéologiques (Renfrew, 1988, p. 438). Malheureusement, il se concentre sur une seule technologie prédominante dans le domaine indo-européen et propose *une relation causale entre la diversification des langues et le développement de l'agriculture* : c'est le « farming/language dispersal model ». Les explications sont variables : suprématie d'une élite, effondrement du système et vagues d'avancées. Comme l'ont remarqué de nombreux critiques, on est loin de tous les aléas possibles de l'histoire qui concernent bien d'autres activités que l'agriculture. On garde toutefois une certaine prudence: il se pourrait que le langage évolue si rapidement qu'aucune conclusion ne saurait être tirée antérieurement à une période de 5 000 ans (Renfrew, 1992b, p. 473).

Il y a plus d'effets d'annonce dans la « nouvelle synthèse » que de vrais résultats scientifiques. Certes on a trouvé de nouvelles catégories de données, renouvelé les arguments du XIX^e siècle et, parfois, trouvé un biais pour avancer comme l'a fait la paléontologie humaine, qui constitue la véritable nouveauté. Il est irritant de voir ces continuelles proclamations de « révolutions scientifiques » annulant les connaissances précédentes, comme si l'on avait dans l'histoire l'exemple d'une seule « science » qui se renouvelle de fond en comble à chaque génération qui souhaite faire carrière. On peut expliquer le succès du renouveau du thème par la médiatisation et l'intérêt du grand public pour ce genre de question à conséquences métaphysiques². À un moment, il faut bien en revenir à la valeur cognitive des discours « scientifiques » : les conséquences théoriques de l'interdit de la Société de linguistique sont infiniment plus solides que ne le laissent supposer quelques affirmations tonitruantes de ces trente dernières années.

¹ Que la méthode suivie ait été largement contestée par les spécialistes de génétique des populations ne nous concernera pas ici.

² Je renvoie à la remarquable démonstration de Nicolai (R.). 2000.

Références bibliographiques

- Aarsleff (H.). [1967] 1983. *The Study of Language in England, 1780-1860*. Princeton N.J.
- Adam (L.). 1883. *Les idiomes négro-aryens. Essai d'hybridologie linguistique*. Paris.
- Auroux (S.). 1979a. « La querelle des lois phonétiques ». in : *Linguisticae investigationes* 3, pp. 1-27.
- Auroux (S.). 1979b. « La catégorie du parler et la linguistique ». in : *Romantisme*, 25-26, pp. 157-178.
- Auroux (S.). 1979c. *La sémiotique des encyclopédistes*. Paris : Payot.
- Auroux (S.). 1983a. « La première société de linguistique : Paris 1837 ? ». in : *Historiographia Linguistica*, X.3, pp. 195-219.
- Auroux (S.). 1983b. « Le signe et la culture dans la linguistique des Lumières ». in : *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 36, pp. 509-513.
- Auroux (S.). 1984. « Linguistique et anthropologie en France 1600-1900 ». in : Rupp-Eisenreich (B.), (ed.). *Histoires de l'Anthropologie : XVI-XIX siècles*. Paris : Kincksieck, pp. 291-318.
- Auroux (S.). 1990. « Representation and the place of linguistic change before comparative grammar ». in : Mauro (T. de) & Formigari (L.), (eds.). *Leibniz, Humboldt and the Origins of Comparativism*. Amsterdam : John Benjamins, pp. 213-238.
- Auroux (S.). 1998. *Le langage, la raison et les normes*. Paris : P.U.F.
- Auroux (S.). 2004. « Victor Henry Kantien ? Entretien avec C. Puech ». in : Puech (C.), (ed.). 2004, pp. 25-38.
- Auroux (S.), (ed.). 1989-2000. *Histoire des idées linguistiques*. Liège : Mardaga [3 volumes].
- Auroux (S.) & Boès (A.). 1981. « Court de Gébelin (1725-1784) et le comparatisme. Deux textes inédits ». in : *Histoire Épistémologie Langage*, III-2, pp. 21-67.
- Auroux (S.), Chevalier (J.-C), Marcello-Nizia (C.) & Jacques-Chaquin (N.), (eds.). 1985. *La linguistique fantastique*. Paris.
- Baggioni (D.). 1988. « La polémique Meillet-Schuchardt sur la parenté des langues (1905-1923) ». in : *Histoire Épistémologie Langage*, X-2, pp. 85-97.
- Balbi (A.). 1826. *Atlas ethnographique du globe ou classification des peuples anciens et modernes d'après leurs langues*. Paris : Rey et Gravier.
- Bateman (R.), Goddard (I.), O'Grady (R.), Funk (V.-A.), Mooi (R.), Kress (J.) & Cannell (P.). 1990. « Speaking of Forked Tongues : The Feasibility of Reconciling Human Phylogeny and the History of Language ». in : *Current Anthropology*, 31-1, pp. 1-24.
- Bellwood (P.) & Renfrew (C.), (eds). 2003. *Examining the farming/language dispersal Hypothesis*. Cambridge : McDonald Institute for Archeological Research.
- Benware (W.-A.). 1974. *The Study of Indo-European Vocalism in the 19th Century, from the beginnings to Whitney and Scherer*. Amsterdam : John Benjamins B.V.
- Beyli (V.-V.). 1997. « Rafinesque's Linguistic Activity ». in : *Anthropological Linguistics*, 39-1, pp. 60-73.
- Bickerton (D.). 1981. *Roots of language*. Ann Arbor : Karoma.
- Bickerton (D.). 1984. « The Language bioprogram hypothesis ». in : *Behavioral and Brain Sciences*, 7-2, pp. 173-221.
- Bickerton (D.). 1988. « Creole language and the bioprogram ». in : Newmeyer (F.-J.), (ed.). *Linguistics : The Cambridge Survey. II : Linguistic Theory : Extensions and Implications*, pp. 268-284.
- Bickerton (D.). 1990. *Language and Species*. Chicago : The University of Chicago Press.
- Blench (R.) & Spriggs (M.), (eds.). 1998. *Archeology and Language. Correlating archeological and Linguistic Hypothesis*. Londres : Routledge.
- Boé (L.-J.) & al. 2003. « When Ruhlen « mother tongue » theory meets the null hypothesis ». in : *Communication au 15^e ICPHS*. Barcelone.
- Bourdier (F.). 1978. « Les recherches de Renan sur l'origine de la parole (1841-1858) » ; suivi d'un texte inédit : « Essai sur la formation du langage (1845) ». in : *Cahiers du cercle E. Renan*, 107, pp. 138-170.
- Bréal (M.). 1876. « Les racines indo-européennes ». in : *Journal Asiatique* [repris dans 1877].
- Bréal (M.). 1877. *Mélanges de mythologie et de linguistique*. Paris.
- Bréal (M.). 1897. *Essai de sémantique*. Paris [cité d'après l'édition de 1904, qui contient également le texte suivant].

- Bréal (M.). 1900. « Les commencements du verbe ». in : *Mémoires de la Société de linguistique* IX [voir le précédent].
- Brinton (D.-G.). 1887. « The Rate of Change in American Language ». in : *Science*, 10, p. 274.
- Brinton (D.-G.). 1891. *The American Race*. New York : Hodges Publisher.
- Broca (P.). 1861. « Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie ». in : *Bulletin de la Société anatomique*, 36, pp. 330-357.
- Broca (P.). 1863. « La linguistique et l'anthropologie ». in : *Bulletin de la Société d'Anthropologie*, 3, pp. 264-319.
- Broca (P.). 1865. « Du siège de la faculté de langage articulé dans l'hémisphère gauche du cerveau ». in : *Bulletin de la Société d'Anthropologie*, 6, pp. 377-393.
- Brosses (C. de). 1765. *Traité de la mécanique des langues ou des principes physiques de l'étymologie*. Paris.
- Brugmann (K.). 1897. *The Nature and Origin of the Noun Genders in the Indo-European Languages*. New York.
- Burggraff (P.). 1865. *Principes de grammaire générale, ou Exposition raisonnée des éléments du langage*. Liège.
- Burnett (J.), Lord Monboddo. 1773-1792. *Of the Origin and Progress of Language*. Edimbourg [6 volumes].
- Campbell (L.). 1986. « A Comment on The Settlement of Americas : A Comparison of the Linguistics, Dental and Genetic Evidence by Greenberg et al. ». in : *Current Anthropology*, 27-5, pp. 488-489.
- Campbell (L.). 1988. « Review of Languages in the Americas by Joseph Greenberg ». in : *Language*, 64, pp. 591-615.
- Carrau (L.). 1978. « L'expression des émotions et l'origine du langage ». in : *Revue des deux mondes*, XXIX, pp. 175-195.
- Cassirer (E.). 1923. *Philosophie der symbolischen Formen*, T. I., *Die Sprache* (1923). Paris [1972, cité d'après la t. f.].
- Cavalli-Sforza (L.). 1996. *Gènes, peuples et langues*. Paris : Odile Jacob.
- Cavalli-Sforza (L.), Piazza (A.), Menozzi (P.) & Mountain (J.). 1988. « Reconstruction of Human Evolution : Bringing together Genetic, Archeological and Linguistic Data ». in : *Proceedings of the National Academy of Science*, 85, pp. 6002-6006.
- Cavalli-Sforza (L.), Piazza (A.), Menozzi (P.), Mountain (J.). 1989. « Genetic and Linguistic Evolution ». in : *Science*, 244, pp. 1128-1129.
- Chavée (H.-J.). 1849. *Lexiologie indo-européenne* [...]. Paris.
- Chavée (H.-J.). 1862a. *Les Langues et les races*. Paris : Chamerot [version remaniée de *Moïse et les langues*, 1853].
- Chavée (H.-J.). 1862b. « Sur le parallèle des langues sémitiques et des langues indo-européennes ». in : *Bulletin de la Société d'anthropologie de Paris* III, pp. 198-205.
- Chavée (H.-J.). 1878. *Idéologie lexicologique des langues indo-européennes*. Paris.
- Curtius (G.). 1867. *Zur Chronologie der indogermanischen Sprachforschung*. Leipzig [t.f. par A. Bergaigne, Paris 1869].
- Chervel (A.). 1979. « Le débat sur l'arbitraire du signe au XIX^e siècle ». in : *Romantisme*, 25/26, pp. 3-33.
- Christy (C.). 1983. *Uniformitarianism in Linguistics*. Amsterdam.
- Collinge (N.-E.). 1985. *The Laws of Indo-European*. Amsterdam : John Benjamins.
- Daniel (G.). 1981. *A short History of Archaeology*. Londres [cité d'après l'édition Paperback de 1983].
- Desmet (P.). 1996. *La linguistique naturaliste en France (1867-1922). Nature, origine et évolution du langage*. Leuven, Paris : Peeters.
- Dessales (J.-L.). 2000. *Aux origines du langage. Une histoire naturelle de la parole*. Paris : Hermès.
- Derrida (J.). 1962. Introduction et traduction de *l'Origine de la géométrie de Husserl*. Paris : P.U.F.
- Derrida (J.). 1967a. « La linguistique de Rousseau ». in : *Revue Internationale de philosophie*, 82-4, pp. 443-462.
- Derrida (J.). 1967b. *De la Grammatologie*. Paris : Minuit.
- Droixhe (D.). 1978. *La linguistique et l'appel de l'histoire*. Genève : Droz.
- Dubois (C.-G.). 1970. *Mythe et Langage au seizième siècle*. Bordeaux.

- Dumont d'Urville (J.). 1834. *Voyage de découverte de l'Astrolabe (...). Philologie, seconde partie (...)*. Paris : Ministère de la Marine.
- Egger (V.). 1881. *La parole intérieure ; essai de psychologie descriptive*. Paris [thèse faculté des lettres de Paris].
- Farrar (F.-W.). 1860. *An Essay on Origin of Language based on modern researches especially on the Work of M. Renan*. Londres.
- Farrar (F.-W.). 1865. *Chapters on Language*, Londres.
- Farrar (F.-W.). 1869. *Families of Speech*, Londres.
- Farrar (F.-W.). [1877] 1883. *Language and Languages*. Londres [= 1865 et 1869].
- Fauchet (C.). 1851. *Recueil de l'origine de la langue et poésie française [...]*. Paris.
- Fodor (J.-A.). 1983. *The Modularity of Mind*. Cambridge, MA : M.I.T. Press.
- Fox (A.). 1995. *Linguistic Reconstruction. An Introduction to Theory and Method*. Oxford : Oxford University Press.
- Gaugengigl (I.). 1846. *Der geistliche Ursprung der Sprache*. Passau.
- Gébelin (A. Court de). 1773-1782. *Monde primitif analysé et comparé avec le monde moderne*. Paris : l'Auteur [9 volumes in 4°].
- Geiger (L.). 1868-1872. *Ursprung und Entwicklung der Menschlichen Sprache und Vernunft*. Stuttgart [2 volumes].
- Geiger L. 1878. *Der Ursprung der Sprache*. Stuttgart.
- Gerber (G.). 1884. *Die Sprache und das Erkennen*. Berlin.
- Gessinger (J.) & Rahden (von W.), (eds). 1989. *Theorien vom Ursprung der Sprache*. Berlin, New York : Walter de Gruyter.
- Gilliéron (J.). 1919. *La faillite de l'étymologie phonétique*. Neuveville.
- Girard de Rialle (J.). 1875. « Le transformisme en linguistique ». in : *Revue Scientifique*, 3 mars 1875.
- Gopnick (M.). 1990. « Feature blind grammar and dysphasia ». in : *Nature*, 344, 715.
- Gopnick (M.), (ed.). 1997 *The inheritance and innateness of grammars*. Oxford : Oxford University Press.
- Gopnick (M.) & Crago (M.-B.). 1991. « Familial aggregation of a developmental language disorder ». in : *Cognition*, 39, pp. 1-50.
- Greenberg (J.-H.), Turner (C.) & Zegura (S.). 1986. « The Settlement of the Americas : A Comparison of the Linguistic, Dental and Genetical Evidences ». in : *Current Anthropology*, 90-4, pp. 1029-1030.
- Greenberg (J.-H.). 1987. *Language in the Americas*. Stanford.
- Grimm (J.). 1652, *Ueber den Ursprung der Sprache*. Berlin.
- Gudschinsky (S.). 1956. « The ABC's of Lexicostatistics (Glottochronologie) ». in : *Word*, 12, pp. 175-210.
- Hagège (C.). 1985. *L'homme de paroles. Contribution linguistique aux sciences humaines*. Paris : Fayard.
- Hales (H.). 1886. « The Origin of Language and the Antiquity of Speaking Man. An Address before the Section of Anthropology of the American Association for the Advancement of Sciences ». in : *Proceedings of the American Association for the Advancement of Sciences*, 35, pp. 279-323.
- Harnad (R.), Setkalis (H.-D.) & Lancaster (J.), (eds.). 1976. *Origins and Evolution of Language and Speech*. New York : Annals of the New York Academy of Science, 280.
- Harris (R.), (ed.). 1996. *The Origin of Language*. Bristol : Thoemmes Press [choix de textes d'auteurs du 19^e siècle sur la question].
- Harvey (J.). 1984. « L'évolution transformée : positivistes et matérialistes dans la Société d'anthropologie de Paris du Second Empire à la III^e République ». in : Rupp-Eisenreich (B.), (ed.). *Histoires de l'Anthropologie : XVI-XIX siècles*. Paris : Kincksieck, pp. 387-410.
- Henry (V.). 1883. *De sermonis humani origine et natura M. Teurentius Varro quid senserit*. [thèse complémentaire de la Faculté des Lettres de Paris].
- Henry (V.). 1888. « cr de Regnaud 1887 ». in : *Revue Critique*, 5 mars 1888.
- Henry (V.). 1896. *Antinomies linguistiques*. Paris.
- Heyse (K.-W.-L.). 1856. *System der Sprachwissenschaft*. Berlin [ed. par Steinthal].
- Hjelmslev (L.). [1963] 1966. *Le langage*. Paris : Minuit [trad. fr.].
- Hoijer (H.). 1961. « Anthropological Linguistics ». in : Mormann (C.), (ed.). *Trends in European and American Linguistics 1930-1960*, pp. 117-127.
- Hombert (J.-M.), (ed.). 2005. *Aux origines des langues et du langage*. Paris : Fayard.

- Hourwitz (Z.). 1798. *Polygraphie ou l'Art de correspondre, à l'aide d'un dictionnaire dans toutes les langues, même celles qui ne possèdent pas seulement les lettres alphabétiques*. Paris [An IX].
- Hourwitz (Z.). 1807. *L'origine des langues ; suivi d'une notice par J.-D. Lanjuinais*. Paris.
- Hovelacque (A.-A.). 1869. « cr de L. de Rosny 1869 ». in : *Revue de Linguistique et de Philologie comparée*, III-I, pp. 102-107.
- Hovelacque (A.-A.). 1873. « La linguistique et le précurseur de l'homme ». in : *Compte rendu de la 3ème session de l'association française pour l'avancement des sciences, 1873*, pp. 613-619.
- Hovelacque (A.-A.). [1876] 1877. *La linguistique*. Paris.
- Hovelacque (A.-A.) & Vinson (J.). 1878. *Études de linguistique et d'ethnographie*. Paris.
- Humboldt (W. von). 1820. *Ueber das vergleichende Sprachstudium in Beziehung auf die verschiedenen Epochen der Sprachentwicklung* [cité d'après *Werke* in 5 Bänden, t. III, Berlin 1963].
- Humboldt (W. von) 1821. *Ueber das Entstehen der grammatischen Formen* [cité d'après la trad. fr. de A. Tonnelat [1869], Bordeaux : ed. Ducros, 1969].
- Hymes (D.). 1973. « Lexicostatistics and Glottochronology in the Nineteenth Century ». in : *Lexicostatistics in Genetic Linguistics : Proceedings of the Yale Conference*. La Haye : Mouton/de Gruyter, 1, pp. 122-176.
- Hymes (D.), (ed.). 1964. *Language in Culture and Society. A Reader in Linguistics and Anthropology*. New York, Londres : Harper & Row.
- Jablonski (N.) & Aiello (L.), (eds.). 1998. *The Origin and Diversification of Language*. San Francisco : Memoirs of the California Academy of Science, 24.
- Jehan (L.-F.). 1858. *Dictionnaire de linguistique et de philologie comparée. Histoire de toutes les langues mortes et vivantes* [Encyclopédie théologique de l'Abbé Migne, t. 34].
- Jespersen (O.). 1822. *Language : its Nature Development and Origin*. London.
- Katz (J. -J.). 1981. *Language and Other Abstracts Objects*. Totowa : Rowman and Littlefield.
- Kelle (J.-N. von). 1856. « Gedanken über den Ursprung der Sprache ». in : *Archiv für das Studium der neueren Sprachen und Literaturen*, XX, pp. 297-332.
- Koerner (E.-F.-K.). 1978. *Toward a Historiography of Linguistics, Selected Essays*. Amsterdam : John Benjamins.
- Koerner (E.-F.-K.). 1982. *Ferdinand de Saussure. Genesis J Evolucion de su pensamiento en el marco de la linguística occidental*. Madrid [trad. de l'anglais Braunschweig, 1973].
- Kruse (O.-F.). 1827. *Freimethige Bemerkungen über den Ursprung der Sprache oder : Beweiss dass die Sprache nicht menschlichen Ursprungs sey*. Altona.
- Kurylowicz (J.). 1927. « i indoeuropéen et h hittite ». in : *Symbolae Grammaticae in honorem Ioannis Rozwadowski*. Cracovie, I, pp. 95-104.
- Kurylowicz (J.). 1935. *Études indo-européennes I*. Cracovie.
- La Calle (A. de). 1881. *Glossologie [...]*. Paris.
- La Grasserie (R. Guérin de). 1888. *Des divisions de la linguistique*. Paris.
- La Grasserie (R. Guérin de). 1895. *De l'origine et de l'évolution première des racines des langues*. Paris.
- Lamberterie (C. de). 1988. « Meillet et l'arménien ». in : *Histoire, Épistémologie, Langage*, X-2, pp. 217-234.
- Laming-Emperaire (A.). 1964. *Origines de l'archéologie préhistorique en France. Des superstitions médiévales à la découverte de l'homme fossile*. Paris.
- Leroi-Gourhan (A.). 1964. *Le geste et la parole*. Paris : Albin Michel [2 volumes].
- Lindeman (F.-O.). 1987. *Introduction to the « Laryngeal Theory »*. Oslo : Norwegian University Press.
- Mellars (P.). 1998. « Neanderthals, Modern Humans and the Archeological Evidence for Language ». in : Jablonski & Aiello, pp. 89-115.
- Meillet (A.). 1908. *Les dialectes indo-européens*. Paris : Champion.
- Meillet (A.). 1914. « Le problème de la parenté des langues ». in : *Scientia*, XV [repris dans Meillet, 1921, pp. 76-101].
- Meillet (A.). 1920. « Les parentés de langue ». in : *Bulletin de la Société de linguistique de Paris*, XXI [repris dans Meillet, 1921, pp. 102-109].
- Meillet (A.). 1921. *Linguistique historique et linguistique générale*. Paris : Champion.
- Meillet (A.). 1937. *Introduction à l'étude comparative des langues indo-européennes*. Paris : Hachette.

- Merian (A.-A., Baron von). 1828. *Principes de l'étude comparative des langues, suivi d'Observations sur les racines des langues sémitiques par M. Klaproth*. Paris.
- Métoz (L.). 2005. *De la glottochronologie à la « Nouvelle Synthèse »*. in : Métoz (L.). *Histoire de la linguistique comparative américaine de 1950 à nos jours : concepts et méthodes*. Lyon [Thèse de Doctorat de l'École normale supérieure Lettres et Sciences humaines].
- Mikch (J.). 1886. « L'idée et la racine ». in : *Revue de linguistique et de philologie comparée*, XIX, pp. 189-206.
- Milner (J.-C.). 1878. *L'amour de la langue*. Paris.
- Milner (J.-C.). 1984. « La constitution du fait en linguistique ». in : Achard (P.) & al., (eds.). *Histoire et linguistique*. Paris, pp. 177-190.
- Monboddo [voir Burnet].
- Mortillet (G. de). 1873. « Le précurseur de l'homme ». in : *Compte rendu de la 3ème session de l'association française pour l'avancement des sciences*, pp. 607-613.
- Mounin (G.). 1967. *Histoire de la linguistique des origines au XX^e siècle*. Paris.
- Müller (M.-F.). 1861. *Lectures on the Science of Language delivered at the Royal Institution of Great Britain (1861)*. Londres.
- Müller (M.-F.). 1863. *idem (1863)* [cité d'après la trad. fr. par Harris (G.) & Perrot (G.). *Nouvelles Leçons sur la science du langage*. Paris. 1867-1868 [2 volumes].
- Muller (J.-C.). 1984. « Quelques repères pour l'histoire de la notion de vocabulaire de base dans le pré-comparatisme ». in : *Histoire Épistémologie Langage*, VI-2, pp. 37-43.
- Muysken (P.). 1988. « Are Creoles a special type of Language ? ». in : Newmeyer (F.-J.), (ed.). *Linguistics : The Cambridge Survey. II : Linguistic Theory : Extensions and Implications*. pp. 268-284.
- Nerlich (B.). 1984. *La pragmatique : tradition ou révolution dans l'histoire de la linguistique française*. Düsseldorf [thèse dactylographiée].
- Nichols (J.). 1998. « The Origin and Dispersal of Languages : Linguistic Evidence ». in : Jablonski & Aiello (eds.). pp. 117-170.
- Nichols (J.). 1999. *Linguistic Diversity in Space and Time*. Chicago : University of Chicago Press.
- Nicolaï (R.). 1990. *Parentés linguistiques (à propos du songhai)*. Paris : Éditions du CNRS.
- Nicolaï (R.). 2000. *La traversée de l'empirique. Essai d'épistémologie sur la construction des représentations de l'évolution des langues*. Gap : Ophrys.
- Nodier (C.). 1808. *Dictionnaire raisonné des onomatopées françaises*. Paris.
- Nodier (C.). 1810. *Archéologue ou système universel et raisonné des langues. Prologomènes*. Paris.
- Noiré (L.). 1877. *Der Ursprung der Sprache*. Mainz.
- Noiré (L.). 1879. *Max Müller und die Sprachphilosophie*. Mainz.
- Olender (M.). [1989] 2002. *Les langues du paradis. Aryens et Sémites : un couple providentiel*. Paris : Seuil.
- Olender (M.). 2005. *La chasse aux évidences. Sur quelques formes de racismes entre mythe et histoire*. Paris : Galaade Éditions.
- Paul (H.). 1880. *Principien der Sprachgeschichte* [cité d'après la 9ème édition, Tübingen 1975].
- Paul (H.). 1907. « Der Ursprung der Sprache ». in : *Beilage für allgemeine Zeitung*, 13, pp. 14 s.
- Pictet (A.). 1859-1863. *Les origines indo-européennes, ou, les Aryas primitifs ; essai de paléontologie linguistique*. Paris [2 volumes].
- Pinker (S.). 1994. *The Language Instinct. How the Mind Creates Language*. New York : W. Morrow and C^o [cité d'après l'édition HarperPerennial, 1995].
- Pinker (S.). [1984] 1996. *Language Learnability and Language Development. With New Commentary by the Author*. Harvard University : Harvard University Press.
- Poliakov (L.). [1971] 1987. *Le mythe aryen. Essai sur les sources du racisme et des nationalismes*. Bruxelles : Éditions Complexe.
- Porset (C.). 1970. Édition, introduction et notes de *l'Essai sur l'origine des langues* de J.-J. Rousseau. Bordeaux : Ducros.
- Porset (C.). 1977. « L'idée et la racine ». in : *Revue des sciences humaines*, 166, pp. 185-204.
- Pott (A.-F.). 1884. « Einleitung in der Sprachwissenschaft, 1.1. Ursprung der Sprache ». in : *Internationale Zeitschrift für allgemeine Sprachwissenschaft* [cité d'après le reprint de Koerner (K.). 1974. Amsterdam].

- Powell (J.-W.). 1880, *Introduction to the Study of Indian Languages, with words, phrases and sentences to be collected*. Washington : Bureau of American Ethnology.
- Puech (C.), (ed.). 2004. *Linguistique et partages disciplinaires à la charnière des XIX^e et XX^e siècles : Victor Henry (1850-1907)*. Louvain, Paris : Éditions Peeters [Avec une bibliographie de V. Henry par M. Décimo, Bibliothèque de l'Information grammaticale, B.I.G. 55].
- Regnaud (P.). 1886. *Essais de linguistique évolutionniste, application d'une méthode générale à l'étude du développement des idiomes indo-européens*. Paris.
- Regnaud (P.). 1887. *Les lois phonétiques sont-elles absolues au sens où l'entendent les néogrammairiens ? Non*. Paris.
- Regnaud (P.). 1888a. « La question de la restitution de la langue mère indo-européenne ». in : *Revue de linguistique et de Philologie comparée*, XXI, pp. 174-182.
- Regnaud (P.). 1888b. *Origine et philosophie du langage ou principes de linguistique indo-européenne*. Paris. [seconde édition augmentée, Paris, 1889 (nos citations sont faites d'après cette édition)].
- Regnaud (P.). 1889. « Le système de l'agglutination devant la logique et devant les faits ». in : *Revue de linguistique et de Philologie Comparée*, XXII, pp. 60-65.
- Regnaud (P.). 1904. *L'origine des idées éclairée par la science du langage*. Paris.
- Renan (E.). 1858. *De l'origine du langage (1848)*. Paris [seconde édition revue et augmentée].
- Renan (E.). 1878. « Des services rendus aux sciences historiques par la philologie ». in : Conférence prononcée à la Sorbonne [citée d'après O. C., Calmann-Lévy, 1958, t. VIII, pp. 1213-1232].
- Renfrew (C.). 1987. *Archeology and Language : The Puzzle of Indo-European Origins*. Londres.
- Renfrew (C.). 1989. « Models of Change in Language and Archeology ». in : *Transactions of the Philological Society*, 87, pp. 103-155.
- Renfrew (C.). 1992a. « World Language and Human Dispersals : A Minimalist View ». in : Hall (J.-A.) & Garwey (I.-C.), (eds.). *Transition to Modernity : Essays on Power, Wealth and Belief*. Cambridge, pp. 11-68.
- Renfrew (C.). 1992b. « Archeology, Genetics and Linguistic Diversity ». in : *Man*, 27, pp. 445-478.
- Renfrew (C.). 1998. « The Origins of World Linguistic Diversity : An Archeological perspective ». in : Jablonski & Aiello (eds). 1998. pp. 171-192.
- Ringe (D.-A.). 1992. « On Calculating the Factor of Chance in Language Comparison ». in : *Transactions of the American Philosophical Society*, 82-1.
- Romaine (S.). 1988, *Pidgin & Creole Languages*. Londres, New York : Longman.
- Rosny (L.-L.-P. de). 1869. *De l'origine du langage*. Paris.
- Rosny (L.-L.-P. de). 1901. « De l'étude de la linguistique comparée du point de vue de l'ethnographie ». in : *Mémoires de la société d'ethnographie*, I, 2, pp. 141-172.
- Rosny (L.-L.-P. de). 1904. *L'Amérique précolombienne, études d'histoire, de linguistique et de paléographie sur les anciens temps du nouveau monde*. Paris.
- Rupp-Eisenreich (B.), (ed.). 1984. *Histoires de l'Anthropologie : XVI-XIX siècles*. Paris : Klincksieck.
- Rulhen (M.). 1987. *A Guide to World's Languages*, vol ; 1 : *Classification*. Stanford : Stanford University Press.
- Rulhen (M.). 1994a. *The Origin of Language. Tracing the Evolution of the Mother Tongue*. New York : John Wiley & Sons Inc.
- Rulhen (M.). 1994b. *On the Origin of Language: Studies in Linguistic Taxonomy*. Stanford : Stanford University Press.
- Salmons (J.). 1992. « « Global Etymology » as Pre-Copernician Linguistics ». in : *California Linguistic Notes*, 25-1, pp. 6-15.
- Sapir (E.). 1905. *Herder's Ursprung der Sprache*. Columbia University [Master thesis, repr. *Historiographia Linguistica*, XI.3, 1984, pp. 355-388].
- Sapir (E.). 1921. *Language. An Introduction to the Study of Speech*. New York.
- Saussure (F. de). 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipzig.
- Saussure (F. de). 1910. *Cours de linguistique générale* [cité d'après l'édition critique de Engler R. en 4 fascicules, Bern, 1968-1974].
- Saussure (F. de). 2002. *Ecrits de linguistique générale*. Paris : Gallimard [texte établi et édité par Simon Bouquet et Rudolf Engler].

- Schlegel (F.). 1808. *Ueber die Sprache und Weisheit der Indier*. [rep. Amsterdam : John Benjamins, 1977].
- Schleicher (A.). 1861-1862. *Compendium der vergleichenden Grammatik der indo-germanischen Sprachen* [...].
- Schleicher (A.). 1863. *Die Darwinsche Theorie und die Sprachwissenschaft* [...]. Weimar.
- Schleicher (A.). 1865. *Ueber die Bedeutung der Sprache für die Naturgeschichte des Menschen*. Weimar (trad. fr. avec le précédent, Paris, 1868).
- Schleicher (A.). 1868. « Ein Fabel in indo-germanischer Sprache ». in : *Beiträge zur vergleichende Sprachforschung*, 5, pp. 206-208.
- Schreyer (R.). 1985. « The Origin of Language: A Scientific Approach to the Study of Man ». in : *Topoi*, 4-2, pp. 181-186.
- Schuchardt (H.). 1922. *Hugo Schuchardt-Brevier. Ein vademecum der allgemeine Sprachwissenschaft*. Bonn.
- Schuchardt (H.). 1979. *The ethnography of variation. Selected writings on pidgin and Creoles of H.S.* Ann Arbor [ed. par Markey (T. -L.)].
- Starobinsky (J.). 1966. « Rousseau et l'origine des langues ». in : *Europäische Aufklärung, Festschrift für Herbert Dieckman*. Munich : Fink Verlag, pp. 281-300.
- Steinen (K. von den). 1892. *Die Bakairi-Sprache. Wörter verzeichniss, Sätze, Grammatik. Mit Beiträgen zu einer Lautlehre der karäibischen Grundsprache*. Leipzig : Koelher (K.-F.).
- Steinthal (H.). [1851] 1858. *Der Ursprung der Sprache im Zusammenhang mit den letzten Fragen alles Wissens*. Berlin.
- Stringer (C.) & McKie (R.). 1996. *African Exodus : The Origins of Modern Humanity*. Londres : Jonathan Cape.
- Swadesh (M.). 1951. « Diffusional Cumulation and Archaic Residue as Historical Explanations ». in : *Southwestern Journal of Anthropology*, 7, pp. 1-21.
- Swadesh (M.). 1959. « Linguistic as an Instrument of Prehistory ». in : *Southwestern Journal of Anthropology*, 15, pp. 20-35.
- Swadesh (M.). 1971. *The Origin and Diversification of Language*. Londres : Routledge & Kegan Paul.
- Sweet (H.). 1901. *The History of Language*. Londres.
- Thurot (F.). 1797. *Tableau des progrès de la science grammaticale* [Introduction et Notes de A. Joly, Bordeaux, 1970].
- Timmermans (A.). 1890. *Traité de l'onomatopée ; ou, Clef étymologique pour les racines irréductibles*. Paris.
- Trask (R.- L.). 1996. *Historical Linguistics*. Londres : Arnold.
- Tromblin (J.-B.). 1997. « Epidemiology of Specific Language Impairment ». in : Gopnik (M.), (ed.). 1997. pp. 91-110.
- Tylor (E.-B.). [1871] 1873. *Primitive Culture ; researches into the development of mythology, philosophy, religion, language art and custom*. Londres [trad. fr., 2 volumes, Paris 1876-1878].
- Vendryes (J.). 1921. *Le langage. Introduction linguistique à l'histoire*. Paris.
- Vendryes (J.). 1855. « Première société de linguistique. La Société de linguistique de Paris (1865-1955) ». in : *Orbis*, pp. 7-21 [sur la question de la primauté et la remise en question de la thèse de Vendryes, voir Auroux 1983a].
- Wackernagel (J.). 1876. *Ueber den Ursprung und die Entwicklung der Sprache*. Basel.
- Wedgwood (H.). 1866. *On the Origin of Languages*. Londres.
- Whitney (W.-D.). 1865. « cr. De Müller 1863 ». in : *North American Review*, pp. 565-581.
- Whitney (W.-D.). 1870. « On the present condition of the question as to the origin of language ». in : *Transactions of the American Philological Association*, pp. 84-94 [1869-1870].
- Whitney (W.-D.). [1867] 1870. *Language and the Study of Language*. Londres.
- Whitney (W.-D.). [1875] 1880. *The Life and Growth of Language*. Londres [trad. fr. Paris, 1875].
- Whitney (W.-D.). 1880a. « Logical consistency in views of language ». in : *The American Journal of Philology*, 1-3, pp. 327-343.
- Whitney (W.-D.). 1880b. « On inconsistency in views of language ». in : *Transactions of the American Philological Association*, pp. 92-112.
- Whitney (W.-D.). 1892. *Max Müller and the Science of language : a criticism*. New York.
- Zaborowski-Moindron (S.). 1878. *L'origine du langage*. Paris [cité d'après la quatrième édition, Paris, 1885].



Origine du langage et origine des langues : réflexions sur la permanence et le renouvellement d'un questionnement des Lumières

Par Robert Nicolai

Université de Nice et Institut universitaire de France,
Nice, France

Mai 2006

0. Introduction

On aura constaté depuis quelques années, que la question de « l'origine des langues » et de « l'origine du langage » est revenue au goût du jour en se déplaçant d'un « non-espace » de réflexion linguistique vers un carrefour interdisciplinaire où elle rencontre d'autres problématiques ayant leurs exigences propres. Cela demande réflexion. Il faut ainsi tracer quelques frontières au sein des positions, propositions et conjectures avancées dans ce lieu où s'articulent à la fois des approches comparatives linguistiques, des recherches typologiques, pragmatiques, sociologiques, éthologiques, génétiques et archéologiques. Où se confrontent des propositions conjecturales, des modélisations informatiques et des représentations empiriquement fondées.

Dans ce champ très ouvert il importe non seulement de ne pas mélanger les perspectives de connaissance à propos des deux thématiques indépendantes que sont celle de l'origine du langage et celle de l'origine des langues mais aussi, face au désir latent de connaissance et à l'effort de construction de nouveaux champs disciplinaires, il s'agit de mieux apprécier les « objets de description » et se donner quelques limites pour mieux distinguer le probable du possible puisque, en l'état, les propositions dans ce domaine ne peuvent qu'échapper à la « réfutation » tout en se donnant pour « scientifiques ».

En confrontant l'approche de Ch. De Brosses qui au 18^{ème} siècle développait une étymologie dans laquelle il montrait que nos langues dérivait d'un langage primitif ancré dans nos réflexes physiologiques et une hypothèse récente de ce siècle, comme celle du protolangage de D. Bickerton, on pense contribuer à une réflexion générale sur la façon dont, en raison de sa forte teneur conjecturale, un thème s'inscrit dans la « logique » de la construction des connaissances du moment.

On pense aussi contribuer à cette réflexion générale qui semble parfois manquer en mettant ensuite en regard l'initiative de Süßmilch qui, en marge de ses travaux précurseurs pour la démographie, par une étude quantitative et sans analyse particulière des propriétés générales des langues se proposait de démontrer à partir du lexique « *la convenance de la langue celtique et de celles d'Orient* », et les approches actuelles de l'origine des langues dérivées de la mouvance post-greenbergienne ainsi que leur récupération formaliste. Mais avant, il n'est pas sans intérêt de broser le cadre intellectuel de l'époque.

1. Origine du langage.

Au siècle des Lumières nos repères linguistiques actuels (grammaire comparée, organisation structurale des systèmes...) n'avaient évidemment pas de pertinence. La théorie darwinienne de l'évolution des espèces n'est pas née et ce que l'on pourrait entendre par « évolution des langues » n'a pas été profilé au moule de la reconstruction historique et des contraintes théoriques de l'évolution par division et innovations. Les régularités des lois phonétiques n'ont pas été établies. On ne sait rien de la structure du cerveau. Les théories de la génétique sont encore à naître. L'Homme-Machine est bien évidemment sans rapport avec la recherche computationnelle. L'histoire de la nature a tout juste commencé à livrer ses premiers secrets. La présence des fossiles commence seulement à faire sens pour l'évolution des espèces, mais l'homme n'y trouve pas encore sa place. La paléontologie ne fait que naître.

Par rapport à l'époque présente nous sommes dans un *monde parallèle* : on s'intéresse au « génie des langues » et au « génie des peuples », on réfléchit à ce que peut signifier pour le

langage et pour ceux qui l'utilisent la distinction entre l'ordre libre des mots et l'ordre contraint qui distingue le latin du français ; on suppose que le climat à une incidence sur l'évolution des langues comme sur le comportement des populations. On ne croit plus très fort à la leçon de la Genèse, ou bien on a appris à la relativiser (l'isoler) en sachant distinguer ce qui relève de la leçon du dogme et ce qui relève des lois de la nature. La place de l'homme est toujours au centre des processus explicatifs et l'universalité de la raison est l'un des piliers de la réflexion connaissante. On dispute sur l'origine matérielle ou non de l'esprit humain autant que sur celle de la langue. L'intérêt envers les populations sauvages et barbares et les langues exotiques est grand : on mesure l'écart qui nous sépare d'eux, on lit les relations de voyages mais l'ethnologie et l'anthropologie ne sont pas nées. On cherche à comprendre les mythes et les coutumes¹ étrangers et l'on veut appréhender la part de l'universel et celle de la raison dans ce monde de différences. On ausculte les phénomènes mécaniques et naturels : ils piquent la curiosité de tout le monde savant. La métaphysique reste une référence stable dans l'élaboration de la connaissance. On utilise naturellement des métaphores anthropomorphiques (langue mère, filiation, etc.) ou vitalistes (germe, racine, etc.). Les Académies fonctionnent, les communications se succèdent...

À Paris comme à Berlin l'on discute de tout : de la quadrature du cercle, de la théorie des mouvements des fluides, du minimum de cire des alvéoles des abeilles, du mouvement progressif, du centre de gravité de tout le système solaire, des richesses de Sparte, de l'origine des langues... Les outils retenus pour l'étude des phénomènes relèvent tout autant de la réflexion philosophique et logique que de l'étude empirique et, à l'aune de nos critères, les procédures de description ne distinguent pas toujours nettement dans l'explication entre des arguments introspectifs, des généralisations contestables, des hypothèses spéculatives et des faits objectivés et vérifiés². C'est dans ce contexte intellectuel que nous nous situons : La perspective de Ch. de Brosses, pas plus que celle de Süßmilch que j'aborderai ensuite n'est exempte de ces caractéristiques d'époque.

1.1. La mécanique de Charles de Brosses.

6 mars 1765. Conformément au Règlement de 1723 et ultime démarche de la procédure, le Syndic Le Breton inscrit sur le Registre XVI de la Chambre Royale et Syndicale des Libraires et Imprimeurs de Paris, sous le numéro 42, Vol. 267, le Privilège du Roi accordé le 27 février 1765 au sieur Vincent, Libraire à Paris³ qui depuis bien deux ans déjà, désirait faire imprimer et donner au Public un ouvrage qui avait pour titre, *Traité de la Formation mécanique des Langues et des Principes de l'Étymologie*. Charles De Brosses, Président au Parlement de Bourgogne, vient d'obtenir le droit de publier deux tomes (489 pages et 533 pages) d'un ouvrage dont le manuscrit était déjà largement passé de mains en mains. Il entend montrer la nécessité physiologique de la naissance du langage chez l'homme et la façon dont il va se transformer à travers son usage, jusqu'à masquer l'évidence de cette origine physiologique ; d'où l'intérêt, pour lui, de l'étude étymologique. Il s'intéresse à la question de l'origine des langues non pas pour elle-même mais en tant que témoignage de la formation de la pensée et il pose quelques-unes des questions rhétoriques de son temps :

Y a-t-il eu une langue primitive & qu'elle est-elle ? Où chercher cette langue primitive ? Quelle est cette langue primordiale ? Subsiste-t-elle encore ? Parmi les langages connus entre lesquels on peut le plus vraisemblablement choisir, y en a-t-il un qui par de justes raisons doive obtenir la préférence⁴ ?

Bien sûr, les réponses vont de soi : on ne peut pas penser que la langue primitive n'a pas évolué. Conséquemment on ne peut pas espérer la reconnaître parmi les langues existantes

¹ Dans cet esprit, on pourra lire avec intérêt l'ouvrage « *Du culte des dieux fétiches ou parallèle de l'ancienne Religion de l'Égypte avec la Religion actuelle de la Nigritie* » que De Brosses a publié en 1760.

² Notons que de tout temps, on a compris que les hypothèses contestables et mal fondées ont fait avancer la science, ce qui est une façon d'interpréter la falsifiabilité poppérienne.

³ Le « Privilège du Roi » était l'autorisation à paraître soumise à la censure d'un corps de « censeurs royaux » qui décidaient de l'autorisation. Mais l'on pouvait toujours, à ses risques et périls, publier en Hollande ou ailleurs, comme le fut, par exemple, « *La philosophie de l'Histoire* » publiée à Amsterdam par ce neveu d'un « Feu l'Abbé Bazin », plus connu sous le nom de Voltaire.

⁴ J'ai bien évidemment modernisé l'orthographe dans l'ensemble des citations.

aujourd'hui et l'on ne peut pas non plus songer à remonter jusqu'à elle¹ par le simple jeu de la comparaison des langues existantes. En effet :

[s]i l'on supposait que la langue quelconque [...] que l'on croira devoir mériter la préférence, était encore restée dans les premiers principes de l'opération de la nature tels qu'ils sont partout [...] on pourrait peut-être espérer de réussir dans une telle recherche. Mais dès lors sans doute, elle était déjà fort chargée d'altérations, de dérivations & d'ornements. C'est-à-dire que les premiers principes y seraient déjà très difficiles à reconnaître, & les germes de la nature fort dépravés. N'est-ce donc pas une chimère que de croire [...] qu'à force d'examen & de comparaisons des langues actuelles, on puisse les ramener toutes à la seule langue primitive que les hommes parlaient avant le déluge. Outre que d'autres langues primitives ont pu se former par l'abandon de quelques enfants dans les déserts, assez naturel à supposer au milieu de tant d'émigrations qui ont suivi de près le déluge, l'intervalle de tant de siècles a tellement travaillé sur les langues, & les a dénaturées de cette primitive à tel point, que ce serait un projet absurde que de prétendre les y ramener.

Il lui faut donc procéder autrement. L'objet « langue primitive », certes, est inatteignable mais cependant par une réflexion appropriée l'on peut espérer appréhender sa trace dans les langues du monde. On va donc changer le plan de la recherche en essayant non plus de « retrouver » la langue primitive mais, à travers une *expérience de pensée*, tenter de saisir la façon dont elle a pu se former et ce par quoi elle a pu se manifester.

Nous voici devant un premier « programme de recherche » à visée « cognitive » – dirait-on peut-être dans le langage d'aujourd'hui – qui demande que l'on imagine ce moment particulier de la naissance du langage où l'homme va commencer à être ce pourquoi on le tient. Nous sommes dans la conjecture :

[r]amenons... la chose à ses premiers principes : considérons-là en elle-même seulement comme si elle en était à son origine. Supposons même pour un moment l'étrange hypothèse de quelques anciens philosophes, qui prétendaient que l'homme des premiers temps de l'humanité vivait isolé dans les bois à la manière des brutes, sans savoir encore faire un usage utile de sa faculté de parler ; & que ce ne fut que petit à petit & par développement qu'il commença d'inventer & de dresser les signes de la parole...

À partir de là, deux directions :

- l'une qui, du constat de la nature physiologique des organes utilisés pour la phonation, conduit à l'élaboration d'une certaine phonétique articulatoire : « *Voyons abstraction faite des langages usités sur la terre comment il en peut éclore un du premier germe des organes, et de la faculté naturelle donnée à l'homme d'en varier les articulations* » ;
- l'autre qui tente saisir de la fonctionnalité de cette voix physique pour manifester les sensations et sentiments humains élémentaires de l'individu à travers la création des premiers signes : « *cherchons à présent à saisir l'instant où les premiers mots naissent des premières sensations. Voyons nos sentiments & nos premières perceptions créer par l'organe de la voix leurs signes représentatifs, tels qu'ils peuvent convenir aux choses signifiées, & autant qu'il est possible à la voix d'effectuer cette convenance, selon ses facultés naturelles* ». Et là, portant son attention au premier apprentissage du langage par les enfants, De Brosses se proposera d'identifier les premières causes qui « *excitent la voix humaine à faire usage de ses facultés* » : les sentiments ou les sensations intérieures et pas les objets extérieurs. Là où la voix « fait sens ».

1.1.1. Approche physique des mots et ordres phonétiques.

À la première étape, son étude phonétique le conduira à :

poser pour principe que dans tous les langages de l'univers, dans toutes les formes quelconques de prononcer, il n'y a qu'une voyelle², « que six consonnes correspondantes à autant d'organes servant à la parole ». [...] « Chaque organe qui est dans la bouche a sa figure et son mouvement propre formant une lettre qui lui est particulière : il y a autant de lettres ou consonnes que d'organe : & il n'y en a pas plus. Ce sont 1. les lèvres. 2. la gorge. 3. les dents. 4. le palais. 5. la langue. Il y en a un sixième, à savoir, le nez, qui doit être regardé comme un second tuyau à l'instrument ».

¹ Parler de « reconstruction » serait un anachronisme ici.

² Je n'épiloguerai pas sur cette analyse qui fait la part belle aux consonnes parce que ce n'est pas l'objectif. Tout en soulignant qu'elle est évidemment « en phase » avec les modes de représentation des langues sémitiques que cet érudit de De Brosses connaissait.

La multiplication des consonnes est alors analysée comme l'effet du mouvement plus fort ou plus faible dans chaque organe, pris comme référence d'une classe naturelle articulatoire et ses résultats l'amènent à élaborer un outil spécifique : *l'alphabet organique & universel* composé d'une voyelle et de six consonnes. En transcrivant les mots des langues dans cet alphabet l'on fait apparaître les rapports phonétiques qui restaient cachés et l'on perçoit l'unité sous la différence introduite par les contingences de l'évolution et les écarts des hommes. Nous avons donc affaire ici non pas à l'élaboration d'une simple représentation des connaissances mais à la construction d'un outil pratique dont la fonction est de mettre en évidence les points d'articulation consonantiques que manifestent les mots (ainsi les sons *p, b, f, v, m*, par exemple, renvoient à un même procès articulatoire qui souligne la pertinence de l'organe des lèvres ; ils seront représentés par un même symbole). L'intérêt heuristique qu'il y trouve se conçoit dans un contexte de description où la systématisme des recherches comparatives n'est pas encore retenue et où la notion de structure d'une langue n'a pas encore de sens bien précis. Ainsi qu'il le montre, il lui suffit de transcrire les mots des langues dans cet alphabet organique pour faire apparaître des rapports qui, sans cela, restaient cachés.

Pour la deuxième étape, l'approche physique des mots lui fera reconnaître une organisation en six ordres qualitativement distingués. Tous liés à la sphère physiologique à l'intérieur de laquelle l'homme évolue nécessairement dans ses mouvements les plus primordiaux. Les voici :

- *Premier ordre* : les interjections qui expriment le sentiment. En effet :

les interjections sont courtes : elles partent du mouvement machinal & tiennent partout à la langue primitive. Ce ne sont pas des mots mais quelque chose de plus, puisqu'elles expriment le sentiment qu'on a d'une chose ; & que par une simple voix prompte, par un seul coup d'organe elles peignent la manière dont on s'en trouve intérieurement affecté.

À partir de cette base la plus physique de la manifestation de la voix une première typologie des « *rapports généraux entre certains sentiments et certains organes* » peut être dégagée.

- *Deuxième ordre* : les mots nécessaires nés de la conformation de l'organe indépendamment de toute convention (les racines labiales, les mots enfantins).

Suivons les premières manifestations de la voix humaine par l'examen des enfants au berceau. Tous en quelque pays que ce soit ayant pour premier mouvement plus facile d'ouvrir la bouche et de remuer les lèvres, forment la voix pleine et articulent la lettre labiale. [...] Les mots Baba, Papa, Mama, [...] sont des racines primordiales nées de la nature humaine, & dont la naissance est une conséquence absolue de cette vérité physique, l'homme parle.

- *Troisième ordre* : les mots presque nécessaires : les noms donnés aux organes de la parole, tirés de l'inflexion même de l'organe (cf. Gorge, Langue, Dent, Bouche).

Mais puisqu'il existe des mots nécessaires dont la structure est absolument liée aux idées de l'âme et à tout le système de l'organisation humaine ; puisque ces mots sont les premiers et les plus originaux, n'est-il pas raisonnable de penser qu'à mesure que l'organe se développe, le même progrès mécanique a sa continuation, quoique plus difficile sans doute à discerner lorsqu'elle s'éloigne et s'étend ? N'est-il pas juste de suivre la route tracée par la nature dans la recherche de l'origine des noms imposés aux choses, & dans l'examen de la naissance des racines de la langue primitive ? [...] L'homme forme volontiers les noms qu'il donne à chaque organe de sa parole sur le caractère ou l'inflexion propre à cet organe...

- *Quatrième ordre* : les mots qui tiennent au physique de l'objet : les onomatopées. On abandonne là l'expression du corps, mais on se sert de ses capacités pour imiter les objets du monde.

L'homme est par nature porté à l'imitation : on le remarque de la manière la plus frappante dans la formation des mots. S'il faut imposer un nom à un objet inconnu, & que cet objet agisse sur le sens de l'ouïe dont le rapport est immédiat avec l'organe de la parole, pour former le nom de cet objet l'homme n'hésite, ne réfléchit, ni ne compare ; il imite avec sa voix le bruit qui a frappé son oreille, & le son qui en résulte est le nom donné à la chose.

[...] Tous les mots de ce genre peuvent donc être regardés comme nécessaires ; leur formation étant purement mécanique & absolument liée au physique des choses, sans que l'arbitraire y ait aucune part ; quoique les hommes puissent d'ailleurs donner à leur guise d'autres noms à ces mêmes choses.

- *Cinquième ordre* : les mots consacrés par la nature à l'expression de certaines modalités des êtres (cf. ST « stabilité », SC « profondeur, creux », FL « liquidité »). Sans doute que la problématique générale et toujours mal cernée du symbolisme phonétique se dessine ici¹.

Toutes les observations ci-dessus prouvent qu'il y a des figures de mots, des caractéristiques de sons liés à l'existence des sensations intérieures ; qu'il y en a de liés à l'existence des objets extérieurs [...]. D'autres observations paraissent nous montrer qu'il y en a aussi de liées à certaines modalités des êtres ; sans qu'il soit possible de démêler nettement le principe de cette liaison entre des choses où l'on n'aperçoit aucun rapport [...] Par exemple, pourquoi la fermeté et la fixité sont-elles le plus souvent définies par le caractère St ? [...] comme pour désigner le creux & la cavité on emploie le K ou le G, ou lettre de gorge, le plus creux & le plus cave des six organes. [...] De même FL, caractère liquide est affecté aux fluides, soit ignée, soit aquatique, soit aérien.

- *Sixième ordre* : les accents qui servent d'appendice au Premier ordre.

Les accents sont une espèce de chant joint à la parole [...] Ils lui donnent une vie et une activité plus grande. Ils sont de la langue primitive, étant chez les hommes l'expression pure et première de la nature : ils y forment un sixième ordre de sons primitifs ; ou pour mieux dire, n'étant pas des mots, ils doivent être joints au premier ordre qui est celui des interjections ; car ils sont comme elles l'expression du sentiment intérieur.

Comme le souligne De Brosses cette stratification est une hiérarchie. Au premier niveau de leur expression les « sensations intérieures » sont tout simplement expectorées. Dans un deuxième temps logique – et lorsque c'est possible – l'homme imite par l'onomatopée auditive autant que par des procédés de synesthésie dynamique ce à quoi il se trouve confronté dans le monde. Dans un dernier temps enfin ce donné vocal initial sera modulé et développé dans un processus de construction sémiotique global. L'émergence du langage ne semble donc pas être fondée sur le préalable d'une pressante fonction de communication sociale et d'échange d'information. Autrement dit, pour lui ce n'est pas de la nécessité d'un besoin de communication sociale que le langage naît, même si c'est grâce à elle qu'il se développe ensuite ; c'est tout simplement de l'expression naturelle physique – contrainte et nécessaire – du rapport de l'homme au monde. Le seul présupposé est probablement, qu'à la différence de beaucoup d'autres espèces il se trouve que l'homme est un être bruyant. Ce qu'il montre généralement dès l'instant de sa naissance dans la « primalité » de son cri.

Sa systématisation en six ordres de ce qu'il identifie comme des primitifs et des radicaux en rapport avec la dénomination va lui ouvrir la possibilité d'analyser le domaine des données lexicales des langues en l'appréhendant non plus comme une totalité non structurée et plus ou moins floue, mais à travers une classification et une organisation susceptible d'être fonctionnalisée comme critère dans son projet de description. Pour lui, les mots des langues ainsi formés en rapport avec ces six ordres « *sont véritablement primitifs et radicaux. Il n'y a aucune langue ancienne ou moderne qui n'en possède plusieurs, lesquels ont des dérivés dans d'autres langues voisines. C'est en rassemblant de chaque langue tous les mots ainsi formés qu'on aurait une langue véritablement primitive...* ». Mais attention ! Cela ne conduit pas pour autant à la construction de cette langue. Nous avons seulement réussi par l'analyse intellectuelle, à dégager des critères qui nous permettent de reconnaître dans le lexique, par des opérations immédiates ou médiatees, ce qui peut *raisonnablement* être référé à des racines et des radicaux primitifs, eux-mêmes dérivés des *germes* que sont les six ordres qui viennent d'être inventoriés. *Germes* et *racines* qui *in fine* fonctionnent comme des *clefs* au sens où un caractère chinois est une clef de ce qu'il représente dans les différents dialectes de cette langue. Ainsi « *nous aurions, par abstraction, une langue primitive que personne ne parlerait, ni n'aurait jamais parlé, du moins dans tout son contenu, quoique tout le monde en ait en soi tous les germes primitifs* ». Puis, bien évidemment, à partir de cette *première fabrique des mots nécessaires* une ébauche de système de dérivation va se mettre en place, lequel opacifiera finalement la nécessité de la relation initiale.

La langue primitive est donc lisible et identifiable dans tous les états des langues du monde mais elle ne saurait être « reconstructible » au sens d'une tradition comparative

¹ Cf. Peterfalvi (1970) pour une version psycholinguistique, Guiraud (1967) pour une version structuraliste, Fonagy (1992) pour une version cognitivo-phonétique.

historique – qui n'est d'ailleurs pas encore née. Elle a pour particularité d'être toujours là, toujours présente, toujours active, matrice de création potentielle actualisée continûment dans toute étape de formation du langage comme dans toute instance de profération de mots. La perspective étymologique de l'auteur ne doit pas tromper, qui avec plus ou moins de bonheur en passe par la collation et l'étude méticuleuse de tous les mots des langues disponibles dont il a connaissance – et de ce point de vue il est très érudit. Nous sommes loin ici de la positivité d'une étude diachronique¹ élémentaire et l'étymologie n'a pas pour but de retrouver quelque langue originelle que ce soit mais plutôt de fournir des éléments de méthode pour aider à comprendre comment, dans la nécessité et dans la contingence, l'homme a agi et transformé le langage et les mots de ses langues.

1.1.2. Formes primitives, arbitraire et social de la langue

La reconnaissance de ces six ordres est essentielle, mais on se doute bien qu'ils ne permettent pas de couvrir toutes les nécessités de la dénomination élémentaire car il existe beaucoup de réalités à nommer dans le monde qui sont perceptibles par la vue sans mettre en jeu le sens de l'ouïe. Comment les nommera-t-on ? Cela est une nouvelle étape et passera par des opérations mentales de comparaison et d'approximation appliquées sur les mots disponibles : l'homme « *saisit mécaniquement l'objet visible par la face ou la qualité apparente qui lui donne le plus de jeu pour en former le nom par comparaison ou par approximation avec d'autres mots naturels déjà faits. Il y a de l'arbitraire à la vérité dans cette méthode ; cependant la nature, à ce qu'il semble ; n'en souffre que le moins qu'elle peut* ». Autrement dit sa reconnaissance implicite de l'existence d'une stratification temporelle dans la dynamique de l'émergence du langage lui permet de comprendre les développements subséquents en pensant à la fois la création des mots à partir des dérivations ultérieures par rapport à un système primitif limité, et la structuration lexicale par des procès de comparaison et d'approximation pour étendre le processus de la dénomination à toutes les parties du lexique qui n'auraient pu être mises directement en relation avec l'organe de l'ouïe.

Finalement on retiendra :

- que la langue primitive aura donné les *racines* des termes habituels qui servent à exprimer nos idées ou à dénommer les objets qui tombent sous nos sens et que rien en ce domaine n'est arbitraire. Tout est déterminé par les contraintes et les caractéristiques de la physiologie humaine et de la réalité physique du monde ;
- que l'homme parle pour faire connaître à un autre homme ce qui est en lui ou ce qui est hors de lui. Soit donc ce qu'il *sent*, ce qu'il *perçoit* ou ce qu'il *a perçu*. Mais que le mouvement de « faire connaître » n'implique pas nécessairement un irrépressible désir préalable d'une communication sociale : mes pleurs font connaître ma douleur et son intensité, je peux désirer la manifester, la faire partager, voire l'imposer au monde qui m'entoure mais au départ, ils ne sont pas nécessairement instrumentalisés en « signes ». Même si plus tard je saurais jouer la douleur à toutes fins utiles. Même si plus tard encore je saurais par comparaison me servir de ces signes pour renvoyer à d'autres réalités : il n'y a même aucune assurance qu'il y ait à l'œuvre à ces stades ultérieurs supposés la manifestation d'un quelconque processus de sémiotisation².

En ce qui concerne les *causes* de l'imposition des noms, De Brosses distinguera entre les *causes immédiates*, celles qui sont rapportées à la peinture ou à l'imitation de la chose à nommer et qui conduisent au langage primitif qui vient d'être présenté, et les *causes médiate*s qui se manifestent par la dérivation de mots déjà existants et qui sont une élaboration subséquente sur ce langage primitif qui, tout en développant le plan du langage contribue à brouiller l'identification des racines premières. Finalement il déduira de cette dynamique l'existence préalable de trois qualités : *la volonté, l'intelligence et la mémoire*, sans lesquelles le développement de ces pratiques en langage n'auraient pu être possible. Les opérations de *comparaison* et d'*approximation* permettent d'élargir le champ de la dénomination, celles de *dérivation* des formes qui le structurent et le transforment demandent ces qualités préalables tout en

¹ L'analogie avec les recherches de P. Guiraud et de tous ceux qui se sont intéressés aux questions de motivation du lexique est ici évidente. Je ne la développerai pourtant pas afin de ne pas risquer de trop sortir du thème.

² Cf. Nicolai (sous presse) pour des remarques sur les processus de sémiotisation.

laissant indéterminée l'appréciation du seuil qualitatif où se fera – parce qu'il se fait aussi – le passage aux signes établis et fonctionnalisés dans leur arbitraire. Mais ce seuil a-t-il une signification hors d'une théorie particulière ?

Cependant cette approche à la fois phonétique, conjecturale et empirique – dans la mesure où De Brosses trouve dans l'examen des mots des langues réelles les justifications de son propos qui « *a permis de reconnaître certains principes mécaniques et nécessaires de la formation du langage, conformes à la construction organique de l'instrument vocal tel qu'il est donné à l'homme par la nature* » – reste insuffisante pour rendre compte de la réalité des langues. De cette langue primitive qui vient à naître aux langues dans l'enfance, aux langues adolescentes, adultes¹, élaborées dans leur génie, puis en déclin le cheminement est long. Comme il l'a suggéré, les causalités physiques finissent par s'estomper dans le même temps que la syntaxe se formera et que les normes se constitueront dans la pratique linguistique de l'homme « être social ». Toutefois l'étude du développement syntaxique n'est pas l'affaire de De Brosses.

Une deuxième expérience de pensée, très « dix-huitièmiste », va alors être proposée qui rend la langue au social. C'est sur elle qu'il fonde le véritable développement du langage, avec ses fonctions de communication ordinaire, de grégarité et de support au développement de la pensée.

« *Supposons qu'un certain nombre d'enfants ont dès leur bas âge été abandonnés loin de tout commerce humain dans un climat désert, où ils ont trouvé le secret de se conserver jusqu'à l'âge adulte* ». Le cadre de l'expérience est présenté il ne reste plus qu'à « étudier » ce qui se passe ; l'hypothèse à vérifier étant « *qu'une troupe d'enfants, supposés mis ensemble, & abandonnés à la nature, se fera pour elle-même une langue propre & primitive, qui dans la suite par le développement et l'extension des idées sera sujette à son progrès & à ses variations* ». Et De Brosses avance que « *sans que l'expérience nous ait distinctement montré ce qu'il en serait, on peut assurer que les choses arriveraient ainsi, aussi hardiment qu'on assurera que ces enfants marcheraient ; puisque l'un est ainsi une suite naturelle de leur constitution primitive* ».

Il argumente :

Voyez ensemble trois ou quatre petits enfants instruits dans la langue vulgaire ; vous ne les entendez pas : cependant ils s'entendent à merveille entre-eux ; ils se sont déjà fait un petit jargon [...]. Que l'on suppose un homme vivant seul dès son enfance, & absolument isolé de toute société, il ne fera pas, ou il ne fera que très peu d'usage de sa faculté de parler. Elle ne sert qu'à communiquer ces idées à autrui [...]. Le commerce avec les hommes donne occasion non seulement de parler pendant la conversation présente, mais encore de réfléchir sur les conversations passées, & de préparer celles à venir. Dans l'hypothèse ci-dessus qui prend les choses au premier pas où il soit possible de les considérer, il n'y a point de langage qu'on puisse appeler discours, mais une espèce de vagissement presque inarticulé qui forme néanmoins quelques mots sans suite. Mais supposons deux ou plusieurs enfants mis ensemble ; alors le naturel, le besoin, l'habitude mettent en jeu les facultés. Chacun profite des inventions de l'autre, & les accroît en continuant d'opérer sur ce premier fond [...]. La puissance physique qui manque aux animaux pour faire certains progrès ne manquant pas à nos enfants, ce petit germe poussera de profondes racines, & jettera un jour des branches infinies sur le plan donné par la nature.

Parallèlement, la généralisation de la caractéristique *arbitraire* qui fait passer ce langage naissant et/ou primitif aux langues constituées dans la *motivation* est perçue comme la conséquence naturelle de son utilisation tout en retenant aussi que, dès son état premier, l'arbitraire en tant que possible n'est jamais exclu dans la fabrication des mots. Ce ne serait donc que par un procès de réduction particulièrement illégitime que l'on rendrait compte de sa perspective à travers les œillères de l'aporie scolastique d'une dichotomie « arbitraire *versus* motivé ».

1.2. Le protolangage de Dereck Bickerton

Un siècle après De Brosses la grammaire comparée développe son cadre méthodologique nouveau et s'attache à des études empiriques concernant le domaine des langues indo-européennes ; corrélativement, se ferme la porte aux approches conjecturales sur l'origine des langues.

¹ Terminologie d'époque, naturellement.

Un siècle de plus encore et les chemins de la grammaire comparée sont non seulement tracés, balisés et scellés mais jugés « dépassés » par certains. La fin du 20^{ème} siècle voit ressurgir le questionnement ancien sur l'origine des langues et du langage ; la parenthèse structuraliste est close, le générativisme linguistique a implosé ; les questionnements cognitivistes ont émergé. Et c'est ainsi que de la part de certains responsables éditoriaux des numéros spéciaux de revue sont orientés sur ce thème, que de la part de certains responsables des politiques de recherche du moment des initiatives d'appels d'offre¹ ont été lancées qui, entre autres choses, ambitionn(ai)ent de croiser les recherches linguistiques, anthropologiques, biologiques, paléanthropologiques, archéologiques mais aussi et surtout les neurosciences, la génétique moléculaire et la génétique des populations.

Aujourd'hui, la question de l'émergence du langage est considérée comme relevant d'une étude complexe qui concerne l'évolution des hominidés dans l'ensemble des dimensions qui ont conduit à *Homo Sapiens*, laquelle est ainsi référée à une échelle temporelle dont la pertinence se définit à la mesure des procès de spéciation dont la paléontologie nous suggère les étapes et la durée. C'est une problématique médiatisée et à la mode : on trouve sur Internet un nombre indéterminé de pages dédiées à ce sujet, rédigées pour tous les niveaux de culture et tout particulièrement pour une vulgarisation rapide et enthousiaste (!).

C'est une problématique ouverte : elle en appelle à plusieurs disciplines et la linguistique dont on pourrait penser qu'elle retrouve là un de ses terrains légitimes, finalement, n'est probablement pas la plus pertinente dans l'affaire. Les théories de l'évolution, (néo)darwinienne en général, la biologie, les neurosciences ainsi que la paléontologie délimitent et dominent le domaine et l'on conçoit que dans un champ aussi vaste l'on atteigne très vite les limites de son incompétence² car l'on n'a pas toujours la maîtrise de ce que l'on pense qu'il faudrait savoir. Au sein d'une telle conjoncture interdisciplinaire l'on est en effet très souvent dépendant et conduit à accepter comme allant de soi des propositions dont, par méconnaissance de la culture disciplinaire qui les a fait naître, l'on évalue mal la fragilité/validité³, sauf à réellement travailler en collaboration. Et même dans ce cas la collaboration ne signifie pas pour autant une (volonté de) compréhension suffisante des arrière-plans théoriques respectifs qui vont se croiser dans la recherche partagée. Alors que cette compréhension est le *sine qua non* du renouvellement de la problématique. C'est donc en partant de la charge métaphorique de « mots » et de leur stabilisation en tant qu' « essences nominales » au sens de Locke que le plus souvent des cristallisations se feront (attracteurs, complexité, chaos, catastrophes, etc.), et que des dynamiques d'idées se diffuseront⁴.

Ceci reconnu et en choisissant de se placer du côté des linguistes, une perspective parmi les plus intéressantes⁵ est sans doute celle ouverte depuis quinze ans par Bickerton (1990, 1995, 2000). S'intéressant à la question de l'émergence du langage il considérera quelques

¹ Cf. la thématique d'un appel d'offre récent à l'initiative du CNRS « *Origine de l'homme, du langage et des langues* », repris ultérieurement dans le cadre des actions de l'ESF (Eurocores Programme : *The Origin of Man, Language and Languages* (OMLL), <http://www.esf.org/eurocores/>). Les références ici retenues : Renfrew pour l'archéologie ; Greenberg, puis Ruhlen pour la linguistique et Cavalli-Sforza pour la génétique des populations sont censées ouvrir la voie (Cf. *la « nouvelle synthèse » de Colin Renfrew*). Quatre « axes thématiques » ont été proposés et financés : (1) langues et gènes ; (2) langage et archéologie/paléontologie ; (3) langage et esprit/cerveau ; (4) langage et société. Ce type d'initiative montre sans beaucoup d'ambiguïté le type de liberté, le type de contrainte et le type de direction qui s'actualise aujourd'hui dans la communauté des chercheurs.

² Et c'est bien ce que souligne Bickerton (2001, p. 581) qui fait ce commentaire dans le compte rendu de deux ouvrages récents : « *The ideal scholar in this field should combine a professional training in linguistics, paleoanthropology, evolutionary biology, neurology, psychology and primatology, at the very least. But human life is finite and too short, so the best one can hope for is that researchers should combine a thorough grounding in at least one of these disciplines with a fairly extensive knowledge of, or at least respect for, all the others. Alas, few if any meet even this standard. And the problem is not so much a defect of intellect or knowledge as one of attitude* ».

³ La vulgarisation simplifiée et répétitive par Internet interposé est-elle un moyen de résoudre ce problème structurel ?

⁴ Bien évidemment on peut remplacer cette remarque volontairement « datée » par une remarque équivalente fondée sur les observations de la sociologie de la science.

⁵ L'une des plus primées au *Hit-Parade* médiatique d'Internet.

unes des situations limites de son emploi et de son apprentissage. Tout particulièrement les quatre suivantes :

- le développement des capacités grammaticales dans les premiers stades de l'apprentissage chez l'enfant (avant deux ans),
- les tentatives d'enseignement du langage aux grands singes qui peuvent apprendre après entraînement un certain nombre d'associations entre signes et catégories et s'en servir ensuite (le bonobo Kanzi),
- les exemples d'apprentissage tardif du langage (cas de la fillette Genie),
- l'émergence des pidgins.

C'est en référence à ces situations limites qu'il a développé l'hypothèse¹ que le langage tel que nous le connaissons aujourd'hui a dû être précédé par un *protolangage* donné pour fonctionnel dans son monde dont des traces fossiles perdurent dans le langage actuel². Ce *protolangage* aurait été le mode de communication normal de *Homo Erectus*, il y a quelques 200 000 ans.

Cela ouvre à une réflexion spéculative sur deux plans :

1. Que peut-on supposer des conditions corrélatives de l'émergence de ce protolangage compte tenu de ce que l'on sait et/ou croit savoir aujourd'hui de l'évolution des espèces de primates et des théories actuelles de l'évolution ?
2. Quelle devait être la nature « communicationnelle³ » de ce protolangage et en quoi est-il censé se distinguer du langage humain actuel ? La différence est-elle qualitative ou quantitative ?

Et dans ce domaine, aujourd'hui pas plus qu'au 18^{ème} siècle, les propositions de réponse ne peuvent pas faire autrement que passer par l'élaboration d'expériences de pensée. Elles sont seulement réactualisées et étalonnées aux théories de l'évolution actuellement retenues.

1.2.1. Question d'émergence

Plusieurs questionnements donc. Par exemple : puisque le langage, ou même une forme un peu élaborée d'un protolangage semble n'avoir jamais existé que dans notre espèce, lesquelles parmi nos caractéristiques humaines sont susceptibles d'être mises en rapport avec ce développement spécifique ? Est-ce le développement d'une « intelligence sociale » que certains voient comme la condition d'émergence du langage au sein de l'espèce qui doit être considéré comme le déclencheur de cette capacité particulière ou bien cette « intelligence sociale » n'a-t-elle tout simplement fait que profiter du déclenchement du langage pour se développer ? Dans quelle mesure cette capacité de protolangage que nous avons dû posséder (*versus* que nous possédons toujours mais n'utilisons plus que dans les situations limites dont les cas présentés ci-dessus sont donnés comme des exemples prototypiques) a-t-elle été un facteur déterminant dans le processus de la sélection des espèces qui a conduit à la stabilisation de *Homo Sapiens* ?

On le voit, en rapport avec la théorie de l'évolution ces questions sont de « grandes questions » qui débordent largement les recherches sur les langues c'est pourquoi, compte tenu du niveau de nos connaissances, les réponses apportées restent hautement spéculatives. Nous voici loin des pratiques laboratoires et de la collation des données empiriques, ramené à construire des « expériences typiques » dont on assurera la *raisonnabilité*⁴ par rapport aux postulats et aux modèles théoriques du moment. À titre d'exemple voici l'une de ces expériences spécialement construite pour argumenter dans la discussion sur la place à donner

¹ Après avoir proposé l'hypothèse du « bioprogramme » au tout début des années '80.

² Comme c'est généralement le cas dans l'évolution des espèces.

³ Le terme « linguistique » ici serait un abus de langage.

⁴ On ne parlera pas de « vraisemblance » car on n'étalonne ce que l'on construit à aucune « réalité » ni à aucune « vérité » : c'est pourquoi je préfère le néologisme « raisonnabilité » qui étalonne ce que l'on construit tout simplement à ce que l'on croit qu'il est « acceptable/raisonnable » de tenir pour « acceptable » sur une base essentiellement discursive. Comme tous les êtres « raisonnables » aussi bien Bickerton que De Brosses emploieront le terme « raisonnable » pour justifier certaines hypothèses de leurs expériences de pensée.

à l'intelligence sociale dans son lien à l'émergence du langage. Par rapport à notre rationalité contemporaine et au cadre d'analyse que l'on s'est donné, elle décrit, explicite et analyse les comportements que l'on doit prêter à nos petits ancêtres d'il y a plus de 200 000 ans :

[...] Now suppose the group has just found a store of honey and is exploiting it, getting stung in the process. One member happens to look up and sees vultures circling just above the horizon. There's dead meat there, maybe some dead megafauna that would feed the group for days. He jumps up and down, pointing. They're too busy with the bees and the honey to take any notice. What's the advantage to him if he can get their attention ? If he's right, and they make a big food find, he'll be a hero, the others will look up to him, his status in the pecking order (and his access to mates) will be substantially enhanced. If he does this sort of thing consistently, he may get to be the leader. But how can he make them understand ? If only he had a real word or two, something more specific than follow-me !

Peut-être d'aucuns penseront qu'il doit bien y avoir une intention maligne cachée à reprendre un tel exemple. Peut-être ont-ils raison, peut-être pas. Mais montrer du doigt cette particularité de la réflexion aujourd'hui conduite dans ce domaine n'est pas en soi une stigmatisation négative de la procédure qu'elle met en oeuvre : il est tout simplement *évident* que l'on ne peut / sait guère procéder autrement. Et Bickerton le souligne sans ambiguïté, qui termine son chapitre en précisant que (2000, chapitre 9) :

This chapter is, I know, largely speculative, and may never be proven or disproven. But nobody knows how much we may yet be able to learn about our remote ancestors. If we do learn more, it won't be just a question of amassing facts. We can't do without facts, but facts by themselves, alas, never say enough. All facts are subject to interpretation, and can take on completely different appearances depending on the lens through which they are viewed. Speculation forms a vital component in science ; it helps to interpret facts and to guide future research, but if the research it has guided turns up things that it didn't predict B even things quite incompatible with it B that's par for the course. Provided you don't turn the goal into a religion, you get more out of looking with a goal in mind than out of blind fishing expeditions. Speculation is light luggage, it's easy to junk it and try again.

Mais il importe de le préciser pour, à l'instar de Bickerton, ne pas se tromper sur la fonction, l'intérêt et la valeur de ces outils nécessaires que sont les expériences de pensée¹.

1.2.2. Nature de la protolangue

Quel est donc le rapport du protolangage au langage ? La réponse la moins troublante serait : aucun. Et cela tout simplement parce que dans le cadre darwinien, l'hypothèse retenue est celle d'une *discontinuité* entre les deux : le langage proprement dit (c'est-à-dire le mode de communication caractéristique de *Homo Sapiens*) serait une évolution non-nécessaire² de ce protolangage dont il se distinguerait par des nouveautés cognitivement importantes telles que :

- l'ordre des mots significatif et contraint dans le langage ;
- l'existence d'éléments inexprimés (pronoms nuls) à certains endroits dans la phrase (afin de compléter les verbes) obéissant à des règles précises ;
- le nombre fixe d'arguments pour chaque verbe, devant nécessairement être exprimés (éventuellement par des éléments nuls) ;
- la construction récursive de phrases plus complexes à partir de phrases plus simples ;
- omniprésence des items grammaticaux, morphologiques ou syntaxiques.

Soit donc : (1) la présence d'items grammaticaux, (2) une structure d'argumentation (ou thématique) et enfin (3) l'existence d'une organisation syntaxique.

Notre langage résulterait ainsi d'une évolution non nécessaire en rupture avec l'usage du protolangage. Elle aurait permis de répondre à des finalités qualitativement différentes, sans

¹ Dans d'autres domaines, de Galilée à Einstein, les expériences de pensée ont nourri la science. De fait elles nourrissent tous les humains : scientifiques et non scientifiques. La distinction à faire gît donc ailleurs.

² Il faut, bien évidemment, distinguer ici entre la nécessité « rétroactive » qui rend nécessaire pour le descripteur la succession qui conduit à son présent en fonction de ses hypothèses, et la nécessité « prospective » qui rendrait nécessaire, indépendamment du descripteur, une évolution sur la base de son fonctionnement dans sa propre contemporanéité.

lien direct avec la communication élémentaire et située d'événements saillants pour laquelle le supposé protolangage était bien adapté. Ce seraient des nécessités distinctives suffisamment élaborées ou suffisamment différenciées qui auraient justifié un « besoin de langage » dans une niche écologique reconstituée et en rupture ; la complexification qui introduit le langage est marquée par l'émergence de la syntaxe en tant que capacité cognitive.

Pour résumer à l'extrême Bickerton et tenter de ne pas le trahir je le citerai se résumant lui-même (2000) :

Before you had syntax, all that existed was a kind of protolangage. If you want to know what that protolangage was like, you can get some idea by looking at the productions of apes who have been taught to use signs or other symbols, or at early-stage pidgin languages (at about the « Me Tarzan, you Jane » level of development), or at the speech of children under two. I say « get some idea », because of course there will be differences between then and now. We can assume our early ancestors talked about more things than apes do and that some of those things were different from the things apes talk about. We know that speakers of any pidgin speak at least one natural human language fluently, and there has to be some carry-over (though if you look at samples of pidgin speech, it will amaze you to see how little), at least in the range of things that can be discussed. We know that children, especially if they are learning an inflected language like Spanish or Italian, will pick up the odd grammatical feature you won't find among apes or early stage pidgin speakers, and probably wouldn't have found among our remote ancestors, either. All protolangage varieties :

- can only string together a small handful of words at a time ;
- can leave out any words they feel like leaving out ;
- often depart from the customary word order unpredictably and for no obvious reason ;
- cannot form any complex structures, whether these be complex noun phrases or sentences more than a clause long ;
- contain, if they have any at all, only a tiny fraction of the inflections and the « grammatical words », things such as articles, prepositions and the like that make up 50 percent of true language utterances.

Et encore (2001, p. 581) :

To me it seems that explaining how language evolved turns on the answers to two crucial scientific questions. The first is, how and why one particular primate species, or one primate line of descent, came to develop a system of communication that allowed the transfer of (potentially) unlimited factual information, and the basic principles of which differed from those of all previous systems of communication. The second is how such a system acquired the very specific characteristics that the syntax of modern human languages exhibits. If one abbreviates these questions to « How did meaningful units (words or signs) evolve ? » and « How did syntax evolve ? », little is lost.

L'idée de la *discontinuité* est importante car elle s'insère harmonieusement dans les schémas théoriques de l'évolution darwinienne (« classique » ou revue dans la perspective des « équilibres ponctués ») et que la théorie de la présence fossile y faisait sens. Linguistiquement elle suggère une réflexion à faire sur les coexistences de systèmes actifs qualitativement différents dans le procès de communication, laquelle est sans surprise dans ce domaine large. Toutefois elle pourrait peut-être devenir plus intéressante si on la poussait à sa limite¹ en explorant cette gestion du discontinu hors de la clôture d'un espace neuro-cognitivement fermé.

1.3. Pourquoi De Brosses ? Pourquoi Bickerton ?

Nous avons vu que, parti de la considération du lexique des langues, De Brosses qui voulait montrer la nature fondamentalement non arbitraire du langage primitif le dérive naturellement des caractéristiques physiologiques primaires de l'homme en s'appuyant sur les traits articulatoires de la phonétique et sur les premiers strates supposés de la constitution du lexique. Avec la notion de langage primitif il a cherché à identifier une forme de communication élémentaire extrêmement limitée et marquée par l'absence de syntaxe (il s'agit de mots, sans plus) dont il supposait qu'elle avait *nécessairement* dû exister avant que l'on ne parvienne à l'état complexe des langues d'aujourd'hui. Il supposait pouvoir en retrouver les « traces » empiriques dans le premier langage des enfants ou dans les contextes de déprivation linguistique. Sa perspective est apparemment continuiste, mais en apparence

¹ Cf. les questions récemment abordées de feuilletage nécessaire et de pluricodisme en liaison avec la problématique de la dynamique des langues appréhendée au plan « cognitivo-culturel » et non pas « neuro-cognitif », Nicolai, 2003, 2005, sous presse.

seulement car si nous analysons son approche dans le détail nous remarquons que la continuité n'est que dans l'entassement et la stratification des formes de la langue : c'est une « successivité » ; elle n'est pas dans les processus de leur construction car les processus dérivationnels et syntaxiques qui sont censés prendre le relais sont de nature différente que l'initial processus de nomination. Ainsi l'idée que l'on obtient « *par abstraction, une langue primitive que personne ne parlerait, ni n'aurait jamais parlé, du moins dans tout son contenu, quoique tout le monde en ait en soi tous les germes primitifs* » est intéressante du point de vue de la compréhension et de l'analyse de l'évolution des ressources lexicales¹ : le « primitif » est à la fois renvoyé à l'origine et retenu activement dans le présent.

Pour ceux qui aujourd'hui s'intéressent à l'origine du langage il s'agit globalement (1) de replacer le comportement langagier dans le contexte de l'évolution des espèces ; (2) d'étudier sa structure pour la relier à une fonction biologique ; (3) de trouver des conditions qui ont pu rendre cette fonction avantageuse. L'hypothèse du protolangage de Bickerton est riche de ce point de vue. Le protolangage est défini² comme n'ayant pas de syntaxe ; il correspond à une modalité limitée de communication proto-humaine dont les propriétés sont qualitativement et fonctionnellement différentes de celles du langage proprement dit qui lui, se définit par sa syntaxe en tant que propriété constitutive. Il coexiste à l'état fossile avec le langage non seulement dans les données matérialisées des langues en tant que formes résiduelles mais aussi en tant que processus actif ou dormant. On le suppose donc toujours présent en arrière-plan de la construction continue des langues et de l'apprentissage du langage et l'on s'attend ainsi à ce qu'il surgisse à travers des usages communicationnels contraints³ tels ceux que Bickerton cite pour appuyer son hypothèse. C'est là que le lien entre le « langage primitif » de De Brosses et le « protolangage » de Bickerton est envisageable en dépit de tout ce qui, à l'évidence, les sépare.

Que tirer de cette rencontre ? De cette fausse collusion ? Tout d'abord ce constat que l'idée sous-jacente à la notion d'un protolangage est une idée *simple* – ce qui ne veut pas dire qu'elle est inintéressante. Elle est *simple* et transcende les siècles⁴ car dès qu'il est question de s'intéresser aux *origines* on est toujours tenté de supposer un quelconque « proto- » : c'est de l'ordre du *topos*. Il y a donc un premier présupposé *linguistico-logique* fondé partiellement sur des faits d'expérience qui fait précéder tout « X » éventuel d'un « protoX » dont il dérivera, et un deuxième présupposé corrélatif, *idéo-logique*, qui avec une grande régularité attribue une fonctionnalité réduite au « protoX » (plus « élémentaire »/« simple ») face au « X » (plus « élaboré »/« complexe »).

Ensuite que c'est une idée *problématique*, et cela à plusieurs niveaux. Tout d'abord si on retient la notion du protolangage, alors il nous faut bien le placer quelque part dans l'évolution des espèces puisqu'à la différence de ce qui était le cas au 18^{ème} siècle, le niveau de précision de nos connaissances nous oblige à le faire. Et ce sera *Homo Erectus*. Mais là (en tenant compte du degré d'exigence supplémentaire que l'on se doit de s'imposer en raison de cette « augmentation » des connaissances et des exigences de la recherche qui caractérise notre siècle) il n'est pas sûr que les cas empiriques contemporains choisis comme illustration d'un succédané de protolangage pour étayer l'hypothèse (langage enfantin, pidgins, état de déprivation linguistique, etc.) soient qualitativement suffisants. Aucun de ces cas n'est assimilable exactement à ce qui serait un passage au langage à partir d'un état de non-langage du temps de *Homo Erectus* car aucun des phénomènes retenus ne se manifeste aujourd'hui sans un contact *Homo Sapiens* et sans au moins la perception d'un langage évolué⁵

¹ C'est en fait elle que Guiraud (1967) a repris ou retrouvé dans ses études d'étymologie structurale en fondant le dynamisme impliqué dans l'immanence de la langue.

² Mais n'oublions pas que ces coupures qualitatives : langage, protolangage, voire prélangage avec *Homo Habilis* ne sont que le fait de décisions théoriques pour mettre en cohérence des hypothèses sans beaucoup de support empirique.

³ Laissons de côté la question de l'enseignement aux grands singes que De Brosses ne pouvait probablement pas imaginer comme « raisonnable ».

⁴ Condillac et le langage d'action, Vico et la langue des héros, autant de références qui sans développer l'idée du « protolangage », contiennent des éléments qui la croisent. Une étude dans ce sens serait à faire.

⁵ Sans compter que, même après mutation, il n'est pas certain qu'un ci-devant *Homo Erectus* se comporte comme un petit *Homo Sapiens* de 200 000 ans plus âgé !

dans son environnement. Le langage (évolué) est toujours « déjà co-présent » à ces réalités-limites et un état de langage non maîtrisé et conjoncturellement non partagé défini dans ce contexte ne saurait « simuler » correctement un état de langage émergeant sur fond de protolangage¹.

Enfin je dirais que c'est une idée *clôturée* en raison de la contrainte structurelle propre aux opérations de pensée censées la conforter. En effet, l'hypothèse se soutient de deux procès.

Le premier est un procès *d'homéostasie* normal dans la construction scientifique comme dans la vie « de tous les jours » : il s'agit de croiser l'hypothèse avec l'ensemble de toutes les autres propositions corrélativement retenues et de vérifier dans quelle mesure une cohérence interne est atteinte, conservée ou améliorée par cet apport. Il s'agit là de ce procès normal (et normalisateur !) qui est à l'œuvre dans toute construction de connaissance, savante ou non, et qui n'engage pas la réflexion sur le contenu de l'hypothèse.

Le second procès est celui de l'expérience de pensée. Celle-ci a pour fonction de montrer (illustrer) la nécessité et/ou la *raisonnabilité* logique et matérielle de l'hypothèse dans les conditions de l'expérience supposée. Dans certains cas *l'expérience de pensée* sera un outil efficace de recherche conjecturale, sauf que dans celui qui concerne l'hypothèse du protolangage il n'est pas sûr que nous soyons capable de construire effectivement une *expérience de pensée* dont la finalité serait de permettre d'imaginer la « logique » des comportements proto-humains d'il y a 200 000 ans afin de comprendre les phénomènes d'émergence du langage dans ce contexte.

Si une expérience de pensée se fonde sur la *raisonnabilité* ainsi que je l'ai suggéré précédemment (et il ne peut pas en être autrement) alors *rien* ne nous donne l'assurance que les faits, comportements et attitudes supposés que cette expérience est censée « *raisonnablement mimer* » aient quelque justification de se conformer à ce *mime raisonnable* que l'on introduit sur la base de la « raison » d'aujourd'hui et de conjectures. Et cela non pas parce que les faits considérés sont éloignés dans le temps mais tout simplement parce que nous n'avons aucune certitude de disposer de toutes les variables pertinentes susceptibles d'intervenir dans l'expérience, sauf à présumer un uniformitarisme des comportements cognitifs déduit de notre fonctionnement *Homo Sapiens* actuel. Il apparaît que si je me proposais de *construire* une expérience de pensée, aussi « farfelue » fût-elle, ce serait – malgré tout – avec les données maîtrisées du *connu* et du *raisonnable* que je le ferais². Elle ne saurait contenir davantage que ce que je sais déjà, car sans dériver de la rationalité élémentaire, je ne peux mettre en scène prospectivement que ce que je peux déduire ou induire de l'état présent et de mes connaissances.

Une expérience de pensée possède une fonction heuristique ; elle est efficace pour toutes les configurations de recherche où ce qui est en jeu est soit sans lien avec une description de faits empiriques comme c'est le cas pour soutenir un raisonnement philosophique ou toute autre élaboration de nature strictement conceptuelle ; soit en rapport avec un état du monde dont il s'agit de rendre compte, mais tel qu'il corresponde à un comportement prévisible le plus souvent cerné dans un protocole d'expérience dont les variables sont clairement définies pour un cadre théorique donné. La non-conformité éventuelle entre le comportement constaté et le résultat « modélisé » dans l'expérience de pensée est alors un phénomène à expliquer, ce qui fera avancer la recherche dans ce contexte de description d'un état du monde. L'expérience de pensée ne prend sa valeur que parce qu'elle introduit la possibilité de cet écart.

Or, dans le cas qui nous occupe nous n'avons tout simplement pas de données empiriques concernant l'état du monde que nous considérons tout en agissant comme si nous étions dans un contexte de description empirique. Nous n'avons aucune certitude sur le fondement « empirique » de notre « expérimentation » ; nous n'avons surtout aucune assurance que tous

¹ Ce que Bickerton n'omet pas de souligner, tout particulièrement en ce qui concerne les pidgins.

² Il est intéressant de ce point de vue de constater combien les illustrations des voyages dans la Lune pensés au 18^{ème} siècle souscrivaient à la poétique et à l'imaginaire de leur temps ; et combien les illustrations de Jules Vernes renvoyaient au « technologisme » mécanique de son siècle. Je laisse aux jeunes lecteurs d'aujourd'hui le soin d'évaluer le rapport de raisonnabilité concernant les voyages intergalactiques en hyperspace du 21^{ème} siècle que les séries télévisées du type « *Stargate* » nous donnent à digérer.

les phénomènes empiriques pertinents dans le processus évolutif que nous souhaitons décrire sont *raisonnablement* pensables (et pensés) dans la clôture de raisonnabilité que nous avons nécessairement introduite. Et bien évidemment, aucune opération d'induction ou de déduction ne peut nous permettre de prévoir l'imprévisible ou de dériver sans prémisses.

Dans ces conditions donc, aucune confrontation à l'état du monde n'étant possible c'est sur la base du seul procès d'homéostasie que sera garantie la « légitimité » des hypothèses et propositions élaborées. Et cette base, à la fois non-philosophique et non-ancrée sur l'expérience empirique, est ici tout simplement, « discursive »¹. Il s'agit d'un ordre de légitimité possible mais si on le retient il importe de reconnaître qu'il ressortit aux critères d'une « scientificité particulière », ce qui pourrait, finalement, contribuer à enrichir les débats.

L'on peut concrètement espérer de parvenir à un état d'homéostasie qui permettrait d'actualiser une « histoire » sans garantie de justification empirique avec l'espoir d'atteindre une plus grande proximité supposée avec des faits objectivement inconnaissables dont l'on est censé se rapprocher², mais on ne saurait, ici comme ailleurs, briser la clôture introduite par la raisonnabilité³ sans avoir pour sanction de sauter des fictions de la science à la science-fiction, changeant ainsi le mode d'être des propositions avancées.

Pour synthétiser : dans le cadre d'une expérience de pensée je ne fais jamais que « reproduire » et « faire jouer » sur fond de « raisonnabilité » ce que je juge bon d'avoir posé au départ. Sans plus. Le contexte qui serait celui de l'actualisation d'un protolangage est encore connu de façon trop incertaine et trop élémentaire pour qu'une expérience de pensée garantisse quelque chose à son propos. J'en conclus cependant – peut-être parce que la science sans une pincée d'irrationnel serait trop triste – que l'hypothèse du protolangage n'en est pas moins une hypothèse intéressante : il importe seulement que l'on précise sans ambiguïté la qualification du type d'information qu'elle nous fournit.

Mais on se tromperait à interpréter ces commentaires comme une critique négative, ils se veulent tout simplement une recherche de lucidité. Notons que, quelle que soit la valeur d'une hypothèse comme celle du protolangage et son ordre de légitimité, son intérêt est évident dans ce qu'elle donne à « réfléchir » et dans ce qu'elle conduit à « structurer ». Mentionnons encore que les idées qui provoquent réflexion et enthousiasme sont parfois le symptôme d'une certaine qualité et que de ce point de vue il est rare que le renouvellement d'une réflexion sur le langage – et le questionnement sur son origine en est un – n'ait pas d'effets positifs, éventuellement hors du domaine initialement visé. Quand bien même il ne serait voué qu'au constat de ses limitations.

Au-delà d'une incitation à la lecture ou à la relecture directe de textes parfois trop vite jugés désuets, ce qui aura encore été intéressant dans cette exploration concerne la dynamique du langage dans toute sa généralité : les considérations de De Brosses comme celles de Bickerton, bien qu'incommensurables les unes aux autres, conduisent à mettre l'accent sur un phénomène important qui est la probable coexistence de « mécaniques », de « processus » parallèles et interconnectés qui participent à la construction des langues et à la mise en signification de leurs formes. Le « moteur » de ce *langage primitif* dont « *tout le monde [a] en soi tous les germes primitifs* », continue à produire nonobstant le développement postérieur de la dérivation et la naissance de la syntaxe ; les formes fossiles du supposé *protolangage* continuent à se manifester à l'occasion. De même qu'on peut reconnaître dans la matérialité des langues une stratification des formes référées à des « origines » différentes qui composent entre-elles, on peut supposer dans la dynamique de leur fabrication un entrecroisement de

¹ C'est bien ce que montre l'inflation prospective et conjecturale des propositions, théories et autres élaborations en recherche « d'inter-cohérence » autour de cette question.

² Ce qui est une « option » philosophique.

³ Finalement – et c'est heureux – la nature et les hommes ne sont pas *nécessairement* raisonnables. Et il est probable que beaucoup de découvertes, si elles n'avaient dû être amenées que par la raisonnabilité des expériences de pensée n'auraient jamais vu le jour. Ce qui, précisons-le, ne minimise en rien l'intérêt des expériences de pensée. L'on peut alors, bien évidemment, souscrire à l'hypothèse du *protolangage*, pourquoi pas, car il s'agit d'une hypothèse séduisante, mais dans le même temps il est intéressant de constater comment elle persiste entre des mondes intellectuels sans rapport tel le 18^{ème} siècle des Lumières et le 21^{ème} siècle des cognitivistes.

processus agissants issus aussi de différentes « origines » mais qui, finalement, composent entre eux pour les constituer dans leur matérialité et ont un rôle dans leurs fonctionnalisations en tant qu'organisations sémiotiques. Le phénomène « langue », pas plus que le procès de l'émergence du langage n'aurait d'autre homogénéité que celle que nous lui réassignons continûment.

Qu'est-ce que cela veut dire ? N'y aurait-il pas ici quelque chose d'important autour de cette notion de stratification de pertinences indépendantes à mettre en rapport avec le processus de sémiotisation interne à la fonctionnalisation du langage qui lui, ne se fonde pas uniquement sur des pertinences physiologiques et neuro-cognitives mais sur la construction contextualisée des symbolismes dans leur historicité ? Rien ne permet de faire de tels ponts. Mais les ferait-on qu'ils ne seraient pas plus dangereux, et pourraient être aussi riches de perspectives que ceux qui font l'enjambement du protolangage au langage.

2. Généalogie des langues

Tant que la paléontologie n'avait pas fourni les données qui ont permis de faire reculer de quelques centaines de milliers d'années le questionnement sur l'apparition de l'homme par rapport aux dates supposées de la Genèse, s'intéresser à l'origine du langage et s'intéresser à l'origine des langues revenait à peu près à la même chose, mais la reconnaissance d'une linéarité évolutive, intégrant ou non des mutations intermédiaires qui, passant des ancêtres d'*Homo Habilis* conduit à *Homo Erectus* pour aboutir à *Homo Sapiens Sapiens* change tout. S'intéresser à l'origine du langage revient à considérer des transformations à l'échelle de l'évolution des espèces et à envisager des conjectures adaptées à cette échelle tandis que s'intéresser à l'origine des langues demande de considérer l'émergence et le développement du phénomène dans une « niche écologique » bien définie à l'intérieur de laquelle aucune mutation n'est envisagée. L'échelle n'est pas la même, les outils et les méthodologies de recherche sont qualitativement différents. Le questionnement sur la généalogie et *l'origine des langues* relève ainsi d'une problématique beaucoup plus « localisée » que celle de *l'émergence du langage* car il se situe à une échelle temporelle incomparablement plus réduite et ne met en jeu (de façon nécessaire) aucune spéculation à propos de la genèse et de l'évolution de nos capacités cognitives.

On s'intéressera ici à quelques aspects de la recherche sur la généalogie des langues et l'on se demandera si la nouvelle irruption du questionnement sur l'origine des langues dans le siècle se fait en rupture avec les approches antérieures, ou bien si elle témoigne d'une certaine continuité. Et de quel ordre.

2.1. Le programme de Süßmilch

J'ouvrirai le débat en prenant pour thème un mémoire présenté à l'Académie Royale de Berlin au milieu du 18^{ème} siècle. Son auteur se proposait de démontrer la parenté du celte (identifié à travers ses évolutions supposées : l'allemand, l'anglais et le hollandais) et des langues orientales (appréhendées à travers l'hébreu, l'arabe, le chaldéen, le syriaque... et le persan¹). Il ne s'agissait donc pas là de remonter à l'origine des langues du monde² mais, modestement, de tenter de prouver une communauté d'origine entre les langues citées ; hypothèse souvent reprise au cours du siècle.

Tout d'abord de qui s'agit-il exactement ? Qui est Süßmilch ? C'est un théologien luthérien, ignoré de la plupart des linguistes car ses travaux sur les langues n'ont pas laissé beaucoup de traces, mais il est très bien connu dans le monde des historiens de la démographie pour avoir écrit le premier grand traité de démographie en langue allemande³. Il utilisera les listes de décès, de naissances et de baptêmes pour déterminer le nombre des vivants à différentes époques, dans les villes (Berlin, Paris, Londres) comme dans les campagnes et il présentera

¹ La généalogie des langues indo-européennes n'ayant évidemment pas encore été élaborée cela explique probablement que le persan se retrouvait avec les langues sémitiques.

² La perspective adamique sort rarement de l'univers de la Genèse.

³ *Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, Tod und Fortpflanzung des selben erwiesen* est daté de 1741 dans sa première édition mais il a été modifié et transformé ultérieurement (1761-1762). La version de 1741 a été traduite et publiée par l'Ined (1998). *L'ordre divin dans les changements de l'espèce humaine démontré par la naissance, la mort et la propagation de celle-ci* qui par cette publication entendait rendre disponible en français un auteur considéré comme un des pères de la démographie, avant Malthus.

une somme de données chiffrées jamais encore rassemblée. Ses analyses en ce domaine sont loin d'être naïves : il introduit l'approche arithmétique à l'anglaise, il prend en compte la spécificité des conditions humaines et sociales en rapport avec ses données, il montre l'utilité des statistiques concernant les maladies, il développe une réflexion sur les obstacles à la multiplication de l'espèce humaine, sur la nuptialité, la mortalité différentielle, etc. Autant de thèmes qui, pour les spécialistes, font de son travail une étude pionnière. Parallèlement, pasteur marqué par la philosophie de Leibniz et de Wolff (qui préfacera son grand ouvrage), il recherche les régularités du monde et les lois de Dieu à travers ce travail. Autrement dit, en s'intéressant à l'étude de la naissance et de la mort et à la propagation de celle-ci, il veut saisir – ainsi que l'explique le titre de son ouvrage – cette trace de l'*ordre divin* qui justifie ce monde comme le meilleur des possibles¹.

De plus, comme la plupart des savants du 18^{ème} siècle, rien du connaissable de son époque ne lui est étranger et c'est en tant que membre de la Classe Philologique de l'Académie Royale des Sciences de Berlin à laquelle il appartient en raison de ses recherches érudites sur les langues qu'il a présenté le mémoire auquel nous nous intéressons. Pour apprécier son étude trois points sont à considérer : (1) les *arrière-plans théoriques* concernant la nature et le comportement des données de langue, (2) l'*approche méthodologique* et le *système de démonstration* et enfin (3) le détail des comparaisons lexicales proposées.

On s'intéressera aux deux premiers qui par leur généralité conservent une importance aujourd'hui. Je ne m'attacherai pas au troisième car il est à la fois trop spécialisé et bien évidemment suffisamment dépassé par l'état actuel des connaissances pour qu'il n'y ait pas lieu de s'appesantir sur la pertinence de ces comparaisons. Cependant on retrouvera en annexe l'ensemble des attestations qu'il a fourni dans son mémoire à l'appui de sa thèse.

2.1.1. Procédure et méthode

En 1746 donc, un pasteur luthérien du nom de Süßmilch présentait à l'Académie Royale des Sciences et des Belles Lettres de Berlin un mémoire dans lequel il entreprenait de démontrer « *la convenance de la Langue Celtique, et en particulier de la Teutonique avec celles de l'Orient* » et de prouver que la « *langue Teutonique est matériellement contenue dans les Langues Orientales, et qu'elle en descend* ». Il reprenait ainsi à son compte ce « fait » supposé de notoriété publique que « *la multitude de mots des langues occidentales ou celtiques, dont le son & la signification s'accorde avec les Langues Orientales est si grande, qu'elle conduit ceux qui n'y penseraient pas d'eux mêmes à former sur ce sujet des conjectures très probables* ». Il ne restait plus alors qu'à démontrer par le menu ce qui n'était qu'un sentiment non encore soutenu par des éléments organisés en faisceau de preuves : que « *la langue celtique, ou du moins l'une de ses filles, est entièrement contenue dans les langues orientales, que par conséquent elle en procède, ou qu'elles viennent toutes d'une source commune* ». L'hypothèse corrélatrice est que les anciens Celtes dont les Teutons faisaient partie, parlaient autrefois les langues orientales. Encore fallait-il le « prouver ». Ce qu'il fera par l'examen attentif des langues que parlent les descendants supposés des Celtes et Orientaux.

Mais comment procéder ? Comment passer du « *montrer* » au « *démontrer* » ? Pratiquement, au départ il importait de faire un premier tri entre les mots d'emprunt et les mots indigènes. L'auteur s'en explique : « *Je cherchai des mots qui fussent propres à notre langue, et que la Navigation ou le commerce n'y eussent pas introduits* » puis il s'attelle à une première comparaison : « *Les épreuves que je fis sur ces mots ne manquèrent jamais* ». Ensuite pour aller plus loin, il choisira arbitrairement un échantillon de mots caractérisé par une propriété phonétique particulière et il le soumettra à la comparaison : « *je choisis une longue suite de mots, qui commençassent tous par Sch ou le Schin des Orientaux* ». Enfin il étendra son étude en traitant exhaustivement un second échantillon de langue, il s'agit de l'ensemble des mots commençant par la lettre R et finalement, il inférera des résultats de son analyse à l'ensemble des langues considérées. Et le tour sera joué.

¹ Bien évidemment ce n'est pas l'objectif ici de développer une monographie sur l'auteur, l'on peut se référer à l'édition de *L'ordre divin...* par l'Ined ou pour quelque idée de la réflexion critique de l'auteur dans son approche des données, on peut aussi renvoyer à la lecture de son court *Essai dans lequel on se propose de déterminer le nombre des habitants de Londres et de Paris*, Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin, 1759. S. 453-463. - Berlin : Haude et Spener, 1766.

Süßmilch fonde sa démarche sur un raisonnement inductif :

Si je puis prouver de tous, ou au moins de la plupart des mots d'une Lettre de l'alphabet, qu'ils s'accordent avec les Langues Orientales, il en résulte que l'on peut prouver la même chose à l'égard de toutes les autres Lettres de l'Alphabet. La vérité de ma conséquence est fondée sur ce principe incontestable : Ce qui peut être affirmé sous certaines conditions d'une partie de quelque Tout, peut être affirmé sous les mêmes conditions des autres Parties, & par conséquent être attribué au Tout entier. La vérité de cette proposition est confirmée par l'Expérience, & lui sert de Règle

et, puisqu'il vérifie son hypothèse par l'interprétation qu'il fait de ses données, il sera amené à conclure qu'« [i] est par conséquent vrai que la Langue Allemande se trouve contenue dans les Langues orientales, qu'elle en dérive, ou du moins que les unes & les autres découlent d'une source commune ». Mais plus précisément, quelle proportion de résultats positifs soutient sa comparaison des langues ? Il présente ses résultats en trois catégories. Et ils sont « écrasants », les voici :

1. *Les comparaisons certaines.* L'échantillon de langue qui comporte la lettre R renferme autour de 100 racines. Dans cet échantillon il établit qu'environ 70 mots montrent une conformité parfaite entre la langue celtique ou allemande et les langues orientales, c'est-à-dire près des trois quarts : « *Non seulement les mots en eux-mêmes, & les lettres dont ils sont composés... mais encore la forme, c'est-à-dire leur signification est parfaitement la même dans la Langue Celtique ou Allemande, et dans les Langues Orientales* ». Comme il ne s'agit cependant pas d'identité entre les mots comparés il précise que « *la légère différence qui se trouve souvent entre un mot dans une Langue & ce même mot dans l'autre, ne peut faire aucune peine à ceux qui sont versés dans les langues* », et pour appuyer davantage son argument sur l'aspect « ordinaire » de ces légères variations il renvoie aux différenciations dialectales que l'on constate à l'intérieur des langues elles-mêmes. Ainsi : « *Que l'on pense seulement combien diffèrent pour l'ordinaire le Haut & le Plat Allemand, ou l'ancien Saxon et le Franconien. Par exemple un paysan de la Marche dit : Watt bebb ick met guw tu dubn, tandis qu'un Berlinoise s'exprime ainsi, was babe ich mit euch zu tbun ; et malgré ces variétés c'est la même langue* ».
2. *Les comparaisons probables.* Le reliquat de l'inventaire est constitué dans sa quasi-totalité de mots pour lesquels « *on ne trouve pas dans la matière & dans la forme cette convenance qui a lieu dans les précédents, mais dont on peut pourtant découvrir la dénomination, sans l'aller chercher trop loin, ni faire aucune violence* ».
3. *Les comparaisons rejetées.* Enfin « *Il y a peut-être entre quatre ou cinq mots [...] qui sont restés en arrière et que je n'ai pu faire venir d'Orient d'une manière très vraisemblable [...] Mais il n'en peut résulter aucun argument contre ma proposition. Quiconque est au fait des grands changements qu'éprouvent les langues, ne se laissera point arrêter par si peu de chose* ». Et il justifie ce dernier commentaire en renvoyant à la relativité des ressources dictionnaires et à la parfois rapide modification des unités lexicales des langues.

Finalement il résume ainsi :

Je pose qu'une langue soit composée de 100 mots. Si je démontre que de ces cent mots, 70 sont certainement, & 25 vraisemblablement, en tout 95, contenus dans une autre langue, de sorte qu'il n'en manque que 5. Personne ne saurait contester que la Langue qui consiste dans ces cent mots ne soit contenue dans l'autre, et qu'ainsi elle n'en procède, ou bien qu'elles sont toutes deux filles d'une même Mère.

Deux siècles plus tard le titre même de sa communication pourrait prêter à sourire mais toutefois l'étude de la démarche intellectuelle de l'auteur peut encore nous apprendre beaucoup sur nos pratiques.

Bien évidemment, une telle approche demande à être replacée dans son siècle pour mieux apprécier sa méthodologie et son système de démonstration en rapport avec les arrière-plans théoriques de l'époque.

2.1.2. Arrière-plans théoriques concernant la nature et le comportement des langues

On ne pourrait pas rendre compte des phénomènes de l'évolution des langues si l'on n'avait pas *a priori* quelques idées sur ce qui est pertinent dans leur dynamique et, sur ce point, les idées de Süßmilch n'ont pas beaucoup d'originalité par rapport à son siècle. L'inventaire qui suit en énumère quelques-unes :

- L'évolution des langues peut être appréhendée à travers la métaphore anthropomorphique d'une filiation mais cela ne permet cependant pas de l'interpréter comme une généalogie à strictement parler. Ou plutôt il semble aller de soi que l'on n'a pas les moyens de décider sur ce point, ou encore on retient l'image mais en la prenant pour ce qu'elle est.
- En cohérence avec la métaphore anthropomorphique généralement utilisée pour traiter de leur évolution il va de soi que les langues peuvent naître, devenir adulte, dégénérer et mourir.
- Aucune notion de système, de loi ou de pression structurale active n'est reconnue.
- L'action des hommes sur la langue est un facteur important dans leur transformation.
- L'importance de l'emprunt lexical, des échanges interculturels et des contacts interlinguistiques de toutes sortes est considérée.
- L'idée de langues mélangées susceptibles de résulter des contacts et échanges au cours du développement des langues, et intégrant toutes les modifications envisageables est aussi considérée.
- Les langues peuvent être caractérisées par leur génie mais celui-ci n'implique pas nécessairement une vision téléologique de leur évolution.
- Le caractère des peuples tout autant que les facteurs climatiques peuvent déterminer certains aspects de leurs modifications.
- Il n'y a pas de problème particulier en ce qui concerne le rapport entre les sons et les lettres, entre la phonétique et les orthographes. Celles-ci, selon leurs différents niveaux d'évolution et leurs différentes modalités de représentation réfèrent aux mots au même titre que celles-là.
- Les différences phonétiques que l'on perçoit entre les mots des langues comparées s'expliquent par l'action de facteurs humains mais elles ne sont pas censées souscrire à des régularités particulières qui permettraient de les figer en lois.

Ces idées sur la filiation globale, l'incidence du facteur humain et la référence à une causalité éclatée qui porte sur les éléments et non sur le système des langues se retrouvent en effet en arrière-plan de tous les travaux du siècle. Ainsi, par exemple, à la même Académie de Berlin, l'Abbé Denina (1794), que l'on pourrait qualifier de « philologue antédarwinien », et qui distinguait entre les approches *conjecturales* sur « l'origine des langues » et les approches *documentées* sur des langues particulières¹, notait que :

[o]n peut [...] poser pour principe que les dix ou douze langues qui se parlent & s'écrivent en Europe, ainsi que leurs différents dialectes, sont sorties d'un langage très ancien qui se parlait dans l'Asie occidentale & dans l'Europe orientale [...]. Les causes de ces différences se réduisent à deux classes, dont une peut se dire physique, l'autre morale ou métaphysique. (p. 72) La première consiste dans le changement des éléments de la parole, c'est-à-dire des lettres tant voyelles que consonnes ; & cette variation est l'effet de l'organisation, comme celle-ci l'est du climat ou de l'air. L'autre vient du sens différent que l'usage, le caprice, les circonstances donnent aux mots primitifs.

¹ (p. 68) *Toutes les langues ont leur commencement & le premier fond de leurs mots, dans les premiers monosyllabes que prononcent les enfants & les hommes sauvages, qui, à cet égard, sont toujours enfants. [...] (p. 70) Deux auteurs très célèbres parmi ceux qui ont écrit sur la formation des langues, trouvent qu'il y a du rapport entre la forme des éléments de la parole, voyelles ou consonnes, & les mouvements que l'on fait en les exprimant ... (p. 71) Mais toutes curieuses & fines que soient ces observations que nous lisons dans le volumineux ouvrage de Court de Gebelin, & dans l'excellent Traité du président de Brosses (qui servent sans doute à nous faire remonter à l'origine des langues en abstrait), elles ne nous aident que peu ou point pour apprendre les langues que l'on souhaite de savoir. Pour les rendre utiles en pratique, il faut, ce me semble, faire abstraction de la langue primitive, qu'on ne saurait absolument trouver, puisque toutes celles que nous pouvons connaître ont été certainement précédées par d'autres plus anciennes ; & partir de quelques-uns de ces idiomes anciens connus, & dont on a lieu de croire que ceux que l'on parle & que l'on écrit encore actuellement en Europe son sortis.*

2.1.3. Approche méthodologique et système de démonstration

Mais les hypothèses sur la nature et le comportement des langues ne deviennent intéressantes que si on leur adjoint – explicitement ou non – une méthodologie de la découverte pour aider à répondre aux questions que l'on (se) pose. Ainsi on admettra que :

- Le lexique est le matériau privilégié pour établir que deux langues ont un rapport généalogique entre elles (disons « qu'elles procèdent l'une de l'autre » ce qui est moins spécifique) ; les mots sont donc des objets suffisamment stables et fiables pour qu'on puisse les utiliser pour l'élaboration d'hypothèses : c'est par leur comparaison qu'on peut évaluer les rapports qui existent entre les langues. Corrélativement, les autres propriétés et caractéristiques linguistiques telles que la morphologie et la syntaxe n'ont qu'une place limitée dans l'élaboration de la preuve.
- On se doute que les ressemblances phonétiques et sémantiques retenues dans les comparaisons peuvent dans certains cas résulter du hasard mais on considère que si l'on procède à une étude quantitative raisonnée portant sur un grand nombre d'unités cela permet d'éviter que les hypothèses soient faussées par ce facteur. Corrélativement le chercheur est *légitimement* le « grand décideur » de la reconnaissance de la ressemblance et de la pertinence des rapprochements. C'est son rôle.
- C'est par une pratique de l'étude des mots, détaillée, documentée et qui croise l'ensemble des informations linguistiques et historiques disponibles que l'on envisage de proposer des conjectures sur le rapport des langues. On tente d'appréhender leur composition à partir de l'inspection des sources variées que l'on peut percevoir à travers leur matériel lexical. Le facteur quantitatif est donc pris en compte à ce niveau et il en découle que pour l'essentiel, les langues ne sont pas considérées comme le résultat d'une pure filiation.

Face à ces pratiques, l'approche de Süßmilch tranche par sa radicalité : anachroniquement, on dirait aujourd'hui qu'il exporte une méthodologie et un mode de raisonnement et pose de ce fait le problème de l'interdisciplinarité. Tout se passe ainsi : l'auteur va conserver l'ensemble des postulats de son époque concernant les langues en général. Spécialiste averti et critique, pionnier dans un champ disciplinaire à la création duquel il contribue fortement (celui de la démographie), il applique avec naïveté au domaine linguistique dont il ne perçoit pas (ou ne retient pas) la spécificité, des procédures qu'il a l'habitude d'utiliser dans son domaine d'excellence. Il introduit donc, pour « démontrer » son hypothèse, une procédure de raisonnement utilisée dans les expérimentations et fondée sur le principe d'induction (où de la caractéristique d'un sous-ensemble linguistique particulier préalablement défini par une propriété indépendante et arbitraire par rapport à la question posée l'on s'autorise à généraliser à l'ensemble total). Cela permet de penser que de son point de vue, hors l'emprunt évident, toutes les unités de la langue ont des caractéristiques et des comportements identiques quant à leur développement possible.

Il met ainsi sur le même plan deux types de phénomènes hétérogènes : ceux qui dans leur matérialité ne relèvent d'aucune latitude interprétative (tel que la mention positive sur un registre d'une naissance, d'un décès ou d'un baptême¹) et ceux qui dans leur matérialité relèvent d'une évidente interprétation (tel le degré de ressemblance phonétique et sémantique entre deux mots appartenant à des langues différentes) ; et il n'a pas le même degré de rigueur critique dans sa collation des données de langues que celui dont il fait montre lorsque par exemple il analyse et interprète le potentiel gonflement des chiffres qu'il trouve dans les registres concernant la population de Londres par rapport à celle de Paris (1776).

Nous identifions ici un premier problème concernant la connaissance préalable de la nature de l'objet à décrire et un deuxième concernant l'exportation d'une méthode hors du domaine « naturel » de son application.

2.2. La généalogie des langues aujourd'hui

Lorsqu'on s'intéresse aux recherches contemporaines sur ce thème on constate qu'elles renvoient à deux univers intellectuels distincts et à deux approches méthodologiques

¹ Bien qu'à un autre niveau d'analyse l'interprétation de la notation et de son contexte puisse être importante.

différentes tout en souscrivant à un même cadre théorique en ce qui concerne les modalités générales de l'évolution des langues.

- La première approche née au 19^{ème} siècle avec le développement de la linguistique indo-européenne a acquis ses lettres de noblesse mais elle est aujourd'hui moins productive en questionnements nouveaux : c'est le domaine de la grammaire comparée, de l'étude des correspondances phonétiques et de la reconstruction des langues. Corrélativement, c'est au développement de cette approche-là que les travaux conjecturaux sur l'origine des langues et du langage ont dû de perdre leur crédit.
- La seconde, née dans le courant du 20^{ème} siècle, correspond à un effort pour trouver des méthodes pratiques susceptibles de mettre en évidence des apparentements généalogiques lorsque la méthode classique n'est pas applicable (entendons par là : lorsque la nature des langues à comparer et les données disponibles ne permettent pas de dégager les types de régularité qui sont censés établir la preuve d'un apparentement généalogique). Elle tend aujourd'hui à se subdiviser selon deux directions qui malgré l'apparence d'un objet de recherche partagé se distinguent nettement par leurs méthodologies et leurs objectifs.

Je ne présenterai pas l'approche bien connue du 19^{ème} siècle et me contenterai d'esquisser le cadre de la seconde¹ que je diviserai en une approche *empiriciste*² et une autre *formaliste*.

L'*approche empiriciste* est généralement conduite par des linguistes, elle ne se fonde pas sur la recherche de correspondances et l'élaboration d'une grammaire comparée au sens de la linguistique historique mais sur trois postulats :

- que les apparentements généalogiques peuvent se manifester à travers la mise en évidence de ressemblances à la fois phonétiques et sémantiques entre les mots d'un « *vocabulaire de base* » préalablement établi de telle façon qu'il soit censé éviter au maximum le risque de statuer sur des mots empruntés ;
- que la comparaison de ce vocabulaire doit porter sur le plus grand nombre possible de langues x afin de limiter les risques de rapprochements intempestifs dus au hasard ;
- qu'en raison même du nombre des langues comparées les grands sous-ensembles de langues vont se dessiner de façon « évidente »³.

Ainsi une comparaison lexicale globale peut *montrer* que l'ensemble des langues italiques se distingue par son lexique d'un ensemble des langues slaves qui lui-même se distinguera d'un ensemble des langues germaniques, etc. Les positions limites de cette approche ont conduit aux hypothèses du type « *proto-Human* » (cf. Ruhlen) ou « *nostratique* » (cf. Dolgopolski).

L'*approche formaliste* est plus récente, elle renvoie à des travaux où se croisent des disciplines formelles, empiriques et expérimentales (modélisation informatique, recherches cladistiques, taxonomies, génétique des populations, paléo-anthropologie, archéologie,

¹ La linguistique historique fait encore partie des disciplines universitaires, on en trouvera des présentations dans de nombreux manuels. La deuxième approche, plus récente et contestée par certains, fait partie du bagage linguistico-médiatique de tous ceux qui s'intéressent aujourd'hui aux questions de l'origine des langues. Pour des développements dans cette voie on peut citer M. Ruhlen (1994). Pour une critique méthodologique, épistémologique et théorique de cette approche et une certaine vision de la controverse qui a accompagné sa naissance je renvoie à Nicolai (2000), mais bien évidemment, beaucoup d'autres auteurs pourraient être cités. Voir aussi Boë (L.-J.), Bessière (P.) & Vallée (N.). (2003) pour une critique fondée sur l'argument probabiliste.

² J'introduis le terme « *empiriciste* » (qui connote à la fois un « rapport aux données » et une catégorisation en *-iste /-isme* avec ses qualités et ses défauts) parce que, bien que cette approche s'appuie sur des données empiriques, ses choix méthodologiques ne permettent pas de la qualifier de « *empirique* » ; Cf. Nicolai (2000) pour une réflexion sur ce thème.

³ C'est cette « évidence » qui est censée se manifester à tout un chacun qui inspecterait le moindre tableau d'exemple, rendant triviale la construction des grands regroupements ; au même titre que Süßmilch notait que cette capacité à reconnaître le rapport entre les langues d'origine « celtique » et les langues orientales était si évident « *qu'elle conduit ceux qui n'y penseraient même pas à former sur ce sujet des conjectures très probables* ».

biologie, linguistique...)¹. Son objectif semble être de développer des procédures de description et des modélisations concernant la représentation des apparentements généalogiques saisissables à travers des analyses formalisées, et le cas échéant de serrer au plus près les phénomènes en appréhendant les dynamiques transversales telles celles de l'emprunt. Les questions posées portent autant, et peut-être davantage, sur les propriétés formelles et la valeur heuristique supposée de ces types de modèles pris en eux-mêmes que sur leur pertinence et leur limite dans la description des phénomènes de l'évolution des langues. Les études effectuées portent indistinctement sur la modélisation des représentations arborescentes indo-européennes bien connues ou sur celles d'autres familles linguistiques, et s'il y a un comportement-limite à mentionner, il ne conduit pas à une dérive des hypothèses sur l'origine des langues mais à une hypertrophie vers la construction de modèles pris pour objets de la recherche en lieu et place des réalités empiriques qui sont censées justifier leur emploi : on cherche davantage à affirmer les potentialités classificatoires du modèle qu'à discuter sa pertinence théorique pour l'utilisation que l'on en fait, et la critique des pratiques de la présentation et de la construction des données n'est pas envisagée. Les travaux s'appuient sur des données lexicales déjà élaborées tirées de ressources documentaires et de bases de données informatiques existantes ce qui constitue un matériau homogénéisé et bien adapté à un traitement formel. La participation des linguistes, bien que nécessaire, reste le plus souvent minimale. Pour mieux rendre compte de l'esprit de ce type d'approche je présente ci-dessous – à titre de symptôme – le résumé d'un article récent écrit par trois informaticiens/mathématiciens et un linguiste indo-européaniste connu (Erdem, VI. Lifschitz, L. Nakkhleh, D. Ringe) qui se proposent de modéliser les effets de contacts (soit donc un problème linguistique intéressant en lui-même) en rapport avec le développement arborescent « ordinaire » de l'évolution des langues. En s'en tenant à ce seul résumé, on trouvera à la fois cette attention envers la modélisation prise comme objectif (objet véritable ?) et la prospective de son application à l'organisation généalogique des langues :

The evolutionary history of languages can be modeled as a tree, called a phylogeny, where the leaves represent the extant languages, the internal vertices represent the ancestral languages, and the edges represent the genetic relations between the languages. Languages not only inherit characteristics from their ancestors but also sometimes borrow them from other languages. Such borrowings can be represented by additional non-tree edges. This paper addresses the problem of computing a small number of additional edges that turn a phylogeny into a "perfect phylogenetic network". To solve this problem, we use answer set programming, which represents a given computational problem as a logic program whose answer sets correspond to solutions. Using the answer set solver SMOODELS, with some heuristics and optimization techniques, we have generated a few conjectures regarding the evolution of Indo-European languages.

2.2.1. Modalisation du questionnement : constance et variation

Si, ce tour d'horizon achevé, l'on tente d'appréhender dans sa continuité le processus de recherche sur l'origine et la généalogie des langues, tout se passe comme si l'approche née au 19^{ème} siècle formait une vaste parenthèse à l'intérieur de laquelle une problématique s'est développée et a été traitée. De grandes questions ont été posées et elles ont été remodelées dans le cadre théorique qui s'est contextuellement construit avec la recherche de réponses les concernant ; puis les grands questionnements se sont tus, il n'est plus resté que le travail de fond qui toujours se poursuit dans le champ disciplinaire bien balisé. Parallèlement une démarque des réponses scientifiques apportées à ces grands questionnements a trouvé sa place dans

¹ Tendanciellement, les travaux des tenants de cette approche sont publiés non pas dans des revues de spécialité concernant les langues ou le langage mais dans des revues dédiées aux sciences expérimentales, sciences de la vie, ou tout autre support réputé diffuser les éphémérides de la recherche « de pointe » (Cf. *Nature*, etc.). Tendanciellement aussi les articles présentés sont cosignés par plusieurs auteurs manifestant ainsi la réalité sinon d'une réflexion de synthèse du moins d'un travail partagé en équipe et de la conjonction disciplinaire sur le thème. L'on se trouve formellement dans la configuration sociologique qui est le standard de référence de la production scientifique « de masse » effectuée par des chercheurs travaillant en laboratoire sur des projets collectifs financés, et tenus à une condition de production déterminante pour la qualité de leur avenir. On trouvera en référence quelques articles produits dans cette perspective : Cf. Rexová & al., 2003 ; Holden, 2002 ; Gray (R.) & Atkinson (Q.-D.). 2003 [*Nature*, vol. 426, 27 nov. 2003] ; Foster (P.) & Toth (A.). 2003 [*PNAS*, July 22, 2003, vol. 100, N° 15]. Bien évidemment beaucoup d'autres travaux de ce type pourraient trouver place ici.

la connaissance culturelle d'arrière-plan de celui que l'on aurait appelé au 17^{ème} siècle¹, « l'honnête homme ». Finalement, le domaine étant donné comme forclos, l'on a fermé la parenthèse.

Mais l'on ne revient jamais à l'identique sur un ancien questionnement : la parenthèse aura laissé des traces. Aujourd'hui, le retour à la problématique de la recherche de l'origine des langues, dans le même temps qu'il a conduit au dénigrement stratégique² des acquis de la linguistique historique va utiliser ses acquis conceptuels. Il n'y aura plus désormais de flou : l'évolution des langues sera pensée (et modélisée) à travers une arborescence généalogique dont le 19^{ème} siècle a généré le modèle et dans un formalisme strict. Beaucoup plus strict que ce qu'autoriserait à construire le détail des travaux des philologues confrontés à l'analyse et à l'interprétation de leurs données et conduits à concevoir la possibilité de l'indétermination des faits.

Corrélativement dans les approches empiricistes, l'attention aux mots, à leur histoire et à leurs transformations a disparu. La recherche pré-philologique est évidemment obsolète, la recherche philologique, stigmatisée, n'est pas reprise : l'on abordera – tout simplement – le problème par des généralisations intuitives sur des pourcentages de mots partagés dans des listes lexicales de qualité variable ; généralisations censées être justifiées par des procédures qui se veulent de nature probabiliste³ mais qui n'ont qu'exceptionnellement reposé sur des calculs de probabilités.

Dans les approches formelles, la collaboration des linguistes et non-linguistes peut parfois donner lieu à d'étranges dialogues où les non-linguistes soit reçoivent de linguistes « sûr d'eux » des matériaux directement utilisables pour faire fonctionner leurs modélisations (mais alors l'intérêt linguistique réel de l'opération est souvent faible), soit reçoivent de linguistes plus « précautionneux » des matériaux aux contours plus indéterminés et aux attributs plus problématiques qu'ils ont du mal à utiliser. On peut entendre alors différentes variantes de ce reproche type que, finalement, « *les linguistes ne savent pas faire quelque chose d'aussi simple que de fournir des mots correctement caractérisés pour qu'on puisse travailler avec !* ». Et curieusement, on remarquera que les problèmes posés par la tentative de dépassement sous-jacente à cette approche formelle ne sont pas sans liens avec ceux posés par l'initiative d'exportation de méthodologie à laquelle s'est livré Süßmilch.

À part cela, les mêmes présupposés théoriques qui avaient cours avant le 19^{ème} siècle sont conservés au 20^{ème} siècle : les « mots » du lexique restent le support privilégié de l'étude même si, conséquence résultant de la « parenthèse », la référence à l'éminence des rapprochements morphologiques peut être *discursivement* privilégiée. Ces « mots » sont toujours appréhendés à partir d'une ressemblance globale entre la forme et les significations et la tâche du chercheur est toujours d'identifier ces ressemblances, de rechercher et de décider des rapprochements⁴. Ils sont toujours donnés pour des « objets » empiriques qui, une fois saisis dans un dictionnaire ou dans une base de données, deviennent comptables, sans retour critique possible sur la réalité de leur pertinence et sur la décision de leur présence. Et l'on pourrait se demander si cet accord sur la fonctionnalité attribuée aux mots dans le modèle d'analyse ne contribuerait pas à renforcer l'interaction positive qui existe entre les tenants de l'approche formaliste et ceux de l'approche empiriciste car ces derniers fournissent aux formalistes des matériaux pour faire fonctionner leurs modèles tandis que les formalistes offrent un « lifting scientifique » aux empiricistes.

Non seulement on ne revient jamais exactement à un ancien questionnement mais la reprise n'a de sens que si on entend le dépasser, c'est là l'une de ses conditions. Ici le

¹ N'y aurait-il pas là, pour les amateurs d'évolutions darwiniennes et d'équilibres ponctuels, une ressource à exploiter ?

² Il est amusant à cet égard de lire l'ouvrage de M. Ruhlen (1994).

³ Précisons qu'il ne s'agit pas là d'une exigence de méthode, d'un *sine qua non* de la recherche mais plus simplement d'une intention, disons d'une référence métaphorique à un principe reconnu et généralement non appliqué sauf pour contester le détail d'une étude particulière.

⁴ La construction, finalement, « s'autodétermine » de sa propre mise en œuvre, renvoyant à une circularité de construction et à un processus d'homéostasie puisque c'est tout d'abord le chercheur qui décidera de la ressemblance et que c'est de la co-construction d'un accord intersubjectif sur la pertinence de son opération qu'elle sera validée.

dépassement sera tenté avec une *radicalisation* de la problématique généalogique¹ qui a conduit au développement de l'hypothèse de la monogenèse des langues du monde et tente de se donner les moyens de retrouver les formes de l'origine, relevant ainsi cette thématique que le 18^{ème} siècle des Lumières avait abandonné dans ses travaux les plus réfléchis pour la laisser au bénéfice du dogme. Cet avatar du mythe de Babel se soutient de conjonctions disciplinaires récentes et l'on postule alors qu'il a dû exister une langue-mère dont toutes les langues actuelles seraient dérivées selon une procédure de division dont le modèle darwinien de la spéciation fournit les références (*proto-Human*). La recherche consiste alors à tenter de « remonter » à cette langue originelle qui n'est plus réservée au mythe ni au dogme, qui n'est plus donnée comme une simple abstraction logique, qui n'est plus présentée comme une nécessité physiologique dont on peut conjecturer les constantes dans les débats d'idées (cf. De Brosses). Elle est posée comme une « réalité » dont on doit rendre compte dans le détail de ses formes lexicales à partir d'inférences comparatives (cf. les racines mondiales de Ruhlen).

Pratiquement le développement de cette hypothèse et sa radicalisation se sont autorisés de la mise en rapport de deux représentations arborescentes qui semblent se recouvrir : celle de la filiation des langues (Greenberg, Ruhlen) et celle de la transmission de certains gènes dans les populations (Cavalli-Sforza). De là à conclure qu'elles se corroborent l'une par l'autre le pas est vite franchi². Or, cela n'est probablement pas si simple, et sans même se prononcer sur la potentielle « véracité » d'une telle hypothèse (Qui sait ! La monogenèse ? Pourquoi pas !) des difficultés surgissent en ce qui concerne la possibilité de situer la problématique qui en traite dans le champ du « « scientifiquement » connaissable ». Difficultés qui, indépendamment de la « complexité » et de l'indétermination générée par la stratification des croisements possibles que le 18^{ème} siècle avait déjà souligné, portent sur l'interprétation des ressemblances entre les mots, la « logique » du modèle arborescent et tout simplement sur l'impossibilité formelle de prouver l'unicité de l'origine des langues.

L'identification et l'interprétation des « ressemblances » est problématique en raison de la difficulté à définir la notion même de *ressemblance* et ensuite en raison de la difficulté qu'il y a à construire une structuration indépendante du projet pré-établi du linguiste, lequel étant l'évaluateur de la ressemblance et le maître de l'hypothèse, est à la fois juge et partie.

De par sa nature, le travail de comparaison linguistique consiste à référer une multiplicité de formes à une représentation qui les subsume. Le modèle arborescent est ainsi déterminé dans son formalisme par cette *logique interne* laquelle, indépendamment de toute justification scientifique, oriente une fuite vers la « protoforme », la « proto-langue » ou tout autre « proto-X » que l'on voudra. On rejoint ici un *topos* prégnant dont il importe d'apprécier l'incidence.

De plus, même s'il s'avérait que l'ensemble des langues du monde puisse être référé à une même langue-mère (et qu'on en acceptât l'idée) cela ne conduirait qu'à l'hypothèse de la monogenèse *de l'ensemble des langues aujourd'hui attestées* (ou référées à des langues attestées). Ici, c'est la clôture de l'univers appréhendé qui est en question³. Cette impossibilité

¹ Etrangement (« cocassement » lorsque l'on sait les *junggrammatischen* diatribes lancées à l'encontre de la grammaire comparée), n'y aurait-il pas là une manifestation du même processus qu'a illustré la « radicalisation » néo-grammairienne ? Mais sans le support empirique qui lui assurait sa « scientificité ».

² Il est évident, et c'est généralement un fait positif dans une analyse, que la convergence d'un ensemble d'hypothèses indépendantes vers un résultat particulier est un point fort pour la validation de ce résultat. Cependant cela n'a de sens que lorsque les hypothèses ne sont pas de simples conjectures, sinon la validité du résultat est obérée par la somme des risques d'erreur que chacune des hypothèses implique. Ce qui ne donne pas nécessairement un statut positif à leur mise en faisceau. Corrélativement, avancer une hypothèse en prenant appui sur le *maillon faible* d'une analyse ou d'une description est fallacieux *a priori* parce que l'hypothèse globale dépendra ainsi de la valeur de ce maillon faible ; ce qui suggère une « circularité » de raisonnement où une cohérence va être fondée sur l'ensemble des points les plus litigieux des domaines de connaissance considérés.

³ Autrement dit : il est clair que travailler sur un échantillon de haricots blancs, même tirés d'un sac dont on saurait, pour une raison conjoncturelle indépendante, qu'il ne contient que des haricots blancs ne me permet pas d'inférer de leur couleur à l'impossibilité de haricots noirs dans l'univers des sacs de haricots ; même s'il se trouve (en supposant cet univers fini) qu'il est effectivement aujourd'hui impossible de trouver le moindre haricot noir. Résultat probable d'une catastrophe létale sur les espèces de haricots : les souris qui logent dans mon grenier (dotées d'un tropisme vers la couleur noire) les ont tous dévorés.

« d'assurer » l'hypothèse de la monogenèse a même été modélisée, ainsi Dessalles (2000) a proposé une modélisation de la dynamique de propagation en partant *a priori* d'une polygenèse¹ et il constate dans son modèle que :

[L]orsqu'une langue disparaît le sous-arbre, fictif bien entendu, de toutes les langues qu'elle aurait pu engendrer « disparaît » avec elle [...]. Au bout d'un temps suffisamment long, les hasards de filiations réussies feront que toutes les langues du moment auront le même ancêtre. Si l'on inverse le raisonnement, le fait que l'on puisse remonter à une langue mère, ancêtre de toutes les langues parlées actuellement, ne prouve pas que cette langue était la seule parlée à son époque.

2.2.2. Continuité et rupture : vers un dépassement

Mais pourquoi diable être « remonté » au 18^{ème} siècle et s'être intéressé à Süßmilch dont il y a fort à parier que même des linguistes contemporains versés en histoire du 18^{ème} siècle n'ont jamais entendu parler ? Tout simplement pour mieux percevoir les effets de *rupture* et de *continuité* qui se manifestent à travers les questionnements que nous élaborons à propos des phénomènes du monde.

En termes de rupture, il y a le « conjoncturel » : on ne peut évidemment pas, au tournant dépassé du 20^{ème} siècle proposer des analyses qui ne tiendraient pas compte des théories, des modèles et des méthodologies disponibles ou « émergentes », ou encore de l'état des connaissances du moment.

Indépendamment des transformations théoriques, *la modélisation* et *la méthodologie* sont des déterminations conjoncturelles. Les choix en ce domaine ont une forte incidence sur la direction des recherches et donc sur la construction des connaissances y afférant qui sont étroitement déterminées (contextualisées) par la disponibilité des pratiques et la mise en œuvre des outils méthodologiques existants. Corrélativement – et très pratiquement – le degré d'exigence que l'on *peut* avoir quant à la nature des questions posées et à la valeur des réponses fournies sauf à être perçu comme « irréaliste », voire dogmatique ou fantaisiste, s'étalonne à un principe de réalité qui demande de ne pas avoir d'exigences qui rendraient *impossible* l'avancement des recherches et de ne pas poser des questions pour lesquelles on n'envisagerait pas au moins quelque premier élément de réponse empiriquement et provisoirement justifiable.

L'état des connaissances, également « contextuel », est tout autant incomparable : le monde du 18^{ème} siècle n'est pas notre monde. Au plan de la linguistique la parenthèse de la linguistique historique et la reconnaissance de la systémativité des langues a changé la donne. Ainsi le caractère aujourd'hui désuet du titre du mémoire de Süßmilch relève à la fois de l'écart entre ce que l'on suppose avoir thésaurisé dans notre connaissance contemporaine à propos des relations entre les langues du monde, et de l'écart par rapport à ce que l'on pense aujourd'hui acceptable de proposer comme « démonstration » dans le cadre d'une comparaison interlinguistique limitée à une centaine de mots.

En termes de continuité, il y a quelques questions (ou « non-questions ») de fond. Peut-être plus dérangement de n'être que rarement abordées – et éventuellement d'être contraintes à n'être pas posables sans quitter l'arène de la recherche considérée² – sont celles qui portent sur *la nature de l'objet* décrit et sur *la place du chercheur* dans son projet de description. J'ai souligné que du 18^{ème} siècle à notre époque la réflexion sur la généalogie des langues s'est rigidifiée : entendons par là que le modèle de sa représentation a tout simplement été poussé dans ses limites sans que la notion de *langue* (qui le concerne de part

¹ Pour des raisons que je ne saurais pas déterminer cela ne conduit pas pour autant l'auteur à mettre en doute l'hypothèse Ruhlen & Cavalli-Sforza ! Dessalles se justifie de cette « rumeur publique » qui affirme que (p. 42) « *Les linguistes s'accordent sur le fait que les langues actuellement parlées dérivent les unes des autres selon une généalogie arborescente, comme les espèces dans l'analogie de Darwin* » et encore « *Le fait que toutes les langues actuellement parlées sur terre, à l'exception peut-être des créoles, dérivent d'une même langue originelle semble apporter du crédit à l'hypothèse d'une création culturelle du langage* » [hypothèse qu'il ne retiendra pas]. L'on a là un exemple manifeste de la « limite d'incompétence » que j'invoquais plus haut, et à laquelle nous nous heurtons tous – y compris moi-même, bien évidemment.

² Ce qui peut conduire à une double contrainte (*double bind*) dans laquelle le chercheur, pour des raisons tout autant de développement de la recherche que prise en compte des réalités professionnelles sociologiques, ne souhaite pas toujours se trouver enfermé. Il y a censure, et autocensure.

en part) n'ait été critiquée dans la perception objectivée et construite *a posteriori* à laquelle elle donne lieu. Ce qui est en référence, c'est toujours une perception idéalisée, décontextualisée, caractérisée par son homogénéité. C'est la *langue* organisme / structure immanente dans sa forme, la *langue* référence identitaire et transcendante à ses actualisations, qui est retenue. Aucune tentative de théorisation n'a été faite de ce qui se passe et de ce qui fonctionne dans les situations réelles de communication au sein desquelles nous construisons et transformons nos codes linguistiques, ou plutôt celles qui ont été faites (tout particulièrement dans le domaine sociolinguistique) ont bien ouvert quelques voies mais sans parvenir à une reconsidération suffisante de cette question de « la-langue ». Par exemple, aucune tentative d'élaboration théorique n'a été tentée pour intégrer les dynamismes linguistiques en rapport avec l'évidence du plurilinguisme et du pluridialectalisme dans la communication. Finalement, la recherche contemporaine dans ce domaine marqué à la fois par la rigidité des modèles évolutionnistes et structuralistes au sens large semble avoir omis d'approfondir la compréhension des phénomènes qu'elle s'est donnée à étudier en remplaçant le procès d'analyse et de construction théorique par un procès mimologique de modélisation, au risque de substituer à son objet un ersatz dont la description plus aisée en masque les dimensions problématiques.

Quelques commentaires visant à prendre un peu de distance par rapport à ce donné d'évidence que semblent être les « phénomènes-langues » ne peuvent donc pas être inutiles et je vais tenter par un chemin buissonnier, de revenir sur la nature de cet objet linguistique qu'il nous est donné de considérer, et plus particulièrement sur sa dynamique et sa fonctionnalité.

2.3. Pour conclure

2.3.1. Le dit de Stephen Gould

D'entrée de jeu j'ai souligné que le traitement des questions concernant l'*émergence du langage* et celui de l'*évolution des langues* se situent sur des échelles incommensurables mais je crois intéressant de lier ce constat avec celui, très voisin, que Gould (1997, p. 269) a fait, hors du domaine du langage, en distinguant les caractères de l'évolution des espèces et ceux de l'évolution culturelle. Comme on le verra, cela peut conduire à des remarques intéressantes à propos de la nature de l'objet linguistique.

Pour Gould :

L'évolution darwinienne est une histoire de prolifération continue et irréversible... Les espèces ne s'unissent pas entre-elles... ; l'évolution naturelle est un processus permanent de séparation et de différenciation [...]. En revanche, le changement culturel est puissamment stimulé par la fusion et l'anastomose de différentes traditions. Un voyageur intelligent peut découvrir la roue dans une contrée étrangère, rapporter l'invention chez lui et changer radicalement et définitivement la culture de sa propre société [...]. Cette mise en commun des traditions, moteur extrêmement puissant du changement culturel, est un mécanisme totalement inconnu dans l'univers plus lent de l'évolution darwinienne.

Ainsi donc :

[L]e changement culturel [...] est un processus potentiellement lamarckien. Tout savoir acquis par une génération est susceptible d'être transmis à la génération suivante grâce à ce que nous désignons du très noble mot d'éducation. Si j'invente la première roue, ma trouvaille n'est pas condamnée à l'oubli par une intransmissibilité héréditaire (comme l'est tout perfectionnement purement corporel). Il me suffit d'apprendre à mes enfants, à mes apprentis, à mon groupe social, comment faire cette roue.

Et encore :

[L]e changement culturel, en revanche, est potentiellement progressif ou auto complexifiant, car l'hérédité lamarckienne provoque l'accumulation, par transmission directe, des innovations favorables, et les emprunts entre traditions permettent à toute culture de choisir et d'adopter les inventions les plus utiles de plusieurs sociétés distinctes.

2.3.2. La nature des « langues »

C'est un même modèle évolutionniste¹ qui est retenu qu'il s'agisse de la diversification des espèces ou de celle des langues, et c'est toujours lui qui le plus souvent – et indépendamment

¹ Cf. Schleicher (A.). [1863] 1980. « Die Darwinsche Theorie und die Sprachwissenschaft ». in : Tort (P.). Évolutionnisme et linguistique. Paris : Vrin [La Théorie de Darwin et la science du langage, trad. M. de Pommayrol].

de la méthode choisie pour établir la division¹ – est en arrière-plan de la recherche sur l'évolution des langues. Or sa légitimité ne va pas de soi et la précédente référence gouldienne prend ici sa valeur en fournissant un intéressant parallélisme : s'il semble évident que l'échelle de préhension de l'évolution du langage est compatible avec celle de l'évolution des espèces, il est tout aussi évident que l'échelle de l'évolution des langues, elle, est du même ordre que celle qui est pertinente pour appréhender l'évolution de la culture².

Le parallélisme ne s'arrête pas là car le fait que l'on trouve des cultures différenciées, tout comme celui que l'on trouve des langues différenciées, n'implique pas que cette différenciation résulte du seul fait d'un procès de division continue – loin de là. Et ce que l'on est tenté d'admettre pour les *cultures* est tout aussi valable pour les *langues*. Comme les cultures, les langues et/ou des éléments de la langue peuvent s'emprunter, se diffuser, se partager, se modifier au contact. Les sociétés peuvent être plurilingues tout comme elles sont pluriculturelles ; elles peuvent être conduites à changer de langue comme de culture, à modifier profondément leurs façons d'être et leurs valeurs comme à transformer drastiquement leurs façons de parler et leurs règles linguistiques. Les langues actualisées dans le présent (dans tous les présents) peuvent résulter d'évolutions indépendantes et souscrire correctement aux régularités phonétiques attendues, mais elles peuvent aussi manifester les « mixtes » les plus divers. Et *les acquis peuvent être transmis*. Ajoutons encore que, pas plus qu'il n'existe de groupes humains sans « culture(s) » qui les lient, il n'en existe sans « langue(s) » qu'ils partagent³.

Dans cette perspective prendre la partie pour le tout, c'est-à-dire se focaliser sur *la-langue* et vouloir rendre compte du tout de l'évolution des langues par une théorie de l'évolution régulière par division (sans théorisation des déterminations contextuelles du phénomène linguistique à travers son usage, ses fonctionnalités, sa dynamique et sa multicodecité) est sans doute une erreur théorique⁴. C'est aussi la plus répandue.

2.3.3. Le contact au cœur des langues

Ces considérations sur la perméabilité des codes des langues et sur leur sensibilité potentielle à l'emprunt, à la diffusion, à la transformation sous l'effet du contexte et sur la nécessité de les théoriser ouvrent une voie à l'étude empirique du contact des langues qui cesse dans cette perspective d'être marginale pour constituer le cœur même d'une réflexion sur la dynamique des langues. C'est ainsi que l'approche concrète de situations de mélanges linguistiques, de formations créoles tout autant que l'ensemble des travaux sur les plurilinguismes et sur les faits de convergence finissent par montrer la nécessité de s'intéresser autrement que de façon marginale aux considérations sur le contact des langues et de retenir *la-langue* hypostasiée non plus comme un cadre intangible de tout usage communicationnel mais comme un simple objet symbolique, constructible et restructurable, normalement négociable en contexte communicationnel. Et de ce fait très concrètement présent dans le détail de ses manifestations⁵.

Mais l'on peut aller plus loin dans la préhension du contact et développer l'hypothèse que, à travers une production linguistique aussi monolingue et homogène qu'elle puisse paraître se développe aussi une *stratification continue* résultant de l'usage même des codes, ne serait-ce que par le simple processus autonymique du « pointage » de concrétisations linguistico-langagières particulières antérieurement manifestées dans le discours. Stratification qui génère un *contexte de « contact »* au cœur des langues elles-mêmes. L'on est alors dans une nouvelle perspective (prospective ?) : non plus celle d'une reconstruction des langues mais celle d'une reconstruction de l'objet de l'étude.

¹ Cf. Nicolai, 2000, pp. 61 et sv.

² Ce qui est sans lien direct avec l'hypothèse avancée concernant l'origine du langage, qui attribuerait son émergence au développement culturel.

³ Notons toutefois que cette homologie entre les deux domaines est « globale » : il n'existe pas pour autant un quelconque rapport de dépendance entre des traits/comportements culturels et des traits/comportements linguistiques.

⁴ Le « modelage » qui s'ensuit et les conséquences de la tendance hypostasier la construction de « la langue » du milieu linguistico-langagier dans lequel elle prend sa forme n'ont pas été mesurées.

⁵ Toutes les études sur les dynamiques normatives supposent ce type de visibilité.

3. Le point de vue du chercheur

Le point de vue du chercheur n'a aucune transcendance : ce n'est pas le point de vue de Dieu aurait-on pu lancer au 18^{ème} siècle si en ce temps-là les chercheurs avaient existé. Le point de vue du chercheur n'a aucune exterritorialité : ce n'est pas un point de vue depuis Sirius aurait-on pu aussi tenter à la même époque. Le point de vue du chercheur est intrinsèquement et irrémédiablement interdépendant de la construction de son objet et du contexte de cette construction, peut-on essayer aujourd'hui. À quoi cela nous conduit-il ? Tout d'abord à prendre quelque distance envers ce que nous assertons et ici l'effectivité d'un processus permanent d'homéostasie dans la construction, la perception et la diffusion de propositions (ici à vocation scientifiques) est probablement importante à prendre en compte.

Et qu'assertons-nous ici ? D'une part des propositions d'existence et des caractéristiques formelles que nous attribuons à des objets partiellement saisissables dans le monde sensible et partiellement construits dans l'intersubjectivité ; et d'autre part des propositions d'interprétation à propos de ces phénomènes. Concrètement ces objets sont les « langues », les « mots » : ils ont une réalité empirique mais celle-ci est dépendante d'un accord intersubjectif les concernant. Ainsi on peut être d'accord sur la proposition « il y a des langues », mais au-delà de cet accord ce qui est reconnu « langue » ici peut ne pas être tenu pour tel ailleurs ; ce que je reconnais comme « caractéristique intrinsèque » de la langue peut ne pas être tenu pour tel dans un autre cadre théorique ou dans un autre contexte : on peut avoir une intuition de ce qu'est un mot mais cela n'implique pas que l'on en fasse la même analyse.

Une partie du fondement de ces propositions d'existence est vérifiable objectivement dans le monde sensible mais une autre partie n'est vérifiable que par l'élaboration d'un accord, d'une homéostasie dans un processus de construction discursive plus ou moins bien indexé sur le monde extérieur.

Alors, dans ce contexte, que signifient les « ressemblances » ? Que signifient les questions de réduction (et donc d'aveuglement au sens propre) dont l'évidente nécessité nous conduit à structurer nos langues ? Y a-t-il des formes « prototypiques » des langues et des mots ? Ou bien cela résulte-t-il simplement d'un accord conduisant à une perception stéréotypée que l'on stabilise de fait dans l'intersubjectivité, puis à partir des propriétés desquelles on construit un cadre théorique ? Ou bien les deux composent-ils ?

Lier la nécessité de la description des objets que l'on se donne (les langues, les mots) et l'effectivité (efficience ?) de cette description non seulement à la référence de leur matérialité concrète et de leur existence conceptualisée et théorisée mais aussi à *la nécessité de reconnaissance du type de processus mis en œuvre dans cette pratique qui les fait exister* me semble être un pas important pour faire avancer cette pratique elle-même.

Cela ne résout aucun problème, mais cela aide à comprendre pourquoi, n'étant ni Dieu, ni habitant de Sirius, la prise en compte non pas de la relativité des réalités empiriques mais celle des constructions intersubjectives que nous faisons à leur rencontre est un point aveugle de notre approche. Et les points aveugles de cette nature semblent perdurer à travers les siècles. Ils sont co-extensifs à la construction des mythes. Celui de la recherche sur les origines des langues en est probablement un.

Références bibliographiques

- Boë (L.-J.), Bessière (P.) & Vallée (N.). 2003. « When Ruhlen's « mother-tongue » theory meets the null hypothesis ». Communication au *Congrès International des Sciences Phonétiques*.
- Bickerton (D.). 1981. *Roots of Language*. Ann Arbor : Karoma.
- Bickerton (D.). 1990. *Languages and Species*. Chicago : Chicago University Press.
- Bickerton (D.). 2001. « Linguists play catchup with evolution ». in : *Journal of Linguistics*, 37, pp. 581-591 [review article].
- Calvin (W.-H.) & Bickerton (D.) 2000. *Lingua ex Machina : Reconciling Darwin and Chomsky with the Human Brain*. Cambridge Massachusetts : M.I.T. Press.
- De Brosses (Ch.). 1760. *Du culte des dieux fétiches ou parallèle de l'ancienne Religion de l'Égypte avec la Religion actuelle de la Nigritie*. Paris.
- De Brosses (Ch.). 1765. *Traité de la Formation mécanique des Langues et des Principes de l'Étymologie*. Paris.
- Cavalli-Sforza (L.). 1996. *Gènes, peuples et langues*, Paris : Odile Jacob.
- Court de Gebelin (A.). 1776. *Histoire naturelle de la parole ou Précis de l'Origine du Langage & de la Grammaire Universelle. Extrait du Monde Primitif*. Paris.
- Denina (C.). 1794/95. [1799]. « Sur l'origine commune des langues allemande, esclavonne, ou polonaise & latine ; et sur l'origine de la langue italienne ». in : *Mémoires de l'Académie Royale des Sciences et Belles-Lettres*, Classe de belles-lettres, S. 68-112. - Berlin : Decker.
- Dessalles (J.-L.). 2000. *Aux origines du langage, une histoire naturelle de la parole*. Paris : Hermès.
- Erdem (E.), Lifschitz (V.), Nakkhleh (L.) & Ringe (D.). 2003. « Reconstructing the Evolutionary History of Indo-European Languages using Answer Set Programming ». in : *Lecture Notes in Computer Science*, Springer-Verlag, Heidelberg, pp. 160-176.
- Fonagy (I.). 1991. *La vive voix*. Paris : Payot.
- Foster (P.) & Toth (A.). 2003. « Toward a phylogenetic chronology of ancient Gaulish, Celtic, and Indo-European ». in : *PNAS*, July 22, vol. 100, N° 15, pp. 9079-9084.
- Gould (St. J.). [1996] 1997. *L'éventail du vivant. Le mythe du progrès*. Paris : Éditions du Seuil.
- Gray (R.) & Atkinson (Q.-D.). 2003. « Language-tree divergence times support the Anatolian theory of Indo-European origin ». in : *Nature*, 426, pp. 435-438.
- Guiraud (P.). 1967. *Les structures étymologiques du lexique français*. Paris : Larousse.
- Holden (C.-J.). 2002. « Bantu language trees reflect the spread of farming across sub-Saharan Africa : a maximum-parsimony analysis ». in : *Proceedings, Biological sciences/The Royal Society*, 269, 1493, pp. 793-799.
- Nicolai (R.). 2000. *La traversée de l'empirique, essai d'épistémologie sur la construction des représentations de l'évolution des langues*. Paris : Ophrys.
- Nicolai (R.). 2001. « La « construction de l'unitaire » et le « sentiment de l'unité » dans la saisie du contact des langues ». in : *Traverses*, pp. 359-85.
- Nicolai (R.). 2003. *La force des choses ou l'épreuve nilo-saharienne. Questions sur les reconstructions archéologiques et l'évolution des langues*. Köln : Köppe.
- Nicolai (R.). 2005. « Language processes, theory and description of language change, and building on the past : lessons from Songhay ». in : Frajzyngier (Z.), Hodges (A.) & Rood (D.-S.), (eds.). *Linguistic Diversity and Language Theories*. Amsterdam, Philadelphia : J. Benjamins, pp. 81-104.
- Nicolai (R.). [Sous presse]. « Du discours aux effets du contact des langues : réflexion sur la fonctions des contraintes anthropologiques dans la dynamique de l'élaboration des Sprachbünde ». in : *Travaux du Cercle linguistique de Prague*, 5 [Nouvelle série], Amsterdam, Philadelphia : J. Benjamins.
- Peterfalvi (J.-M.). 1970. *Recherches expérimentales sur le symbolisme phonétique*. Paris : CNRS.
- Rexová (K.), Frynta (D.) & Zrávý (J.). 2003. « Cladistic analysis of languages : Indo-European classification based on lexicostatistical data ». in : *Cladistic*, 19, pp. 120-127.
- Ruhlen (M.). [1994] 1997. *L'origine des langues*. Paris : Belin.

- Schleicher (A.). [1863] 1980. « Die Darwinsche Theorie und die Sprachwissenschaft ». in : Tort (P.). *Évolutionnisme et linguistique*. Paris : Vrin [*La Théorie de Darwin et la science du langage*, trad. M. de Pommayrol].
- Süßmilch (J.-P.). [1741] 1998. *L'ordre divin dans les changements de l'espèce humaine démontré par la naissance, la mort et la propagation de celle-ci*. Paris : Ined.
- Süßmilch (J.-P.). 1746. « Réflexions Sur la convenance de la Langue Celtique, et en particulier de la Teutonique avec celles de l'Orient, par lesquelles on démontre que la Langue Teutonique est matériellement contenue dans les Langues Orientales, et qu'elle en descend ». in : *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin*. Berlin : Haude et Spener, pp. 188-203 [1745].
- Süßmilch (J.-P.). 1766. « Essai dans lequel on se propose de déterminer le nombre des habitants de Londres et de Paris ». in : *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin*. Berlin : Haude et Spener, pp. 453-463 [1759].
- Tort (P.). 1980. *Evolutionnisme et linguistique*. Paris : Vrin.
- Warnow (T.), Evans (St.), Ringe (D.) & Nakhleh (L.). 2004. « A Stochastic Model of Language Evolution that incorporates Homoplasy and Borrowing ». Paper presented at the *MacDonald Institute*.

Annexe

Relevé des données présentées par Sűbmilch concernant la lettre R¹, in : Réflexions Sur la convenance de la Langue Celtique, et en particulier de la Teutonique avec celles de l'Orient, par lesquelles on démontre que la Langue Teutonique est matériellement contenue dans les Langues Orientales, et qu'elle en descend - *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin*, 1745. S. 188-203. - Berlin : Haude.

Voces Orientales				Voces Celticae & inprimis Teutonicae			
1	<i>gourabon</i>	a	corvus	corbeau	<i>Rabe</i> ; Belg. <i>Rave</i>	corvus, abjecto gain maxime gutturali ac difficili.	corbeau ; ...
2	<i>renck</i>	p	fraus, imposture, techna, ludificatio	imposture, ruse, fourberie, mystification	<i>Raencke</i> (h.e) ; Belg. <i>Ranken</i>	technae ac fraudes.	ruses, fourberies avec fraudes
3	<i>rafaa</i>	a	abstulit, sustulit	emporter, entraîner ; présenter	<i>Raffen</i> , wegraffen, <i>rapen</i> . rapfen. ropfen ; Belg. <i>raapen</i> ; lat. <i>Rapio</i> .	Anglo Sax. raefen, rapina, reafian, rapere, spoliare	entraîner, emporter, enlever de force, dépouiller
4	<i>ragahha</i>	a	propendit, praeponderavit libra, superavit, praestitit, praevaluit	peser, être pesant, prépondérant, peser plus d'une livre, dépasser, prévaloir	<i>Ragen</i> , hervorragen		
	<i>ragih</i>	a	praeponderans, superans	au dessus de, prédominant			
5	<i>ramam</i>	h	extulit se attoli, elevari	s'élever	<i>Ramen</i>	suspendere Gothice	
	<i>ramah</i>	h	excelsum	élevé, haut	<i>hramjan</i> (h.e)	in altum agere	pousser vers le haut
	<i>ram</i> , pl. <i>ramim</i>		excelsi, elati, collis, ara.	élevés, haussés, colline, autel	<i>Ram</i>	crux. ab altitudine & ara & collis	croix en hauteur, autel, colline
					<i>Raam</i>	fuligo, quod alta petat	
6	<i>ramah</i>	h	jecit, jaculatus est	jeter, lancer	<i>Ramen</i>	terminum, scopum petere, figere	atteindre son but, borne, limite
	<i>rama</i>	a	sagitta petivit scopum	flèche qui atteint son but	<i>beramen</i>		
	<i>ramin</i>	a	jaculator, sagittarius	lanceur, accusateur ; archer	<i>Ram</i>	terminus, scopus	limite, cible
					<i>Ramin</i>	nomen gentis nobilis, vertas Schütze	nom d'une famille noble
7	<i>ramah</i>	s	fundamentum jacere, fundare, injicere	jeter les fondements, bâtir, jeter dans, inspirer, susciter	<i>Rammen</i>	arbores in terram adigere ad fundamenta jacienda	enfoncer des arbres dans la terre pour jeter/bâtir des fondations
	<i>ramma</i>	a	aedificavit, reparavit domum	construire, réparer la maison			
8	<i>ranaa</i>	a	flaccus, gracilis fuit	flasque, être maigre	<i>Ran</i> Belg. <i>rank</i>	gracilis, ein rarer Leib. dun, smal, tanger	maigre
9	<i>rawzel</i>	c	pera viatoria & coriacea, cui viaticum imponitur	sac de voyage en cuir que l'on prend pour la route	<i>Ranze</i> , <i>raenzel</i>	sarcina viatoris, levi immunatione	bagage d'un voyageur, de poids léger
	<i>razmaron</i> , pl. <i>razamon</i>	a	Sarcina, vestes in unum colligatae & convolutae	bagage, paquetage, vêtements rassemblés et roulés en un balot			
10	<i>ranisaton</i>	a	extremitas rei, cartilaginis nasi, manicae ora	extrémité d'une chose, du cartilage du nez, attache d'une manche	<i>Ransi</i>	margo, extremitas rei. id. q. Rand, Bort.	marge, frontière, bord, extrémité d'une chose
11	<i>raaseh</i>	h	movit, contremuit, concussus est cum impetu, strepitu	bouger, trembler, être secoué avec force, avec bruit	<i>Rauschen</i> Belg. <i>ruisschen</i> , <i>ruissen</i> . Angl. <i>rush</i> , <i>crush</i> . Anglo Sax. <i>raescettan</i>	cum strepitu moveri. Gr. ποικίλον.	être secoué avec des mouvements bruyants
	<i>raascha</i>	a	tremuit, tremulum reddidit	trembler, produire un tremblement			
12	<i>raaschan</i>	a	celeriter incedens	avançant rapidement	<i>Rasch</i> , <i>risch</i> . Belg. <i>ras</i> , <i>rasselyk</i> , <i>rasheit</i>		
13	<i>raauschon</i>	a	capite tremens	tremblant de la tête	<i>Rausch</i> Belg. <i>roes</i> , <i>roezig</i>	crapula.	ivresse
	<i>raauson</i>	a	capite tremens, celeriter jactans pedes, capite tremens prae somnolentia, incessu vacillans	tremblant de la tête, agitant rapidement les pieds, dodelinant à cause de la somnolence, vacillant dans sa marche (titubant)			
14	<i>rasa</i>	s	delirium, vertigo	délire, tournoiement	<i>Rasen</i> . Belg. <i>raazen</i> , <i>raasbollen</i>		

¹ Dans le relevé original l'auteur a utilisé pour ses gloses la langue véhiculaire de son époque. Les temps ayant changé, j'ai pensé qu'il était utile de proposer une traduction française des gloses latines. Mais n'ayant pas une connaissance suffisante de ce véhiculaire ancien c'est à l'obligeance de Sylvie Mellet que je dois cette traduction. Je la remercie ici de sa gentillesse et du temps qu'elle a bien voulu y consacrer.

15	<i>rasaa</i>	a	stabilis, firmus, immotus fuit, navis stetit ad ancoram	stable, ferme, être immobile, navire à l'ancre	<i>rasten</i> . Belg. <i>rust</i> . Angl. <i>rest</i>		
16	<i>razaz</i> , <i>resas</i>	e, c, s	quassavit, concussit, collisus est	secouer (comme un prunier), choquer l'un contre l'autre frapper contre	<i>rassen</i> , <i>rasseln</i> . Gr ; <i>ράσσω</i> . Belg. <i>rateln</i> , <i>ratelwacht</i>	collidor.	être frappé contre quelque chose, heurter
17	<i>rati</i>	a	doctus, eruditus vir ac doctor, pecul. rerum divinarum	savant, homme érudit, spécialement dans les affaires divines	Rath (h.e) Belg. <i>raad</i>	legum peritus, legum doctor.	connaisseur des lois, docteur des lois
	<i>rad</i>	p	honoratus, magnus	honoré, estimé, grand			
	<i>ratton</i>	a	Princeps loci & Praefectus	chef local, préfet			
18	<i>rata</i>		consilium iniit (malum) s. viam monstravit, direxit, erudit, monuit. Admonitor	prendre une résolution, montrer la voie, diriger, instruire informer, avertir	<i>Raten</i> , Belg. <i>raaden</i>		
19	<i>ratsaa</i>	h	perforavit	trouer, perfore	<i>Ratze</i> Belg. Rot, rat, rot-muis	glis. quod fruges absumat & comminuat.	loir qui détruit, anéantit les moissons
	<i>ratsaa</i>	a	comminuit	mettre en pièces, broyer			
	<i>rezinta</i>	h	curculio, vermis fruges absumens	charançon, vers qui mange les fruits			
20	<i>rabaa</i>	a	abstulit, sustulit	emporter, entraîner ; présenter	<i>Rauben</i> , <i>rauber</i> . Belg. <i>rooven</i> , <i>roven</i> . <i>Roof</i> . <i>Roover</i>		
	<i>ruba</i>	p	raptor, furans	voleur, qui dérobe			
21	<i>rafaa</i>	a	vestem resarcivit, hinc	raccomoder un vêtement	<i>Rauba</i>	Alam. vestis. Gall. & Ital. robe, roba	vêtement germanique, gaulois et italique
	<i>marfaon</i>		vestis preciosa	vêtement précieux			
	<i>raffon</i> , pl. <i>rofoufon</i>	a	vestis pannusque mollis	vêtement en tissu et souple			
22	<i>radaa</i>	a	infecit, inquinavit, depravato colore fuit	imprégner, souiller, ; être d'une couleur sale, décoloré	<i>Raude</i> , <i>reude</i> Belg. <i>ruidigheid</i> , <i>ruidig</i>	scabies, quae corpus inquinat, commaculat.	gale qui gâte (corrompt) et souille le corps
	<i>radaon</i>	a	macula, inquinamentum, vestigium sanguinis in corpore	tache, immondice, ordure, traces de sang dans le corps			
	<i>roudaon</i>		cum totum corpus doloribus cruciatur	être torturé de douleurs par tout le corps			
	<i>muradon</i> (subst.)	a	(a verbo radda) libidinosus, cupidus coeundi	(du mot radda) : libidineux, obsédé d'accouplement	Braut <i>raude</i> . in Saxonia superiori.	vid. Wacht	
	<i>rua</i>	h	malus fuit, displicuit. it. Deformitas, turpitude	être mauvais, déplaisant. it. difformité, laideur, laideur morale	<i>Rau</i> , rave Sitten, Weter, Lufft (h.e) Angl. rough	malus displicens.	mal déplaisant
23	<i>rachhau</i>	a	terra dura, grassior, altiorque	terre dure, trop lourde et trop profonde	<i>Rauch</i> , <i>rauh</i> (h.e) Belg. <i>rouw</i> , <i>rauw</i> , <i>rou</i> , <i>ruww</i> , <i>ruig</i> . Subst. rouheit, rouwheit, ruigheid	asper	rugueux, âpre, raboteux
	<i>rahwon</i>	a	terra lapidosa, eliator	terre pierreuse...	Angl. rough. Anglo Sax. reoh asper, scaber		
	<i>rahu-gon</i> , <i>rachaton</i>	a	mollis, tener.	souple, flexible ; rendre, délicat	<i>Rauch</i> , <i>rauh</i> (h.e)	tactu molle	doux au toucher
	<i>rachwon</i> , <i>ruchwon</i>	a	mollis	souple, flexible			
	24	<i>raa</i>	a	abstinuit à re	s'éloigner d'une chose	<i>Reuwe</i> nunc <i>Reue</i> . Anglo Sax. <i>reowe</i> . Belg.	
<i>arawa</i>			a turpibus respuit	se repentir de ses fautes	<i>rou</i> , <i>rouw</i> , <i>berow</i> , <i>rouwen</i>		
<i>raawon</i>		a	cum quis respiscit, à turpibus ad praeclara se convertit	lorsque quelqu'un se repent, se détourne des turpitudes pour aller vers les choses lumineuses (belles, supérieures, excellentes, etc.)			
25	<i>rama</i>	a	excessit, discessit elongatus suit	sortir, s'éloigner, partir	<i>Reumen</i> , <i>Raum</i> . Belg. <i>ruim</i> , <i>ruimte</i> . Anglo Sax. <i>rumian</i>	cedere, locum dare. it. ryman	laisser la place, céder le terrain
	<i>rajmon</i>		discessus <i>Distantia</i>	éloigné par la distance			
26	<i>ranan</i>	h, c	murmuravit	murmurer	<i>Raunen</i> (h.e)	murmurare	murmurer
	<i>rinnum</i>		murmuratio, querela	murmure, plainte ; querelle	Anglo Sax. <i>Runian</i> Angl. to rowne one in the ear. Belg. <i>roenen</i> , <i>ruynen</i> , <i>runen</i>	mussitare, susurrare.	se taire, garder le silence ; murmurer, bourdonner
27	<i>ranan</i>	a	sonuit	sonner, résonner	<i>rana</i>		
	<i>rananon</i>	a	res aestate semper in aqua sonans	chose (animal ?) qui fait sans cesse entendre son cri dans l'eau pendant la chaleur de l'été			
28	<i>raphat</i>	s	corrosus est veribus.	rongé par les vers	<i>Raupe</i> . Belg. <i>rups</i> , <i>rupse</i>		
	<i>iof rowaon</i> , <i>afrowon</i>	a	eruca, vermis qui in oleribus nascitur	chenille, ver qui naît dans les légumes			

29	<i>raffon</i> , pl. <i>rofuw-</i> <i>fon</i>	a	arcus. res arcuata (2) vestis, pannus mollis. (3) crates viminea	voûte, arc. Objet courbé. (2) vêtement, tissus souple. (3), claie de bois ou d'osier	Reif (h.e) (2) <i>ref</i> . (3) <i>rif</i> , hodie <i>Ribbe</i> . Rippe Belg. <i>Rib</i> , <i>ribbe</i> . <i>Reif</i> , funis, ligamen- tum. Belg. <i>reep-</i> <i>hout</i> , i.e. <i>Reif holz</i>	quod incurvum & arcuatam (2) olim venter. (3), semircu- lus osseus pectus arcua- tum efficiens	qui est incurvé et arqué (2) autrefois : ventre, sein (3) poitrail osseux demi circulaire funis = corde, ligamentum = lien
	<i>ruffon</i> , <i>ruffaton</i>	a	funis ad paleas ligandas	corde pour attacher la paille			
	<i>rifaf</i>	p	funis quo pes ligatur	corde pour lier les pieds			
	<i>rifak</i>	p	funis quo pes ligatur	corde pour lier les pieds	<i>Reif</i>	funis, Angl. a roope. Anglo Sax. <i>rape</i>	corde, câble
30	<i>rafafon</i>	a	molles, tenues, deliciosae vestes	vêtements souples, fins, élégants	<i>Reif</i> , <i>rhaff</i> , <i>ref</i>	olim <i>vestis</i> . vide ante Raube. robe num. 21	autrefois vêtement
	<i>rafyfon</i>	a	vestes sericae, nitore splendentes, pallium quod annulo serreo firmatur & supra dorsum rejicitur	vêtements de soie, resplendissant d'un éclat soyeux, manteau qui est attaché par un anneau de fer et qui est rejeté sur l'épaule	Anglo Sax <i>raefels</i> , vestes, it. <i>reaf</i> , hinc Latino barb. <i>raupa</i> , <i>raubaroba</i> . Angl. <i>robe</i>	propre <i>Vestis</i> <i>pretiosior</i>	à proprement parler, vêtement assez précieux (coûteux)
31	<i>rafyfon</i>	a	(2) <i>resplendens rore</i> <i>arbor</i>	arbre resplendissant de rosée	<i>Reif</i> , ros. Belg. <i>ryp</i> . <i>rief</i>		
32	<i>raefon</i> . it. <i>Raefra-</i> <i>fon</i>	a	deorsum pendentes rami. <i>Arbor ramis</i> <i>promissis</i>	branches pendant vers le bas arbre aux branches pleureuses	<i>Rebe</i> . Belg. <i>Ranik</i> van een Wynstock		
33	<i>refaa</i> , <i>rifaa</i>	a	demessarum frugumin aream comportatio	moyen de transport des blés coupés vers l'aire de battage	Reiffen Belg. <i>ryp</i> , <i>rypen</i> . Angl. & Anglo Sax. <i>ripe</i>	<i>maturescere</i> . <i>maturus</i> .	mûrir, mûr, faucher, moissonner moisson- neur champ de céréales, récoltes
	<i>eyarmür</i> - <i>rifaa</i>	a	tempus ejus rei seu triturationis	moment de cette chose (sans doute du transport ???) ou du battage	Anglo Sax. <i>ripien</i> , <i>ripan</i> , <i>falcare</i> , <i>metere</i> . <i>Rippeare</i> , <i>messor</i> . <i>Ryp</i> , <i>Seges</i> , <i>standing</i> <i>corne</i>	<i>maturescere</i> .	faucher, moissonner moissonneur champ de céréales, récoltes
34	<i>rakam</i>	a	arithmetica, notatie	arithmétique, action de marquer un signe, de noter	<i>Recken</i> . <i>Rechnen</i> . Belg. <i>rekenen</i> . <i>rekenaar</i> , <i>rekening</i>		
	<i>ehli</i> <i>rakam</i>		arithmeticus	mathématicien			
	<i>rakn</i>	a	notare, scribere, pingere	marquer, écrire, représenter			
35	<i>Harac-</i> <i>kon</i> feu <i>harra-</i> <i>chon</i>	a	radix gulae ad partem superiorem	partie supérieure de la racine du gosier	<i>Rachen</i> . Francis olim II <i>racho</i> Anglo Sax. <i>raca</i> . Belg. <i>raak</i>		
36	<i>raechg</i>	p	iracundus	irascible, irritable, emporté	<i>Rache</i> . <i>rachgierig</i> . Belg. <i>raak-gierig</i>		
	<i>rekinon</i>	a	gravis, modestus, constans vir	pondéré ; modéré, mesuré, calme ; homme mûr	<i>Rekkin</i> (h.e) Goth. <i>Reeks</i> Lat. <i>rex</i> . Anglo Sax. <i>Rica</i> . <i>Rick</i> , <i>rik</i>	ol. heroës, bellatores. Landrecker duces & Proceres provinciae Ita quoque <i>veteres</i>	autrefois : héros, combattants ; chefs et personnages éminents d'une province ; roi. On désigne aussi de ce terme les anciens
	<i>Rüknon</i>	a	praesidium, fucimen- tum, robur it. <i>validus</i> . <i>Proceres regni</i>	garde, ??? ; solide et costaud. Les grands d'un royaume	It. Anglo Sax. <i>Rice</i> , <i>regnum</i> , <i>ditio</i> , <i>imperium</i> (nobis <i>Reich</i>) <i>Riesiam</i> , <i>regnare</i> . <i>Recen-</i> <i>dome</i> , <i>regimen</i> , <i>gubernatio</i> . <i>rece-</i> <i>nysse</i> , <i>directio</i>		royaume, empire, domination, autorité, gouvernement
37	<i>rekana</i> (verb.)	a	requievit, nixus est	être reposé, prendre du repos			
	<i>rükkn</i> , pl. <i>erkan</i>	a	praesidium, fulcimen- tum rei & res magni momenti. <i>Proceres</i> <i>regni</i>	garde, soutien, appui, état d'une chose ; choses de grande impor- tance ; grands personnages d'un royaume			
38	<i>arkow-</i> <i>non</i>	a	Princeps (pecul. agricultorum) est à <i>praeced. verbo</i>	chef (notamment des agri- culteurs)		<i>Ἀρχων</i> . ita <i>Castellus</i>	
	<i>raah</i>	h, a	pavit. oves educere, regere, tueri	faire paître les brebis, les élever, les mener, les surveiller, protéger		Hinc (1) <i>Roy</i> , rex. <i>Amercanis</i> <i>Rouë</i> , <i>roë</i> , plur <i>roeed</i> , <i>rouanez</i>	roi
39	<i>roaeh</i>	h	pastor. amicum se praebens	berger se présentant comme un ami		(2) <i>Raugraf</i> <i>germanorum</i> <i>Praefectus</i> <i>Imperatoris</i>	Préfet impérial
	<i>rain</i> , pl. <i>ruaton</i>	a	Pastor. <i>Praefectus</i> <i>cuiuscunque rei</i>	berger. Préposé à quelque chose		(3) <i>Roland</i> inde <i>fortassis</i> <i>explicationem</i> <i>quoque accipit</i>	
	<i>ryaje</i>	a	observare, regere, <i>curare</i>	faire attention, guider/diriger, prendre soin de			
	<i>raey</i>	a	<i>pastus</i> , <i>rectus</i>	nourri, dirigé			
	<i>raeaja</i>	a	<i>subditi</i> , <i>coloni</i>	soumis, colonisés			

40	<i>raka</i>	h	expandit	exposer à l'air, développer, expliquer	<i>Recken</i> (h.e)	expandere, extendere Belg. <i>rekken</i>	peser avec soin, juger, apprécier ; étendre, allonger, élargir
	<i>rikkaim</i>	h,c	extenuationes laminarum	action de rendre minces des feuilles, des lames, des plaques			
41	<i>reka</i>	c	vestem resarcivit	il raccommode un vêtement	<i>Rock</i>	vestis, Gr. Ρακος Belg. <i>rok</i> . Anglo Sax <i>roce</i> , tunica. Hine Latino Barb. roculus, roecus, rochus, rochetum, item Anglicum <i>roctet</i> , rocket	vêtement, tunique
	<i>rekoa</i>	c	maculosum, maculis conspersum	taché, maculé de taches			
	<i>rakua</i>	c	resartum	raccommoeder, réparer			
		s & sam	pannus, vestis, vestimentum	morceau d'étoffe, vêtement			
	<i>rokaton</i>	a	panniculus, assummentum	guenille, lambeau d'étoffe ; pièce (au sens de : morceau de rapiéçage)			
42	<i>rakan</i>	c	abstulit, eripuit, evacuavit, ejecit, expuit	enlever, arracher, évacuer, jeter hors de, cracher/vomir	<i>Reka</i>	Island, ejicere	jeter hors de, chasser
	<i>rekan</i>	c	vacuus, inanis, it. amens	vide, vain ; fou	<i>rek</i>	res naufragae, ejectae	chose naufragée, jetée au loin
	<i>rakijon</i>	a			<i>brach</i>	res abjecta, sputum	chose immonde ; crachat
	<i>raka</i>	s			<i>Wrach</i>	Suecis, ejectamenta maris	fucus???? Rejets de la mer
	<i>rakaa</i>	a	mente laboravit, proscidit dicteris, satyra	travailler intellectuellement ; blesser d'un bon mot, d'un sarcasme	<i>Werck pro Wreck. Racker</i>	stupa, ejectamentum lini	étoupe, détrit en/du lin
43	<i>redah</i>	c	castigare, erudire	réprimander, punir ; enseigner, former	<i>Rede</i>	Belgis Vernunst. Vers-tand	
	<i>red & redd.</i>	p	peritus, intelligens, philosophus	habile, expérimenté, intelligent, sage	<i>redenlos, unvernünftig. redelyk, verstanding, witzig. een redelick Dier</i>	ubi intelligenter ac decenter agitur	quand on agit intelligemment et convenablement
	<i>rede</i>	p	ordo, series, regula, linea	ordre, série, règle, ligne			
	<i>ridaon</i>	a	intelligentia, decus, debitum	intelligence, honneur/décence ; dette	Anglo Sax. redlic	consultus. redeleas, praeceps. & c. Goth. <i>rathjo</i> . Alam. <i>redina</i> Lat. ratio	décision chef raison
44	<i>Raedita</i>	a	(1) respondit, contradixit, refutavit, contenderunt inter se	répondre, contredire, réfuter, discuter/polémiquer	<i>rede</i>	sermo, qui verbis aliorum reponitur, dein generatim qui eunque	
	<i>reed</i>		refutatio	réfutation			
	<i>reddi gewab</i>		respondio	réponse			
45	<i>Raedita</i>		(2) reddidit restituit, retractavit. recuperatus, restitutus est, restitui curavit	rendre, restituer, retirer, repris, rendu	<i>Redden</i> Belg. nobis <i>retten</i> , erretten		
	<i>rididem</i>	a	restitutio redd.	rétablissement, réparation, restitution			
	<i>reta</i>	c	reduxit, reverti, fecit	ramener, revenir, faire			
46	<i>ridaon</i>	a	pallium, vestis, operimentum	manteau, vêtement, couverture	<i>Rideau</i>	Gall	
	<i>Regg. regga</i>	a	movit, agitavit, tremere fecit	faire bouger, agiter, faire trembler	<i>Regen, erregen rege, commotus. Belgae non habent</i>	ciere, commovere,	mettre en mouvement, secouer
	<i>regreg</i>	a	mobilis, tremulus	mobile, tremblant			
	<i>regf</i>	a	commovere & commoveri	remuer et être remué (secoué)			
	<i>regd</i>	a	tremere. terreri	trembler, être effrayé			
47	<i>ergowg aton</i>	Inde	oscillum	balançoire	<i>Ergern, sich ergern</i>	ira in motum conjici & conturbari	
	<i>Ter-guwgaton</i>	a	perturbatio, conturbatio	perturbation, trouble			
	<i>Raegon</i>	a	pluvia aqua, spec. stagnans	eau de pluie, spécialement stagnante	<i>Regen Rigo</i>	pluvia.	pluie. Faire couler
	<i>raegion</i>	a	pluvia frequens	pluie dense	<i>Rege</i> Belg. <i>regen</i>	vox antiqua Celtica de fluviis adhibita.	
48	<i>rigan, plur. rega</i>	a	aquae stagnantes inhibitae	eaux stagnantes, retenues	Angl. rain. Anglo Sax. ren. renboga, iris. renian, pluere, to raing, ren, pluvialis		
	<i>Rege</i>	p	ordo, series, vestigium	ordre, série, trace	<i>Reege, reige, reihe.</i> Latino barb. <i>riga</i> , Gall. <i>raje</i> , raye, Bel. <i>ry, rek, rang</i>	riga = sillon, raie	
49	<i>ick ege diraeht</i>	p	una serie arbores plantare	planter les arbres sur une rangée			

	<i>nisan-dem</i>						
50	<i>Regia, regen</i>	a	ora latusve pecul. putei, regionis, coeli	bord ou côté, en particulier d'un puit, d'une zone, d'une région du ciel	<i>Regio</i>	zone	
51	<i>raehha</i>	a	latiore ungula praedita capra, equa	chèvre munie d'un sabot assez large, jument	<i>Reh</i> Belg. <i>ree, rhee.</i> Angl. a roe, roe buck. Sax. <i>raa</i>	caprea, cerva.	chevreuil, cerf
	<i>rehaon</i>	a	latitudo unguiae	largeur de l'ongle			
52	<i>rumph. pl. rimah et ermah</i>	a	hasta, lancea	lance, javelot	<i>Reem</i> rimyn (h.e) Ruder-Stange. Lat. <i>remus.</i> Belg. <i>riem, roer</i>		rame, aviron
	<i>remmah</i>	a	hastarum confector	fabricant de lances			
53	<i>rejaon</i>	a	mundum illum reddidit	rendre propre	<i>Rein</i> Belg. <i>rein</i>		
54	<i>rehaek</i>	a	attingere, prope esse, accedere	atteindre, être proche, s'approcher	<i>Rehken, reichen,</i> erreichen. Belg. <i>raaken, reiken.</i> Angl. to reach. Anglo Sax. <i>raecan, araecan</i>		
55	<i>rejon, rion</i>	a	fluvii decursus ex alto loco	descente d'un fleuve depuis un lieu élevé	Rhein : Belg. Rhyn.	nomen fluvii	nom d'un fleuve
56	<i>rezz, rüzz pro erüzz</i>	a	Oryza	riz	Reiss. Iral. rifo. Gall. ris.	Oryza	riz
57	<i>rahsa</i>	p	multum peregrinatus	ayant beaucoup voyagé	<i>Reesen</i> Belg. <i>reizen</i>	proficisci.	partir en voyage
	<i>resan</i>	p	perducens, pervenire faciens	qui conduit à destination, qui atteint son but			
	<i>resa, reisa</i>	c	stadium, curriculum	stade, course			
58	<i>Ruz</i>	h	currere	courir	<i>Rosi</i> Belg. <i>ros</i>	equus. a cursu dictus.	cheval de course
	<i>Roz</i>		currens, cursor	courant, coureur			
59	<i>raasa</i>	a	Caput fuit, praefuit, rexit	être à la tête de, diriger	<i>Reisen</i>	surgere, erigi, caput esserre.	surgir, se dresser, sortir la tête, dresser la tête au-dessus des autres
	<i>rejs, rejjis pro reis</i>	a	caput, princeps, praeses, primus	tête, chef, celui qui préside, le premier	Anglis imprimis mansit, to rise, rising. <i>Riese</i> , caput super alios efferens. <i>Riese</i> olim Praefectus militum, der oberste <i>Ries.</i> Belg. <i>reus, reuzen-groose</i>		
60	<i>Reisi dilaweran</i>		militum dux	chef des soldats			
61	<i>Resas</i>	c	rumpi, dirumpi, confringi in minutas partes	être rompu, cassé, brisé en mille menus morceaux	<i>Reissen</i>	Gr. $\rho\eta\sigma\epsilon\iota\nu$	
	<i>resisim</i>	h	ruptarum majores	fractures importantes	<i>Riss</i>	ruptura. Gr. $\rho\eta\sigma\iota\varsigma$	rupture, fracture
	<i>raezz</i>	a	frangere, conterere, comminuere	casser, briser/broyer, détruire	<i>Ritze</i>	ruptura	rupture, fracture
	<i>ress, plur. Risas, res</i>	a	vestis dissoluta, trita	chose, vêtement déchiré(e), broyé(e)	Anglice to raze, rumpere, diruere Belg. <i>ryten.</i> t cum s saepe commutatur		briser détruire
	<i>rissaset</i>	a	teri ac dissolvi	être broyé et dissout			
	<i>Rize</i>	p	minutum contritus	déchiré en menus morceaux			
	<i>ris</i>	p	vulnus, ulcus	blessure, plaie ; ulcère			
62	<i>ress</i>	a	conspergere humore, parum pluviae & sparsim emittere	asperger d'un liquide, arroser d'un peu de pluie légère	<i>Riesen, rieseln,</i> <i>es rieselt,</i> si pluvia sparsim ac leniter cadit ac decidit aestate	guttatim cadere,	tomber goutte à goutte es rieselt = quand la pluie tombe lentement et en gouttes espacée en été
	<i>risas</i>	a	pauca & sparsim irorans aqua	arroser/humidifier de quelques gouttes d'eau dispersées	<i>Ros,roris.</i> Lat.		rosée
	<i>riz</i>	p	fundens, spargens	répandant, aspergeant	<i>Rosee</i> Gall.	pluvia tenuis Restat quoque in Arroser	pluie fine
63	<i>razy</i>	a	placens, gratus, acquiescens, gratum habens, ad assensum pertrahere, consentientem reddere	qui plaît, bienvenu ; qui approuve, qui est reconnaissant	<i>Reitzen</i>	ad adsensum, beneplacitum, acquiescentiam pertrahere	conduire à l'approbation, à l'acquiescement, faire que quelque chose plaise
	<i>ryza</i>	a	beneplacitum	qui plaît bien			
	<i>razah</i>	h	beneplacitum	qui plaît bien			
	<i>razia</i>	h	oblectario	amuser, distraire			
	<i>Arazah, in Hiph. hirza</i>	h	acceptum reddere seu facere ut quis quid velit ratumque habeat	faire accepter ou faire en sorte que quelqu'un veuille quelque chose et le ratifie	<i>Hertzen</i> (h.e). <i>Herz, cor, animus,</i> beneplacitum	facere ut quid cui placeat	faire en sorte que quelque chose plaise à quelqu'un
64	<i>Radah</i>	h	desumere, proprie divellere id quod cohaeret	prendre pour soi, en particulier séparer ce qui est attaché ensemble	Raden, roden		
	<i>redia</i>	h	extractio, avulsio	extraction, arrachement			
	<i>redah</i>	c	castigare, erudire, arare	châtier, dégrossir/ éduquer, labourer			

65	<i>Raedai</i>	a	(1) calcavit terram unguis suis equus modo medio inter cursum & vehementiorem gradum. Sublato uno pede in altero subsultim progressus est	un cheval frappe la terre de ses sabots d'une façon intermédiaire entre la course (= le galop ???) et le pas énergique (= le trot ???)	<i>Reiten riden.</i> Angl. <i>ride</i> , riding, rider. Belg. <i>ryden</i> , <i>ryen</i> .	rustici	
			(2) Percussit, jacto lapide petivit, propugnavit lapidum jactu. rem profligavit, & c.	frapper, atteindre d'un jet de pierre, combattre avec des pierres, lapider	<i>Reiten.</i> Anglo Sax. <i>Rad, rade</i> (1) equitatio, iter equestre (2) incursum, irruptio, inuasio. <i>Rad-here</i> , exercitus equestris	olim etiam bellum gerere, in bellum prosicisci	autrefois aussi : faire la guerre, partir en guerre équitatio, randonnée équestre invasion, envahissement, cavalerie
66	<i>Raetaa</i>	a	coivit cum muliere, congressui idonea, matura fuit	s'accoupler avec une femme idoine à la rencontre, qui est pubère	<i>Rad</i>	Island. res uxoria	biens de l'épouse
	<i>rataon</i>	a	congressus cum muliere	se rencontrer avec une femme	<i>radspell</i>	divortium	divorce
67	<i>Riach, heriach</i>	h. c. s.	odoratus est	être parfumé, odorant	<i>Riechen.</i> Belg. <i>rieken, ruiken</i>		
	<i>reach</i>		odor, odoratus, fragrantia	odeur, odorant, odeur suave	<i>Ruch, rauch, roich.</i> Belg. <i>reuk.</i> (h.e)	odor, quem respirat. Anglo Sax. rec, fumus, recan, vaporare, evaporare, fumare	odeur qu'on respire, exhaler
	<i>rahha</i>	a	odor, odoratus, fragrantia it. odorem de se emitit	odeur, odorant, émettre une odeur suave			
68	<i>Ruaeh</i>	h. a	ventus, it. Spiritus, Anima, Spiritus Dei, Mens, voluntas cogitatio, Prudentia	vent, it. Esprit âme, Esprit de Dieu, intelligence, volonté, action de penser, Prévoyance, prévision	<i>Ruchen</i>	curare restat in ruchloss h.e. mentis ac prudentiae expers. Anglo Sax. recc, cura, recceleas, negligens, impruidus, piger	
	<i>ruwhhon</i>	a	ventus, it. Spiritus, Anima, Spiritus Dei, Mens, voluntas cogitatio, Prudentia	vent, it. Esprit âme, Esprit de Dieu, intelligence, volonté, action de penser, Prévoyance, prévision			
69	<i>Rit, rita</i>	c	virgultum tenue it. juncies species	rameau fragile, qui a l'aspect d'un jonc	<i>Rute.</i> Belg. <i>roed</i>	virgultum	rameau
70	<i>Rud</i>	c	id	idem	<i>Ried.</i> Angl. <i>reed.</i> Belg. <i>riet</i>	juncus, calamus palustris.	jonc, roseau des marais
	<i>ratat</i>	c. h.	tremuit, contremuit	trembler ; se mettre à trembler	Anglo Sax. red, read		
	<i>Ratio-nem</i>		denominationis junci & virgulti à tremore desumtam esse patet. hinc etiam reliqua nomina ex eadem ratione	il est clair que les dénominations du rameau et du jonc ont été tirées du mot "tremblement". De ce même mot aussi tous les autres noms qui ont la même motivation.			
71	<i>Raas</i>	a	trepidare, tremere	s'agiter, se démener ; trembler	<i>Raus</i>	olim arundo	autrefois roseau
	<i>reays</i>	a	trepidans	agité	Latino Barb. <i>Rau-sea.</i> Gall. <i>Roseau.</i>		
	<i>riason</i>	a	tremulae, hastae	tremblantes/oscillantes ; hampes de javelot	<i>Reis, ries</i>	surculus, virgultum tenue	bouture/marcotte ; baguette, fin rameau
72	<i>Raaraa</i>	a	calamus longus	long roseau, longue canne	<i>Ror Isl. reyr</i>		
73	<i>raa-raat</i>	a	agitata fuit aqua, movit se res	l'eau s'agite, une chose bouge	<i>Rüren Angl. to rear.</i> Belg. <i>roeren</i>		
	<i>raraa</i>	a	commovit oculos, motitavit aures, caudam agitavit	rouler des yeux, remuer les oreilles, agiter la queue			
74	<i>Rarariton</i>	a	aqua in campo tenuiter expansa		restat in <i>Rorreis</i>		
75	<i>Rir, rira</i>	c	humor crassus, mucilago	liquide épais ; moisissure/humeur, glaire	<i>Rur</i>	dysenteria a mucilagine intestinorum corrupta & egesta	dysenterie due à une humeur corrompue qui s'épanche des intestins
	<i>ragron, riron</i>	a	medulla tenuis corrupta & nigra, aqua manans ex ore infantis	moelle fine corrompue et noire, eau coulant de la bouche d'un enfant			
76	<i>Iraddon</i>	a	roseo, ruso colore fuit	être d'une couler rose, rouge	<i>Roth, Angl. red. Redden.</i> Belg. <i>reod, root.</i> Anglo Sax. <i>read</i>	insicere colore ruso aut rubro.	colorer en rouge ou en rose
	<i>redon</i>	a	Vestigium croci seu sanguinis. Sanguini impressit seu per eum duxit	trace de safran ou de sang			

77	<i>rathsa</i> seu <i>raththa</i>	a	tritus, dissolutus fuit funis, vestis, marcore, squalore deformis fuit	la corde a été usée par le frottement, rompue ; le vêtement a été abîmé par la pourriture et la saleté	<i>Rotten</i> , verrotten. Angl. <i>to rot</i> , rotten, rottenneff. Belg. <i>Roesten</i> , nobis <i>Rosten</i>		
	<i>rethson</i>	a	res trita, dissoluta	chose broyée, dissoute			
78	<i>ratha</i>	a	miscere, perturbare, pravum inire consilium seu modum	mélanger, remuer, bouleverser, mettre en désordre, prendre une mauvaise décision ou façon de faire	<i>Rotten</i> , zusammen rotten, Rotten und Secten		
79	<i>rataa</i> (a verbo)	a	formantur subst.		<i>Rotte</i> . Belg. <i>rot</i>	turba hominum.	troupe d'hommes
	<i>rotuow</i> , <i>rataon</i> , <i>rottaon</i> , & <i>denique artaon</i>	ar	multitudo & frequentia hominum	foule, rassemblement d'individus			
80	<i>Ratat</i>	h, c	tremuit, tremere fecit	trembler, faire trembler	<i>Rütten</i>	tremere, trepidare, quatere	trembler, être secoué
	<i>Rata</i>	a	tumultum, clamorem excitavit, edidit	susciter, provoquer un désordre, faire naître un grand cri	<i>Rutten</i>	tremor, febris, vertigo	tremblement, fièvre, vertige
	<i>ritaton</i> , <i>retaton</i>		tumultus clamor	cri, tumulte			
81	<i>Radza</i>	a	lentus, crassus evasit	sortir épais, visqueux	<i>Rotz</i>	mucus crassus & lentus narium. it. morbus purulentus equorum	morve épaisse, et visqueuse du nez. it. maladie purulente du cheval
		s	sordes, morbus purulentus & pruginosus	crasse/chassie/purulence ; maladie purulente			
82	<i>Radzanna</i>	a	firmus, stabulis fuit	solide, être stable	Belg. <i>Rots</i> , <i>Steenrots</i>	rupes a stabilitate	rochers
	<i>Radzinnon</i>	a	firmus, stabulis fuit	solide, être stable	<i>Rotz</i>	sepulcra veterum Celtarum à lapidibus ordine circumpolitus	tombeau des anciens Celtes faits avec des pierres disposées en cercle
	<i>ardu-non</i>		lapides recto ordine constituri, alii aliis impositi. it. Sepulcra veterum	pierres destinées à être déposées en ligne droite, posés les unes sur les autres			
83	<i>Rekisim</i>	h	colles, excelsa, salebrosa	collines, hauteurs, rocailles	hinc rupes Graecis <i>ῥώξ</i>		
	<i>riksa</i>	c	lapillus, scrupus	petite pierre, pierre pointue			
	<i>rakasa</i>	a	sitmus, stabilis fuit, invenit mineram	???			
	<i>rikason</i>	a	fossilia terrae	tas de terre			
84	<i>Racha</i>	a	constituit, firmavit rem.	établir, consolider une chose ou une affaire	hinc aliud nomen Rupis Celticum Roc, roche, rocher. Angl. rock	de là vient un autre nom celtique du rocher : Roc	
	<i>raki</i> seu <i>rechi</i>	a	durans, stabile, perenne	durable, stable, pérenne			
	<i>rockon</i>	a	fundamentum, terra dura grossior altiorque	fondation, support. Terre dure, assez épaisse et profonde			
85	<i>rawbon</i> , <i>raubon</i>	a	Collis altior	assez haute colline	inde Latinorum <i>Rupes</i>		
86	<i>raekraka</i> , <i>raekrakon</i>	a	quod huc illuc movetur	chose qui est déplacée, transportée ça et là	<i>Rucken</i> . Belg. rukken	movere & moveri.	bouger et être bougé
	<i>rokrakaton</i>	a	commotiones, agitationes	émotions, agitations			
87	<i>Rykkæet</i> , pl. <i>rukuk</i>	a	terrae tractus ad fluvii latus	escarpement de terre le long d'un fleuve (talus, digue ???)	<i>Ruck</i> , <i>rucken</i>	dicitur	se dit
	<i>Rukk</i> & <i>rykk</i>	a	terra expansa	terre étendue		(1) de cacumine rei, montis, tetrae, animalis, quod ascenditur	(1) du sommet d'une chose, d'une montagne... sur lequel on grimpe
	<i>rakuwon</i> , <i>raekuww</i>	a	accumulatus, tumulus terrae prominens	tas, tumulus de terre proéminent (qui dépasse en hauteur)	Ita Hunsruck	(2) de terra expansa & collibus interspersa.	(2) de la terre répandue et étalée entre des collines
	<i>Rakija</i>	a	ascendit, elevavit, extulit superior evasit	monter, s'élever, dresser, s'en aller par le haut		(3)) Gr. <i>ῥαχία</i> , crepido littoris petrosa, littoris salebrae. <i>ῥάχις</i> spina dorsi vid. num. 37	(3) de la base d'un rivage en pierre, d'une falaise
	<i>rekd</i> & <i>rukud</i>	a	in culmine quasi consistere	s'installer pour ainsi dire au sommet		Belg. rug. vid. supra num. 37	
88	<i>Rikkakon</i>	a	servus, mancipium	servitude, alienation	<i>Ruckh</i> Ol. ein Eighenthum, eigen Land		
	<i>rykk</i>	a	servitus, mancipii status	servitude, aliénation			
	<i>rakijkon</i>	a	res acquisita, quam certo constat ab alio	chose acquise, dont il est établi clairement qu'un autre la possède			

			possideri				
89	<i>Rafaa</i>	a	convaluit, coaluit vulnus	guérir, cicatriser	<i>Rufe.</i> Belg. roef, roof, rove, roove	crusta ulceris.	croûte d'une blessure
90	<i>Rafaa</i>	a	alta voce fuit	... à haute voix	<i>Ruffen, Ruf</i>		
	<i>rufaah-ton</i>	a	elatio vocis	action d'élever la voix	Belg. <i>roepen.</i> Anglo Sax. rof, rofe	clarus, insignis, illustris, praestans	célèbre, illustre, remarquable, éminent, qui excelle
	<i>raefon</i>	a	cum quis inclarescit, cum quid evulgatur	lorsque quelqu'un se distingue, lorsque quelque chose est divulgué (révélé à tous)			
	<i>Rufaet</i>		elatio vocis	voix forte			
91	<i>Raha</i>	a	quievit	se reposer	<i>ruhen</i>		
	<i>rahwon</i>		quies	repos, calme			
	<i>rehin</i>	a	tranquilla vita	existence tranquille			
92	<i>Ramam</i>	h	attolli, elevari	être porté, soulevé	<i>Rümen.</i> <i>Sich rümen,</i> esserre se. Belg. <i>roemen</i>	esserre laudibus.	
	<i>rum</i>	h	Extulit se, altus, exaltatus fuit	s'enorgueillir			
93	<i>Rewn & run</i>	a	extrema consultatio	dernière question ou délibération	<i>Runa</i>	consilium	délibération
94	<i>ruste</i>	p	pagus. it.	bourg, village	Lat. & Gall. <i>Rusticus</i>		relatif à la campagne
	<i>rustak & rustai. rustar</i>	p	paganus, pagi incola	paysan, villageois ; habitant d'un village	<i>rustique, rustaud, rustre</i>		rustique, rustaud, rustre
95	<i>Resed & rused</i>	a	bene dirigi rectoque ductu uti. it. recti studium, directio	être bien dirigé et avoir une conduite droite ; étude du droit chemin	<i>Rüsten</i>	in ordinem redigere, instruere, parare	revenir en ordre ; insérer, assembler, apprêter, arranger
	<i>rüsti</i>	a	directio in bonam viam	mettre dans la bonne voie			
	<i>restet, rüst</i>	p	ordo, series	ordre ; enchaînement, rangée			
96	<i>rus-chchata & ruse-hechon</i>	c	nates	fesses, croupion	<i>Rutschen</i>	per nates moveri	mouvement des fesses



Mai 2006

1. Introduction

Au XIX^e siècle, la linguistique est le lieu d'un étrange rejet concernant la question de l'origine du langage. Étrange, parce qu'il s'agit d'une interdiction d'ordre *institutionnel* : dans l'article 2 des Statuts de la Société de linguistique de Paris (S.L.P.), fondée en 1866, on lit : « La Société n'admet aucune communication concernant, soit l'origine du langage soit la création d'une langue universelle » (S.L.P., I, 1871). Étrange, enfin, parce que la Société reviendra très vite sur ce double interdit et que la plupart de ses membres consacreront en fait de nombreuses pages à l'un et l'autre de ces sujets « tabous »...

Mais le rejet de ces deux questions est en effet en rapport étroit avec l'installation du paradigme historique en linguistique. Celui-ci substitue, on le sait, une conception historico-sociale des langues grevée de contingence à la tradition naturaliste bien illustrée par August Schleicher (1821-1868) qui enferme l'historicité des langues dans une temporalité organiciste, et évolutionniste. Dans ce cadre remis en cause de toute part à la fin du XIX^e siècle (Bréal, Paris, Meyer, Whitney) l'historicité n'avait donc rien d'accidentel. Il s'agissait, au contraire, de nécessité implacable : le changement dans les langues se ferait selon un ordre naturel déterministe implacable.

Évidemment, ces deux perspectives antagonistes sont incompatibles en ce qui concerne la question de l'origine du langage. Comme l'écrit S. Auroux,

[o]u il y a des lois d'évolution linguistique, et certains faits permettent d'inférer quelque chose sur l'origine des langues ; ou les phénomènes linguistiques sont des entités historiquement contingentes, et l'absence de faits datant de leur origine est indépassable (Auroux, 1989, p. 142).

Victor Henry (1850-1907) fournit des arguments précis concernant le rejet dans son ouvrage *Antinomies linguistiques*. Selon l'auteur,

[q]ue le linguiste doive s'interdire toute recherche sur l'origine du langage c'est un point qui semble définitivement acquis, tout au moins parmi les linguistes, si paradoxale qu'en soit la première apparence : l'origine du langage n'est pas, a priori, un problème linguistique, puisque la linguistique ne se propose pour objets que des langues toutes formées, dans leur état actuel, historique ou préhistorique, et qu'il ne lui est donné que de constater l'évolution, jamais la naissance d'un langage (Henry, 1896, pp. 25-26).

Suivant la tradition « historique », le linguiste ne doit s'occuper que de la langue formée, toujours déjà en circulation, ce qui explique aussi bien l'interdiction des études concernant l'origine du langage, que ceux qui concernent les langues universelles.

Ces discussions aboutissent finalement à Ferdinand de Saussure (1857-1913). Pour le linguiste genevois en effet « [o]n peut discuter éternellement sur cette naissance, mais son plus grand caractère c'est d'être parfaitement le même que celui de la croissance » (Saussure, 2002, p. 94).

La dimension sociologique de l'interdit formulé par la Société linguistique de Paris, précédant donc les justifications théoriques, est, pour S. Auroux, évidente. En effet, selon l'auteur, « [s]ans l'article II, la théorie saussurienne [...] n'aurait sans doute pas été ce qu'elle est » (Auroux, 1989, p. 146). J. Trabant, pour sa part, observe que

[c]'est dans l'espace du culturel et de l'historique que la linguistique a trouvé son autonomie. La linguistique est discipline autonome en tant que discipline historique. Puisque les langues, ces choses culturelles de l'humanité, sont son objet, elle ne s'occupera pas des choses trans-historiques, ni du début ni de la fin de l'histoire des langues (Trabant, 2004, p. 41).

Ceci posé, nous voudrions toutefois soulever une question : Saussure, en rejetant la question de l'origine du langage dans son projet de constitution d'une linguistique autonome, n'a-t-il pas assumé une position théorique originale ? S'il est ainsi quelle position a-t-il véritablement assumé ? Nous souhaiterions, dans cet article, revenir sur la mise en place par Saussure de l'alternative origine/fonctionnement à la fois dans le *Cours* rédigé et les Sources manuscrites disponibles, dans la mesure où il s'agit à notre avis d'un « topos » essentiel (quoique souvent implicite) de la conception par Saussure d'un rapport très particulier entre langue et temps, souvent oublié, voire mal compris, par des commentateurs post-saussuriens, surtout dans les années 1960/70, aspect pourtant fondamental de la linguistique saussurienne (nous pensons ici aux linguistes attachés surtout à la synchronie et non aux historiens de la linguistique et théoriciens – comme H. Meschonnic – plus sensibles à l'opposition entre les deux paradigmes que nous venons d'évoquer).

C'est lors d'une discussion sur la nature de l'objet de la linguistique que Saussure fait tout d'abord référence à la question de l'origine du langage, dans un dialogue explicite avec le linguiste américain William Dwight Whitney (1827-1894). Étant celui qui a ouvert une discussion d'ordre épistémologique, Whitney joue un rôle important dans la mise en cause de la tradition naturaliste en linguistique. Son point de vue est toutefois ambigu : on peut dire qu'il oscille entre un retour au XVIII^e siècle et l'ouverture d'une perspective nouvelle. Nous voudrions, ainsi, commencer en précisant ce point de vue et les enjeux qui y sont impliqués.

2. Le langage comme institution humaine chez Whitney

Pour Whitney, le caractère essentiel du langage c'est qu'il est une institution. Il faut ajouter : une institution parmi d'autres. Selon le linguiste américain,

if we are to give language a name which shall bring out its essential character most distinctly and sharply, and even in defiance of those who would make of it an organism, we shall call it an INSTITUTION, one of the institutions that make up human culture (Whitney, 1871, p. 51).

Le principal problème qui hante les linguistes du XIX^e siècle est sans doute celui concernant le changement linguistique. Whitney, en s'opposant à la tradition de Schleicher, explique le changement par l'action ou la volonté humaine, autrement dit par un tout autre biais que le naturalisme. Selon l'auteur, le concept de langue comme organisme est un dogme construit par Schleicher, qui

does not attempt any proof of [it], or even let us see clearly the grounds on which it rests in his own mind. [...] If voluntary action of men has anything to do with making and changing language, then language is so far not a natural organism, but a human product. And if that action is the only force that makes and changes language, then language is not a natural organism at all, nor its study a natural science (*op. cit.*, p. 37).

Ainsi, la volonté humaine est placée au premier plan et la notion d'histoire qui se dessine est à comprendre dans le sens le plus large, c'est-à-dire, l'activité de l'esprit dans sa liberté. Pour Schleicher il s'agit du domaine de la philologie, et non pas la science du langage ou glottique. Pour Whitney au contraire, le langage est « une institution d'invention humaine » (Whitney, 1875, p. 28, nous soulignons).

En critiquant l'idée schleicherienne selon laquelle

speech is the sole exclusive characteristic of man, and that any given anthropoid ape who should be able to speak would be called by us a man, [and a brother] however unlike a human being he might be in other respects (Whitney, 1871, p. 57).

Whitney argumente :

[a]s to this last assertion, it is so easy to speculate where the test of fact can never be applied, that I will not take the trouble to contradict it, although my own conviction is strongly against it, and I cannot but doubt whether Schleicher himself would have proved equal to fraternizing with his fellow-man if the case had been realized. But certainly, speech is so far from being man's sole distinctive quality that it is not a quality at all, in our author's sense ; it is only a possession. The capacity of forming and acquiring speech is a quality, and one among the many which constitute the higher endowment of man ; but let the child of the most gifted family of the most highly cultivated race grow up untaught, in solitude, and he will no more employ a language than he will build a temple, paint a picture, or construct a locomotive [...] ; because speech, like the other elements of civilization, does not go down by inheritance, but by the process of teaching and learning (*op. cit.*, p. 58).

En effet, en ce qui concerne le rapport du langage à l'homme, pour Whitney, « man was man before the development of speech began ; he did not become man through and by means of it » (*op. cit.*, p. 58). L'auteur croit en une continuité humaine et donc « the rise of language had nothing to do with growth of man out of an apish stock, but only with his rise out of savagery and barbarism » (*op. cit.*, p. 58).

Si l'arbitraire du signe et son caractère conventionnel constituent les piliers de la perspective de Whitney, il s'agit là, apparemment, de soutenir d'un côté l'arbitraire comme un principe rationnel, et, de l'autre, l'idée d'un contrat originel, ce qui renvoie bien le débat à la question de l'origine du langage. L'idée de convention chez Whitney peut ainsi renvoyer à une définition du langage comme nomenclature et présupposer une primauté à la pensée. Il s'agit alors d'une conception téléologique du langage : étant une institution d'invention humaine, il implique une raison originelle régnant continûment.

2.1. Whitney entre raison et déraison

Une telle lecture de Whitney n'est toutefois pas unanime parmi les historiens. Pour J.-E. Joseph (2004), si Whitney autorise une interprétation du signe comme principe rationnel, cela n'est pas vraiment ce que le linguiste américain pensait. Si « his wording did not always make this clear » c'est probablement parce que « the view he was contesting was the powerful racist one, not the almost forgotten Enlightenment notion of languages as deliberated logical creations » (Joseph, 2004, p. 69). En effet, pour Whitney, « la double possibilité des changements internes et externes indépendamment les uns des autres, repose sur la nature du lien entre le sens et la forme, *nature toute accidentelle et arbitraire* » (Whitney, 1875, p. 65, nous soulignons). Dans son premier ouvrage sur la linguistique générale, datant de 1867, l'impossibilité d'une action tout à fait libre de l'individu semble claire :

[w]e did not make our own tongue, or any part of it ; we neither select the objects, acts, mental states, relations, which should be separately designated, nor devised their distinctive designations. We simply received and appropriated, as well as we could, whatever our instructors were pleased to set before us. Independence of the general usages of speech was neither encouraged nor tolerated in us ; nor did we feel tempted toward independence. Our object was to communicate with those among whom our lot was cast, to understand them and be understood by them, to learn what their greater wisdom and experience could impart to us. In order to this, we had to think and talk as they did, and we content to do so. Why such and such a combination of sounds was applied to designate such and an idea was to us a matter of utter indifference ; all we knew or cared to know was that others so applied it (Whitney, 1867, p. 14).

Ce fragment est particulièrement important sous deux aspects : d'abord, il subordonne l'action individuelle à l'usage commun. Deuxièmement, il indique l'insignifiance d'une investigation sur l'origine : « why such and such a combination of sounds was applied to designate such and such idea was to us a matter of utter indifference ».

Whitney réaffirme ces points de vue dans son ouvrage de 1875, intitulé *Life and growth of language*, et par là même toutes les contradictions que nous venons de présenter. Comme l'observe C. Normand, dans la courte préface (non datée) qu'elle a écrite pour l'édition de la version française de cette œuvre, (*La vie du langage*, traduite par Whitney lui-même), il s'agit là d'une « perspective assez confuse où l'intention individuelle se distingue mal de l'influence de la société » (Normand, 1988).

Suivant le point de vue de C. Puech et A. Radzynski (1978), Whitney, sans pourtant le noter lui-même, ouvre une voie originale en remettant en cause, avec l'idée d'une historicité sans finalisme, le caractère rationnel présupposé dans l'hypothèse de la langue comme institution. Le rôle de l'action humaine a été relativisé par l'auteur, en opérant un bouleversement lors de la mise au premier plan de l'accident historique, lequel croyons-nous, constitue l'indice d'un ordre propre du langage. Il s'agit là pour C. Puech et A. Radzynski, d'un « glissement de niveau » (Puech & Radzynski, 1978) : Whitney passe d'une réflexion juridique sur les fondements du langage (l'origine) à une réflexion sur les conditions et les modalités de son existence :

Whitney substitue pratiquement, à une problématique de la légitimité et du fondement, une réflexion d'un autre ordre, qui ressortit à l'histoire. C'est une histoire conçue comme développement ouvert et sans finalisme. Elle est donc en rupture avec les conceptions organicistes qui prévalaient encore à l'époque de Whitney et amenaient à ne pouvoir envisager la vie des langues que comme la pleine actualisation d'un patron préconçu, se révélant à lui-même dans le déploiement substantiel et naturel de son être. C'est l'historicité ouverte du langage qui, chez

Whitney, le fait être avant tout une institution sociale, ou encore, comme il précise, « un corps d'usage », ce qui implique qu'on peut l'envisager avant tout comme un instrument non naturel, une réalité qui n'est ancrée dans aucune nature humaine. Il serait donc vain, nous avertit l'auteur, d'aller chercher un modèle théorique, et encore moins une explication opérante des lois régissant la vie du langage du côté de la psychologie. Car il n'y a pas de rapport immédiat entre la pensée et le langage (*op. cit.*, p. 53).

Selon les auteurs, ce n'est pas la définition donnée par Whitney d'institution en soi-même, mais plutôt l'enjeu théorique qu'elle implique, à savoir, « la détermination d'une spécificité différentielle du langage et plus particulièrement à la psychologie » (1978, p. 54) qui permet de comprendre sa prise en compte par Saussure, dans l'édification de son projet de « linguistique autonome ».

3. Saussure et le principe de l'arbitraire : l'accident originel

Dans son « Coup d'œil sur l'histoire de la linguistique », texte qui ouvre le *Cours*, Saussure rend hommage à Whitney (et aux néogrammairiens) pour avoir placé la linguistique sur son axe véritable : « [g]râce à eux, on ne vit plus dans la langue un organisme qui se développe par lui-même, mais un produit de l'esprit collectif des groupes linguistiques. Du même coup on comprit combien étaient erronées les idées de la philologie et de la grammaire comparée » (Saussure, 1985, p. 19). Le linguiste genevois continue en note :

[I]a nouvelle école, serrant de plus près la réalité, fit la guerre à la terminologie des comparatistes, et notamment aux métaphores illogiques dont elle se servait. Dès lors on n'ose plus dire : « la langue fait ceci ou cela », ni parler de la « vie du langage », etc., puisque la langue n'est pas une entité, et n'existe que dans les sujets parlants (1985, p. 19, note 1).

Dans ses « Notes pour un article sur Whitney », Saussure écrit :

Quelques illuminés on dit : le langage est une chose tout à fait extra-humaine, et en soi organisée, comme serait une végétation parasite répandue à la surface de notre espèce.

D'autres : le langage est une chose humaine, mais à la façon d'une fonction naturelle.

Whitney a dit : le langage est une institution humaine. Cela a changé l'axe de la linguistique (Saussure, 2002, p. 211).

Néanmoins, si, pour Saussure, la langue ne relève pas de la nature, elle n'est toutefois pas une institution parmi les autres : « [p]our bien faire sentir que la langue est une institution pure, Whitney a fort justement insisté sur le caractère arbitraire des signes », mais le linguiste américain « n'est pas allé jusqu'au bout et n'a pas vu que ce caractère arbitraire sépare radicalement la langue de toutes les autres institutions » (Saussure, 1985, p. 110).

La critique de Saussure envisage la mise en cause d'une possible conception téléologique du langage chez Whitney. Certes, le signe est arbitraire, mais il n'est fondé sur aucune raison : « le langage n'est pas contenu dans une règle humaine, constamment corrigée ou dirigée, corrigable ou dirigeable par la raison humaine » (Saussure, 2002, p. 214). L'auteur continue plus bas :

[l]'institution du mariage selon la forme monogame est probablement plus raisonnable que selon la forme polygame. Cela peut philosophiquement se discuter. Mais l'institution d'un signe quelconque, par exemple σ ou s pour désigner le son s , ou bien de *cow* ou de *vacca* pour désigner l'idée de vache est fondée sur l'irraison même (*op. cit.*, 214).

L'idée d'une « correction continue de l'esprit » qui découlerait « depuis l'origine, d'une harmonie visible entre l'idée et le moyen d'expression » (Saussure, 2002, p. 219) est pour Saussure absurde et si l'auteur parle d'un contrat, il ne faut pourtant pas considérer ce mot « comme une législation à la manière des philosophes du XVIII^e siècle, comme dépendant de notre volonté » (Saussure, 1997, p. 12). Selon l'auteur,

[l]e moment où l'on s'accorde sur les signes n'existe pas réellement, n'est qu'idéal ; et existerait-il qu'il n'entre pas en considération à côté de la vie régulière de la langue. La question de l'origine des langues n'a pas l'importance qu'on lui donne. <Cette question n'existe même pas> (question de la source du Rhône – puérile !) Le moment de la genèse n'est lui-même pas saisissable, on ne le voit pas. Le contrat primitif se confond avec ce qui <se> passe tous les jours dans la langue, <avec les conditions permanentes de la langue > (*op. cit.*, p. 12).

Si, d'un côté, Saussure refoule la question de l'origine, car il s'agit là d'une question « puérile », « le moment de la genèse n'[étant] lui-même pas saisissable », d'autre, l'auteur affirme que « [l]e contrat primitif se confond avec ce qui <se> passe tous les jours dans la langue, <avec les conditions permanentes de la langue> ». En effet, pour l'auteur, « le

phénomène primordial du langage est l'association d'une pensée à un signe » (Saussure, 2002, p. 47). Il s'agit là d'un accident, l'accident originel, un accident qui se poursuit d'ailleurs indéfiniment, conférant à la langue son statut « tourbillonnaire », son statut indissociablement historique et sémiologique.

Le débat autour de l'origine joue bien un rôle important chez Saussure : c'est en réponse à la problématique whitneyenne autour de la nature de l'objet de la nouvelle science que le maître genevois pose la seule question concernant la linguistique : celle du rapport signe/pensée. Plutôt que la reconstitution de l'émergence du langage, c'est, ainsi, le rapport signe/pensée qui importe. La question de l'origine chez Saussure conduit bien ainsi à celle sur « l'essence même du langage ».

Ce rapport signe/pensée n'est toutefois pas encore « la loi tout à fait finale du langage » : si « la première expression de la réalité serait de dire que la langue (c'est-à-dire le sujet parlant) n'aperçoit ni l'idée *a*, ni la forme *A*, mais seulement le rapport *a/A* » (*op. cit.*, 39), cette expression est encore « tout à fait grossière ». En effet, « [i]l n'aperçoit vraiment que le rapport entre les deux rapports *a/AHZ* et *abc/A*, ou *b/ARS* et *blr/B*, etc. » (Saussure, 2002, p. 39). Ainsi,

la loi tout à fait finale du langage est qu'il n'y a jamais rien qui puisse résider dans un terme [...], que *a* est impuissant sans le secours de *b*, celui-ci de même sans le secours de *a* ; ou que tous deux ne valent donc que par leur réciproque *différence*, ou qu'aucun ne vaut, même par une partie quelconque de soi (je suppose « la racine », etc.), autrement que par ce même plexus de différences éternellement négatives (Saussure, 2002, pp. 218-219, souligné dans le texte).

Les éléments du système n'ont leur signification ou sens ou, encore, leur valeur¹, qu'en rapport avec les autres éléments du système, de façon à ce qu'ils soient vides d'un sens ou signification propre, ce qui renvoie à une perspective non substantialiste.

Voilà les conditions permanentes du langage. Celles qui, pour Saussure, relèvent d'une même problématique que celle de l'origine. Une telle affirmation est loin d'être évidente et nous pensons qu'elle indique, en effet, un rapport très particulier, chez Saussure, entre la langue et le temps, un rapport paradoxal, et souvent mal compris qui a fait, d'ailleurs, l'objet de nombreuses critiques (*cf.* entre autres Bakhtine, 1929 ; Foucault, 1966 ; Calvet, 1967 ; Pêcheux, 1969, 1975 ; Milner, 1978). Nous voudrions dans les lignes suivantes, revenir ainsi à ce rapport, qui est, croyons-nous, l'un des noyaux durs de la théorie saussurienne de la langue.

4. Le système et le temps chez Saussure

L'une des conséquences de l'arbitraire du signe chez Saussure est que « la langue est radicalement impuissante à se défendre contre les facteurs qui déplacent d'instant en instant le rapport du signifié et du signifiant » (Saussure, 1985, p. 110). Lors de sa première conférence à l'Université de Genève, en 1891, Saussure se demande : jusqu'à quel point les études linguistiques ont-elles leur place dans une Faculté des Lettres et non dans une Faculté des Sciences ? La réponse pour Saussure est très claire et ne laisse aucun doute : la linguistique relève d'une science historique car

tout dans la langue est histoire, c'est-à-dire qu'elle est un objet d'analyse historique, et non d'analyse abstraite, qu'elle se compose de faits, et non de lois, que tout ce qui semble organique dans le langage est en réalité contingent et complètement accidentel (Saussure, 2002, p. 149, souligné dans le texte).

L'immobilité absolue n'existe pas en matière de langage et « au bout d'un certain laps de temps la langue ne sera plus identique à elle-même » (Saussure, 1985, p. 273). Cela doit, selon Saussure, se vérifier même à propos des langues artificielles : « [c]elui qui en crée une la tient en main tant qu'elle n'est pas en circulation ; mais dès l'instant qu'elle remplit sa mission et devient la chose de tout le monde, le contrôle échappe » (*op. cit.*, p. 111). C'est une « loi fatale » et « [p]assé le premier moment la langue entrera très probablement dans sa vie sémiologique » (*op. cit.*, p. 111).

L'étude des légendes germaniques est particulièrement éclairante sur ce caractère constitutif de la langue. Dans ce travail, développé parallèlement aux cours de linguistique à

¹ Saussure, en effet, l'affirme : « nous n'établissons aucune différence sérieuse entre les termes valeurs, sens, signification, fonction, ou emploi d'une forme ; ni même avec l'idée comme contenu d'une forme ; ces termes sont synonymes » (Saussure, 2002, p. 28).

l'Université de Genève, Saussure, en partageant le point de vue des historiens selon lequel il y a des rapports entre la matière des légendes et les événements historiques, veut démontrer que l'origine de la Chanson de Nibelungen, contrairement à ce qui est admis en général, est helvétique et non allemande. Saussure ne tarde toutefois pas à se confronter avec des impasses fondamentales : malgré les concordances considérables entre chronique et légende, une preuve du lien direct ne s'établit pas, non pas à cause de l'insuffisance des données, mais en raison de l'absence de critère d'assurer des concordances non fortuites entre les deux. Selon Saussure,

[l]'exercice qui consiste à rechercher une « identité » entre un personnage de la légende et un personnage de l'histoire ne saurait avoir, d'avance, qu'une portée très limitée. <C'est une chose qui>¹ demanderait en tout cas une méthodologie <spéciale>, rien que pour savoir en quoi consistera l'identité, à quel signe nous devons la connaître et la proclamer. Le nom à lui seul ne signifie rien c'est évident <certain>. Sont-ce donc les actes du personnage, ou son caractère, ou son entourage, ou []², ou quoi encore qui constituent le critère de l'identité ? C'est un peu tout cela et rien de tout cela <parce que tout peut avoir été à la fois transformé et transporté de A à B.> Plus on étudiera la chose, plus on verra que la question n'est même pas de savoir où réside plutôt qu'ailleurs, l'identité, mais s'il y a un sens quelconque à en parler (Saussure, cité par Fehr, 2000, p. 89).

La recherche concernant les légendes germaniques aboutit donc au constat de son impossibilité même. En revanche,

ce constat [met] en pleine lumière le fait que les problèmes avec lesquels se [débat] Saussure à *partir des légendes* [débordent] largement le cadre initialement délimité de son projet et [soulève] des problèmes intéressant la théorie linguistique » (Fehr, 2000, p. 96, souligné dans le texte).

La possibilité d'un tel rapport n'échappe pas à Saussure. Les symboles, qui sont les personnages de la légende, « sont soumis aux mêmes vicissitudes et aux mêmes lois que toutes les autres séries de symboles, par exemple, les symboles qui sont les mots de la langue » (Saussure, cité par Fehr, 2000, p. 96). Ainsi, les « mots de la langue » n'existent que parce qu'ils sont en circulation et le cas d'un idiome qui se trouve dans un état d'immobilité et de repos n'existe pas (sur ce point cf. également Turpin, 2003 ; Normand, 2004).

Les réflexions autour de la mobilité de la langue préoccupent Saussure, cependant, au moins depuis 1891, quand l'auteur, lors de sa deuxième conférence à l'Université de Genève, parlait déjà du principe du « mouvement de la langue dans le temps » (Saussure, 2002, p. 157). C'est là un principe absolu : « Nous posons donc le principe de la transformation incessante des langues comme absolu. Le cas d'un idiome qui se trouverait en état d'immobilité et de repos ne se présente pas » (*op. cit.*, 158).

Ce principe peut susciter toutefois une question : étant donné que « le fleuve de la langue coule indéfiniment » (Saussure, 1985, p. 193), comment se fait-il qu'on puisse se faire comprendre, ce qui d'ailleurs constitue, selon Saussure (2002, pp. 178-179), le but, la fin et l'effet de toutes les langues ? Cette discussion est loin d'être simple et relève d'une problématique autour de la question de l'identité. Saussure fournit un exemple très éclairant :

Un original nommé Boguslawski a dernièrement fait annoncer dans une ville de Russie l'ouverture d'une exposition d'un nouveau genre : c'étaient simplement 480 portraits photographiques représentant tous la même personne, lui, Boguslawski, et identiquement dans la même pose. Depuis vingt ans avec une régularité admirable, le premier et le quinze de chaque mois, cet homme dévoué à la science se rendait chez son photographe, et il se trouvait maintenant en état de faire profiter le public du fruit accumulé de ses labeurs. Je n'ai pas besoin de vous dire que, si dans cette exposition on prenait sur la paroi deux photographies contiguës quelconques, on avait le même Boguslawski, mais que si l'on prenait le n° 480 et le n° 1 on avait deux Boguslawski. De même, si l'on avait pu non pas photographier mais phonographier au jour le jour dès l'origine tout ce qui a été exprimé en parole sur le globe ou sur une partie du globe, on aurait des images de langue toujours ressemblantes d'un jour à l'autre, mais considérablement différentes et parfois incalculablement différentes de 500 ans en 500 ans, ou même de 100 ans en 100 ans (*op. cit.*, p. 157).

¹ Le signe diacritique <> dans les notes cité par Fehr indique « tout ce qui se donne à lire dans les paragraphes immédiats (en inscription supérieure, inférieure, adjacente) des lignes tracées à la main (interpolations, reformulations, rectifications, annotations en marge, etc.) » (Fehr, 2000, p. 286).

² Le signe diacritique [] dans les notes cité par Fehr indique des blancs, des interruptions.

Cette comparaison renvoie, dans une certaine mesure, à la mise en cause de la notion de *changement*. Si la langue est dans un processus de transformation incessante, la conscience des sujets parlants ne connaît que la dimension du présent, là où « la langue [...] est à chaque moment l'affaire de tout le monde » (Saussure, 1985, p. 107), « une chose dont tous les individus se servent toute la journée » (*op. cit.*, p. 107), ce qui rend impossible toute forme de révolution. Nous revenons ainsi, à nouveau, à la mise en cause par Saussure du concept whiteneyen de la langue comme institution dans la mesure où :

les prescriptions d'un code, les rites d'une religion, les signaux maritimes, etc., n'occupent jamais qu'un certain nombre d'individus à la fois et pendant un temps limité ; la langue, au contraire chacun y participe à tout instant et c'est pourquoi elle subit sans cesse l'influence de tous. Ce fait capital suffit à montrer l'impossibilité d'une révolution (1985, p. 107).

Il s'agit là du « principe de la continuité de la langue dans le temps », ou de l'« ininterruption forcée » (Saussure, 2002, p. 151), objet de la première conférence de Saussure à l'Université de Genève. Ce principe consiste dans le fait que « jamais et nulle part on connaît pas historiquement de rupture dans la trame continue du langage » (*op. cit.*, p. 152). Si l'on considère un certain état de langue et un certain état de langue antérieur, nous sommes frappés par la grande distance entre eux et, ainsi, nous tendons à y trouver deux choses, l'une succédant l'autre. Qu'il y ait succession, c'est pour Saussure une évidence. Mais qu'il y ait deux choses dans cette succession c'est « faux, radicalement faux, et dangereusement faux » (*op. cit.*, p. 152). En effet, « chaque individu emploie le lendemain le même idiome qu'il parlait la veille et cela s'est toujours vu » (*ibid.*).

Ces deux principes, du mouvement de la langue dans le temps et de la continuité dans le temps, « loin d'être contradictoires, se trouvent dans une corrélation si étroite et si évidente que, aussitôt que nous sommes tentés de méconnaître l'un, nous faisons injure à l'autre » (Saussure, 2002, p. 157). Par là Saussure répond à la question sur la possibilité de la communication face au principe du mouvement de la langue dans le temps : il s'agit de présupposer un ordre dans la langue où, pour la conscience des sujets parlants, les changements pourtant réels ne sont pas aperçus. Cet ordre qui relève du présent de la langue est pourtant radicalement historique et se fonde sur le fait social, là où la dimension de l'accident semble trouver une raison relative. Il ne s'agit pas là d'une raison *a priori*.

Dans ce point, Saussure se démarque nettement des philosophes du XVIII^e siècle aux prises avec le thème de l'origine. Pour Johann Gottfried Herder (1744-1803), le lauréat de l'Académie de Berlin en 1769¹ et l'un de derniers auteurs de cette période à écrire sur ce sujet par exemple, l'homme, étant privé de l'instinct, jouit d'une liberté absolue par rapport à l'animal, ce qui d'ailleurs explique chez lui l'origine du langage. Toutefois, contrairement à ce que pense l'auteur lui-même en soutenant un rapport constitutif entre l'humain et le langage, cela pose une antériorité de l'humain par rapport au langage. Ainsi, l'homme est libre avant l'apparition du langage, ce qui signifie postuler une raison primitive, gouvernant son fonctionnement. La liberté se fonde, ainsi, sur une faculté humaine, ce qui conduit Herder à un paradoxe : en soutenant la constitution de l'homme par le langage, comment se fait-il qu'il soit déjà homme depuis toujours ? C'est-à-dire, avant même l'apparition du langage ?

Il en va autrement chez Saussure. Le fonctionnement linguistique se fondant primordialement sur le principe de l'arbitraire, la liberté humaine, si elle existe, s'expliquerait par le fait que le sujet parlant, n'existant que dans la langue, est soumis à un fonctionnement tout à fait arbitraire. La raison qui le gouverne est relative, dans la mesure où elle ne trouve sa base que dans le fait social. Les règles de la langue sont historiques et sont toujours en train de se constituer : « la loi synchronique est générale, mais elle n'est pas impérative » (Saussure, 1985, p. 131). Bien qu'elle s'impose aux individus par la contrainte de l'usage collectif, il ne s'agit pas là d'une obligation : « *dans la langue* aucune force ne garantit le maintien de la régularité quand elle règne sur quelque point » (*op. cit.*, p. 131, souligné dans le texte). Saussure parle, ainsi, d'un côté par lequel la langue n'est pas complètement arbitraire, où une raison relative semble régner. Une telle raison relève du fait social, impliqué *dans* la langue.

¹ L'Académie de Berlin dans cette année là met en concours la question suivante : « En supposant les hommes abandonnés à leurs facultés naturelles, sont-ils en état d'inventer le langage ? Et par quel moyen parviendront-ils à cette invention ? On demande une hypothèse qui explique la chose clairement et qui satisfasse à toutes les difficultés » (Pénisson, 1977, p. 16).

Selon Saussure,

à aucun moment, et contrairement à l'apparence, [la langue] n'existe en dehors du fait social, parce qu'elle est un phénomène sémiologique. Sa nature sociale est un de ses caractères internes (*op. cit.*, p. 112).

Si l'arbitraire du signe implique un ordre propre de la langue, il implique dans le même temps une détermination sociale constitutive du système linguistique :

le choix qui appelle telle tranche acoustique pour telle idée est parfaitement arbitraire. Si ce n'était pas le cas, la notion de valeur perdrait quelque chose de son caractère puisqu'elle contiendrait un élément imposé du dehors. Mais en fait les valeurs restent entièrement relatives, et voilà pourquoi le lien de l'idée et du son est radicalement arbitraire.

À son tour, l'arbitraire du signe nous fait mieux comprendre pourquoi *le fait social peut seul créer un système linguistique*. La collectivité est nécessaire pour établir des valeurs dont l'unique raison d'être est dans l'usage et le consentement général ; l'individu à lui seul est incapable d'en fixer aucune (*op. cit.*, p. 157, nous soulignons).

Ainsi, la dimension du présent – autrement dit, la synchronie, où le phénomène primordial du langage ne cesse pas de se constituer dans une trame indéfinie – est, chez Saussure, radicalement historique.

Le principe du mouvement de la langue dans le temps et celui de la continuité sont ainsi étroitement liés. En effet, comme l'affirme Saussure,

en dernière analyse, les deux faits sont solidaires : le signe est dans le cas de s'altérer parce qu'il se continue. Ce qui domine dans toute altération, c'est la persistance de la matière ancienne ; l'infidélité au passé n'est que relative. Voilà pourquoi le principe d'altération se fonde sur le principe de continuité (Saussure, 1985, pp. 108-109).

En effet, dans la première conférence à l'Université de Genève, Saussure observe « qu'on ne peut faire naître nulle part une langue » (Saussure, 2002, p. 155), et donc « [t]outes les langues se parlant à la même époque sont de même âge » (*ibid.*). Réduire l'étude de la langue à l'étude de l'évolution des sons, ce qui fut d'ailleurs le but de la grammaire historique et comparée, signifie oublier la dimension du sujet, pourtant, constitutif du fonctionnement linguistique. Le phonétisme est toujours miné par le phénomène de l'analogie, ce qui amène Saussure à comparer la langue à une robe, « couverte de rapiécages faits avec sa propre étoffe » (Saussure, 1985, p. 235).

Supposer un rapport d'opposition entre le système et l'histoire est pour la perspective saussurienne un point de vue tout à fait impossible. Il a été pourtant souvent attribué par des commentateurs post-saussuriens, pour qui l'ordre synchronique relève d'un domaine anhistorique. Il s'agit là des lectures intéressées, dont la caractéristique la plus marquante c'est l'absence d'une mise en perspective historique. Comme l'observent J.-L. Chiss et C. Puech, ces lectures se laissent en général,

fasciner par une représentation de la production des connaissances trop exclusivement attachée à une logique ponctuelle de la connaissance scientifique où priment la simultanéité, la ponctualité du « moment de connaissance » unique et perpétuellement conservé identique à lui-même parce que définitivement inscrit dans un champ de vérité réputé intemporel (Chiss & Puech, 1999, pp. 54-55).

Ce malentendu étant posé, nous nous demandons donc en quoi consiste exactement cet ordre synchronique. Cela fera l'objet de la partie suivante de notre exposé, dont le titre réfère au texte, récemment publié « De l'essence double du langage » (*cf.* Saussure, 2002).

5. « L'essence double du langage »

Pour Saussure, « [i]l est faux (et impraticable) d'opposer la *forme* et le *sens* » (Saussure, 2002, 17, souligné dans le texte). Le point de vue plus juste serait plutôt « d'opposer la *figure vocale* d'une part, et la *forme-sens* de l'autre » (*op. cit.*, p. 17). Nous croyons que le germe des malentendus dont nous venons de parler repose surtout sur une lecture du texte saussurien qu'y voit justement une séparation entre la forme et le sens, ce qui est « faux » et « impraticable ».

La distinction entre ces deux points de vue de la « figure vocale » et de la « forme-sens » correspond, respectivement, à la séparation entre diachronie et synchronie. Il s'agit là de deux domaines d'études possibles, tous deux légitimes, mais radicalement opposés. Cela, il convient de le remarquer, ne représente en aucune manière une contradiction par rapport à ce que

nous avons vu au-dessus. Voyons ainsi en quoi consiste une telle distinction entre les points de vue de la diachronie et de la synchronie.

Selon Saussure,

- Ou bien SIGNE et *suite de temps* mais alors point d'IDÉE dans le signe. C'est ce qu'on appelle la phonétique.
- Ou bien SIGNE et IDÉE : mais alors inversement point de *suite de temps* ; nécessité de respecter complètement l'instant et uniquement l'instant. C'est le domaine de la morphologie, de la syntaxe, de la synonymie, etc. (Saussure, 2002, p. 54, souligné dans le texte).

Le premier aspect relève de la diachronie ; il enveloppe l'étude de la « forme », ou du « signe » dans une « suite de temps » et ne concerne pas le sujet parlant. Il ne s'agit là en effet que de l'étude des rapports entre termes qui se succèdent en se substituant les uns aux autres dans le temps, « en dehors de toute intention » (Saussure, 1985, p. 122), c'est-à-dire, en dehors du rapport signe/idée. On reste ainsi enfermé dans le domaine de la phonétique et « le caractère diachronique de la phonétique s'accorde fort bien avec ce principe que rien de ce qui est phonétique n'est significatif ou grammatical, dans le sens large du terme [...] » (*op. cit.*, p. 194). On ignore ainsi le sens d'un mot et ne considère que son enveloppe matérielle. Il s'agit d'y « découper des tranches phoniques sans se demander si elles ont une signification » (*op. cit.*, p. 194). La perspective diachronique est le domaine d'étude de la tradition historico-comparatiste. Le fait synchronique, au contraire, « ne connaît qu'une perspective, celle des sujets parlants » (*op. cit.*, p. 128) ; il est ainsi toujours significatif. On étudie les rapports coexistants d'un état de langue.

Pour illustrer l'opposition entre synchronie et diachronie, Saussure présente quelques faits. Par exemple, bien qu'aujourd'hui les sujets parlants établissent un rapport entre « décrépi » et « décrépît » dans les expressions « un mur décrépi » et « un homme décrépît », ces mots n'ont rien à voir l'un avec l'autre du point de vue historique. Le radical français « crép- » – dans les verbes « crépir », pour « recouvrir de mortier et décrépir, enlever le mortier » – est fourni par le latin « *crispus* », « ondulé », « crêpé », tandis que « décrépît » est le résultat de l'emprunt au latin du mot « *dēcrepitus* » signifiant « usé par l'âge ». Or, ce point de vue diachronique échappe entièrement à la conscience des sujets parlants, qui ne connaissent que la dimension du présent, c'est-à-dire, le rapport entre deux termes coexistants dans la langue. Comme l'affirme Chiss,

l'hétérogénéité saussurienne du synchronique et du diachronique ouvre à une conception théorique du synchronique comme présent de la connaissance, qui subordonne l'intelligence du passé de la langue à la prise en compte de la synchronie comme système articulé de concepts (Chiss, 1978, p. 102, repris dans Chiss & Puech, 1997).

Saussure différencie à ce propos l'analyse objective de l'analyse subjective. La première est fondée sur l'histoire, tandis que la seconde ne l'est pas et relève du point de vue du sujet parlant. En d'autres termes, l'analyse objective est l'analyse du linguiste et échappe à la conscience des sujets alors que l'analyse subjective est justement celle à laquelle les sujets parlants procèdent lors de leur usage quotidien de la langue. Ainsi, suivant un exemple de Saussure, si l'on prend la forme grecque « *híppos* », le grammairien distingue trois éléments : une racine, un suffixe et une désinence « *hípp-o-s* », tandis que le sujet parlant n'en aperçoit que deux « *hípp-os* ». Selon Saussure, le premier

est souvent tenté de voir des erreurs dans les analyses spontanées de la langue ; en fait l'analyse subjective n'est pas plus fautive que la « fautive » analogie [...]. La langue ne se trompe pas ; son point de vue est différent, voilà tout (*op. cit.*, p. 251).

L'auteur illustre cette distinction à travers une comparaison :

Le mot est comme une maison dont on aurait changé à plusieurs reprises la disposition intérieure et la destination. L'analyse objective totalise et superpose ces distributions successives ; mais pour ceux qui occupent la maison, il n'y en a jamais qu'une (*op. cit.*, p. 252).

Ainsi l'analyse « *hípp-os* » examinée plus haut, n'est pas fautive. Elle est tout simplement anachronique, car établie par la conscience des sujets parlants et se reporte à une autre époque. La conclusion de Saussure est que si « *hípp-o-s* » ne contredit pas « *hípp-os* », il ne faut cependant pas juger ces formes de la même façon et « cela revient à poser une fois de plus *la distinction radicale du diachronique et du synchronique* » (*op. cit.*, p. 252). Les faits appartenant à la série diachronique ne sont pas du même ordre que ceux de la série synchronique, car le changement phonétique se produit en dehors de toute intention. Le fait de synchronie, quant à lui, est toujours significatif.

Ces deux points de vue correspondent donc bien à deux ordres d'études possibles et légitimes mais incompatibles. Néanmoins, cela ne signifie pas que le point de vue synchronique soit an-historique. Comme l'a bien observé C. Normand,

[o]n a cru cependant de façon trop simple que le problème du temps se trouvait réglé par le renvoi à une étude entièrement indépendante de celle des états ; cette abstraction, tout en étant nécessaire à une description rigoureuse des faits, risque, à ne retenir qu'elle, de masquer la véritable nature de la langue (Normand, 2004, p. 133).

En effet, l'opposition entre les domaines diachronique et synchronique ne correspond pas à une opposition entre un domaine historique et un domaine anhistorique. Il s'agit d'une opposition d'ordre méthodologique qui sert à distinguer un domaine qui relève plutôt du physiologique (la diachronie) d'un domaine qui relève de la conscience des sujets parlants (la synchronie). Nous pensons avoir bien montré lors de notre discussion autour du principe du mouvement de la langue dans le temps et celui de la continuité dans le temps que le rapport signe/idée, d'ordre synchronique, est radicalement historique ; il s'agit d'un ordre toujours en train de se constituer, dans un processus continu. Le système et l'histoire ne sont pas deux notions opposées. Elles apparaissent en fait essentiellement liées. Il n'y a pas de système en dehors du fait social et les restrictions impliquées dans la chaîne signifiante sont d'ordre historique et non pas individuel ou universel. Un état de langue est pour Saussure, en effet, « toujours le produit de facteurs historiques » (Saussure, 1985, p. 105).

Pour terminer, nous voudrions faire une remarque concernant la notion de *fait social* chez Saussure. Bien qu'il s'agisse, à la fin du XIX^e siècle, d'un lieu commun, cette notion renvoie cependant à des sens très différents. Pour Saussure, il s'agit, nous l'avons vu, d'un caractère interne à la langue. Et par là, le linguiste se démarque de la perspective sociologiste de son élève Antoine Meillet (1866-1836), chez qui l'influence d'Émile Durkheim est évidente, ce qui, d'ailleurs, semble placer la nature de l'objet linguistique sur un tout autre terrain. La notion d'arbitraire, par exemple, semble jouer un rôle peu important chez le linguiste français, tandis que chez Saussure, il s'agit d'un principe, le « premier principe » (*op. cit.*, 100). L'auteur observe en effet, à propos de ce principe, que

il est souvent plus aisé de découvrir une vérité que de lui assigner la place qui lui revient. Le principe énoncé plus haut [de l'arbitraire] domine toute la linguistique de la langue ; ses conséquences sont innombrables. Il est vrai qu'elles n'apparaissent pas toutes du premier coup avec une égale évidence ; c'est après bien des détours qu'on le découvre, et avec elles l'importance primordiale du principe (*op. cit.*, p. 100).

Ce point là est en rapport étroit avec la distinction faite par Saussure concernant les deux ordres d'étude linguistique, mais autrement posée : juste après avoir établi la distinction entre les points de vue diachronique et synchronique, Saussure l'affirme : « il y a lieu de distinguer dans la langue les phénomènes *internes* ou de conscience et les phénomènes *externes*, directement saisissables » (Saussure, 2002, p. 17, souligné dans le texte). À ce premier ordre de phénomènes correspond le point de vue de la synchronie ; au second, celui de la diachronie. La linguistique de Meillet est essentiellement externe et pose une primauté de l'ordre sociale, à la différence de Saussure, pour qui le principe de l'arbitraire met en cause une telle primauté en plaçant la langue sur un tout autre terrain, ce qui fait de cette science une science autonome.

6. Conclusion

Nous espérons avoir montré comment la position de Saussure à l'égard du problème de l'origine reflète un point de vue tout à fait particulier en ce qui concerne l'essence de l'objet de la linguistique. Aujourd'hui les études saussuriennes sont en vogue à nouveau ; elles se veulent une réaction contre les représentations que le XX^e siècle a produites.

Dans cet article, nous avons voulu nous concentrer sur un point spécifique, celui concernant la question du rapport système/temps. Chez la plupart de commentateurs post-saussuriens, Saussure aurait exclu l'aspect historique de la linguistique, en concevant la langue comme un système fermé ou même, dirons quelques auteurs, comme une structure. Il faut remarquer que le terme *structure* ne figure pas dans l'œuvre de Saussure et qu'il s'agit là d'anachronisme. En effet, de plus en plus on distingue Saussure du structuralisme.

Nous avons vu à plusieurs reprises que, pour Saussure, le problème de l'origine ne diffère pas de celui des conditions permanentes. Dans cette affirmation, nous trouvons le réflexe parfait d'un aspect tout à fait central de la théorie saussurienne : le rapport constitutif entre système et histoire. Loin d'y voir une incompatibilité, Saussure considère le système comme radicalement historique. Fondé sur aucune raison en dehors de l'histoire, le fonctionnement linguistique est orienté par une raison relative toujours en train de se produire et dont le seul but est la préservation de la dimension du sujet parlant, c'est-à-dire, la continuation de ce qui se produit ; il s'agit là de l'ininteruption forcée dont parle Saussure.

Le caractère radicalement historique du système ne doit pas se confondre avec les deux ordres d'études que Saussure pose comme légitimes, mais essentiellement opposés : les points de vue synchronique et diachronique. Il s'agit là de séparer un ordre d'étude qui relève de la dimension du sujet parlant et du sens, d'une autre dont le seul aspect envisagé est le son pris dans une suite temporelle.

Pour conclure, nous voudrions faire une brève remarque : il nous semble que les Sources manuscrites ne présentent pas un tout autre Saussure. Les aspects soulevés ici concernant le point de vue saussurien sur l'objet de la linguistique, quoique d'une manière moins claire, se présentent déjà dans le *Cours* rédigé. Rappelons la lecture du philosophe Maurice Merleau-Ponty, qui, n'ayant pas eu accès aux Sources manuscrites, et donc, n'ayant eu comme référence que le *Cours*, affirme en 1953 que « Saussure pourrait bien avoir esquissé une nouvelle philosophie de l'histoire » (Merleau-Ponty, 1953 et 1960, p. 56). Encore, il faut ajouter que la plupart des Sources manuscrites sont disponibles depuis 1968, ce qui peut ouvrir un tout autre débat : pourquoi a-t-on lu Saussure en excluant tant d'aspects cruciaux ? Une telle problématique relève, croyons-nous, du domaine ouvert par Chiss et Puech (1999), à savoir : « l'histoire de l'émergence disciplinaire » (Chiss & Puech, 1999, p. 53). Une telle perspective nous amène moins à la construction d'une nouvelle lecture de Saussure, que l'étude du rôle du *Cours* dans la constitution des sciences du langage au XX^e siècle. Mais là il s'agit d'une toute autre question, qui appelle une recherche à part entière.

Références bibliographiques

- Auroux (S.). 1989. « La question de l'origine des langues : ordres et raisons du rejet institutionnel ». in : Gessinger, (J.) & Radhen (von W.). *Theorien vom Ursprung der Sprache*. Berlin/New York : Walter de Gruyter, pp. 122-150.
- Bakhtine (M.). 1929. *Le marxisme et la philosophie du langage*. Paris : Minuit.
- Calvet (L.-J.). 1985. « Postface : Lire Saussure aujourd'hui ». in : Saussure (F. de). *Cours de linguistique générale*. Paris : Éditions Payot & Rivages, pp. 505-513 [Cours publié par Charles Bally et Albert Sechehaye avec la collaboration d'Albert Riedlinger, édition critique préparée par Tullio de Mauro. Édition critique originale italienne de 1967 ; notes et commentaires de T. de Mauro trad. par L.-J. Calvet, 1972].
- Chiss (J.-L.). 1978. « Synchronie/diachronie : méthodologie et théorie en linguistique ». in : Normand (C.), (ed.). *Langages*, 49, pp. 90-111.
- Chiss (J.-L.) & Puech (C.). 1997. *Fondations de la linguistique. Études d'histoire et d'épistémologie*. Louvain-la-Neuve : Duculot [2^e édition].
- Chiss (J.-L.) & Puech (C.). 1999. *Le langage et ses disciplines XIX^e-XX^e siècles*. Paris, Bruxelles : De Boeck & Larcier, Département Duculot.
- Fehr (J.). 2000. *Saussure entre linguistique et sémiologie*. Paris : P.U.F.
- Foucault (M.). 1966. *Les mots et les choses*. Paris : Gallimard.
- Joseph (J.). 2004. « The linguistic sign ». in : Sanders (C.). *The Cambridge companion to Saussure*. Cambridge : Cambridge University Press, pp. 59-75.
- Merleau-Ponty (M.). 1953-1960. *Éloge de la philosophie*. Paris : Gallimard.
- Milner (J.-C.). 1978. *L'amour de la langue*. Paris : Seuil.
- Normand (C.). 1988. « W.-D. Whitney et la linguistique générale ». Préface à la traduction de Whitney (W.-D.). [1875] 1988. *La vie du langage*. Paris : Didier.
- Normand (C.). 2004. *Saussure*. Paris : Les belles lettres [2^e édition].
- Pêcheux (M.). 1969. « Analyse automatique du discours ». in : Mالدیدیر (D.), (ed.). 1990. *L'inquiétude du discours*. Paris : Éditions de Cendres, pp. 97-132.
- Pêcheux (M.). 1975. « Les vérités de la Palice » : in : Mالدیدیر (D.), (ed.). 1990. *L'inquiétude du discours*. Paris : Éditions de Cendres, pp. 175-243.
- Pénisson (P.). 1977. « Introduction ». in : Herder (J.-G.). *Traité sur l'origine de la langue*. Paris : Aubier, pp. 5-37.
- Puech (C.) & Radzynski (A.). 1978. « La langue comme fait social : fonction d'une évidence ». in : Normand (C.), (ed.). *Langages*, 49, pp.46-65.
- Saussure (F. de). 1968. *Cours de linguistique générale*. Wiesbaden : Otto Harrassowitz [Édition critique par Engler (R.). T.I.].
- Saussure (F. de). [1916] 1985. *Cours de linguistique générale*. Paris : Éditions Payot & Rivages [Cours publié par Charles Bally et Albert Sechehaye avec la collaboration d'Albert Riedlinger, édition critique préparée par Tullio de Mauro, postface de Louis-Jean Calvet. Édition critique originale italienne, 1967 ; notes et commentaires de T. de Mauro trad. par L.-J. Calvet, 1972].
- Saussure (F. de). 1997. *Deuxième cours de linguistique générale : 1908-1909*. Oxford, New York, Séoul : Pergamon press [d'après les cahiers d'Albert Riedlinger et Charles Patois].
- Saussure (F. de). 2002. *Écrits de linguistique générale*. Paris : Gallimard [Textes établis et édités par Simon Bouquet et Rudolf Engler avec la collaboration d'Antoinette Weil].
- Société de Linguistique de Paris. 1871. « Statuts de 1866 ». [Site internet consulté en janvier 2006 : <http://www.slp-paris.com>, rubrique « Historique »].
- Trabant (J.). 2004. « La linguistique face à l'origine du langage ». in : Puech (C.), (ed.). *Linguistique et partages disciplinaires à la charnière des XIX^e et XX^e siècles : Victor Henry (1850-1907)*. Louvain, Paris, Dudley, MA : Éditions Peeters, Coll. « Bibliothèque de l'Information Grammaticale », 55, pp. 39-51 [Avec une bibliographie de V. Henry par Marc Décimo].
- Turpin, (B.). 2003. « Légendes – Mythes – Histoire. La circulation des signes ». in : *Cahiers l'Herne*. Paris : L'Herne, pp. 307-316.
- Whitney (W.-D.). [1867] 1870. *Language and the study of language*. London : N. Trübner [3^e ed.].
- Whitney (W.-D.). 1871. « Strictures on the Views of August Schleicher respecting the Nature of Language and kindred subjects ». in : *Transactions [« puis » : and proceedings] of the American philological association*. Hartford puis Cambridge : Haverford, pp. 35-64.
- Whitney (W.-D.). [1875] 1988. *La vie du langage*. Paris : Didier.



Mai 2006

1. Introduction : l'anatomie d'une transition²

Le comportement langagier humain repose sur plusieurs capacités cognitives propres à notre espèce. Les animaux n'utilisent pas de système de communication qui soit directement comparable au nôtre. Les tentatives pour leur enseigner différents aspects du langage humain (Premack & Premack, 1983 ; Savage-Rumbaugh & Lewin, 1994 ; Pepperberg, 1999) révèlent que certaines des capacités cognitives qui nous permettent de manier le langage semblent ne pas être présentes chez les animaux testés. L'une de ces capacités, maintes fois soulignée, concerne l'usage d'une syntaxe. Certes, les oiseaux, par exemple les rossignols, sont connus pour produire des chants présentant plusieurs niveaux de combinatoire (Hauser, 1996, p. 286). Cependant, rien n'indique que les différentes structures ainsi produites servent à exprimer des significations différenciées. Les primates entraînés produisent parfois spontanément des groupements de deux mots sémantiquement reliés au contenu concret exprimé, avec une tendance à observer un ordre systématique (Savage-Rumbaugh & Lewin, 1994, p. 161). Ce type de performance, qui reste éloignée de ce que les êtres humains produisent avec le langage, amène certains auteurs à voir là l'expression d'une capacité protolangagière (Bickerton, 1990, 1995).

Le protolangage a été défini, par extrapolation à partir de l'observation des pidgins, comme une forme d'expression consistant à assembler quelques mots dans un énoncé court, sans support grammatical : absence de mots grammaticaux, absence de dépendance à distance dans la phrase, absence de marques de flexion, absence d'ordre défini (Bickerton, 1990). Le protolangage, vers lequel nous régressons spontanément dès que nous sommes en difficulté linguistique, est ainsi présenté par Bickerton comme un précurseur du langage, une sorte de capacité intermédiaire entre la communication spontanée des primates et le langage proprement dit qui est universellement pratiqué dans notre espèce. Une expression moderne du protolangage peut s'observer sur les moteurs de recherches utilisés sur la toile mondiale. Bien que ces moteurs de recherche permettent de s'exprimer à l'aide d'expressions booléennes et reconnaissent des groupes nominaux, les utilisateurs ont tendance à s'exprimer par des suites de mots sans liens grammaticaux entre eux (Guichard, 2002).

Quelle que soit l'attitude que l'on adopte vis-à-vis de la réalité du protolangage, la question de la transition menant, au cours de l'évolution, à la forme de langage que nous connaissons, ne peut être ignorée. Certaines de nos aptitudes cognitives nous permettent de manier des constructions langagières signifiantes et de réaliser des opérations non triviales sur ces structures, comme l'interrogation, la mise au passif, le changement modal ou temporel, etc. Sans prendre ici parti sur la question de savoir si ces dispositions cognitives sont spécifiques au langage ou correspondent à des fonctions plus générales (Piattelli-Palmarini, 1979), nous cherchons dans la structure du langage des indices susceptibles de nous renseigner sur leur ordre d'apparition. Pour tenter de reconstituer l'anatomie de la transition vers le langage, nous nous intéressons particulièrement à l'interface entre la syntaxe et la sémantique, en adoptant

¹ Je remercie *Laleh Ghadakpour* qui a bien voulu commenter une version antérieure de cet article. Je voudrais exprimer ma gratitude aux personnes qui sont à l'origine de l'initiative « Origine de l'Homme, du Langage et des Langues » du CNRS, qui ont permis à la problématique des origines du langage de retrouver une juste place dans le débat scientifique français.

² Ce travail a été effectué dans le cadre de l'initiative OHLL (Origine de l'Homme, du Langage et des Langues) du CNRS.

le point de vue de la modélisation. En considérant la complexité relative des opérations de synchronisation entre la structure syntaxique et le sens, nous en viendrons à proposer un ordre dans l'apparition de certaines capacités qui sous-tendent la maîtrise du langage.

Dans la suite, nous commençons par poser le problème en reprenant certaines caractéristiques formelles de la syntaxe qui sont généralement invoquées pour démontrer l'originalité de notre mode de communication, notamment la récursivité et les systèmes de flexion. Nous tentons de dégager un modèle simplifié du rôle de ces caractéristiques dans l'expression du sens. Nous tirerons argument de la structure de ce modèle simplifié pour proposer une succession plausible dans l'établissement des capacités cognitives qui sous-tendent ces aspects du langage. Nous concluons en replaçant ces propositions dans le cadre des recherches actuelles sur l'origine de la capacité langagière.

2. Quelques aspects notables de la capacité syntaxique

La variété des langues est grande, mais elle n'est pas infinie. En particulier, toutes les langues spontanées du monde, à l'exception des pidgins récents, possèdent une grammaire qui impose la flexion de certains mots et leur agencement dans la phrase. Aucun peuple au monde ne s'exprime spontanément en groupant, sans ordre particulier, des mots invariables¹.

Certes, les règles syntaxiques diffèrent largement d'une langue à l'autre. Certaines langues comme le latin ou le dyirbal semblent très permissives en ce qui concerne l'ordre des mots dans les phrases simples (Dixon, 1972), alors que d'autres comme le français ou l'anglais sont nettement moins tolérantes sur cet aspect de la grammaire. Certaines langues, comme le lakota, mettent en œuvre un système morphologique complexe affectant principalement le verbe, alors que des langues comme l'anglais ou le chinois sont très économes en flexions (Van Valin & LaPolla, 1997). Malgré ces différences flagrantes, il est possible de faire quelques observations générales sur la structure des langues qui suffiront pour la suite de notre propos.

La première observation porte sur l'existence de dépendances entre les mots. Dans un énoncé protolangagier du type *maison-voisin-feu*, chacun des trois mots entretient un rapport sémantique direct avec la scène perçue décrite. La communication peut fonctionner sans que les mots d'un tel énoncé ne dépendent les uns des autres. Les langues humaines introduisent plusieurs formes de dépendance, qui sont clairement indiquées par la grammaire. Dans la phrase *la maison du voisin est en feu*, les mots entretiennent des relations formelles, indépendamment de leur signification. Ces relations, par exemple, rapprochent la préposition *en* et le nom *feu*, ou les noms *maison* et *voisin*, alors qu'elles laissent *voisin* et *feu* sans lien direct. Ces dépendances priment souvent sur le sens : la grammaire d'une expression comme *la souris mange le chat* impose une relation sémantique qui va à l'encontre de ce que suggère la simple association entre les sens évoqués par les mots *souris*, *chat*, *manger*.

La deuxième observation concerne le fait que les dépendances syntaxiques sont asymétriques. Dans une expression latine comme *domus domini*, le génitif du mot *dominus* indique sans ambiguïté que ce mot est un satellite du mot *domus*. En français, les deux expressions *la maison du maître* et *le maître de la maison* se distinguent par la position relative des mots par rapport à la préposition *de*. Cette dépendance asymétrique s'interprète de diverses manières, comme résultant d'une opération de fusion syntaxique (Chomsky, 1995) ou d'une dépendance sémantique (Van Valin & LaPolla, 1997). Du fait de la présence de dépendances asymétriques, les énoncés langagiers comportent des syntagmes, dont la tête domine, au sens structurel, l'élément dépendant.

La troisième observation concerne la structure « moléculaire » des énoncés langagiers. Comme les relations de dépendance peuvent s'enchaîner, les syntagmes se retrouvent emboîtés les uns dans les autres comme des poupées russes. Dans une phrase simple comme *J'ai moyennement aimé l'acoustique du concert auquel j'ai assisté hier*, on trouve maints emboîtements : le syntagme verbal centré sur le verbe *assister* dépend (indirectement) du syntagme nominal centré sur le mot *concert*, qui fait partie intégrante du syntagme nominal centré sur

¹ Les pidgins de création récente constituent clairement une exception. Toutefois, ils sont utilisés comme langue principale seulement dans des conditions sociales anormales qui imposent un brassage d'individus adultes de communautés linguistiques différentes (Bickerton 1990). L'histoire de la formation des créoles et l'exemple des enfants sourds du Nicaragua (Kegl, Senghas & Coppola, 1999) indiquent que les pidgins sont une forme de communication instable qui disparaît en tant que moyen de communication principal dès que la structure sociale permet aux enfants de moins de sept ans de former des communautés.

acoustique, lui-même inclus dans le syntagme verbal dont la tête est le verbe *aimer*. Les différentes théories syntaxiques peuvent différer par le choix précis des types de syntagmes, mais toutes admettent le principe de l'imbrication. Il en résulte une structure qui peut être qualifiée d'arborescente (si l'on représente le graphe des dépendances), de moléculaire (par analogie avec la chimie des macro-molécules, qui sont constituées d'autres molécules), de fractale (par référence aux structures vivantes ou physiques invariantes par changement d'échelle) ou de récursive. Ce dernier qualificatif fait référence à une propriété de la procédure qui permet de lire ou d'engendrer la structure syntaxique. Les procédures récursives ont la propriété d'appeler une copie d'elles-mêmes au cours de leur exécution. C'est donc la même procédure qui permet d'analyser le syntagme nominal *acoustique du concert auquel j'ai assisté hier* et le syntagme nominal inclus *concert auquel j'ai assisté hier*. D'un point de vue technique, le fait remarquable est que la récursivité à l'œuvre dans le langage humain est une récursivité centrale, c'est-à-dire une récursivité qu'il est impossible de ramener à une simple itération répétitive.

La dernière observation que nous faisons ici à propos des dépendances syntaxiques est qu'elles s'exercent à distance. Dans *l'acoustique du concert auquel j'ai eu la chance d'assister hier*, on comprend que le nom *concert* sert de complément à la fois au nom *acoustique* et au verbe *assister*. Pourtant, *concert* et *assister* sont séparés de sept mots. La distance de ce type de dépendance peut être arbitrairement grande (*le concert auquel la fille du copain que j'ai connu quand j'étais à la fac a eu la chance d'assister avait une mauvaise acoustique*). L'interprétation de ces dépendances à distance varie d'une théorie à l'autre, selon que l'on parle de mouvement (le complément de *assister* migrant en tête de clause) ou de duplication silencieuse.

Pour plusieurs auteurs, certaines de ces propriétés syntaxiques sont indissociables. L'existence de dépendances asymétriques, l'emboîtement des syntagmes et les différentes formes de dépendance à distance iraient de pair, elles constitueraient une sorte de paquet cadeau que la Nature nous aurait offert en nous dotant de la capacité syntaxique. L'ensemble de ces propriétés résulterait de la seule opération de fusion (Berwick, 1998) ou, de manière équivalente, de l'apparition fortuite dans notre espèce d'un accès à la récursivité (Hauser, Chomsky & Fitch, 2002). De fait, il est difficile d'imaginer des intermédiaires dans l'accès à la récursivité. Dans ces conditions, la transition du protolangage (ou du non-langage) vers le langage ne peut être que soudaine. Certains ont pu invoquer l'intervention d'une improbable macro-mutation (Bickerton, 1990). Sa cause aurait été purement fortuite, et ne devrait rien à un quelconque accroissement du pouvoir communicationnel (Chomsky, 1975, p. 75 ; Piattelli-Palmarini, 1989 ; Lightfoot, 2000). La nouvelle capacité récursive, peut-être à la faveur d'un décloisonnement fonctionnel (Hauser, Chomsky & Fitch, 2002), aurait entraîné l'ensemble des propriétés de la syntaxe qu'il nous est donné d'observer dans les différentes langues, avec parmi elles les propriétés de dépendance que nous avons soulignées ici.

La question de la transition vers la capacité syntaxique, s'il s'est agi d'un changement abrupt et fortuit, perd beaucoup de son intérêt scientifique, car elle n'a pu produire qu'un dispositif peu adapté aux utilisations que les êtres humains en font.

Language design as such appears to be in many respects « dysfunctional, » yielding properties that are not well adapted to the function language is called upon to perform (Chomsky, 1995, p. 162).

[...] full-blown evolutionary novelty can also suddenly arise, so to speak, for no reason, because novelty caused by sheer proximity between genes is not governed by function and it, therefore, eludes strict adaptationism (Piattelli-Palmarini, 1989, p. 8).

Le point de vue selon lequel l'évolution pourrait ainsi produire des structures ou des compétences complexes peu fonctionnelles apparaît comme non fondé sur le plan biologique (Desalles, 1996), et ne saurait s'appliquer au langage (Pinker & Bloom, 1990 ; Pinker & Jackendoff, 2005). La théorie de la sélection naturelle et l'étude cognitive du langage suggèrent toutes deux que la syntaxe remplit une fonction précise, qui est celle de l'expression du sens sur le canal de l'expression vocal, et que sa structure est la conséquence évolutive de cette fonction. Il est donc légitime d'étudier la fonction de la syntaxe pour tenter de proposer une reconstitution des grandes étapes par lesquelles est passée l'évolution du langage.

Dans la suite, nous allons nous attacher à remettre en question ces deux propriétés supposées de la transition du protolangage au langage que seraient son caractère abrupt et son caractère fortuit. En observant, du point de vue de la modélisation, l'interface entre la syntaxe

et la sémantique, nous allons suggérer une transition en deux étapes au moins, dont chacune serait motivée, sur le plan biologique, par ses conséquences sur le type de communication qu'elle rend possible.

3. À l'interface entre la grammaire et le sens

La manière dont le sens d'un énoncé est construit à partir de sa structure grammaticale a fait l'objet de nombreux travaux et continue à être l'enjeu de nombreuses recherches. Les paradigmes divergent largement, allant d'une conception mettant en jeu deux constructions isomorphes (Montague, 1974 ; Jackendoff, 1990 ; Carpenter, 1997) à diverses conceptions dans lesquelles la structure grammaticale ne fait qu'imposer des contraintes à une construction sémantique qui se déroule en parallèle (Johnson-Laird, 1977 ; Kamp & Reyle, 1993 ; Talmy, 2000). Pour notre propos, une théorie globale de la construction du sens n'est pas nécessaire. Il nous suffit de considérer certains aspects de la synchronisation entre le code syntaxique et le niveau sémantique.

Le premier aspect de cette synchronisation concerne l'identification de la relation prédicat-arguments. Nous nous plaçons dans l'hypothèse où la représentation sémantique d'un énoncé met en jeu un prédicat logique¹. Cette hypothèse est compatible avec la plupart des modèles de la sémantique². Il est important de noter que l'hypothèse n'implique pas l'existence d'un langage de la pensée, dans la mesure où le prédicat en question peut être une représentation transitoire, construite en contexte à partir d'éléments non symboliques (Ghadakpour, 2003).

Prenons l'exemple d'un prédicat à deux places $P(x,y)$. Ce sera par exemple un prédicat associé à l'énoncé *Pierre frappe Paul*. L'une des fonctions fondamentales de la syntaxe consiste à offrir les moyens de repérer x , y et P dans l'énoncé. En particulier, il s'agit de permettre de distinguer les énoncés associés à $P(x,y)$ et à $P(y,x)$. Les langues utilisent divers moyens pour parvenir à ces identifications des actants. Les plus évidents sont la position et le marquage flexionnel. Ainsi, dans une langue comme le français, la position des mots est suffisamment stricte pour permettre le repérage du prédicat et de ses arguments, et pour distinguer par leur sens l'énoncé *Pierre frappe Paul* et l'énoncé *Paul frappe Pierre*. En latin, la présence des cas nominatif et accusatif rend le même service dans un énoncé comme *Paulus Petrum ferit*. La grammaire d'une langue donnée permet également de repérer arguments et prédicat dans le cas de prédicats à trois places (*Paul vend sa voiture à Jacques*), à une place (*Paul est malade*) ou à zéro place (*il pleut*).

Dans la précédente section, nous présentions l'existence de dépendances asymétriques au sein de la phrase comme une caractéristique fondamentale du langage. L'un des effets fondamentaux de ces dépendances est précisément de permettre l'identification de la relation entre le prédicat et ses arguments. En d'autres termes, le syntagme constitue le reflet visible de la prédication. Appelons fonction argumentale cette propriété de la syntaxe.

Le second aspect de la synchronisation syntaxe-sémantique auquel nous allons nous intéresser est la détermination. Pour être utilisable au niveau pragmatique, la représentation sémantique doit être suffisamment instanciée. Dans certains contextes, ce n'est pas le fait qu'un quelconque être vivant ait mangé quelque chose qui importe, mais bien que le chat de la maison a mangé le rôti destiné aux invités. Dans un autre contexte, par exemple celui d'un malade qui recouvre l'appétit, l'identité du mangeur est pertinente, alors que l'entité mangée peut rester non précisée. Voyons comment la syntaxe se met au service de la détermination des arguments des prédicats.

Les dispositifs syntaxiques permettant la détermination des arguments logiques sont variés. Intéressons-nous à deux d'entre eux, la modification et la relativisation. Dans l'énoncé :

Va me chercher le petit livre que Paul a offert à Jacques

la modification par l'adjectif *petit* et la clause relative *que Paul a offert à Jacques* aident l'interlocuteur à mieux comprendre de quel livre il s'agit. Si l'on représente le sens

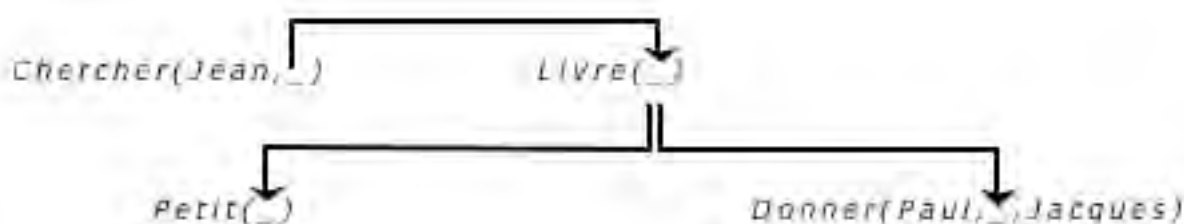
¹ Les prédicats logiques sont le moyen privilégié par lequel le sens des énoncés est interfacé avec les aspects pragmatiques, notamment l'argumentation (Dessalles, 2000).

² En particulier, il est possible de convertir en prédicats logiques les constructions produites dans les modèles qui utilisent des formalismes à base de traits, d'éléments logiques, d'éléments fonctionnels, de graphes ou d'éléments schématiques. Les modèles qui ne font intervenir que des représentations analogiques du sens, par exemple des images, pourraient faire exception.

correspondant à l'aide de prédicats, on obtient une représentation du type $Petit(x) \& Livre(x) \& Offert(Paul, x, Jacques)$, en supposant que l'auditeur de l'énoncé soit capable de former, dans le contexte courant, les prédicats notés *Petit*, *Livre* et *Offert*. Dans ce type de représentation (qui rappelle le langage informatique *Prolog*), il est convenu que deux variables de même nom, au sein de la même clause logique, désignent le même objet. On pourrait imaginer que les êtres humains s'expriment de cette manière, à l'aide de propositions simples et de variables :

{
Va me chercher quelque chose ;
ce quelque chose est un livre ;
ce quelque chose est petit ;
Paul a offert ce quelque chose à Jacques.

Ce mode d'expression par partage de variables entre propositions simples semble lourd à la plupart d'entre nous (bien qu'une minorité d'informaticiens utilisateurs du langage *Prolog* l'affectionne pour exprimer des connaissances précises). La syntaxe des langues naturelles permet le plus souvent de faire l'économie de variables explicites, conformément au schéma suivant.



Dans ce schéma, les flèches représentent la liaison sémantique qui existe entre les syntagmes. Cette liaison est obtenue grâce à un certain nombre de dispositifs syntaxiques, parmi lesquels le principe de liaison sémantique qui stipule que deux syntagmes en relation syntaxique (complémentation ou adjonction) doivent partager une variable au niveau sémantique¹ (Dessalles, 2000). Ce principe correspond aux liaisons entre fragments de discours qui apparaissent dans un formalisme comme celui de la DRT (Kamp & Reyle, 1993).

Le principe de liaison sémantique n'est pas un moyen infaillible d'assurer la liaison entre arguments logiques. En français, le syntagme *le dépassement de la voiture* est sémantiquement ambigu, car on ne sait pas si la voiture en question dépasse ou est dépassée. Ce phénomène s'explique par le fait que le principe de liaison sémantique sous-spécifie, dans ce cas, l'argument connecté dans $Dépasser(_, _)$ & $Voiture(_)$, où l'un quelconque des deux arguments de *Dépasser* peut être lié à l'argument de *Voiture*.

La fonction essentielle du principe de liaison sémantique et des différents dispositifs syntaxiques qui le servent est de permettre la détermination des objets du discours. Si l'on conçoit chacun des prédicats comme une contrainte posée par le locuteur sur le monde perçu, chaque liaison avec un nouveau prédicat limite le champ des possibilités, et donc facilite la compréhension pour l'interlocuteur. Identifier l'objet du discours consiste à résoudre une équation comme $Petit(x) \& Livre(x) \& Offert(Paul, x, Jacques)$, où x est l'inconnue. Un locuteur pertinent s'attachera à donner suffisamment de contraintes prédictives pour que la solution apparaisse comme unique à son interlocuteur (l'utilisation d'un déterminant défini annonce cette unicité pour l'interlocuteur). Dans un contexte déjà contraint, exprimer simplement $Livre(x)$ pourra suffire. Le principe de liaison sémantique permet de mieux comprendre la raison d'être de la récursivité syntaxique. Chaque nouveau prédicat introduit pour spécifier l'argument d'un autre prédicat peut se voir, à son tour, adjoindre un nouveau prédicat qui préciserà ses propres arguments. Il en résulte une structure sémantique arborescente qui lie les prédicats les uns aux autres.

Le principe de liaison sémantique agit localement sur des syntagmes proches, compléments ou adjoints. Paradoxalement, ce même principe de liaison sémantique permet de comprendre la raison d'être de certaines dépendances à distance. Notamment, il confère un rôle au

¹ Ce principe doit s'appliquer avec discernement. Par exemple, il demande de distinguer un emploi prédictif de la proposition *sur, dans elle voit le livre sur la table*, d'un emploi non prédictif comme dans *elle compte sur Paul*. Dans le deuxième cas, il n'y a qu'un seul prédicat, et donc aucune liaison sémantique.

mouvement (ou à la duplication silencieuse) des syntagmes nominaux, lorsqu'un rapprochement doit être obtenu en dépit d'une trop grande distance structurelle. Dans le syntagme *l'acoustique du concert auquel j'ai eu la chance d'assister hier*, le problème consiste à lier les arguments de *Acoustique(_)*, de *Concert(_)* et de *Assister(_,_)*. Or, le mot *concert* ne peut être complément à la fois du nom *acoustique* et du verbe *assister*. Le pronom *auquel*, qui renvoie au complément silencieux de *assister*, permet d'assurer cette double liaison. Le principe de liaison sémantique s'applique entre *acoustique* et son complément *concert*, et entre *concert* et le pronom *auquel*¹. La liaison sémantique est ensuite assurée par le lien direct entre le pronom *auquel* et le complément silencieux de *assister*. Si l'on analyse de manière classique notre exemple comme résultant d'un mouvement en début de clause du complément de *assister*, ce déplacement a pour rôle d'amener le mot *concert* en position d'établir un lien sémantique avec le mot *acoustique*.²

Nous venons d'évoquer deux grandes fonctions assurées par la syntaxe, la fonction argumentale et la détermination. Dans un scénario où la capacité à gérer les relations syntaxiques serait un pur produit du hasard, parachutée, pour ainsi dire, dans le paysage cognitif humain, l'existence de ces deux fonctions, et le fait que la syntaxe les assure plutôt bien, apparaissent comme miraculeux. Nous proposons d'inverser la perspective, et de voir l'aptitude à manier la syntaxe, non comme un événement miraculeux dépourvu de toute cause, mais comme le résultat d'une évolution en plusieurs étapes.

4. Du modèle à la reconstitution évolutive

Du point de vue de la modélisation, la syntaxe apparaît comme un appareil bien agencé, capable notamment de remplir deux fonctions essentielles, l'identification des arguments des prédicats et la détermination des objets du discours par liaison sémantique entre prédicats. Il est naturel de chercher à aller au-delà de ce constat, en s'interrogeant sur l'origine phylogénétique de cet appareil. Si l'on renonce à considérer l'aptitude syntaxique comme une capacité cognitive monobloc fournie clé en main par un hasard de l'évolution, il devient pertinent de tenter une reconstitution de l'ordre dans lequel les composantes de cette capacité sont apparues au cours de la phylogenèse.

La démarche que l'on peut adopter dans ce type de recherche est identique à celle des anatomistes évolutionnistes, qui distinguent les caractères ancestraux des caractères dérivés. Lorsqu'un caractère morphologique ne peut exister qu'en présence d'un autre caractère morphologique, il est nécessairement dérivé. Les ocelles qui apparaissent sur les ailes des papillons sont des caractères dérivés, apparus après les ailes au cours de la phylogenèse. Dans le cas des capacités à manier les relations syntaxiques, il est légitime de se demander laquelle, des deux fonctions que nous avons considérées, a été assurée en premier.

Posée sous cette forme, la question semble devoir recevoir une réponse dépourvue d'ambiguïté. L'identification des arguments des prédicats est sans conteste un préalable à la possibilité de lier les prédicats entre eux. En d'autres termes, la maîtrise des outils syntaxiques qui nous permettent d'identifier les arguments des prédicats *a dû précéder* l'utilisation du principe de liaison sémantique à des fins de détermination.

La conséquence la plus importante de ce constat est que l'aptitude à manier les relations syntaxiques a pu apparaître en deux temps, *la récursivité n'intervenant que dans le second temps*. Le scénario est le suivant. Premier temps, apparaissent les clauses simples n'exprimant qu'une seule prédication. Cette prédication peut être qualifiée de principale : c'est celle qui est utilisée au niveau pragmatique pour l'argumentation (Dessalles 2000). Ainsi, l'énoncé *Pierre a frappé Paul* est directement exploitable par l'auditeur qui peut adopter une attitude conversationnelle, voire s'engager dans une action. À ce stade, la prédication trouve une trace explicite dans le langage, alors qu'elle était absente du protolangage (dans une expression protolangagière du type *combat Pierre Paul*, où les trois mots ne sont ensemble que parce qu'ils se rapportent à la même scène, la constitution d'une éventuelle prédication est entièrement à la charge de l'auditeur, qui doit de plus deviner qui est le frappeur et qui est le frappé).

¹ Le statut du lien entre ce type de pronom et son antécédent varie d'un modèle à l'autre. Il nous suffit ici de considérer que ce lien transmet la liaison sémantique.

² En français, la préposition « migre » avec le complément, alors qu'en anglais, elle peut rester sur place, apparaissant alors comme une postposition.

Dans un deuxième temps, la capacité d'exprimer des prédicats acquiert une fonction dérivée. Les énoncés continuent d'inclure une prédication principale, celle qui est utilisée par le niveau argumentatif, mais ils comportent maintenant des prédications auxiliaires servant à la détermination. Ces prédications auxiliaires apparaissent notamment sous la forme de modifications par adjonction ou complémentation. Grâce au principe de liaison sémantique, les prédications auxiliaires peuvent jouer leur rôle de spécification des objets du discours¹. Le phénomène remarquable est que la spécification des arguments concerne non seulement la prédication principale, mais également les prédications secondaires, dont les arguments demandent eux aussi à être spécifiés. Il en résulte un enchaînement récursif des prédications auxiliaires qui, grâce au principe de liaison sémantique, peuvent s'enchaîner les unes aux autres. C'est donc à ce stade que la capacité de gérer les enchaînements récursifs de syntagmes trouve sa raison d'être.

5. Discussion

Le scénario que nous venons d'esquisser a pour principal mérite de rendre inutile l'hypothèse improbable selon laquelle la capacité syntaxique serait arrivée en bloc par hasard, pour trouver ensuite une utilité dans la communication. La prise en compte de la synchronisation syntaxe-sémantique permet de distinguer deux aptitudes fondamentales, apparues successivement, dans la capacité syntaxique, chacune étant motivée par un rôle fonctionnel. La première a permis d'exprimer la prédication principale, en mettant en œuvre un énoncé simple isolé pour chaque prédication. Il n'est pas besoin d'invoquer un quelconque événement miraculeux ici. Plusieurs études de simulation évolutive ont montré que des formes de syntaxe pouvaient émerger dans un contexte où il est utile d'exprimer des prédications (Batali, 1994, 1998 ; Kirby, 2002). Le rôle de la syntaxe, dans ce cas, est de donner des indices visibles des dépendances asymétriques entre le prédicat et ses arguments. De ce point de vue, les différents systèmes flexionnels (systèmes de cas, de genres, de classes, etc.) apparaissent comme concurrents des systèmes de position (Jackendoff, 1999, 2002). S'il existe deux dispositifs en compétition pour la même fonction, la satisfaction de cette fonction ne tient vraisemblablement pas du miracle, contrairement à ce que supposent les auteurs cités précédemment (cf. section 2).

Ce qui motive certains à penser que la maîtrise de la syntaxe est le fruit d'une innovation unique et improbable est lié au caractère moléculaire des énoncés langagiers. La capacité d'emboîter les syntagmes les uns dans les autres, de manière récursive, ne semble pas pouvoir être acquise de manière progressive. Dans notre scénario, cette capacité apparaît dans la seconde étape, lorsque les prédications auxiliaires sont utilisées comme contraintes de détermination. Nous ne savons pas si l'émergence d'une syntaxe récursive est spécialement improbable lorsqu'il s'agit d'exprimer des prédications auxiliaires sémantiquement liées. Nous pouvons cependant mentionner deux indices qui limitent cette improbabilité. Premièrement, certains chercheurs ont pu faire émerger des syntaxes récursives dans d'autres contextes, pour répondre à un problème d'économie de formes (Nowak, Plotkin & Jansen, 2000) ou de généralisation (Kirby, 2002). Deuxièmement, la syntaxe récursive n'est pas le seul moyen d'exprimer l'enchaînement des prédications auxiliaires. On peut parfaitement imaginer un mode d'expression à la Prolog fonctionnant par partage de variables (cf. section 3). Si un tel système n'a pas émergé pour servir l'expression humaine spontanée, c'est sans doute que le principe de liaison sémantique est apparu plus rapidement, dotant les énoncés d'une structure moléculaire et de mécanismes de dépendance à distance.

Certains travaux présentent la syntaxe comme une innovation ayant permis de répondre aux besoins d'une pression lexicale accrue. Devant l'augmentation du nombre des significations à communiquer, la combinaison syntaxique apparaît comme un moyen de limiter le nombre formes lexicales à mémoriser (Kirby, 2000 ; Nowak, Plotkin & Jansen, 2000). On comprend, d'après ce qui précède, que la pression lexicale justifie, au mieux, l'apparition du protolangage. Les énoncés du protolangage, en combinant les significations de mots juxtaposés, permettent d'évoquer des scènes pour lesquelles les locuteurs ne disposent pas de mots uniques adéquats (Dessalles 2000).

¹ De manière marginale, les prédications auxiliaires peuvent jouer un rôle argumentatif. Dans *J'ai vu la jolie Marie qui était très pressée*, les prédications introduites par l'adjectif *jolie* et par la relative *qui était très pressée* ne servent pas à spécifier les objets du discours, mais expriment des éléments de connaissance à des fins argumentatives. Dans notre scénario, cet usage des prédications auxiliaires est lui-même un usage dérivé par rapport à la détermination.

La syntaxe réalise une performance bien différente. Elle permet d'exprimer la prédication. Selon ce scénario, le protolangage était limité à l'évocation de scènes concrètes. La transition du protolangage au langage a constitué une véritable innovation, par laquelle nos ancêtres, en accédant à la prédication, se sont engagés dans la voie de l'argumentation (Dessalles, 2000).

La capacité sémantique de prédication a souvent été ramenée à l'activité de catégorisation, qui passe par la mise en relation d'un objet localisé avec un ensemble de propriétés (Hurford, 2003). Présentée de cette manière, la prédication semble chose bien banale, et doit être l'apanage de nombreux animaux. Nous avons souligné ailleurs la différence fondamentale qui existe entre la capacité de catégorisation et la prédication argumentative exprimée par le langage¹ (Dessalles & Ghadakpour, 2003). La prédication langagière serait, elle, une acquisition récente de notre espèce, dont l'avènement a entraîné, en deux phases, l'apparition de la capacité à maîtriser les relations syntaxiques.

En proposant ce scénario, nous sommes conscients de renvoyer une partie du mystère de l'évolution des capacités langagières dans le niveau sémantique. Si l'expression syntaxique de la prédication principale, puis des prédications auxiliaires, cessent de sembler miraculeuses, l'improbabilité de la transition du protolangage au langage repose maintenant entièrement sur les capacités cognitives permettant de penser de telles prédications. Nous assumons pleinement ce changement de perspective, que nous pouvons résumer de la manière suivante.

- Premier temps : apparition d'une nouvelle capacité sémantique, permettant aux individus de former des relations prédictives à des fins argumentatives (Dessalles, 2000). Une première forme de syntaxe apparaît alors pour exprimer ces prédications (un énoncé simple par prédication). Les relations entre le prédicat et ses arguments sont assurées par des systèmes comme la flexion (cas, genres, classes...) ou la position.
- Deuxième temps : apparition d'une nouvelle utilisation sémantique de la prédication. Les prédicats peuvent être utilisés comme des contraintes au service de la détermination des objets du discours. De nouveaux outils syntaxiques apparaissent alors pour exprimer ces prédications auxiliaires. Le principe de liaison sémantique est utilisé pour lier localement les prédications entre elles. L'emboîtement récursif des syntagmes et les dépendances à distance émergent comme des moyens efficaces d'exprimer les enchaînements de prédications.

Compte tenu des résultats qui ont pu être obtenus dans les simulations d'évolution langagière, il est permis de penser que l'apparition des outils syntaxiques ne présente pas de caractère miraculeux, dès lors que l'on suppose données les capacités sémantiques correspondantes. En revanche, il est difficile, dans l'état actuel des connaissances, d'estimer les chances d'apparition de la capacité de prédication.

6. Conclusion

La modélisation de la synchronisation entre syntaxe et sémantique nous a permis de distinguer deux fonctions importantes de la syntaxe : la fonction argumentale et la détermination. Nous avons suggéré que ces deux fonctions étaient apparues successivement. La première, obtenue notamment par des mécanismes de flexion et de position, permet l'expression de la prédication principale, utilisée par le niveau argumentatif. La seconde, obtenue par l'emboîtement des syntagmes et les dépendances à distance, permet l'accumulation de prédications auxiliaires qui servent à contraindre les objets du discours pour l'interlocuteur.

Le scénario qui est proposé ici repose sur une modélisation simplificatrice de l'interface entre syntaxe et sémantique. La syntaxe comporte bien d'autres raffinements, que nous avons volontairement ignorés. Notre objectif était de proposer un cadre simple, un scénario en deux étapes, sur lequel il est ensuite possible de greffer d'autres hypothèses concernant l'apparition d'autres caractéristiques du langage. Nous avons voulu marquer la différence entre les outils requis pour l'expression de la prédication principale et celle des prédications annexes, en soulignant que la récursivité et les dépendances à distance ne sont utiles qu'à la seconde de ces fonctions.

¹ Une différence fondamentale est que la prédication langagière peut être systématiquement et explicitement niée en contexte : cet objet n'est pas un livre parce qu'il n'est pas en papier. La prédication langagière peut être formée en contexte par une opération de contraste entre représentations (Ghadakpour, 2003).

Le moteur de la transition, selon notre scénario, n'est donc pas l'apparition d'une nouvelle capacité syntaxique, mais l'apparition d'une nouvelle capacité sémantique. En formant des prédicats, en les exprimant par des énoncés simples, puis par des énoncés moléculaires pour atteindre une détermination poussée des objets du discours, nos ancêtres ont pu affirmer des relations et surtout remettre en question les relations affirmées par les autres. Conformément à ce qui est généralement admis, nous avons tout lieu de penser que cette transition correspond à l'avènement d'*homo sapiens*.

Références bibliographiques

- Batali (J.). 1994. « Artificial Evolution of Syntactic Aptitude ». in : Ram (A.) & Eiselt (K.), (eds). *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. Hillsdale, N.J. : Lawrence Erlbaum Associates, pp. 27-32.
- Batali (J.). 1998. « Computational simulations of the emergence of grammar ». in : Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds). *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press, pp. 405-426.
- Berwick (R.-C.). 1998. « Language evolution and the minimalist program : the origin of syntax ». in Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds). *Approaches to the Evolution of Language - Social and Cognitive Bases*. Cambridge : Cambridge University Press, pp. 320-340.
- Bickerton (D.). 1990. *Language and species*. Chicago : University of Chicago Press.
- Bickerton (D.). 1995. *Language and human behavior*. London : UCL Press [reed., 1996].
- Carpenter (B.). 1997. *Type-logical semantics*. Cambridge, MA : M.I.T. Press.
- Chomsky (N.). 1975. *Réflexions sur le langage*. Paris : Flammarion [reed. 1981].
- Chomsky (N.). 1995. *The minimalist program*. Cambridge, MA : MIT Press.
- Dessalles (J.-L.). 1996. *L'ordinateur génétique*. Paris : Hermès [<http://www.infres.enst.fr/~jld/papiers/pap.evol/96061301.html>].
- Dessalles (J.-L.). 2000. *Aux origines du langage : Une histoire naturelle de la parole*. Paris : Hermès-sciences [<http://www.enst.fr/~jld/papiers/pap.evol/99111703.html>].
- Dessalles (J.-L.) & Ghadakpour (L.). 2003. « Object recognition is not predication ». Commentary on James R. Hurford : « The neural basis of predicate-argument structure ». in : *Behavioral and Brain Sciences*, 26, 3, pp. 290-291 [<http://www.ling.ed.ac.uk/~jim/BBSNEURO/dessalles.html>].
- Dixon (R.-M.-W.). 1972. *The Dyrbal language of North Queensland*. Cambridge : Cambridge University Press.
- Ghadakpour (L.). 2003. Le système conceptuel, à l'interface entre le langage, le raisonnement et l'espace qualitatif : vers un modèle de représentations éphémères. Paris : Thèse de doctorat, École Polytechnique [<http://www.enst.fr/~jld/theses/laleh>].
- Guichard (E.). 2002. *L'internet : mesures des appropriations d'une technique intellectuelle*. Paris : Thèse de doctorat de L'École des hautes études en sciences sociales.
- Hauser (M.-D.). 1996. *The evolution of communication*. Cambridge : The MIT Press.
- Hauser (M.-D.), Chomsky (N.) & Fitch (W.-T.). 2002. « The faculty of language : what is it, who has it, and how did it evolve ? ». in : *Science*, 298, pp. 1569-1579.
- Hurford (J.-R.). 2003. « The neural basis of predicate-argument structure ». in : *Behavioral and Brain Sciences*, 26, 3 [<http://www.ling.ed.ac.uk/~jim/newro.htm>].
- Jackendoff (R.). 1990. *Semantic structures*. Cambridge : The MIT Press.
- Jackendoff (R.). 1999. « Possible stages in the evolution of the language capacity ». in : *Trends in cognitive sciences*, 3, 7, pp. 272-279.
- Jackendoff (R.). 2002. *Foundation of language - brain, meaning, grammar, evolution*. Oxford : Oxford University Press.
- Johnson-Laird (P.-N.). 1977. « Procedural semantics ». in : *Cognition*, 5, pp. 189-214.
- Kamp (H.) & Reyle (U.). 1993. *From discourse to logic : Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht : Kluwer Academic Publisher.
- Kirby (S.). 2002. « Learning, bottlenecks and the evolution of recursive syntax ». in : Briscoe (T.), (ed). *Linguistic evolution through language acquisition : Formal and computational models*. Cambridge, MA : Cambridge University Press, pp. 173-203.
- Montague (R.). 1974. *Formal Philosophy*. London : Yale University Press.
- Nowak (M.-A.), Plotkin (J.-B.) & Jansen (V.-A.-A.). 2000. « The evolution of syntactic communication ». in : *Nature*, 404, pp. 495-498.
- Pepperberg (I.-M.). 1999. *The Alex studies - Cognitive and communicative abilities of grey parrots*. Cambridge, MA : Harvard University Press [reed. 2002].
- Piattelli-Palmarini (M.), (ed.). 1979. *Théories du langage - Théories de l'apprentissage*. Paris : Seuil.
- Piattelli-Palmarini (M.). 1989. « Evolution, selection and cognition : From « learning » to parameter setting in biology and in the study of language ». in : *Cognition*, 31, pp. 1-44.

- Pinker (S.) & Bloom (P.). 1990. « Natural language and natural selection ». in : *Behavioral and Brain Sciences*, 13, 4, pp. 707-784 [<http://www.bbsonline.org/documents/a/00/00/04/99/index.html>].
- Pinker (S.) & Jackendoff (R.). 2005. « The faculty of language : what's special about it ? ». in : *Cognition*, 95, pp. 201-236.
- Premack (D.) & Premack (A.-J.). 1983. *L'esprit de Sarah*. Paris : Fayard [reed. 1984].
- Savage-Rumbaugh (E.-S.) & Lewin (R.). 1994. *Kanzi : the ape at the brink of the human mind*. New York : John Wiley & Sons.
- Talmy, (L.). 2000. *Toward a Cognitive Semantics : Concept Structuring Systems*. Cambridge, MA : MIT Press.
- Van Valin (R.-D.) & LaPolla (R.-J.). 1997. *Syntax. Structure, meaning and function*. Cambridge University Press.



Mai 2006

Dans un précédent ouvrage, l'un d'entre nous (Chapouthier, 2001) avait développé une thèse visant à décrire le fonctionnement des êtres vivants à l'aide de la métaphore de la mosaïque. En tant qu'œuvre d'art, on le sait, une mosaïque est une représentation, figurée ou abstraite, où la perception de l'ensemble résulte de la juxtaposition, puis de l'intégration par le système visuel, de petits éléments colorés, qui conservent cependant leur individualité et leurs caractéristiques propres de couleur ou de forme. En un sens plus philosophique, le terme de « mosaïque » peut être appliqué, d'une manière très générale, à tout ensemble où les propriétés globales du « tout » laissent une autonomie aux propriétés des parties qui le constituent. C'est dans ce sens philosophique que nous voudrions appliquer le concept de structure en mosaïque au langage.

Dans un premier temps, nous rappellerons brièvement comment cette image de la mosaïque permet de décrire de manière adéquate les êtres vivants dans leur anatomie comme dans certaines de leurs fonctions essentielles. Dans un second temps, nous voudrions plaider ici pour l'utilisation de ce même modèle en linguistique.

1. Les êtres vivants : des systèmes en mosaïques

À tous les niveaux des organismes vivants, on peut mettre en évidence des structures en mosaïques, au sens qui vient d'être défini, c'est-à-dire des ensembles où les propriétés du tout laissent une autonomie aux propriétés des parties qui le composent.

Au niveau des gènes, on sait que les gènes complexes de mammifères sont des ensembles composés d'« introns » silencieux (ils n'ont pas d'activité biologique immédiate) et d'« exons », responsables du métabolisme, c'est-à-dire de la vie de la cellule. Ces ensembles intégrés laissent cependant une certaine autonomie à leurs parties (les exons et les introns), au point que les théoriciens modernes de l'évolution (Ohno, 1970) pensent que des introns silencieux identiques se multiplient par juxtaposition (appelée duplication) durant les temps géologiques, pour ensuite se différencier les uns des autres par des mutations, intégrer leur fonctionnement et « émerger » comme ensembles d'exons efficaces, amenant à la genèse d'organes ou de systèmes nouveaux.

Au niveau de l'anatomie des animaux, chaque niveau d'organisation peut être considéré comme une mosaïque laissant une large autonomie aux niveaux sous-jacents (Chapouthier, 2001, 2003). Ainsi la cellule peut être considérée comme une mosaïque d'organites, l'organe comme une mosaïque de cellules, l'organisme comme une mosaïque d'organes, laissant leur part respective de causalité aux parties et au tout. De même les sociétés organisées, comme celles des abeilles ou des primates, sont des mosaïques d'individus, gardant leur autonomie (qui dans le cas des sociétés humaines est synonyme de leur liberté) tout en maintenant le fonctionnement d'ensemble de la société. Schématiquement le monde vivant est un ensemble de mosaïques à étages.

Nous avons montré que les mêmes « structures en mosaïque » se retrouvent dans deux des constructions les plus complexes du monde vivant : *l'anatomie du cerveau humain* et *la pensée* qui lui est liée (nous ne nous pencherons pas ici sur le problème philosophique des relations entre cerveau et pensée). L'encéphale humain est composé de différentes parties (fondamentalement cinq), qui gardent des spécificités de fonctionnement, tout en contribuant au fonctionnement harmonieux de l'ensemble. Il en est de même des innombrables aires qui remplissent le cortex cérébral (aires de la vision, de l'audition, de la compréhension du langage oral, de l'expression orale, de la lecture, de l'écriture...). Il en est de même des deux hémisphères cérébraux (à ceci près qu'ici la mosaïque se limite à deux éléments). Dans tous ces cas, le fonctionnement de l'ensemble laisse une large part d'autonomie au fonctionnement de ses parties.

Enfin, comme nous l'avions exposé (Chapouthier, 2001), deux aspects essentiels de la pensée peuvent être appréhendés comme des structures en mosaïques : la conscience et la mémoire. Malgré l'impression que nous en avons, la *conscience* est en fait, à la fois, tout et parties. Les parties peuvent se disjoindre, par exemple, dans des cas pathologiques, chez les sujets dont les deux hémisphères sont dissociés (sujets dits « split-brain »). Chez eux, deux consciences cohabitent dans le même cerveau. Des phénomènes plus « normaux », comme le rêve, peuvent aussi donner une idée d'une autre forme de conscience, capable de cohabiter avec ce qui nous paraît une conscience unique, la conscience du sujet éveillé. Quant à la *mémoire*, qui, elle aussi, nous semble unitaire, on peut montrer qu'elle est un véritable « patchwork » de capacités mnésiques très différentes, acquises par nos ancêtres animaux au cours de l'évolution : habituation, conditionnements, mémoire de l'espace, mémoire cognitive..., capacités qui, toutes, participent de l'ensemble de ce que nous appelons « la mémoire », tout en conservant, chacune, des spécificités de fonctionnement.

Comment se sont formées de telles mosaïques ? En général par l'application répétée de deux grands principes qui traversent toute l'évolution des espèces : tout d'abord la *juxtaposition* d'entités identiques n'ayant pas d'interaction entre elles (introns, cellules, morceaux d'organismes, organismes...), puis *l'intégration* de tels ensembles pour constituer des systèmes où les entités, jadis simplement juxtaposées, acquièrent des interactions entre elles. Bien entendu, dans les systèmes les plus complexes, comme la conscience ou la mémoire, il n'est pas toujours aisé de retrouver un stade de juxtaposition, qui est en revanche évident dans l'analyse des systèmes plus simples (Chapouthier, 2001). Même si on peut y distinguer des parties disjointes, des fonctions comme la conscience ou la mémoire apparaissent comme des mosaïques déjà fortement intégrées.

Venons-en au langage.

2. Application au langage

Lorsque je prononce ou que je lis une phrase, c'est-à-dire lorsque je prononce ou que je lis successivement les différentes unités sémantiques qui la constituent (et qui ont ici le statut de « parties »), la signification générale du message ne se développe que petit à petit, avec parfois des renversements du sens lors de l'arrivée successive de nouveaux éléments sémantiques. Le « tout », c'est-à-dire ici le sens global du message apparaît donc comme une mosaïque de ses parties (les unités successives) qui conservent cependant une certaine autonomie de sens si on les prend indépendamment.

Plus précisément encore, on peut dire que, dans son fonctionnement, le langage présente lui aussi cette double propriété qui caractérise la mosaïque, la juxtaposition et l'intégration des parties dans un niveau supérieur de complexité.

2.1. Juxtaposition formelle et linéarité dans le langage

Le langage est en effet constitué d'unités discrètes qui vont constituer les parties de l'énoncé. Nous ne parlerons ici que des unités de rang supérieur, les morphèmes lexicaux et grammaticaux, même si l'on peut éventuellement parler aussi de mosaïque à propos des phonèmes, ces unités qui s'intègrent en un niveau supérieur pour constituer les morphèmes. Dans le fonctionnement du langage, il y a bien juxtaposition formelle des unités pour construire l'énoncé. Cette juxtaposition présente cependant une particularité, celle d'être linéaire et séquentielle. Dans la parole, la chaîne des unités est prise dans le temps et celles-ci n'apparaissent jamais simultanément. Dans l'écriture, en revanche, les parties coexistent dans l'espace, mais continuent à s'enchaîner séquentiellement. Certains poètes, comme Apollinaire dans ses *Calligrammes*, ont certes joué avec l'utilisation de l'espace pour dépasser la simple forme linéaire mais, même dans ce cas particulier, du point de vue fonctionnel, l'enchaînement des unités qui constituent l'énoncé reste séquentiel. Or, au fur et à mesure du déroulement de l'énoncé, ces parties, formellement juxtaposées, sont constamment soumises à des processus d'intégration, formels (par la construction de relations syntaxiques au sein de l'énoncé) mais surtout sémantiques ; ces processus d'intégration des parties permettent de construire un niveau supérieur de complexité, celui du sens de l'énoncé. Les unités constituant l'énoncé participent ainsi à la construction de l'énoncé tout en gardant une certaine autonomie, d'une part parce que, en tant qu'unités du lexique, elles existent indépendamment de leur mise en énoncé et constituent des parties individuées avec leurs valeurs propres, d'autre part parce que, au sein de l'énoncé, le sens du tout ne se réduit pas à l'addition du sens de ses parties.

Du fait des propriétés de l'antéposition de l'adjectif en français, par exemple, *un bel imbécile* n'est pas un individu qui est à la fois beau et stupide, c'est un individu particulièrement idiot. Il en va de même pour les mots composés : leur signification ne se réduit pas à la somme de leurs parties ; une *pomme de terre* n'est pas une variété de pomme (voir notamment Corbin (1992) sur la morphologie compositionnelle). Précisément parce qu'au sein de l'énoncé, il se produit un travail particulier d'intégration des unités. Ajoutons enfin que, comme dans les autres structures en mosaïque, dans le langage, l'intégration des parties en une structure de niveau supérieur se fait par spécialisation de chacune des parties dans des fonctions spécifiques : il s'agit ici des fonctions syntaxiques que remplissent les unités dans un énoncé donné.

2.2. L'intégration à l'œuvre dans le langage

Quels sont les rapports entre les parties et le tout dans le langage ? Comment se fait l'intégration des unités dans l'énoncé ? Nous allons évoquer rapidement les principaux mécanismes à l'œuvre dans la construction du sens de l'énoncé. Ceux-ci sont bien connus des linguistes, mais ils méritent qu'on y jette un regard extérieur en se posant la question des modalités d'intégration des parties de cette mosaïque particulière. Le principal processus d'intégration réside dans l'intégration des unités en une structure prédicative dans laquelle elles acquièrent des fonctions syntaxiques et sont utilisées pour construire une assertion. En outre, deux phénomènes intéressants nous semblent caractériser cette construction. D'une part, les unités présentent au départ une surcharge sémantique potentielle qui va être réduite au fur et à mesure de la construction de l'énoncé ; d'autre part, le niveau de l'énoncé se caractérise par une non-linéarité des effets de sens. La construction du sens dans l'énoncé est en effet marquée par différents processus non linéaires, propres à l'énonciation, qui vont permettre la spécification progressive du sens de l'énoncé et la stabilisation du sens des unités. On se contentera ici d'en évoquer les principaux mécanismes (pour plus de détails, voir notamment Robert, 1997, 2003 et sous presse).

Quelques mots d'abord sur la surcharge sémantique initiale des unités. Hors contexte, les unités de la langue sont polysémiques : la polysémie est un phénomène général dans les langues qui constitue vraisemblablement un processus d'optimisation du système en permettant à un nombre minimum d'unités de référer à un maximum de choses. Rares sont en effet les unités qui n'ont qu'un seul sens, y compris pour les termes de la vie quotidienne. Ainsi, *un chapeau* est bien entendu un couvre-chef mais c'est aussi l'exergue d'un article, *une règle* est un instrument pour tirer des traits mais c'est aussi une loi, *une baignoire* est un élément de salle de bain mais c'est aussi une loge dans un théâtre. Pour reprendre le bel exemple étudié par Corbin et Temple (1994), *un bleu* en français peut désigner un hématome, un fromage, une tenue de travail aussi bien qu'une jeune recrue, entités qui, à un titre ou un autre, présentent toutes de manière typique la propriété d'être bleues. L'unité linguistique donne ainsi accès à un réseau de valeurs référentielles qui sont reliées entre elles.

On n'étudiera pas ici la nature des relations entre les différents sens du terme (dont les principales sont la métaphore, la métonymie et les relations de schématisation). On soulignera simplement que ce réseau de relations est structuré de manière complexe, pluridimensionnelle, car les relations paradigmatiques entre mots ou entre syntagmes contribuent sans aucun doute à la construction de valeurs de sens diverses et à l'établissement de relations multiples entre les mots et les divers sens des mots (voir le paradigme des couleurs en politique *un blanc* : un royaliste, *un bleu* : un républicain, *un rouge* : un communiste, *un vert* : un écologiste, mais aussi l'opposition entre *un petit blanc* et *un petit rouge*...). La polysémie initiale des unités constitue donc une propriété spécifique des parties qui ne se conserve pas lors du processus d'intégration dans l'énoncé.

En effet, la mise en énoncé produit un frayage contextuel qui permet de construire le cadre de référence. Ce frayage contextuel active l'une ou l'autre des valeurs de sens du terme en jeu, réduisant la polysémie initiale. Ainsi, au fur et à mesure de la construction de l'énoncé, le sens des unités est de plus en plus spécifié, si bien que la fin de l'énoncé est de plus en plus prévisible. La construction du sens de l'énoncé procède de toute une série de mises en relation entre divers composants, mais que l'on peut décrire par un mécanisme fondamental : le « repérage » (Culioli, 1990, 1999). Le repérage est une opération élémentaire et dissymétrique de mise en relation de deux éléments, par laquelle un élément (le repère) est pris comme point d'ancrage pour la localisation (abstraite) d'un autre (le repéré). Cette opération, qui sert à construire la détermination, intervient à différents niveaux structurels et se joue entre des

éléments de nature variée. Chaque mise en relation produit un travail sur les unités en jeu et contribue à la spécification du sens de celles-ci et à la construction du sens de l'énoncé.

Tous les facteurs, contextuels, lexicaux et grammaticaux, interviennent dans la construction progressive du sens de l'énoncé et dans la spécification de la valeur des termes. En effet, lorsqu'elle apparaît dans l'énoncé, une unité est mise en relation, de manière concomitante, avec des éléments de différents niveaux : relation avec un contexte verbal et situationnel préalable, relation avec d'autres éléments du lexique, relation avec des structures syntaxiques. Tout est mise en relation dans le langage et ces mécanismes de mise en relation produisent du sens par une interaction constante entre les éléments en jeu. La mise en énoncé des mots active par là telle ou telle valeur latente d'un terme et produit un frayage contextuel (elle trace un chemin à l'intérieur de la forêt du sens). Toutes les valeurs d'un terme ne sont donc pas présentes dans l'énonciation et les spécifications contextuelles sont telles que le sens prototypique d'un terme n'est pas nécessairement le plus probable.

La situation d'énonciation (au sens large de contexte verbal et situationnel) apparaît comme le repère fondamental et le premier facteur de spécification du sens des unités auxquelles elle sert de point d'ancrage. Ainsi, selon que vous êtes chez le boulanger ou le luthier, le terme de *flûte* sera *a priori* rattaché au domaine sémantique construit par la situation où vous vous trouvez et référera de ce fait soit à une baguette de pain, soit à un instrument de musique. La situation d'énonciation fonctionne donc comme « l'attracteur de sens » par défaut : elle induit un domaine de référence auquel les termes utilisés tendront à être rattachés. Le contexte peut également induire une variation de sens au niveau grammatical qui bouleverse le sens de tout l'énoncé : ainsi, la valeur grammaticale de *vous* dans *je vous coupe la tête ?* est bénéfactive et signifie « pour vous » si vous êtes chez le poissonnier, et attributive signifiant « à vous », si vous êtes sous la menace d'un fou.

À partir du repère cardinal de l'énonciation, différentes mises en relation vont se construire à l'intérieur de l'énoncé, qui vont également contribuer à la spécification de la valeur des termes et à la construction progressive du sens de l'énoncé. Celles-ci mettent en jeu tous les éléments, contextuels, lexicaux, grammaticaux, qui interagissent en permanence au cours de l'énonciation. Ainsi, le contexte verbal va contribuer à spécifier le sens des termes en définissant un champ sémantique auquel les nouvelles unités tendront à être rattachées : *une monture* ne référera pas à la même chose selon qu'on parle d'équitation ou de lunettes ; cette articulation au contexte peut également restructurer de manière variable des groupes entiers d'unités : selon que l'on parle de vacances ou d'architecture, c'est toute la relation prédicative *il a fait le pont* qui va changer de sens.

Le rôle des relations syntaxiques dans la construction du sens de l'énoncé est bien connu : la place et les règles d'agencement propres à la langue vont définir la fonction syntaxique des différents termes dans l'énoncé. Mais ces règles syntaxiques vont aussi contribuer à en spécifier le sens particulier des termes. Pour ne citer qu'un exemple, l'ordre des mots joue déjà un rôle dans la spécification du sens des unités, comme en atteste l'opposition entre *un homme grand* (valeur physique de l'adjectif) et *un grand homme* (valeur appréciative de l'adjectif) ; la place de l'adjectif en français a donc un rôle sémantique qui en contraint le jeu (cf. **la verte herbe / l'herbe verte*) et produit des glissements de sens.

Enfin, la mise en relation d'une unité avec une autre contribue aussi à en spécifier le sens. Ainsi, le terme *tuyau* dans *filer un tuyau* a peu de chance d'être compris dans son sens propre (« tube ») plutôt que dans son sens figuré (« information »). Ces mises en relation ne constituent pas un simple filtrage parmi les valeurs possibles d'une unité, mais elles produisent un véritable « travail » sur le sens des termes qui se construit par une interaction. Ainsi, le mot *tendre* prend un sens différent suivant qu'il s'applique à un *homme tendre* ou à un *steack tendre*. Dans l'expression *les décisions de la Maison Blanche*, la mise en relation de *Maison Blanche* avec un nom d'action contraint l'interprétation métonymique de *Maison Blanche* (le lieu pour l'institution qu'il héberge).

Ces mises en relation peuvent avoir des effets rétroactifs. Ainsi l'introduction de l'adjectif *rouillé* dans un *pied-de-biche rouillé* produit une restructuration du syntagme complétif en un nom composé désignant un instrument en métal. De même pour *une gueule-de-loup fanée*, un *œil-de-perdrix douloureux*... L'articulation au contexte déclenche l'interprétation métaphorique de certaines expressions. Ainsi le sens de la phrase *casser sa pipe* dépend du contexte dans lequel elle est employée qui lui donne un sens propre ou un sens figuré dans lequel la significa-

tion du terme n'est pas compositionnelle et le procédé d'intégration encore plus radical. Ce phénomène s'applique aussi aux constructions grammaticales. Dans une phrase comme *le policier observait l'homme à la jumelle*, le complément à *la jumelle* peut être rattaché au prédicat (*observait*) ou à l'objet (*homme*) ; l'interprétation instrumentale ou associative de ce complément dépend du contexte sémantique et pragmatique dans lequel la phrase est employée.

Cette relation entre les termes d'un énoncé est notamment régulée par un mécanisme fondamental « d'isotopie sémantique » (Greimas, 1966, Rastier, 1987) qui consiste à rattacher le sens d'un terme à l'univers sémantique de celui qui précède pour construire un fil conducteur interprétatif. C'est pourquoi ce n'est pas le sens prototypique de *chapeau* (« couvreur-chef »), de loin le plus courant, qui est convoqué dans l'exemple suivant emprunté à Victorri (1997), mais son sens beaucoup plus rare de « début d'article » :

Il défendait avec assurance cet article, alors qu'il n'en avait parcouru que les quelques lignes qui constituaient son chapeau

Ces processus d'intégration jouent constamment au cours de l'énonciation et rendent la fin de l'énoncé de plus en plus prévisible. Grâce au frayage contextuel, il y a donc dans le processus d'énonciation une prévisibilité de plus en plus grande de l'information qui va suivre. Mais celle-ci n'est pas absolue.

D'une part en effet, il y a toujours possibilité de rupture du fil conducteur, l'introduction d'un terme inattendu à la fin de l'énoncé provoquant alors un effet de surprise dont le poids informationnel est d'autant plus grand qu'il arrive à une place inattendue parce que chargée des spécifications du frayage contextuel (*en avril ne te découvre pas d'un Dim*). C'est là l'un des ressorts de la poésie, de la rhétorique mais aussi de la publicité, comme l'a montré B.-N. Grunig (1990) à qui cet exemple est emprunté. La fin de l'énoncé est ainsi le lieu privilégié de ce qu'on a appelé les « bombes sémantiques » dont les effets ne sont pas additionnels car ils induisent des phénomènes de restructuration, de résonance, de diffusion et de stratification du sens (Robert, 2003).

Ces effets non linéaires tiennent à la fois à des facteurs structurels, l'existence d'unités à portée énonciative, et au caractère permanent des mises en relation au sein de l'énoncé qui permet une rétroaction parfois vertigineuse. Les unités dites « énonciatives » ont une portée sémantique qui dépasse leur portée syntaxique : leur sens porte sur tout l'énoncé et non sur le seul terme dont elles dépendent grammaticalement, comme par exemple la négation, les adverbes *probablement* et *heureusement* ou l'adjectif *admirable*. Ainsi dans *Jean est admirable de travailler ainsi*, du point de vue syntaxique l'adjectif *admirable* est attribut du sujet, mais du point de vue sémantique, il exprime un jugement du locuteur sur l'ensemble de la relation prédicative <*Jean, travailler*>. L'intonation elle aussi entraîne des bouleversements sémantiques qui ne sont pas proportionnels à son poids formel puisqu'elle permet de transformer une assertion en interrogation ou en exclamation et donc d'en bouleverser le sens. Tant que l'énoncé n'est pas achevé les rétroactions et les bouleversements de sens sont possibles, comme dans le slogan suivant d'un garagiste : *Vous pouvez tous crever... Le garage X vous regonflera*.

La poésie joue parfois des restructurations permises par les mises en relation nouvelles qui se produisent constamment au cours de l'enchaînement. Ainsi le haïkou (orthographe volontairement francisée) suivant de Jean Monod (dans Antonini, 2003) :

L'absente de tout
bouquet la voilà me dit
en se montrant l'aube.

On considérera ici les trois versets du haïkou comme des parties du poème, même si celles-ci pourraient, à leur tour, être décomposés, comme des « tous », en éléments sémantiques plus simples. Si l'on s'arrête au premier verset, on pense immédiatement à une femme « absente de tout ». Dès le second verset, puisqu'il s'agit d'un bouquet, on songe que l'absente est sans doute une fleur. Seul le troisième verset donne son sens à l'ensemble : il s'agit en fait de l'aube, qui, cependant, laisse leur autonomie et leur pertinence poétique aux éléments qui l'ont précédée : l'aube garde aussi, en elle, le sourire de la fleur et le parfum de la femme. Mosaïque s'il en est, ce haïkou traduit parfaitement, au sein du langage, la dialectique du tout et des parties, qui avait été décrite pour l'anatomie et le fonctionnement des organismes vivants.

3. Un parallélisme avec le fonctionnement du vivant

On peut même poursuivre le parallélisme avec les structures du vivant en posant qu'il a existé, pour le langage comme pour le vivant, un stade préliminaire de « juxtaposition » des unités. C'est du moins ce que propose Talmy Givón (1998) dans sa distinction entre mode de communication « grammatical » et mode « pré-grammatical ». Selon cet auteur, la langue permet deux modes de communication qui correspondent à deux stades de la phylogénèse mais qui sont aussi à l'œuvre dans certaines situations linguistiques particulières et, avec une importance inverse, dans la communication normale. Le premier mode de communication, le mode prégrammatical qui nous intéresse et qui est supposé avoir pré-existé au mode grammatical dans l'émergence du langage, apparaît à un certain stade de l'acquisition, chez les aphasiques agrammaticaux et aussi dans la communication en pidgin, c'est-à-dire lorsque deux communautés parlant des langues différentes forment une langue de communication basique destinée à répondre à des besoins de communication limités. Dans ce mode de communication, il n'y a pas de morphèmes grammaticaux, la syntaxe est simple et procède par juxtaposition, l'ordre des mots a une valeur pragmatique (et non syntaxique) et les pauses sont nombreuses. Du fait qu'il n'y a pas de grammaire, la nature des relations entre les éléments du lexique n'est pas explicite (on pense à des énoncés d'enfants du type *mémé*, *minou*, *salon*). De fait, les similitudes entre les exemples donnés de pidgin, d'énoncés de patients agrammaticaux ou de récits d'enfant sont frappantes. Il s'agit de simple juxtaposition d'unités à valeur pragmatique. Ce mode de communication est lent, analytique, coûteux cognitivement et sujet à un fort taux d'erreur précisément parce que la nature des relations et la fonction des unités dans l'énoncé ne sont pas codées : il s'agit bien d'un stade de simple juxtaposition. Givón l'oppose au mode grammatical qui est rendu possible par la structure des langues où la morphologie grammaticale est abondante, les constructions syntaxiques complexes et où l'ordre des mots correspond à des fonctions syntaxiques. Ce mode de communication qui correspond selon nous au stade de la mosaïque, est rapide, économique, automatisé mais acquis tardivement.

4. Conclusion

Beaucoup certes reste à connaître sur les origines d'une fonction aussi complexe que le langage humain. Parmi les multiples origines, biologiques ou culturelles, le présent exposé apporte un certain éclairage sur ce qui pourrait être une des facettes proprement biologiques de l'origine du langage.

Nous avons, en effet, défendu l'idée que les êtres vivants sont des mosaïques à étages. Nous avons défendu l'idée qu'ils ont une structure en mosaïque à des niveaux très variés de leur organisation, depuis les gènes ou les cellules, jusqu'à l'anatomie de l'encéphale ou même l'organisation de grandes fonctions mentales, comme la conscience ou la mémoire. Force nous est de constater, à la suite du présent exposé, que cette structure en mosaïque peut également être retrouvée dans l'analyse du langage.

Si on le compare avec les modes de fonctionnements des êtres vivants, le langage possède donc de remarquables similitudes avec le reste des processus du vivant.

Références bibliographiques

- Antonini (J.), (ed.). 2003. *Anthologie du haïku en France*. Lyon, France : Éditions Aléas.
- Chapouthier (G.). 2001. *L'homme, ce singe en mosaïque*. Paris : Odile Jacob.
- Chapouthier (G.). 2003. « L'évolution de la complexité ». in : *Pour la Science*, 314, pp 78-81.
- Corbin (D.). 1992. « Sens et définition : de la compositionnalité du sens des mots construits (Réponse à Claire Vanderhoeft) ». in : *Linguisticae Investigationes*, XVI, 1, pp. 189-218.
- Corbin (D.) & Temple (M.). 1994. « Le monde des mots et des sens construits : catégories sémantiques, catégories référentielles ». in : *Cahiers de lexicologie*, 65, 2, pp. 5-28.
- Culioli (A.). 1990. *Pour une linguistique de l'énonciation*. Paris, Gap : Ophrys [tome I]
- Culioli (A.). 1999. *Pour une linguistique de l'énonciation*. Paris, Gap : Ophrys [tome II]
- Givón (T.). 1998. « The Functional Approach to Grammar ». in : Tomasello (M.), (ed.). *The new psychology of language*. London : Erlbaum.
- Greimas (A.-J.). 1966. *Sémantique structurale*. Paris : Larousse.
- Grunig (B.-N.). 1990. *Les mots de la publicité*. Paris : Presses du CNRS.
- Ohno (S.). 1970. *Evolution by Gene Duplication*. Munich : Springer Verlag.
- Rastier (F.). 1987. *Sémantique interprétative*. Paris : Presses Universitaires de France.
- Robert (S.). 1997. « Variations des représentations linguistiques : des unités à l'énoncé ». in : Fuchs (C.) & Robert (S.), (eds.). pp. 25-39.
- Robert (S.). 2003. « L'épaisseur du langage et la linéarité de l'énoncé : vers un modèle énonciatif de production ». in : Ouattara (A.), (ed.). *Parcours énonciatifs et parcours interprétatifs - Théories et applications*. Paris, Gap : Ophrys, pp. 255-274.
- Robert (S.). [sous presse]. « Des mots dans tous les sens : facteurs de variation et mécanismes de stabilisation dans le langage ». in : Létoublon (F.), (ed.). *Les langages scientifiques*. Paris : Belin.
- Victorri (B.). 1997. « La polysémie : un artefact de la linguistique ? ». in : *Revue de sémantique et pragmatique*, 2, pp. 41-62.



The Mosaic of Language

By Stéphane Robert, CNRS-LLACAN & INALCO, France
and Georges Chapouthier, CNRS
UMR 7593, University of Paris 6, France

May 2006

In a previous book (Chapouthier, 2001), one of the authors developed a thesis describing living beings by using the metaphor of the mosaic. A mosaic, as a work of art, is a representation, either figurative or abstract, and the perception of the total image as a whole is the result of the juxtaposition of small coloured elements, and their subsequent integration when perceived visually, but with the small units maintaining their individual status and characteristics of colour and shape. In a more philosophical sense, the term « mosaic » can, in general, be applied to any structure where the overall properties of the whole do not destroy the autonomy of the properties of the component parts. This philosophical approach will be used to apply the concept of the mosaic structure to language.

A brief review will first show how the mosaic metaphor can provide an adequate description of living beings, covering anatomy and certain essential functions. Secondly, we shall present arguments showing the relevance and usefulness of this same model for linguistics.

1. Living Beings : Mosaic Systems

Mosaic structures can be observed in living organisms and at every level of the structure, as defined above, i.e. structures where the properties of the whole do not cancel out the autonomy of the component parts.

Complex mammalian genes are comprised of silent introns (with no immediate biological activity) and exons actively involved in cell metabolism, i.e. the life of the cell. The integrated structures (the genes) concede a degree of autonomy to their component parts (the exons and introns). In fact, modern theoreticians in the field of evolution (Ohno, 1970) claim that identical silent introns were produced through juxtaposition (referred to as « duplication ») during geological periods, and then differentiated through mutation, integrating their role and appearing as groups of active exons involved in producing new organs and new biological systems.

In animal anatomy, each level of organisation can be considered as a mosaic, with a large degree of autonomy at the lower levels (Chapouthier, 2001, 2003). A cell can thus be seen as a mosaic of organites, an organ as a mosaic of cells, an organism as a mosaic of organs, where each imparts its respective share of determinism to the parts and to the whole. Similarly, organised societies, e.g. bees and primates, are mosaics of individuals that maintain their autonomy (and in human societies, this autonomy means freedom) ; but at the same time they support the overall functioning of society. In a simple pattern, the living world is a set of mosaics on different scales and levels.

We have shown that similar mosaic structures can be used to describe two of the most complex constructions in the living world : *the anatomy of the human brain* and *human thought* ; (the philosophical brain-mind question will not be debated here). The human encephalon is comprised of different parts, basically five, each retaining specific functional features, while also contributing to the harmonious functioning of the whole. The same applies to the large number of areas comprising the cerebral cortex (visual and auditory areas, plus areas devoted to the understanding of oral language, oral expression, reading and writing etc.). The same organisation is found in the two cerebral hemispheres (the difference being that the mosaic has only two pieces). In all cases, the functioning of the whole allows for a large degree of autonomy in the functioning of the component parts.

As was previously shown (Chapouthier, 2001), two essential components of thought can be seen as mosaic structures : consciousness and memory. Contrary to our own personal impressions, *consciousness* is both the whole and the combined component parts. The parts

can be separated, e.g. in pathological conditions affecting humans where the two hemispheres are disconnected. In these subjects, known as « split-brain », two consciousnesses prevail in one brain. More common phenomena, such as dream states, can also produce another form of consciousness that can coexist with what appears to be a single consciousness. In *memory*, a patchwork of highly contrasting mnemonic capacities can be seen ; this is the heritage acquired via our animal ancestors in the course of evolution, through habituation, conditioning, spatial memory, cognitive memory and so on. All these abilities play a role in the whole that we call « memory », while retaining the features specific to their own functioning.

How did these mosaics emerge ? In general, through the repeated application of the two main principles observed throughout the evolution of the species : first *juxtaposition* of identical entities, without any interaction occurring (introns, cells, parts of organisms, organisms etc.), then *integration* of these entities, leading onto the construction of systems where the entities, once simply juxtaposed, reach a level of interaction. In the most complex systems, such as consciousness or memory, it is always difficult to detect the juxtaposition stage which is so obvious in the analysis of simple systems (Chapouthier, 2001). The separate parts of consciousness and memory are not easy to distinguish, as their functions can be seen as strongly integrated mosaics.

2. Application to Language

A person pronouncing or reading a sentence, pronounces or reads the different semantic units in order ; these are seen here as the « parts ». The overall meaning of the message emerges gradually and may undergo some reversal of the meaning when new units are included. Thus the « whole », i.e. the overall meaning of the message, appears as a mosaic of its parts, the successive linguistic units, but these still maintain a certain degree of semantic autonomy when considered individually.

Language in operation can be seen as displaying the dual characteristics of mosaics : the juxtaposition and then integration of its parts to produce a higher level of complexity.

2.1. Formal Juxtaposition and Linearity in Language

Language is made of discrete units forming the parts of the utterance. Here we are only referring to higher level units, i.e. lexical and grammatical morphemes, for while we may also see mosaics of phonemes : phonemes are integrated into higher level units to produce morphemes. Language in operation involves formal juxtaposition of units to build an utterance, and the juxtaposition is both linear and sequential. In speech, the chain of units is produced over time, never concurrently, and while written language has parts coexisting in space, the links are still sequential. A number of poets, for example Apollinaire in his *Calligrammes*, have used spatial effects to surmount the constraints of linearity, but even then, from a functional point of view, the organisation of the units comprising the sentences remains sequential.

During the gradual unfolding of the utterance, these parts, while formally juxtaposed, undergo constant processes of integration into the overall meaning of the utterance ; and these processes are formal (e.g. the construction of the syntactic relations within the utterance) and mostly semantic. The integration of the parts produces a higher level of complexity which is the meaning of the utterance. The units comprising the utterance are the building blocks that form the utterance, but they maintain a certain degree of autonomy, firstly because they exist independently of their use in the utterance, as individual parts with their own values, and secondly because the meaning of the whole utterance cannot be reduced to the sum of the meaning of the component parts. For example, in French, the positioning of an adjective before the noun in *un bel imbecile* (a fine idiot) can change the meaning of the adjective : the expression *un bel imbecile* does not refer to a person who is both stupid and handsome. The same holds true for the compound nouns : their overall meaning can not be reduced to the sum of their parts ; for instance a *pomme-de-terre* (a potatoe, lit. « an apple-of-the earth ») is not a variety of apple (see Corbin (1992) on compositional morphology).

This is because specific integration processes operate within the utterance. In language, as is the case with other mosaic structures, the integration of the parts at a higher level is achieved through the specialization of each part with specific functions, as in the example cited where the units in the utterance have played specific syntactic roles.

2.2. The Integration Process at Work in Language

What are the relationships between the parts and the whole in language ? How are the units integrated into the utterance ? We will briefly review the main mechanisms at work in the construction of meaningful utterances ; these mechanisms are familiar to linguists but need to be considered from an external point of view to see how the parts of a given mosaic are integrated as a whole. The most important phenomenon of integration in language relies in the integration of units in a predicative structure where they acquire syntactic functions and are used to build up an assertion. Moreover, we have observed two interesting features characterizing this construction : first, due to the polysemy of languages, the linguistic units present a potential semantic overload which will be reduced as the utterance is constructed ; and secondly, the utterance level has structural features that shape the meaning in non-linear ways. The overall meaning of a sentence is built through different non-linear processes operating at the level of the utterance, gradually specifying the meaning of the utterance and stabilizing the meaning of the separate units. Only the main processes are reported here (for further details, see *inter alia*, Robert, 1997, 2003 and in press).

First there is the initial semantic overload of the units. In isolation, language units are polysemous, this being a widespread phenomenon in languages, and a facility for optimizing a system as a minimal number of units can refer to a maximum number of entities. Units with only one meaning are extremely rare, even in everyday life ; for example, the French word *chapeau* refers to a hat but also means the introductory paragraph of an article ; *règle* is a rule, an instrument for drawing lines, but it also means a rule or regulation ; *baignoire* is a bathtub, but also denotes a private box in a theatre. A fine analysis by Corbin & Temple (1994) cites the case of « *un bleu* » (blue one) which can mean a bruise, blue cheese, dungarees or even a new recruit : all these entities have, in different ways, the common property of « being blue ». A linguistic unit thus offers access to a network of referential values that are related in one way or another. The present study will not investigate the way the different meanings of a given term are related (mostly by metaphor, metonymy and schematization).

We will only mention that this network of relationships is a complex and multidimensional structure where the paradigmatic relationships between words and phrases contribute to create the various meanings and help build multiple relationships between words and their different meanings ; e.g. the colour paradigm in French politics : a *blanc* (white) being a royalist, a *bleu* (blue) a republican, a *rouge* (red) a communist, and a *vert* (green) an ecologist, yet the red-white distinction is totally different when referring to wine. The initial polysemy of a given unit is a specific property of that unit, but disappears once it is integrated into an utterance.

Using words in an utterance sets up a contextual linkage, creating in turn a frame of reference within which the potential semantic overload of the units can be dealt with. This linkage activates one or another of the latent values for the given term and reduces its initial polysemy. Thus an utterance is formed gradually, with the meaning of the units becoming more specific and therefore the information still to be given becomes more predictable thanks to a mechanism of « anchoring » : the development of the meaning of an utterance emerges from a series of connections between the different component parts ; this is the common, basic mechanism of « anchoring » (Culioli 1990, 1999). Anchoring is an elementary, asymmetrical operation that links two elements : one element (the locator) acts as an anchoring point to locate (in an abstract sense) the second meaning (the located). This operation helps specify the meaning of the units affecting different structural levels and elements of a different nature. Each anchoring relationship helps specify the meaning of the related elements and build up the meaning of the utterance.

All factors, whether contextual, lexical or grammatical, have a role to play in gradually building up the meaning of the utterance and specifying the significance of the terms. When a unit is part of an utterance, it is simultaneously linked (or related) to elements that are anchoring points at different levels, i.e. a previous verbal or situational context, other lexical elements and syntactic structures. Language is nothing more than relationships and the mechanisms establishing the relationships produce meaning through the constant interaction of the related elements. When a word is used in an utterance, it activates one of the latent values of the term and produces a contextual linkage, clearing a path through the thick forest of meaning. Speech does not trigger all the possible meanings of a given term ; sometimes the context is so specific that the standard meaning of a term is not necessarily the most likely.

When speaking, the situation (i.e. the broad verbal and situational context) provides the basic anchoring point or locator and the first factor specifying the meaning of the units located. Depending on whether a French-speaking person is buying bread or a musical instrument, the word *flûte* will immediately be linked to the semantic domain of the situation, denoting either a small loaf of crusty bread or musical instrument. The speech situation operates as the default « meaning attractor » providing the field of reference for the terms used. The situational context may also introduce variations to the meaning of the grammatical units, completely changing the overall meaning of the utterance : for example, the grammatical meaning of *vous* (you) in the question *Je vous coupe la tête ?* will mean to cut the head off a fish « for you » (benefactive) when at the fishmonger's, and will mean to « cut your head off » (attributive) if threatened by a person wielding an axe.

Starting from the cardinal anchoring point set by the speech situation, various relationships are established within the utterance and then help specify the meaning of the words and gradually build up the meaning of the utterance. All the elements — contextual, lexical and grammatical — are involved in establishing the relationships between the parts of the sentence ; for example, the verbal context will help specify the meaning of the terms by defining the semantic field relevant to the subsequent units. The word *monture* — (1) a mount, and (2) a frame — will change according to the subject being discussed, i.e. horse-riding or eye-glasses. The link to the verbal context can then restructure entire groups of units : when discussing architecture or days off work, the whole predicative relation of the expression « *il a fait le pont* » will change : the architect has literally « made the bridge », while the figurative meaning applies to a person who has taken days off work between a weekend and a public holiday.

The role of syntactic relations in the construction of the meaning of a statement has been clearly established : the position and organisational rules governing a given language determine the syntactic role of the different terms in the statement. But these syntactic rules also help stipulate the specific meaning of the terms. One example can show how word order plays an initial role in specifying the meaning of the units, and that is the contrast in French between *un homme grand* (where the adjective placed after the noun denotes a physically tall man) and *un grand homme* (where the adjective before the noun expresses a judgement : a great man). The position of the adjective in French has semantic force, producing shifts in meaning.

The relationship between one unit and another can also help specify the meaning ; e.g. the noun *tuyau* (a pipe or tube), when used in the expression *filer un tuyau*, is meaningless if understood literally, but makes sense in the figurative expression *filer un tuyau* which means to give a tip or convey information. The context does more than just filter semantic values, it creates its own semantic values through interaction. The word *tendre* (tender) changes meaning depending on whether it refers to a person or a steak. The phrase *White House decisions*, establishes a relationship between the *White House* and an action requiring a metonymic interpretation of the words white and house (the site denoting the institution on the site). Such relationships can have retroactive effects.

One interesting example in French is the term *gueule de loup*, literally the « wolf's mouth » ; but if the adjective « wilted » is used — *gueule de loup fanée* — the adjective retroactively changes the expression from a possessive into a compound noun denoting a flower, a snapdragon. In the same way, the meaning of metaphoric expressions emerges through a reanalysis of the phrase triggered by the contextual linkage. For instance, the same phrase *to kick the bucket* has a different meaning according to the context in which it is used : it can mean « to bowl over the bucket » (literal meaning) or « to die », metaphoric reading in which the overall meaning is not compositional and the integration is more drastic. The same holds true for grammatical constructions. In a sentence like *the magician touched the child with the wand*, the syntactic attachment of the complement clause « with the wand » to the noun « child », or to the predicate « touched », as well as the instrumental vs associative meaning of the preposition « with », can only be specified by the semantic and pragmatic context in which this sentence is used.

This relationship between the terms of an utterance are regulated by a basic mechanism known as the « semantic isotopic » (Greimas, 1966 ; Rastier 1987) whereby the meaning of a term is linked to the semantic field established by what precedes it, developing a guiding

thread for the meaning to be understood. For this reason, the prototype meaning of *chapeau* (hat), which is in common usage, is not the meaning applying in an example cited by Victorri (1997) which is a minor usage where the word *chapeau* denotes a brief heading text introducing an article.

Integration processes operate constantly throughout an utterance and gradually make the end more and more predictable. Through contextual linkage, linguistic communication is such that the information to follow becomes increasingly predictable, but this predictability is never absolute.

First, the main thread can always be broken. Unexpected information introduced at the end of an utterance can thus have a surprise effect : the impact of the information is stronger precisely because it is unexpected in relation to the accumulated specifications of the contextual linkage. One example is the advertisement for « Dim » hosiery : *en avril ne te découvre pas d'un Dim*, based on an alliterative French proverb warning against the sudden return of cold weather in springtime, *en avril ne te découvre pas d'un fil* (in April, don't remove a stitch (of clothing)). This sort of reversal of the information curve is a key element of rhetoric, and also of advertising, as B.-N. Grunig (1990), who cited this example, has shown. Many advertising slogans have an impact by using an unexpected term with high informational value and a clash effect when included in a familiar expression. The end of the utterance is therefore the focal point for what has been called « semantic bombs » ; these do not have simply cumulative effects as they induce restructuring, resonance, diffusion and stratification of the meaning (Robert 2003).

These non-linear effects are related to structural factors, to the existence of utterance-modifying units and to the permanent establishing of relationships within the utterance which makes it possible to have a retroactive effect that can sometimes be quite acrobatic. Utterance-modifying units (« enunciative morphemes ») have semantic scope that goes beyond their syntactic role : their meaning affects the entire utterance and not just the term to which they are bound by grammar, e.g. the use of negation, the adverbs *probably* and *fortunately* or the adjective *admirable*.

Thus, in French *Jean est admirable de travailler ainsi* (John is wonderful to work so (hard)), more idiomatically expressed as « it's wonderful of John to work so (hard) », the adjective *admirable* modifies the syntactic subject *Jean*, but it also expresses a judgment on the part of the agent of the speech act (i.e. the speaker or « enunciator ») concerning the predicative relationship <*Jean, travailler*> as a whole. Intonation also introduces radical semantic changes quite disproportionate to the formal impact, as it can transform a statement into a question or an exclamation and totally change the meaning.

Until the utterance is complete, it is always possible to produce retroactive effects and changes in meaning. A slogan used by a French service station offers an example : *Vous pouvez tous crever... Le garage X vous regonflera*, playing on the double meaning of *crever*, either to get a flat tire or to die, and the double meaning of *regonfler*, to re-inflate or to give someone a new lease on life.

Poetry sometimes works on restructuring produced through new relationships established through the different linkages. The haiku quoted here, by Jean Monod (cited in Antonini, 2003), is one example :

L'absente de tout
bouquet la voilà me dit
en se montrant l'aube.

The three verses of the haiku are parts of the poem, even though each of the three could be broken down in turn and seen as whole units, divided into simpler semantic components. In the first verse, the feminine form *absente de tout*, makes the reader think of a woman who is absent ; in the second verse, the reference is a bunch of flowers, *bouquet*, and as the flower, *fleur*, is feminine in French, the reader then thinks that the *absente* must be a flower ; but then the third verse gives the full meaning as the absent figure is the dawn, *aube*, another feminine noun. The dawn does not destroy the autonomy and poetic elements of the preceding references : the dawn can also retain the image of the flower and of the woman. Here is a fine mosaic and this haiku is a fine expression, in language, of the dialectic between the whole and the component parts — the same mosaic observed with the anatomy and functioning of living beings.

3. A Parallel with Living Organisms

The parallel seen with the structures of living organisms could be pursued by postulating that, in the case of language and of living organisms, there is a preliminary stage where the units are juxtaposed. This hypothesis was proposed by Talmy Givón (1998) who made the distinction between « grammatical » and « pre-grammatical » communication. According to Givón, language offers two means of communication that can be seen as the counterparts of the two stages of phylogenesis, but they are also involved in certain specific linguistic situations, and with reverse importance are involved in normal communication. The first means of communication, and which is of interest here, is the pre-grammatical process thought to exist before the emergence of the grammatical mode of language ; it appears at a certain stage in language acquisition, and is seen with agrammatical aphasic subjects and in « Pidgin » speech, i.e. when two communities speaking different languages develop a language for basic communication designed to meet minimal communication requirements. This channel of communication has no grammatical morphemes ; the syntax is simple, operating through juxtaposition ; the word order is pragmatic (and not syntactic) and there are many pauses. Because there is no grammar, no explicit relationships operate between the different elements of the lexicon ; it could be compared to speech by young children, e.g. *Granny... cat... garden*. The similarities between Pidgin, the utterances of agrammatical subjects and a child's speech are quite striking ; it is simply juxtaposition of units for pragmatic effects. As a means of communication, it is slow and analytical, requiring considerable cognitive input and featuring a high error rate because the nature of the linkages and the function of the units comprising the utterance are not coded : it is clearly a stage of simple juxtaposition. Givón sets this against the grammatical mode which is made possible by using the structure of languages with proper grammatical morphology and complex syntactic constructions, and where the word order relates to syntactic functions. This type of communication, which we see as the mosaic stage, is swift, economical and automated, but is only acquired after a considerable period of time.

4. Conclusion

A great deal is yet to be learnt on the origin of such a complex function as human speech. There are many sources, both biological and cultural, and the present paper has endeavoured to cast light on what could be an intrinsically biological facet of the origin of language.

The theory we have put forth is that living beings are mosaics at different stages, that they have a mosaic structure with highly diverse levels of organisation, ranging from genes or cells to the anatomy of the encephalon and even the organisation of key mental functions, such as consciousness and memory. In the present paper we have found that this mosaic structure is also apparent in language when analysed.

When compared with the workings of living organisms, language presents quite striking similarities with the other processes at work in living organisms.

References

- Antonini (J.), (ed.). 2003. *Anthologie du haïku en France*. Lyon, France : Éditions Aléas.
- Chapouthier (G.). 2001. *L'homme, ce singe en mosaïque*. Paris : Odile Jacob.
- Chapouthier (G.). 2003. « L'évolution de la complexité ». in : *Pour la Science*, 314, pp 78-81.
- Corbin (D.). 1992. « Sens et définition : de la compositionnalité du sens des mots construits (Réponse à Claire Vanderhoeft) ». in : *Linguisticae Investigationes*, XVI, 1, pp. 189-218.
- Corbin (D.) & Temple (M.). 1994. « Le monde des mots et des sens construits : catégories sémantiques, catégories référentielles ». in : *Cahiers de lexicologie*, 65, 2, pp. 5-28.
- Culioli (A.). 1990. *Pour une linguistique de l'énonciation*. Paris, Gap : Ophrys [tome I]
- Culioli (A.). 1999. *Pour une linguistique de l'énonciation*. Paris, Gap : Ophrys [tome II]
- Givón (T.). 1998. « The Functional Approach to Grammar ». in : Tomasello (M.), (ed.). *The new psychology of language*. London : Erlbaum.
- Greimas (A.-J.). 1966. *Sémantique structurale*. Paris : Larousse.
- Grunig (B.-N.). 1990. *Les mots de la publicité*. Paris : Presses du CNRS.
- Ohno (S.). 1970. *Evolution by Gene Duplication*. Munich : Springer Verlag.
- Rastier (F.). 1987. *Sémantique interprétative*. Paris : Presses Universitaires de France.
- Robert (S.). 1997. « Variations des représentations linguistiques : des unités à l'énoncé ». in : Fuchs (C.) & Robert (S.), (eds.). pp. 25-39.
- Robert (S.). 2003. « L'épaisseur du langage et la linéarité de l'énoncé : vers un modèle énonciatif de production ». in : Ouattara (A.), (ed.). *Parcours énonciatifs et parcours interprétatifs - Théories et applications*. Paris, Gap : Ophrys, pp. 255-274.
- Robert (S.). [sous presse]. « Des mots dans tous les sens : facteurs de variation et mécanismes de stabilisation dans le langage ». in : Létoublon (F.), (ed.). *Les langages scientifiques*. Paris : Belin.
- Victorri (B.). 1997. « La polysémie : un artefact de la linguistique ? ». in : *Revue de sémantique et pragmatique*, 2, pp. 41-62.



Language Evolution : What Evolved ?

By John Schumann, University of California, Los Angeles (USA)
Donald Favareau, National University of Singapore (Singapore)
Charles Goodwin, University of California, Los Angeles (USA)
Namhee Lee, University of California, Los Angeles (USA)
Lisa Mikesell, University of California, Los Angeles (USA)
Hongyin Tao, University of California, Los Angeles (USA)
Daniel Véronique, University of Paris III (France)
Alison Wray, Cardiff University (Wales)

May 2006

The study of linguistics is at an interesting juncture. The Chomskian claims about the Language Acquisition Device and Universal Grammar are essentially neurobiological claims and evolutionary biological claims. However, until recently our evolutionary and neurobiological knowledge was insufficient to directly address these areas. Now with what we know about the brain and with what we know about the processes involved in evolution, exploration of the biological foundations of language is much more feasible.

The study of the evolution of language was virtually silenced when it was banned by the Paris Linguistic Society at the end of the 19th century. However, the International Conferences on the Evolution of Language, held every two years since the first one in Edinburgh in 1996, have made a huge difference. Most of this recent research has focused on the sentence as the linguistic unit that must be explained in an evolutionary account of language. But is the sentence or the ability to produce, comprehend, and learn sentences really what evolved ?

On November 12-14, 2004 a Roundtable funded by the journal *Language Learning*, was held at UCLA to discuss the appropriate data on which to base evolutionary and neurobiological accounts of language. This report from the Roundtable consists of two parts. The first is a summary of each paper presented, and second is prospective developed by the convener, John H. Schumann, that suggests a possible position/synthesis on the issue of « What evolved ». All the presenters in various ways may agree and disagree with aspects of this position, but they all agreed that one « possible » synthesis was preferable to a disputable pseudo-consensus resulting from line-item approval by each participant.

As mentioned above, the overall question addressed by the roundtable was « What evolved ? » This question is important to ensure that the targets of evolutionary and neurobiological accounts of language are the appropriate ones. Just as we would not want to base an account of rudimentary human motor activity on the physiology of how an Olympic ice skater does a triple Lutz, we would not want to base a biological account of the origins of language on the kinds of forms that are used under the relaxed psycholinguistic conditions made available by reading and writing. What we want to know is what the brain had to be like for language to emerge and what that initial language may have looked like.

To address the overall question, individual presentations were organized around three more basic questions :

1. How did language evolve ?
 - Lee : *What biological developments are likely to have made language possible ?*
2. What exactly evolved ?
 - Favareau : *What semiotic relations would have to have already been in place in order to make the emergence of human language possible ?*
 - Schumann : *How do oral linguistic signs become reliably indexicalized ?*
 - Goodwin : *How autonomous is language from other semiotic systems ?*
 - Wray : *How could a pre-modern protolanguage be viable ?*

3. How does modern language reflect what evolved ?

- Tao : *How much of what we see in modern language is really as old as language ?*
- Véronique : *What clues to language evolution can we gain from Creoles and second language acquisition ?*
- Mikesell : *How complete are complete sentences ?*

1. The Presentations

1.1. How does Modern Language Reflect What Evolved ? How much of what we see in modern language is really as old as language ?

« *A Corpus-Based Analysis of Gapping Constructions in English and Its Implications for Understanding Language Evolution* »

Hongyin Tao (in collaboration with Charles F. Meyer of the University of Massachusetts in Boston) reported at the Roundtable on the implications of Gapping constructions (Ross, 1967-1986) for understanding the evolution of language. They estimate that there have been more than 160 books, papers published on this phenomenon, and the topic is featured in almost all textbooks on formal syntax. Typical Gapping structures discussed in the literature are of the following type : *I ordered a drink and my friend [] an appetizer*. Gapping has been analyzed for verbs, verbs and objects, adverbs, modals, auxiliaries, and determiners.

Since nearly all studies of Gapping have been done on created sentences that were analyzed according to acceptability judgments, Tao and Meyer wanted to examine Gapping in natural discourse. Therefore, they chose as their database the British component of the International Corpus of English (ICE-GB), consisting of one million words (600 000 spoken and 400 000 and written), or 500 two-thousand word samples from genres such as spontaneous conversation, newspapers, fiction etc. The researchers found that among 17 629 examples of local coordination in the corpus, only 120 tokens (or less than 0,7 %) contained gapped structures. This means that while there have been more than 160 published studies of this construction, only 120 tokens of it can be found in the million word ICE-GB corpus.

The authors conclude that in natural discourse, Gapping is an extremely rare syntactic phenomenon. In the ICE-GB corpus Gapping constructions had the following distribution (per thousand words) :

1. Dialogues of any kind (0/360 K = 0),
2. Spontaneous monologue (8/100K = 0,08),
3. Planned monologue (13/120 K = 0,11),
4. Speech-inclined written genres (14/70 K = 0,20),
5. Prototypical written genres (85/330K = 0,26).

This analysis indicates that Gapping is not an oral discursal phenomenon, and therefore, it is unlikely to have been part of what evolved in language evolution. Instead, it appears to be a product of writing, especially journalistic writing, that may have had some minimal influence on dialogic spoken language.

Tannen (1987) and Meyer (1995) both have convincingly demonstrated that discourse favors repetition to mediate its transitory nature and to promote engagement and resonance between speakers. This feature of discourse seems very likely to explain the rarity of Gapping in speech. Additionally, Gapping may be more cognitively demanding because it may increase the amount of new information in a clause. For example, in a sentence such as *I want to try to begin to write a novel, and Mary a play* both « Mary » and « a play » could be two pieces of new information. These cognitive demands are more easily accommodated in written language where revision (in writing) and repetition (in reading) allow remediation of a cognitive failure in either production or comprehension.

The authors conclude that Gapping is a feature of composed language, and is most typical of written genres. It is dispreferred in spoken discourse where it is found only in highly planned oral monologues. With respect to developing evolutionary and oral biological accounts of language, Gapping structures and similar linguistic phenomena are not appropriate candidates for explication. If language evolved 50,000 years ago and writing appeared 5 000 years ago, fully fledged

language would have existed for 45,000 years without structures like English Gapping sentences. Additionally, since Gapping seems to be only part of written language and since written genres are only acquired through instruction, then Gapping constructions are examples of what the brain can do if it is sent school, and they are not an example of what the brain evolved to do in the environment of our evolutionary adaptation. Gapping structures may be typical of many of the forms with which linguistics has concerned itself. In developing biological accounts of language, therefore, it may be wise to check large language corpora to determine whether a particular structure is likely to have appeared prior to the cultural innovation of writing.

1.2. How did Language Evolve ? What Biological Developments are Likely to have made Language Possible ?

« *Biological Substrates that Contributed to Language Emergence* »

Lee and Schumann (2003, see also below) suggest that contrary to Chomsky, humans have no specific genetically based and neurally instantiated representation for grammar (i.e., no UG). Nevertheless, in order for language to evolve, substantial physiological, including neurobiological, preparation may have been necessary.

Namhee Lee, at this Roundtable, outlined these biological underpinnings. Agreeing with Ajaenian (2001), he suggested that for any number of reasons, language, as it is known today, would not have evolved without a vertical vocal tract made possible by bipedalism. In addition, there are some researchers (e.g., Corballis, 2002) who maintain a gesture-origin theory of language evolution. According to this theory, arms that were liberated by bipedalism became available for communication via gesture. Additionally, Lee suggested that a change in cranial base flexion allowing the head to remain upright may have paved the way for a lowered larynx which is essential for speech.

In typical mammals, including primates, Neanderthals, and human babies up to one year old, the larynx is located high in the epiglottis almost reaching the soft palate. Therefore, the pharynx is short allowing only limited vocal sounds to be produced (e.g. [a, i, u] are not possible). In human adults, the lowered larynx increased control over the expiration phase of respiration and thus allowed the production of longer vocalizations and a greater range vowel and consonant sounds. The completion of the descent of the larynx occurred about 150 to 100 thousand years ago (Lieberman, 2000). There have been arguments that modern humans and Neanderthals have the same vocal apparatus based on the discovery in the Kebara 2 Neanderthal of the hyoid bone, which is similar to that of modern humans (Johanson & Edgar, 1996). However, the hyoid bone is not connected to any other bone and floats freely ; therefore, although the hyoid bone in the Neanderthal is similar to that in modern humans, its location may be different in the two species.

MacNeilage and Davis (2001) argue that another biological substrate that may have contributed to the appearance of language is mandibular oscillation. Mandibular oscillation originally served the function of digestion (chewing, sucking, licking). Later the process was exapted by higher terrestrial primates for the purpose of facio-visual vocal communication (lip smacks, tongue smacks, teeth chatters). The oscillation cycle of mandibles paired with regular phonation may have become the basis for syllable formation in hominids (consonants with closed mandible and vowels with open mandible).

Kay, Cartmill, and Balow (1998) suggest that the expansion of the hypoglossal nerve canal may be associated with human language. The hypoglossal nerve is used to control tongue movement, and this control provides better articulation. The hypoglossal nerve canal in modern *Homo sapiens* is 1, 85 times that of common chips, 2, 44 times that of pygmy chips, and 1, 8 times that of gorillas. The canal in *Australopithecus africanus* and *Homo habilis* is about the same size as that of apes. The researchers have found hypoglossal canals that are similar to those of modern *Homo sapiens* in two middle Pleistocene hominids (300 000 years old), in two Neanderthals, and one early *Homo sapiens*.

Brain size or encephalization may also have had an effect on language. There is a positive correlation between brain size, behavioral innovation and social learning capacities in which a larger brain is thought to support better learning (Reader & Laland, 2002). Four million years ago, hominid brains were about 400 cc ; by 2 million years ago brain size had increased to 500-

800 cm³ The hominids living 1, 5 million years ago had a brain size of 752-1 250 cc. By 600 000 years ago encephalization in hominids had increased to 1 500 cc (Lieberman, 2000). The brain size of animals generally correlates with their body mass ; primates brains are about 3 times larger than non-primates having the same body weight. Human brains are three times larger than the brains of chimps (Allman, 1999).

Lee proposes that there may have been a change in humans' cortico-basal ganglionic loops which may have facilitated procedural learning. Patients with Parkinson's disease and Huntington's disease both have basal ganglia damage and consequent linguistic deficits (Ullman, 2001). Lesions in the basal ganglia lead to aphasia characterized by agrammatism (Fabbro, 1999). Individuals with a particular mutation in the Fox P2 gene have basal ganglia with an underdeveloped caudate nucleus and also have language difficulties (Lai & al., 2001 ; Vargha-Khadem & al., 1998). Finally, adults language learners have difficulty with procedural aspects second language acquisition, and Lee speculates that the deficit results from interference from other motor programs implemented in the basal ganglia or from competition with the cortico-hippocampal system involved in declarative memory (Poldrack & Packard, 2003).

Lee also indicates several psychological phenomena related to language that may have biological underpinnings that supported language evolution. The first is the ability to make symbolic reference (discussed below by Favareau and by Schumann). The ability to use signs to refer to other signs (a hallmark of language) may have involved some biological change. Tomasello (2003) suggests that abilities for joint attention, imitation, intention reading, and pattern finding are all crucial to language acquisition and use. The biological substrates for these processes would also constitute part of the biological basis for language. Greenspan and Shanker (2004) suggest that human infants developed special abilities for (1) perception of stimuli with inhibitory control over action with respect to those stimuli, (2) nuanced facial expressions, 3) and extended attention to and interaction with caregivers' emotional expressions. They argue that these abilities are important for language acquisition and, therefore, would have been important for language evolution. Although Greenspan and Shanker regard the attention and interaction abilities as epigenetic and, therefore, learned ; it is also possible that the tendency to attend and to interact with conspecifics may have also involved biological changes leading to an « interactional instinct » rather than a specifically « language [grammar] instinct » as proposed by Pinker (1994).

Seen from this perspective, there is no specific biological substrate for language. Instead, language relies on the complex interaction of numerous physiological substrates and neurobiological regions, systems, and processes all of which support language behavior without representing specific aspects of language such as grammar. This point will be elaborated below the discussion of language as a complex adaptive system.

1.3. What Exactly Evolved ? Does the search to discover the origins of human language start by searching for the earliest instances of what we can already recognize as human language activity - or is this already a « too late » starting point for understanding language as part of an evolutionary continuum ?

« *The Evolving Cultures of Nature* »

Donald Favareau's roundtable presentation expounded the notion that human language was not the birth of publicly shared semiotic systems, but rather, was itself birthed by the pre-existence of such publicly shared semiotic systems, which are ubiquitous throughout the natural world. Favareau's presentation briefly sketched out both the continuities and discontinuities between human language and its predecessor non-human counterpart systems in the experiential worlds of animals, fish, insects, protists and plants.

Inspired largely by the work of Terrence Deacon (1997, 2001) and by work done in the newly emerging interdisciplinary of *biosemiotics* (which is the study of sign processes as they appear variously across the biological spectrum), Favareau noted that - at a time when strictly materialist reductionist explanations of life and its evolution have become increasingly incompatible with what biologists are now conceding is the complex, adaptive and non-linear nature of organization and interaction in the natural world - the conceptual work now taking place under the aegis of the *biosemiotics* (e.g., Barbieri, 2001 ; Deely, 2001 ; Emmeche, 1999, 2000 ; Favareau, 2001, 2002 ;

Hoffmeyer, 1996, 2000 ; Kull, 1998 ; Markos, 2002, Sebeok, 1990, 2001 ; Taborsky, 2001 ; Turovski, 2000) may help us better understand the principles whereby not only our social world and its linguistic systems, but also the very biological world with its species-specific semiotic systems came into being not as a « given » in the furniture of the universe, but as a locally organized, massively co-constructed, context-creating and context-sustaining interactional *achievement* in that universe instead.

Informed primarily by philosopher Charles Sanders Peirce's (1839-1914) semiotic logic of relations, animal ethologist Jakob von Uexküll's (1864-1944) constructivist biology of perception, and physicist and chemist Ilya Prigogine's (1917-2003) explication of self-organizing systems as decentralized accomplishments, the interdisciplinary project of biosemiotics is grounded in the conviction that the living organism, too, must be understood not only in its material organization, but also in the organization of its interactions (both internal and external). Biosemiotics holds that these two *sets* of organizing relations are, in fact, interdependently bicausal, and that *sign relations* (understood in the broadest sense, as below) mediate such activity as gives rise to behavior in the world.

Favareau then went on to briefly describe how *sign processes* permeate the relations of living things. On the most obvious level, he noted, we see sign processes underlying human spoken language and written texts ; primate, canine and reptilian display behavior ; birdsong ; pheromone trails ; and the deceptive scents, textures, movements and coloration of a wide variety of symbiotically interacting insects, animals and plants. Less obviously, perhaps, there are the chemotaxic signs by which single celled animals negotiate the world and upon which the human body's immune system operates ; the chemical and electrical events that constitute the signals and messages of the brain and central nervous system ; and the nucleotide sequences that, when read by cellular mechanisms, give rise to life from the genetic code. All of these phenomena are examples of true sign processes - i.e., substitution relations whereby something is « represented » to an organism by something other than itself - and yet each of these instantiations differ from each other in a number of fundamentally important ways.

Until recently, however, no one discipline has attempted to provide a synthetic explanation of how the cultural processes of sign use and the biological processes of sign use do and do not relate. Biosemiotics employs the semiotic logic of relations developed by philosopher and scientist Charles S. Peirce in order to distinguish the various orders of sign processes ubiquitous to the world of living beings. An architectonic well beyond the scope of this presentation (and this paper), the relevant Peircean categories discussed with reference to the roundtable on language evolution were the fundamental triadic relationship of *sign-object-interpretant* which alone makes sign use possible, and the nested hierarchy of sign types *icon-index-symbol* which, as Deacon (1997) has very thoroughly and convincingly argued, both underlies the ability of human language use as well as establishes its continuity (and singular point of discontinuity) with the sign processes of the rest of the animal kingdom.

A *sign*, for Peirce, is « something which stands to somebody for something... [and] not in all respects, but [only] in reference to... [its] ground » (p. 2228). Simplifying considerably : Peirce's definition reveals that : (1) there are no such independently existing things as « signs » *per se* - instead, there are only independently existing things that are used *as* signs by the agents that act upon them as such. (2) Such triadic action (the taking of thing *x* to « stand for » *y* by agent *z*) *alone* brings the « sign relation » into being, and that (3) such action consists in a living agent (« somebody ») in each and every instance of sign-use *actively* joining the sign vehicle (that thing *x* in the world which it will use as a sign for *y*) to its *object* (not *y* itself, but only those aspects of *y* relevant to the experiential world of agent *z* [cf. Peirce, p. 2229, p. 5401]) so as to result in an effective *interpretant* - or « proper significate effect » (p. 5475) through which agent comes to feel, act upon and ultimately reason about the world.

This nested hierarchy of *feeling* (perceiving) the world through signs, *acting upon* the world through signs and *reasoning about* the world through signs Peirce characterized as relations of *firstness*, *secondness* and *thirdness*, respectively - with the characteristic sign types corresponding to these hierarchical categories being *icons*, *indexes*, and *symbols*. Simplifying greatly (again) : to the extent that a sign « partakes in the *character* of its object, » it is an *icon* ; to the

extent that a sign is « really and in its individual existence *connected* with its individual object, » it is an *index* ; and to the extent that a sign « will be interpreted as denoting the object, in consequence of a *habit*, » convention or law, it is a *symbol* - for the agent to whom the object *is being used as* a sign of something else at all (p. 4531).

Given this understanding of « sign relations » as primarily *not* something psychological, linguistic or even human-specific but rather, as simply those relations that any living organism may stand in towards the objects in its world, biosemiotics makes relevant for our understanding of language evolution Peirce's observation that a sign is « something which stands to somebody for something... [and] not in all respects, but [only] in reference to... [its] ground » (p. 2228) through defining that « ground » in terms of animal ethologist Jakob von Uexküll's (1864-1944) constructivist notions of species-specific perception-action cycles (*Funktionkreis*) and the subjective experiential worlds (*Umwelten*) that such perception-action cycles give rise to in animals over evolutionary time and that, in turn, drive action initiation in ontogenetic time.

Noting that the world inhabited jointly by all species is perceived radically differently by each (e.g., as a world of pure sonar, smell, tactile stimulation, chemotaxis or ultraviolet radiation - or as a world of visual and aural experience wherein the ultrasonic and UV reality perceptible to one's neighboring creatures fails to show up as « reality » at all), von Uexküll's most famous example of a highly limited but perfectly species-preserving *funktionkreis* - and - *umwelt* dynamic is that of the tick.

The tick, noted von Uexküll (1934) lives in a perceptual world consisting only and entirely of butyric acid, tactile pressure and heat - and an tightly conjoint action-response schema towards the subjective experience of each of these three phenomena alone. Thus, the tick hangs deaf, blind and motionless on its branch until the presence of butyric acid (a component of animal sweat) appears in its world of subjective experience, at which point it releases its grip on the branch. Falling on to the source of the butyric acid - i.e., the body of the warm-blooded animal that was passing below the deaf and sightless tick - tactile contact in the upright position initiates running activity in the search for heat (skin). Upon registering the presence of heat, the tick begins burrowing into the animal's skin to feed (1934, pp. 10-12).

Given that the tick has no visual or aural apparatus, and gives no evidence of having even the ability of detecting anything other than the three specific aspects of the world above, it would make no sense to say that the tick « knows » that the blood it feasts on is carried by such unfathomable phenomena as horses, cows and pigs, and that it is the sweat of these animals that carries the butyric acid that alone sets the tick's « function cycle » in action, feeding it and allowing it to survive. And yet : there *are* such things as horses, cows and pigs actually existing in the world and it *is* the sweat of these animals that carries the butyric acid that alone sets the tick's « function cycle » in action, feeding it and allowing it to survive.

What this reveals to us about *sign-processes* thus is crucial : Sign processes are not, foundationally and in their *firstness*, linguaform codes corresponding to psychological conceptual categories - but are rather, just as we have described them above, « substitution relations whereby something is « represented » to an organism by something other than itself » (cf. Favareau, 2000, 2004).

In Peircean terms, the tick's *Umwelt* or experiential world « carves » out of the plenum of possible perceptual experience just the three perceptual phenomena made available to it by its evolutionary heritage (i.e., the heritage of its species' perception-action cycle success over time). Brute, immediate perception or registration of these phenomena in whatever way they are experienced (i.e., not as « butyric acid » but simply as *that* feeling or registration as opposed to *not* that feeling or as opposed to some *other* feeling) constitute the *icons* of that sign relation for the tick - the feeling or immediate registration of butyric acid, tactile pressure or temperature detection, however those things may be experienced by the tick.

What is *not* critical for the survival of the tick is for the tick to have internal « labels » for these percepts *nor* for the tick to have the psychological understanding that these percepts operate as the *indexical signs* for the objects that they represent (i.e., co-present animals, those animals' flesh and the blood meal underneath that flesh, respectively). What *is* critical is that the tick *act upon* the butyric acid, tactile sensations, and temperature changes as the *signs* for animal presence, covering

flesh and underlying blood meal. And it is the evolution of its species perception-action cycle that guarantees the veridical conjoining of these agentive actions, objects and signs.

Biosemiotics argues that what is true of the tick is true of all other living organisms, all of whom have to somehow come to « know » the world and to act in it successfully using only those signs made available to it by way of the perceptual apparatus with which it has become evolutionarily endowed (cf. Sebeok, 1977). Primitive sea creatures, as Llinas (2001) has pointed out, could survive using only the most grossly discriminating photoreceptive patches to distinguish extremes of dark and light. Yet armed with these just these two iconic distinctions, these animals could successfully exploit the corollary *indexical* relations that were « really and in their individual existence *connected* » with these icons. Food - but also increased exposure and its attendant dangers - were both physically (and thus indexically) connected with the ocean surface represented to the animal through the photoreception of its iconic light. Shelter and relative safety from predators - but also radically diminished feeding opportunities - were both physically (and thus indexically) connected with the ocean bottom represented by its iconic dark. And here again we see that what constitute successful *sign* relations in the first instance does not have as its fundament human minds or language-mediated thinking, speaking, or writing practices - but, rather, the triadic joining of objects by the agents of the world through substitution relations grounded and vetted in successful action, or *use*.

As evolution endows animals of increasing complexity with correspondingly fine-grained perceptual apparatuses, such apparatuses, in turn, allow the animal to engage in more complex and fine-grained interactions with the world - a world not only of objects, but of other agents also. Yet before we see the kind of classic « animal communication behavior » of mating calls, dominance displays, territory marking or even pheromone trailmaking, sign use for survival crosses not only individual, but also species, *umwelten* in the mindless, brainless morphology of animal camouflage and in the multisensory mimicry of plants. In both these latter cases, successful survival for organism A is predicated on its exploitation of the icons and the indexes integral to sign processes employed for survival by organism B.

An elegant example of this is the morphology of a thermogenic Mediterranean lily called *Helicodicerus muscivorus* or, more commonly, the Dead Horse Arum (Angioy & al., 2003). Found on gull colonies where rotting bird corpses and their attendant carrion blowflies are abundant, these gruesome smelling arums precisely mimic the smell, sight, texture and even temperature of a rotting corpse in order to attract the blowflies into its prison-like chambers, deposit its pollen on them, and then finally release them to carry its seed and thus reproduce (Kite, 2000 ; Stensmyr & al., 2002).

Examples of such fine-grained mimicry abound in nature, but what is most relevant to our present discussion is the acknowledgement that the Dead Horse Arum cannot in any psychological or conceptual way « know » what a rotting gull corpse looks like, smells like, feels like, or what its body temperature is upon recent expiration - an important « sign » that the arum uses in exploiting the carrion fly (Angioy & al., 2003). Nor can it subjectively experience, « see » or in any self-reflectively cognitive sense, « know » even of the existence of the carrion flies - much less « understand » their role in the process of disseminating pollen. Yet because (non-psychological, non-conceptual) *sign relations* are integral to the blowflies' successful negotiation of the world - the success of the plant's survival is predicated not on its (non-existent) endosemiotic « psychology »--- but on the exosemiotic action in the world that results as its evolved biology interacts with the subjective sign experience of the fly.

Favareau then went on to argue for the ubiquity of sign processes in nature using a variety of examples of semiotically interacting animals, fish, insects and plants. He maintained, with Hoffmeyer (1996), Stjernfelt (2002) and Kull (2000) that all organisms are born into an unlabeled world of things and must use some of those things in the world as signs by which to know how to live and to survive in that world. Knowledge is built from signs, claims Favareau, and communication is a way that agents use signs to build knowledge together.

For by carving up the unlabeled world of time and space into iconic relations and in setting up indexical relations across entities, organisms begin effecting the material world causally based on the immaterial mediating relationship of using things as signs. Indexes can be chained to other

indexes so as to result in incredibly complex long chains of purposively adaptive behavior - and internal states such as hunger and exhaustion come to manifest in the organism's phenomenology as icons which can be brought into indexical relations with other icons.

Such acts of semiotic mediation take place recursively not just at the locus of the individual, however, but perhaps most generatively on the level of aggregate, interacting agents. There, such relations can themselves be embedded in even high-order systems of substitution relations - and the history of human culture, it has been argued (Deacon, 1997 ; Deely, 2001 ; Donald, 1991 ; Tomasello, 1999) consists in just this recursive representative strategy wherein « each subsequent representation in the semiotic chain represents the prior object-sign relation, *taken itself* as a higher-level semiotic object » (Parmentier, 1996, p. 5).

Deacon's (1997) discussion of the *culture* of symbolic reference underlying human language use makes it clear that the primary phenomenon to be accounted for in an evolutionary account of language is not so much the faculty of a human brain or the linguistic facility of a species (much less of an individual), but rather, the development of a multiply-embedded semiotic way-of-being in the world. Characterized by a Peircean notion of *thirdness*, this way of being is predicated on a network of sign relations that are being held for use in perpetuity *outside* the agent - i.e., in a public domain of interactively-constituted sign-exchange whereby meanings can be created, negotiated and co-operatively sustained. *Participation* in this system alone enacts and enables « meaning, » claims Favareau - both here and in the animal world - and in this sense, it is the natural history of agents and their actions in the world that is the proper starting point for undertaking a natural history of signs - only one of whose multitudinously various end-products is human language.

1.4. What Exactly Evolved ? How could Premodern Language be Viable ?

« *Units of Meaning at the Dawn of Language* »

Alison Wray addressed the topic of « What evolved ? » from the point of view of the role that pragmatics plays in making good shortcomings in different types of communication. In most kinds of animal communication there appears to be an absence of specificity in calls : a food signal almost never includes referential information about what or where the food is (the bee dance is an exception), nor does a warning signal normally specify the nature of the threat (except possibly in vervet and Campbell's monkeys). In short, the onus is on the receivers to make good the missing information, so as to understand how the call relates to them. This is achieved by applying knowledge of the world, past experience and immediate contextual information. Without such interpretative augmentation, the value of the message is much reduced. Although natural selection would be expected to settle on a reasonable balance between the burden on sender and receiver, the holistic nature of animal calls is a substantial constraint on the sender's potential to be explicit.

Humans also use pragmatics to handle absence of specificity in formulaic utterances (Wray 2002a). If we hear « look out ! » we cannot take appropriate evasive action until we have ascertained the nature of the hazard. When we encounter input like 'Could you put the whatd'yamcallit onto the what's it' we similarly use context and pragmatics to complete comprehension. Humans also use pragmatics for a different kind of message completion. If we read *Dock monk bolts* we will use our contextual knowledge to extrapolate a meaningful relationship between the words. If it is a newspaper headline we might understand that a monk undergoing a legal trial has absconded, whereas in a technical machine assembly manual the same words might constitute an instruction.

Do animals also have the capacity to use pragmatics to disambiguate relationships between referents ? If the predator calls of monkeys are translated as names, then we might infer that the « Eagle » call requires the hearer to work out why the eagle has been named. However, this denies the functionality of the call, as a warning/instruction, e.g. « behave in a way consistent with my having seen an eagle ». Only if the item really is a « name » can there be ambiguity as to why it might be being uttered. Kaminski, Call & Fischer (2004) have shown that dogs are capable not only of associating names with objects but also an unfamiliar name with an unnamed object. Again, however, dogs associate the name with a command to fetch the named object, so there is no opportunity for pragmatics to play a role in disambiguating the purpose of the naming event. Rather, the « name » entails the instruction to find and fetch that item.

Bonobos in captivity (e.g. Savage-Rumbaugh, Shanker & Taylor, 1998) have shown a remarkable ability to understand spoken English, but very little ability to produce more than single words or agrammatical pairs. It could be that their comprehension relies heavily on pragmatics to infer likely relationships between referents. If so, bonobos, like humans, have the capability to use pragmatics both to make good reference in a known manipulative message (in the wild) and to make good the relationship between referents when these are given (when comprehending English). If bonobos can do both, then we can reasonably infer that our ancestors could. In both cases, the latter role for pragmatics would have to remain latent until naming began. Thus, the emergence of naming is key to releasing pragmatics into the new role.

Of three published models of human protolanguage, two already have naming, and one does not. Bickerton's (1998) scenario gives concatenations like « run mammoth » and « man stone ». These require pragmatics to disambiguate the relationship between the referents, since there is no word order and the arguments of protoverbs are not compulsory. Meanwhile, such a protolanguage, being based on referential items, appears to relieve pragmatics of its more ancient role in disambiguating reference, though, since we have it today, it could not have disappeared.

Wray's (1998, 2000, 2002b) protolanguage has no independent words for things or actions, only holistically expressed complete messages. Wray proposes different holistic forms for « give this to him », « give this to her », « give that to her », etc. but less semantic differentiation is also possible. Forms meaning « give something to someone » ; « stay away from someone/thing », etc. would effectively be equivalent to verb frames. Without the pronouns of Bickerton's scenario, however, pragmatics would remain necessary to provide information about the actor, patient, etc. Wray's scenario is more « primitive » than Bickerton's, and fits more seamlessly with a progression from a single use of pragmatics to the modern dual use.

Jackendoff's (2002, p. 246ff) protolanguage has naming and a protogrammar in which linear position signals semantic relations : Agent First, Focus Last. Naming and grammar between them provide both types of disambiguation, reducing the role of pragmatics to its modern level.

It seems reasonable to propose that the capacity to use pragmatics to disambiguate relationships between named entities and actions is part of 'what evolved'. If it is a latent ability in bonobos, its emergence may not need independent explanation, being contingent on the emergence of naming. Wray excludes naming from protolanguage, and has argued that referential underspecification would constitute a steady state inhibiting innovation. In contrast, it might be conjectured that in Bickerton's and Jackendoff's models, the introduction of pragmatics to disambiguate the relationships between named objects and actions would pressure the communication system into developing more explicitness about these relationships, rendering the protolanguage transient. Whether Jackendoff's scenario supplies sufficient explicitness to create a steady state is difficult to judge. Absent in his protolanguage are phrase structure, embedding, and the constraints contingent on embedding : traces, subadjacency, etc. However, structural embedding is not universally attested in unwritten languages (see Wray & Grace forthcoming for a full account). If Jackendoff is right, then naming and order came early, and phrase structure later. In contrast, Wray's account builds up semantically complex messages to bursting point, so that when naming finally emerges, grammar immediately follows, as forms map onto existing semantic representations. Wray's position makes it clear that an evolutionary account of language must consider holistic language and not simply analyzed units composed into sentences. Therefore, we must know how the brain chunks linguistic input.

1.5. How does Modern Language Reflect what Evolved ? What Clues to Language Evolution can we Gain from Creoles and SLA ?

« *Second Language Acquisition and Creolization : A Small Window on Language Evolution* »

Daniel Véronique addressed the relevance of second language acquisition and creolization for the issue of « what evolved ? » He noted the characteristics of protolanguage : utterances are organized by functional considerations ; null elements are recoverable from knowledge of the situation and from world knowledge ; verbal arguments may not be expressed ; complex structures are absent and there is a lack of function words and inflectional morphology. It is interesting to observe that with the exception of the final characteristic (lack of functors), these

properties correspond to what Mikesell (see below) and others have described of normal conversational interaction in fully-fledged language. If protolanguage and conversational discourse operate in similar ways, then the study of conversational interaction may provide some clues to how language evolved.

As Lee & Schumann note (see Part 2 below), early second language acquisition and early pidginization may provide, at least, oblique reflections on the evolution of language, Véronique's Roundtable presentation allows us to also include creolization because all three forms of language development can occur among adults. This notion reflects a new perspective on creolization in which some creoles are now viewed as having developed without an antecedent pidgin (Parkvall, 1995 ; Arends, 1995 ; Mufwene, 1996). Atlantic as well as Indian Ocean French related creoles appear not to have had a preceding pidgin stage. They developed under conditions in which the contributors to the Creole were adult speakers of nonstandard varieties of the target language or non-native speakers of the target language. Each speaker provided an idiolect with features that differed to some extent from other idiolects. Mufwene (2001) argues that this collection of idiolects provided a feature pool out of which forms were selected for broader use within the community. As Lee and Schumann have argued, the vetting process inherent in the interaction among speakers of these idiolects would allow for the selection of those forms that were most efficiently producible, comprehensible, and learnable (Kirby, 1998). Thus the language would emerge as a complex adaptive system from the massive interaction among varying idiolects.

The early stages of second language acquisition have been studied extensively, and the language produced during this period is referred to as the Basic Variety (Klein & Perdue, 1992, 1997). This period in L2 learning is characterized by utterance structure in which the entity with highest control comes first (generally an agent) and the information that is in focus comes last. Véronique argues that both the early stages of French as a second language (FSL) and French based creoles (FBC) are characterized by iconicity (Croft, 1990) which is defined as an attempt to maintain a one-to-one relationship between form and function such that syntactic structure directly reflects semantic structure. In Peircean terms, this may be viewed as a search for an indexical relationship in which words, as directly as possible, refer to actions and things rather than accruing their meaning by reference to other words or morphemes. This concept-word indexicality is used to refer *inter alia* to existence, possession, and time in both FSL and FBC.

In this discussion, the traditional terminological distinctions among early SLA, early pidginization, and early creolization have been maintained. However, Véronique indicates that for him and others in the field (Becker & Veenstra, 2003), all three processes may be one and the same thing. Such unification affords a conceptual perspective on adult language development that may facilitate thinking about adult-adult interaction in language evolution.

The following examples of existential structures and of structures for expressing temporality from French as a second language and from French based creoles offer some perhaps oblique insights into the kinds of processes that may have involved in the evolution of language.

1.5.1. Existentials in FSL and FBC.

1.5.1.1. Il y a (There is) and c'est (It is) in FSL

One of the phrasal constraints posited by Klein and Perdue 1997 for the Basic Variety runs as follows :

- {V } - NP2.
Cop

It is in this type of string that *il y a* (there is) and *c'est* (it is) are found in FSL. These items share propositional functions such as :

- clefting,
- acting as a carrier of negation and of temporal information, and discursive functions such as :
- topic introduction,
- contrasting foreground-background in texts.

il y a expresses mainly existence and possession whereas *c'est* is identificational (Duff, 1993).

a- In early grammar

One of Véronique's subjects, Malika B. (MB), a speaker of Moroccan Arabic L1 produces the following utterance in a picture description. The utterance involves code switching between her L1 and French (see 1994) :

[1] : I = Investigator
I : et ça qu' est-ce que c'est ça ? [What is this ?]
MB : /se/ al bajda [It's eggs]

I : voilà où est le chat ? [where's the cat ?]
MB : /e/ le table [is the table]

/e/ + X is a productive device in early FSL grammar.

/jãna/ + X (there is) is used to mark existence and possession (see Duff, 1993).

[2]
I : tu travailles chez Madame F ? [You work at Madam F.]
MB : Madame F + Madame F [Madam F + Madam F]
I : oui [yes]
MB : moi /li/ bar [me (the /she) bar]
I : hm hm
MB : /li/ bar moi +++ toi /jãna li/ café [(the) bar me you have (the/ she) coffee]

1.5.1.2. *Il y a* and *c'est* in learner language development

il y a or rather *jãna* is found in all learner varieties. It marks possession as well as existence, although the existential value tends to predominate in the more advanced learner varieties. In 3, the first token of *jãna* is an existential and the second expresses possession :

[3]
AE : voilà /jãna/ deux personnes /iveni/ /terte/ avec quelqu'un /jãna/ des drogues +
[Ok there are two personnes (they) come search with somebody has drugs]
voilà /eveni/ euh voilà /ja/ rien /itrap/ avec l'autre qui /mãz/ à côté de
[Ok come there's nothing catch the other who is eating besides]
(Cycle 2, encounter 5 = 2.5).

jãna is used extensively for clefting in more advanced learner varieties leading to proto-relatives (see Véronique, 1997), as in the following example where the contrast between *jãna* and *se* is quite striking :

[4]
AB : /jãna une femme /se/ une vieille /iladi/ non /se/ pas le monsieur /se/ le fille (3.5)
[there's a lady it's an old (woman) (she) says no it's not the man it's the girl]

1.5.1.3. *Il y a* [There is] and *c'est* [It is] in FBC

In early texts from Mauritius, the marker expressing existence and possession is *y en a* [there is/have]. First tokens of the item are found in texts dating from ca. one century after the settlement of the island (1721).

[5] : 1804
Mo y en a femme
[I have a woman] (Chaudenson, 1981, p. 81).

[6] : 1804
Y en a ça qui bon y en a ça qui mauvais
[There are those who are good, there are those who are bad] (Chaudenson, 1981, p. 81).

In the course of evolution, *Mau.y en a* is eroded to *iéna* (Baissac, 1880) and to contemporary *ena* (Chaudenson 1981, p. 59).

Se [c'est] is quite rarely found in texts from Mauritius, or Réunion :

[7] : 1828

Péchés qui y-en-a dans licaire, sé quand moi oblié Bon Dieu
[Sins I have in the heart, it's when I forget God] (Chaudenson, 1981, p. 109).

It is more frequent in Seychelles Creole, although this marker is different from se in Atlantic FBC (Baker & Corne 1982, pp. 40-41). Note that the equivalent of *Mau.ena* in Seychelles Creole is *ana*.

The use of *y en a* and *ana* in Indian Ocean FBC provides an example of reanalysis in creole genesis. Subsequent erosion provides a clue as to the grammaticalization of the reanalyzed forms.

Se in Atlantic FBC occurs in equational sentences and seems to have been reanalyzed along different lines from Indian Ocean Creoles. In contemporary Atlantic FBC, it is analyzed either as a copula (Valdman, 1978, pp. 237-240) or a resumptive pronoun (DeGraff, 1992, pp. 168-209).

1.5.2 Temporality in FSL and FBC

1.5.2.1. The development of time reference in FSL : from pre-basic varieties to post-basic varieties

i)- *Temporality in pre-basic and basic varieties*

Previous work shows that pre-basic varieties make use of lexico-pragmatic means and discourse principles before any emergence of verbal morphology (Andersen (ed.), 1984 ; Pfaff (ed.), 1987 ; Bardovi-Harlig, 1999). Despite progress in L2, these first indirect means are maintained.

- a) Recourse to scaffolding by the TL speaker and reliance on contextual inference
- b) Use of the iconic principle of natural order

[8] (Abdelmalek recounts his arrival to France, 1.1)

- I : tu peux/oui alors tu peux me dire quand c'est/depuis combien de temps tu es là ?
[Can you/yes O.K can you tell me when/since how long you are her ?]
- AE : comment le problème comme /ãtre/ la France ?
[how the problem for coming to France ?]
- I : ouais par exemple ouais
[Yes eventually yes]
- AE : ah ouais parce que moi /liãtre/ la France /jana/ pas de passeport /jana/ pas de rien
[Oh yes because me (I) enter France there's no passport there's nothing]
- I : ouais
[Yeah]
- AE : parce que /ãtre/ la France /e/ la montagne
[because enter France (is/and) the mountain]
- I : tu es passé par la montagne ?
[You went through the mountain ?]
- AE : ouais
[Yeah]
- I : ah
- AE : /jana/ cinq jours /e/ la montagne après /lãtre/ la France /lepase/ la douane de
[There's five days (in) the mountain then come to France go through the customs of
France /komjes/ ? quinze kilomètres
[France how much ? Fifteen kilometres]
- I : à pied ?
[On foot ?]
- AE : ouais /lapje//e/ après /ilaparte/ l'autoroute /jana/ pas des sous /jana/ rien après
[yeah on foot after go with the speed way there's no money there's nothing then]
/jana/ /e/ le stop après /leveny/ le gendarme
[there's hitch-hiking then come the police]

Explicit reference to a time span preceding time of speech (+ Anterior) is mainly marked through the questions and the scaffolding provided by the investigator. The iconic principle of natural order applies and [+ anterior] time intervals are marked by connective *après* (after). Chronological order is also fostered through the use of calendaric information /jana/ cinq jours (there is five days) and through indirect lexical means like motion verbs /ãtre/ (enter), /lepase/

(pass), /ilaparte/(go away) /leveny/(come) and specific nominals montagne, douane, autoroute. The action verbs exhibit long forms (li/e) +Ve. Only one stative verb /jana/. (there is) is present in this extract. The verbs exhibit formal variation either in pre-verbal position Ø/le/ila/li (which corresponds to the position of the clitic pronoun and the auxiliary in TL) or in post-verbal position (corresponding to the verbal ending in TL) Ø/e/i/, without any explicit correlation between form and meaning (see also Dietrich, Klein & Noyau, 1993, pp. 171-172).

ii)- *Complexifying the expression of temporality*

More grammaticalised means may also be used such as :

a- *Expression of TT (Topic Time) through temporal subordinate clauses as in 9 :*

[9]

- Z : ton mari ta copine /eleparti/ le Maroc à vacances le/ la dame ne /se/ pas
 [(the/ your) husband of (your/ my) friend has left for Morocco on holidays the lady does not know]
 /e/ tous les papiers à la maison /e/ ton mari la femme tous les papiers passeport
 [and all the official documents at home and (your/ the) husband the woman all the documents passport]
 /e/ après /kølø/ /parti/ le Maroc
 [and after when he arrives in Morocco]

 /kølø/ /ilparti/ au Maroc /e/ après le /kominis/ /e/ après /saje/
 [when he arrives in Morocco and then (he will visit) the judge and then it's over]

b- *Verbal morphology*

Inflection is explored from the start by learners but they master TL verbal morphology very slowly. By the end of the third cycle of data collection, most Moroccan Arabic learners have identified the copula, the auxiliary e / a and at least two basic verb forms V+Ø et V+e, which express different aspectual and temporal values. In 10,

[10]

- I : vous n'y allez pas souvent ? (à l'école)
 [you do not visit (school) often ?]
 Z : une fois pour hanan /tõbe/ à l'école (...) mais /jãna/ les os /gõfil/ comme ça
 [Once for hanan fallen in school (...) but bones swollen like this]
 après moi /parti/ avec hanan pour /levwar/ l'institutrice
 [then me go with hanan to meet the teacher]

 Z : après moi /le parti/ mon docteur pour /il a fe/ les radios pour ma fille
 [then me go (to) my doctor for he x-rays my daughter]

1.5.2.2. Temporality in FBC

i)- *Verb morphology*

In the first sentences of Atlantic FBC noted by travellers, colonists or priests, verbs are written in the infinitive form like tenir in this sentence from Saint-Christophe (Mongin Ca., 1682 quoted by Jennings, 1995, p. 72).

[11]

Si moi pas tenir Louis moi mourir de faim
 [If I did'nt have Louis, I would die of hunger]

or in an ambiguous Ve form as *iurer*, *dérober* and *aller* in the following text from Pelleprat (1655) :

[12]

Seigneur, toy bien sçave que mon frère luy point mentir, point luy iurer, point dérober, point aller luy à femme d'autre
 [Lord, you know very well that my brother hasn't lied, hasn't sworn, hasn't] stolen hasn't been with woman of another]

Note also in this text *sçave* which exhibits a Vstem form.

The same holds true for Indian Ocean creoles where in the first written quotations of the language (see Chaudenson for 1981 for IOC texts). verbs appear in V stem form or in long form Ve or in the infinitive.

[13]

Moi voulé baiser ça négresse la
[Me want (to) make love (to) that negress] (1777).

As the last written sentence suggests modals like voulé are found earlier in verbal predicates in creole texts than other pre-verbal markers (see Véronique, 1995 ; Hazaël-Massieux, 1996, pp. 209 -220).

As far as it can be assessed, it would seem that in the first FBC texts temporality is mainly expressed by the very means identified for pre-basic and basic learner varieties in SLA : reliance on context and pragmatic means, on diagrammatic iconicity (the principle of natural order) and the lexical meaning of verbs. Possibly, aktionsart explains the distribution of V stem and Ve verb forms as in learner varieties. The development of the expression of tense, aspect and modality is a late development where both reanalysis and grammaticalization played their part.

ii)- FBC pre-verbal markers

Time reference in FBC is mainly expressed through a set of pre-verbal markers derived from periphrastic verbal means in use in dialectal varieties of French (Goodman, 1964 ; Gougenheim, 1929). However, different French based creoles have reanalyzed these forms in differing ways, leading to partially different systems for conveying temporal information (Goodman, 1964 ; Valdman, 1978 ; Hazaël-Massieux, 1996).

a- être/été (à) > te, ti

b- The grammaticalization of French periphrastic verbs and its subsequent evolution
pour/être pour > pu and aller/va > alé/va

These data, Véronique argues, support the idea that the organizational principles that underlying the Basic Variety may have also been active in the FBC continuum. In their early grammars, FSL and FBC share common properties under the partial influence of the iconicity leading to a simple form-function relationships in the domains of reference to existence, possession, and time inter alia.

Thus, based on what occurs in a creolizing community, we can imagine a possible scenario in which large numbers of hominid speakers of disparate protolanguages may have, at some time in history, come together and interacted intensively, the result being the emergence of communication system which was similar enough across idiolects to constitute a community language. Since creole languages are the result of convergence among idiolectal interlanguages, a window on language evolution may be derived from the study of adult second-language acquisition and incipient creolization.

However, because meaning and structure- building work in language evolution is done in conversational interaction among speakers, the best window on language evolution would come from the study of conversational interaction among second-language speakers within and across a communities. This would at least allow a micro inspection of the vetting process that takes place as various forms are proffered from this linguistic selection pool provided by the idiolects of interlanguage speakers.

1.6. What Exactly Evolved ? How Autonomous is Language from other Semiotic Systems ?

« The Multiplicity of Semiotic Resources in Conversational Interaction »

Charles Goodwin offers a perspective on language based on semiotic analyses in discourse. He points out that both traditional linguists and laymen unproblematically consider language a system, process and/or mental faculty that produces a rich and novel collection of complex symbolic objects such as sentences. Linguists argue that this entity is the proper and distinctive subject matter of their field. They generally believe that it and its evolution can be analyzed as isolated self-contained systems quite separate from social organization. Thus language is seen as simply a psychological process rather than a social one. Goodwin argues that language is not just

a symbolic system but a primordial form of human social organization ; indeed, the distinctive social organization that defines us as species. The basic unit required for the evolution of human language is not linguistic structure as a self-contained system, but instead a multi-party, multi-modal construction of distinctively human forms of action. This process encompasses not just language structure but an ecology of interacting sign systems clustered around the bodies of multiple participants building action together by attending to each other, the detailed organization of the talk in progress, the larger activities interlocutors are pursuing together, and frequently the relevant structure in the environment (Goodwin, 1981, 1994, 2000, 2003b).

Goodwin's study of the interaction among an aphasic person and normal interlocutors within a family brings into stark relief the multiple semiotic resources that are available for the brain to use as participants build action together within situated human interaction (Goodwin, 1995, 2003a). Because of a stroke in his left hemisphere Chil, can produce only three words : **Yes**, **No**, and **And**. **Yes** and **No** are what conversation analysts call Second Pair Parts (Sacks, 1995 ; Sacks & al., 1974), utterances that respond to a prior action constructed by someone else. Chil's vocabulary thus presupposes that he inhabits a world that has already been structured by the semiotic activities of others. By indexically using both the structure of prior talk, and the field of action and meaning it has made relevant (Schegloff, 1968) as a point of departure for his own action, Chil is able to get others to produce the talk that he needs to say what he wants to say.

The following, analyzed in more detail in Goodwin (in press) provides an example. Chil's son Chuck and daughter-in-law Candy are talking with him about the amount of snow the winter has brought to the New York area where Chil lives. After Candy notes that not much has fallen « this year » (which Chil strongly agrees with in talk omitted from the transcript), in line 11 she proposes that such a situation contrasts markedly with the amount that fell « last year ». Initially, with his « °yeah- » Chil seems to agree (in the interaction during the omitted talk Chil was strongly agreeing with what Candy was saying, and thus might have grounds to expect and act as though that process would continue here). However, he ends his agreement with a cut-off (thus visibly interrupting and correcting his initial agreement) and moves to strong, vivid disagreement in line 13. Candy immediately turns to him and changes her « last year » to « the year before last. » Before she finishes Chil (line 15) affirms the correctness of her revised version.

The participants have been discussing snow in the area where Chil Lives

1 Candy: You haven't had **that** much this year have you.

...((8 lines ommitted))

10 Candy:	But la [st year. Whooh!
11 Chuck:	[°mm
12 Candy:	[In the l [ast year-
13 Chil:	[°yeah- [No No. No.:
14 Candy:	er the year before la [-st.
15 Chil:	[Yes.



Despite his severely impoverished language Chil is able to make a move in the conversation that is both intricate and precise : unlike what Candy initially proposed in line 10, it was not « last year » but the « the year before last » when there was a lot of snow. Chil says this by getting someone else to produce just the words that he needs. The talk in line 14 is semantically and syntactically far beyond anything that Child could say on his own.

Despite his sparse vocabulary Chil in fact has access to a wide and important range of meaning making resources. These include gesture, body posture which can be used to construct participation frameworks with others, gaze, and frequently structure in his surrounding environment, as well a wealth of knowledge and experience shared with consequential others such as his family. He also uses affective prosody over nonsense syllables to display meaningful stance.

Intersubjectivity is accomplished by Chil and his interlocutors as practical action, that is as methods for constructing the shared meanings required to build in concert with others those actions that make up the lifeworld of a local community. This is done within interactively organized multi-party, multimodal participation frameworks, within which different kinds of signs can mutually elaborate each other. Misunderstanding and repair (Schegloff & al., 1977), within the framework of collaborative action, provide resources for establishing and holding each other accountable for congruent public interpretation of the signs used in interaction. Thus communication is achieved by embedding all available semiotic resources to build an ecology of meaning making processes. This and similar research, of course, does not capture the evolution of language, but it does allow the observation of the evolution of a communication system from which it may be possible to accrue at least an oblique understanding of what was involved in language evolution.

Goodwin demonstrates how eye gaze and phrasal breaks serve as semiotic resources in conversational interaction (Goodwin, 1981). Speakers' use listeners' eye gaze as an index of attention. If a speaker detects the absence of an attentional gaze and looks at the nongazing hearer, he/she will typically interrupt the prosodic contour of the utterance and abandon its emerging syntax. The hearer will then begin to move his or her gaze to the speaker who then begins a new utterance. If the speaker has not yet gazed at the nonattending listener, the speaker will normally pause, and the hearer will then move his or her gaze to the speaker.

Goodwin also shows how repair in conversational interaction may serve as a semiotic resource for language learning. In self repairs such as « we went t- I went to » ; « if he-(0.4) could- if you could » several facts about the language are implicitly displayed. From the rephrasing, information is given in about what forms may appear in a particular position and how these forms may contrast. In the repair, « my ;,son- my oldest son » a noun phrase is highlighted and information is provided about how it may be expanded with an optional adjective. Goodwin points out that in traditional linguistics such repairs would be discarded as performance failures, and if the utterance were to be used in a linguistic argument, it would be sanitized to remove the repair. But such « degenerate input » may in fact provide parsing information essential for acquiring the language. Thus, from this perspective, repairs are not noise ; they are semiotic resources that facilitate both language acquisition and language use (Goodwin, 1981).

Goodwin's perspective shows that brains involved in conversational interaction have many resources for making meaning. Brains are not required to interpret isolated sentences. For producing and comprehending conversation, brains can use eye gaze, facial expression, body movement, phrasal breaks, repair, and the environmental surround. A biological account of language that disregards these semiotic resources is by definition an inadequate account.

1.7. How does Modern Language Reflect what Evolved ? How Complete are « Complete Sentences » ?

« *Examining Argument Structure in Conversation : A Matter of Index Grounding* »

Lisa Mikesell's presentation concerns the presence or absence of overt arguments in conversational utterances. It is very much in the tradition of the research on oral discourse undertaken by Chafe, Hopper, Thompson, and others (discussed in the Convener's Perspective in Part II below). It also deals with issues of formulaicity presented in at the Roundtable by Alison Wray. Additionally, Mikesell's analysis reflects Goodwin's position that brains engage multi-semiotic resources to produce and to comprehend language.

Mikesell notes that English argument structure is generally assumed to consist of subject/agent-verb-object in which the subject and verb must be expressed, and if the verb is transitive, the object must be overtly expressed. This appears to be the case when the unit of analysis is the isolated sentence. However, when the language is examined in its discourse context, it becomes apparent that any of these forms may be absent/elided. What seems to be required is that the referent for the elided item be understood from context. Here the context might be prior speech in the discourse, the ecological surround in which the discourse takes place, gesture, eye gaze and/or shared background knowledge. Such recovery of reference is referred to as the indexical ground.

Mikesell offers a position on argument structure that is based on three tenets. First, the language that constitutes relevant data must be embedded in context. Second, the representation of semantic roles is considered available before grammaticalization occurs. Third, these representations can exist in the context outside the utterance. In the following example we see subject/agent ellipsis.

[1] JS : II61 [Pomerantz, 1984, p. 67]

01 E : ...n she said she- depressed her terribly
 02 J : Oh it's [terribly depressing.
 03 L : [Oh it's depressing.
 04 E : Ve[ry
 05 L : [But it's a fantastic [film.
 06 J : [It's a beautiful movie.

In line 01 of this excerpt, the event or object that is causing the depression is not realized. Because this excerpt is incomplete, it cannot be determined whether the elided element is indexically grounded in the context. However, the excerpt is interesting because we are able to clearly see how the participants in the conversation have no difficulty understanding what the referent is or how the elided element may be indexically supported in context, even if it is not clearly apparent for an outsider. The speaker changes the utterance mid-stream to a construction containing a verb, *depress*, which requires a subject, i.e. something that causes the depression. Once again, the elision of the subject goes unnoticed by the participants, and it is evident to observers as the conversation continues that the elided subject refers to a film.

It is important to note that the « elided » element in this particular excerpt is a minimally phonated element ; thus it is possible that it may have been uttered but not heard by the transcriber. While this is unlikely given the expertise and meticulous training these analysts receive, from the perspective of the hearer, which is our primary interest, the element might as well have been unspoken. Further, the presumed « elided » element is uttered in lines 02, 03, and 04. It was suggested that a possible reason for the « elision » of the element in line 01, but the occurrence of it in lines 02 through 04, may be that in these latter occurrences there are two morphemes, i.e. *it's* rather than one, *it*.

In English, an argument in object position can also be elided :

[2] Carsales 5 [Ono & Thompson, 1995, p. 18]

1 D : I got to leave her on her own,
 2 .. let her see=
 3 you know.
 4 G : ...(9) yeah,
 5 D : ...even if she goes out with other men,

In this excerpt, the verb *see*, should take an object, yet there is no object present ; however, one can be inferred from context, i.e., it has indexical grounding. While it is not entirely clear what the object of the verb is or should be, the scene can be filled by the interlocutor who can interpret the appropriate meaning without the complete structure that we are accustomed to in written discourse or that we interpret in linguistic research. D is discussing how he should let his ex-girlfriend find out what she really wants without his interference. When he says, *I got to leave her on her own, let her see=*, from context we can provide a sufficiently accurate index, perhaps *let her see D as the man for her* or *let her see her decision as a bad one*. While the argument is not realized in the structure, an argument can be provided with indexical support. Thus, while it is true that many verbs occur in scenes that seem inappropriate from a theoretical framework, context often provides the indexical ground needed to fulfill these scenes, emphasizing again that structure should not be our sole focus when it comes to the study of language, or at the very least, the study of argument structure.

This instance of the verb *see* is different from intransitive *see*, which is fairly common in English. For instance, the following intransitive uses of *see* are structurally acceptable : *let me see, let's see, wait and see, I'll have to see*, etc. These uses of *see*, however, are not of the same category ; they seem to be formulaic in nature and highly restricted. This claim could be explored

in future research by examining *let me see* and *let her see* in a corpus in order to determine the frequency with which both constructions occur.

The following examples illustrate topic elision.

[3] [Nair, 2003, p. 43]

576 a : I'll tell you a friend of yours I met recently...
577 C : you **TOLD** me
578 a : oh

[4] Joyce & Stan [Schegloff, 1995]

35 S : [Any changes of uh : party affiliation er °>anything like
that(t) ? <
36 J : Uh : not at this moment. (u)When do I haftih : **tell** you by ?
37 S : Well the last day is May fifth. Fer the Ju :ne uh fourth
primary.

In excerpts, 3 and 4, we notice an extension of the usual topics locally selected by the verb *tell* which requires a subject (the teller), and object (the person being told), and a topic (the thing being told). In both instances, the topic is elided but it is supported indexically in line 576 in excerpt 3 and line 35 in excerpt 4.

Similar to the verb *see* discussed above, the verb *tell* also occurs both intransitively and as a formulaic expression : *how can you tell ?*, *time will tell*, *you're telling me*, etc. The *tell* that occurs in line 576 is such an instance where *I'll tell you* is a fixed expression often used to open up a conversation or change of topic. Again, these instances are highly formulaic and restricted in use and could be examined in corpora. *You told me* in excerpt 3 could also arguably be an instance of a formulaic realization whereby the topic is no longer required structurally and is therefore only « elided » semantically. However, whether this instance is viewed as a structural « elision » or as a formulaic realization, the point is still made that structure is flexible and dictated by use as long as indexical ground is maintained throughout discourse, i.e., it is not the structure that is most crucial but the indexical ground that the structure (or lack thereof) provides.

What Mikesell's discussion suggests is that an evolutionary or neurobiological account of language is not necessarily constrained to assume that structure is fixed. Semantic completeness may be achieved without full explicitness in the linguistic output. Furthermore, if semantics drives structure, rather than the reverse, it may not be necessary to account for 'missing' material in terms of a departure from some structural template. Rather, structure is contingent on what needs to be expressed. This approach will naturally result in inherent fuzziness regarding what the structures are in their own right, since their realization will be subject to contextual conditions. The brain will rely on indexical grounding through context to infer arguments. And if, as stipulated earlier, conversation is the primordial form of language, then in doing conversation, the brain does not operate under a requirement that subject and object arguments be realized. When context is removed, the structure may be necessary for meaning to be conveyed. When there is support by ecology, it is simply a matter of maintaining the indexical ground that is necessary for the communication to be effective.

If oral language for conversational interaction is what evolved, then the research reported here by Tao and Mikesell and earlier research by Hopper, Chafe, Thompson, Kalmar, and Mithun (see Part II, Convenor's Perspective, below) demonstrates that the language which we are responsible to in an evolutionary account is very different from the language that has been the prime concern of most mainstream linguistics.

1.8. What Exactly Evolved ? How do Oral Linguistic Signs become Reliably Indexicalized ?

« *Indexicalizing Language* »

John Schumann's Roundtable presentation is concerned with how we can come to know the nature of the invisible, nonmaterial phenomenon called language in order to understand its evolution and its neurobiological underpinnings. Language is an aspect of human behavior, and as Handy & Harwood (1973) argue, « human behavior involves transactions wherein something is regarded as

standing for or referring to something else. This process we name « sign behavior », or simply « sign ». Note that « sign » is not the name of the thing that is regarded as standing for something else ; « sign » is the name of the transaction as a whole » (i.e., is the short name for « sign process ») (p. 21). Understanding all phenomena means understanding via signs. An account of language, a system of signs, requires the use signs to refer to language itself. Thus in this section, we turn again to a discussion of sign relationships based on the work of Charles Sanders Peirce.

Peirce distinguished among three broad categories of sign : icons, indexes, and symbols. As Favareau has shown, sign behavior takes place across the biological spectrum, but here our analysis is limited to human sign use in order to simplify the discussion. Icons emerge as the human infants learn to distinguish items in their unlabeled world. An entity (a chair, a ball, a cookie) is a primary icon because it is totally similar to itself. Icons involve relationships not only of identity but also of similarity. Something may be interpreted as iconic of something else because of similarity in shape, color, or function to the original object. A child might, for example, see the stopper of an olive oil cruet as an ice cream cone (the rounded ahead and canonical base of both objects). Clouds are often interpreted as being iconic of other objects if the cloud formation resembles that object. The relationship of iconicity involves an interpretation of similarity between objects or events.

An index is a sign that has become associated habitually with something in the world. In biology, a peahen examines details of the peacock's tale as an index of the quality of the male's genes. In language, a pronoun indexes a previously expressed noun or a known person, place, or thing. A red traffic light indexes the requirement that traffic stop. The word « table » or « neuron » indexes the entity, table or neuron.

A symbol, in Peircean terms, is a sign that refers to another sign. For example, the word motivation refers other words such as desire, goal, incentive. Thus words can be indexes or symbols depending on whether they are interpreted as referring to things in the world or other words. Deacon (1997) has pointed out that when signs refer to other signs, and most particularly when words derive their meanings via reference to other words, a virtual world is created that moves beyond the physical world. This level of existence, where words refer to other words, has been called the symbolosphere (Schumann, 2004). Symbolic relationships consisting of word-word relationships are inherently relativistic because the words themselves may have multiple meanings and connotations that refer for their interpretation to other words with multiple meanings and connotations.

The most successful sciences appear to be those in which the phenomena of interest are easily indexicalized. If a phenomenon can be observed with the naked eye or with some visual prosthesis (e.g., telescope, microscope), and if words can unambiguously refer to the phenomenon, then facts can be accrued and some degree of certainty can be attained. Therefore, in science we are always looking for indexicalization ; we are trying to move from symbolic conceptualizations to indexical reference where the words we use can, as directly as possible, refer to phenomena and processes in the world. The kind of indexicalization we seem to prefer is through vision ; if we can see it, we are more likely to consider that we have truth than if we hear it, feel it, or taste it, or know it merely via definitions built on other definitions (e.g., symbolic conceptualizations).

An example of indexicalization through vision leading to a verifiable fact and therefore truth can be found in the history of neurobiology. At one time there was a debate as to whether the cortex was composed of a reticulum in which all units were connected without breaks between them and an opposing theory in which it was hypothesized that individual units were separated from others via small gaps. Golgi represented the former theory and Cajal the latter. Both received the Nobel Prize and presented their Nobel lectures proclaiming opposite points of view. Some years later the light microscope was invented and then the electron microscope that allowed us to visualize (i.e., see) actual neurons, and it was discovered that Cajal was correct. The brain was not a reticulum ; it consisted of billions of individual units (neurons) that were connected across spaces called synapses.

In physics a major task has been to « see » subatomic particles. This was first made possible in the 1950s and 60s by bubble chambers. These chambers were large tanks filled with hydrogen in which atoms collided and the particles which comprise them created trails of bubbles that could

then be photographed (Amato, 2003). This allowed an indexicalization of subatomic particles in which the tracks of the bubbles served as reliable indexes of the particles that they produce.

Moving in the other direction, telescopes have allowed us to indexicalize (i.e., see) large bodies at great distances. Speculations about the structure of the solar system have been settled by telescopic observations. The amazing images provided to astronomers via the Hubble Space Telescope are well known. Amato (2003) describes a less well-known telescope called The Chandra X-ray Observatory that can capture x-rays that penetrate the clouds and dust of the Milky Way. This telescope can detect x-rays from galaxies beyond the dust of the Milky Way. Thus the x-rays indexicalize for the scientists the existence of astronomical bodies at enormous distances and through galactic regions obscured by clouds and dust.

When we move from science to engineering, indexicalization becomes virtually automatic. Any technology that we produce is visible one way or another. Thus we can generate words to unambiguously refer to the various components of any piece of technology. Therefore, with the right expertise, it is possible to understand completely how some technological instrument operates. We make it from visible parts, and we observe its visible processes, and therefore, the technology is thoroughly indexicalized, and we can say we know it ; we understand it. If we confuse the production of technology and the certainty we have about its components and processes with what should be expected of all scientific endeavors, we become greatly disappointed. In the social sciences it is very difficult to get clear indexicalization ; frustration with this fact, and implicit comparisons of social scientific endeavors with those of engineering, frequently leave us feeling that the enterprises of sociology, anthropology, and psychology are either less deserving of the label « science » or that they must be done more rigorously so that they can behave like the sciences and technologies that are lucky enough to be easily indexicalized.

An example of a construct in the social sciences that is difficult to indexicalize is « motivation ». How do we point to motivation. We can certainly point to behavior that we assume is motivated, but we will still have not identified motivation. The problem with indexicalizing something like « motivation » is that the term can only be understood as it relates to and participates in a web of other terms such as : intention, incentive, desire, goal, reward, approach, avoid, motive, incentive motive, action tendency, wanting, liking, emotion, arousal, valence, appraisal, attitude. We come to understand motivation in relation to some or all of the things just listed. In other words, « motivation » makes symbolic reference to these words and these words have intrinsic symbolic relationships with « motivation ».

In second language acquisition it is recognized that motivation is important in becoming bilingual. Therefore, much research has been devoted to trying to understand motivation in second language acquisition. Researchers have investigated motivation from many perspectives with respect to second language learning : intrinsic motivation, extrinsic motivation, instrumental motivation, integrative motivation, value expectancy motivation and so on. They do this by devising questionnaires that elicit subject responses to items that the researchers believe reflect motivation. These questionnaires typically ask about the subjects' behavior, imagined behavior, or beliefs. For example a subject might be asked « If you had the opportunity, would you go to see a French movie ? », « Do you believe that French culture has contributed to North American life ». Based on responses to such questions, researchers attempt to quantify and categorize the subject's motivation. As is easily seen, in this process motivation is never indexicalized ; it is never seen. A motivation is merely inferred from the subjects expression of actual behaviors, imagined behaviors, beliefs, desires, and preferences. The phenomenon at issue remains embedded in webs of symbolic reference and inference. In these situations, as we have seen in all studies of motivation, closure or final answers are virtually impossible. And therefore research on motivation in second language acquisition continues ; each well-designed study provides an illumination of the phenomenon in an oblique way, but it never provides a final answer. Questionnaires cannot allow us to see motivation ; they simply provide another symbolic representation of it. No indexicalization occurs.

Now the question becomes how do we indexicalize language ? This is rarely seen as a problem. It is implicitly assumed that if one writes down something they want to say or that they believe they heard, then that representation provides an indexicalization of authentic language

which can be used for scientific inquiry into the nature of language. Such activity assumes that writing and spoken language have a strong correspondence. However, as we have seen, extensive research on spoken language by discourse and conversational analysts has shown that spoken language is quite different from that seen in written texts. It is less complex, often lacking subordinate clauses. Units of spoken language do not necessarily correspond to phrases and clauses as they are defined in written language. Spoken language largely consists of fragments, formulaic expressions, backups, repetitions, rephrasings etc.

Many of the characteristic aspects of oral language have been dismissed as disfluencies of performance. In other words, they are viewed as poor reflection of more systematic « grammatical » abilities. Thus normal linguistic behavior is seen to be deficient and degraded manifestation of an underlying competence. But in most of the history of linguistics, a distinction between spoken language and written language has not been maintained and in fact has become confounded. The primacy of the written form is characteristic of many linguistic traditions. For example, in Russian grammar texts, utterances that are illustrative of particular grammatical principles are taken from literature and are referenced by author (e.g. Tolstoy, Chekov, Lermontov etc.). The obvious implication here is that the written language contains the grammar of the language, and the spoken language is what is authentic.

Now how do modern linguists try to capture language, i.e., try to indexicalize it? Field linguists typically elicit utterances from participants in order to find instances of particular grammatical functions. Generative linguists create a sentence and then ask themselves whether it is well formed. In these two instances, we have language that either is elicited or created by the researcher. But how do we get language that actually occurs? Access to authentic language only became available with the development of audiorecording technology in the middle of the 20th century. However, the initial technology was large, cumbersome and expensive and was not amenable to the recording of natural conversational interaction. The technology began with wire recorders, moved on to reel-to-reel tape recorders, then to portable cassette recorders, and now video recorders that can capture both sound as well as visual aspects of the context including eye gaze, body position and the surrounding physical environment.

Indexicalizing language involved creating ways to visualize language, to see it, not just to hear it. In the '70s methodologies were created for transcribing data from the audiorecordings to written transcripts. These methods included microanalysis of the audio information through the careful transcription of utterances, repetitions, after thoughts, filled clauses, false starts, pauses, and incomplete phrases. All these so-called « performance problems » or deficits were maintained in the indexicalizations and not removed as they would be in writing. The question then becomes which is the canonical form of language. Which is the form of language that we must responsible to in neurobiological and evolutionary accounts of language? During the last half-century, the argument has been made that the sanitized sentence constitutes the linguistic competence appropriate for the proper study of language. But since that enterprise has not yielded satisfactory results, modern discourse analysts and conversational analysts start with carefully transcribed speech of actual language use with all its warts, and then ask the question, « what in this linguistic performance allows appropriate reference to be made by the speaker and to be understood by the hearer ».

This effort, however, merely indexicalizes language behavior. What we also want to do is to indexicalize language processing. Efforts in this direction are now being made using various neural imaging technologies. However, for those efforts to be successful, we must first be sure we have indexicalized the appropriate behavioral phenomenon. The relevant indexicalization is recorded natural conversation transcribed according to the procedures of conversational analysis.

2. Convener's Perspective

In attempting to answer the question, « What evolved in language evolution? », we must distinguish between what the brain evolved to do vis-a-vis language, and what it can do. The brain is capable of reading and writing, but clearly it did not evolve to do those things. The most reasonable guess is that the brain evolved to engage in conversational and perhaps narrative discourse. Language appears very different when researchers look at it horizontally and when they

look at it vertically. From the horizontal perspective, what has to be explained is the evolution of the sentence and its grammar. From the vertical perspective what has to be explained is the evolution of conversational or discourse structure. The content of common talk might reasonably be seen to involve exchanging information about social relations and about the states, attitudes, and properties of oneself and others. Such conversational interaction would be the primordial form of language, and much of this conversational interaction would also involve formulaic speech and simple topic-focus structure.

Lee and Schumann (2003, and see full discussion below) expand on an emergentist view of language evolution with the following scenario : Hominids interact verbally using a protolanguage lexicon. Attempts to express meanings consistently over time generate a complex adaptive system that structures utterances of the language. Through the vetting process inherent in verbal interaction among agents, grammatical structure is altered to fit the human brain. This obviates the need for the genetic assimilation of the UG. What ensues is the production of a series of cultural artifacts (technologies) : spontaneous oral language, composed oral language, and written language. The relaxed psycholinguistic conditions that exist during the production of composed language (oral or written) allow for the development of complex morphology and syntax. Spontaneous oral conversational interaction then is the primordial form of language, and it is what must be accounted for by theory of the evolution of language.

2.1. What does Oral Language look like today ?

In a review of two recent books on the evolution of language, Bickerton (2001) makes note of those « who write about language evolution in blissful ignorance of what it is that evolved » (p. 590). What evolved is a crucial question in the study language evolution. But it is not clear that at this point that we have an answer to this question.

I hypothesize that oral conversational interaction is what evolved, but since many linguists are unfamiliar with spontaneous oral language (« language in the wild »), we will begin with a short review of the literature which characterizes such language. Conversational analysts see the basic unit of language as the turn constructional unit. Turns are not isomorphic with sentences or with any unit of grammar, although sentences and their concomitant grammar can appear within turns and across them. So to understand what it is that evolved, it is necessary to explain the ability to do conversation. From the perspective of complex adaptive systems (Lee & Schumann (2003), and below), studying conversation is studying the interaction from which grammar is an emergent property. Grammar emerges as an epiphenomenon of conversational interaction.

Hopper (1998) sees grammar as emergent in the sense that it is the product of communicative interaction and not the cause of it. In conversational interaction certain forms (words, morphemes, word order) become frequent. These forms, in Hopper's view, are temporary solutions to the communication of certain meanings. They are always open to change, although such change may be diminished as a result of arbitrary standards set by a literary language, grammars, and dictionaries. What constitutes grammar are the frequent but temporary patterns in the language.

According to Hopper, much of grammar is formulaic, prefabricated, and routine. The patterns in language are subject to local variations and future alteration. Forms are not wired into the individual but exist among individuals as past and present patterns of communication. In discourse, utterances are not necessarily sentences, but instead they are frequently concatenations of words that have been encountered by the speaker frequently in the past and those that have been used by the speaker and his interlocutor previously in the current discourse. From Hopper's perspective, learning a language involves acquiring such routines, holistic constructions, patterns and their variations and coming to be able to use them effectively in discourse -- stringing them together, mixing them, creating them, altering them in different genres, and incorporating them from one genre to another. The process is never completed ; there is no end state ; grammar is never « in ».

Based on an analysis of dinner table conversations, Chafe (1985) argues that spoken language consists of idea units made up of about 7 items, which therefore closely matches short-term memory capacity. He notes that words per idea unit in writing average about 11. His comparison of oral and written texts demonstrates that in spoken language there is relatively little subordination, and

idea units are either expressed independently or are linked by coordinating conjunctions. He observes that sentences are, in fact, difficult to identify in spoken language. Mithun (1984) in an analysis of several languages found that subordination is rare in the discourse of those languages that have not developed literacy. Kalmar (1985) observed that the Eskimo language, Inuktitut, subordination (i.e. relative clauses) only appeared as it developed a writing system.

Thompson and Hopper (2001) present evidence that in conversational discourse, utterances are low in transitivity and frequently have only one participant, and those clauses with two participants are low in transitivity as measured by a set of ten transitivity parameters. The authors' data indicate that English conversation is overwhelmingly characterized by three types of constructions : intransitive verbal clauses (« I don't remember. I was belly-aching ».), copular clauses (« Trish is pregnant again. That's the whole point. She's still at home »), epistemic/evidential clauses (« I guess we are, I don't think it's workable ») (pp. 38-39). The authors argue that the structures that are typically considered in current linguistic research are extremely rare in normal conversation. Most ordinary talk is not about events or actions, but instead it is about states, attitudes, properties of people and situations, and assessments of people and situations.

In trying to determine what it is that evolved, it is necessary to imagine language before it was influenced by writing. Written language is language transformed by a cultural technology, and may be similar to foods that have been genetically altered (McWhorter, 2001) to have a longer growing season, to resist insects, or to last longer after harvesting. If sometime in the future, scientists were to try to explain how such food evolved and were unaware of the artificial genetic alteration, their accounts of the food's evolution would be inaccurate. The capacities needed to produce language for conversational interaction may be very different from those necessary to produce isolated grammatical sentences. Nevertheless, such sentences appear today in human speech, but they may be the result of the influence of writing on speech (Kalmar, 1985). Human brains may have evolved the psycholinguistic abilities to do conversation, but they may be also able to incorporate into the spoken language forms developed under the very different processing constraints that obtain in reading and writing. This is especially true when society provides individuals with years of exposure to and training in the altered forms of language - via grammar school, high school, university, and pervasively through the media.

Kalmar (1985) has pointed out that language can be altered not just by writing, but also by oral texts that have been composed (verbal rituals, chants, prayers, songs, spells, incantations, narratives, epics, etc.). Such composed texts are held in memory and could be grammaticalized overtime in ways that make them different from spontaneous speech. The altered forms in such nonspontaneous language would then gradually influence spoken language.

The upshot of this research is that oral language as it evolved in conversational interaction is very different from language that is planned, composed and written. Therefore, it would appear the kind of language that is mediated by the brain and that constituted the earliest forms of linguistic behavior would have been oral. Current linguistic research that does not strictly distinguish between oral and written language may provide accounts of grammatical structures that are perhaps marginally relevant to an evolutionary perspective and a neurobiological perspective on language.

The challenge facing scholars of language evolution is to find current situations of language use that more closely match what we imagine the conditions of language use were in its earlier evolution. Kalmar (1985) suggests that analyses be made of unwritten languages in societies that have not generated large numbers of composed texts. Pidgin languages and signed languages are obvious candidates, but in both cases the analyses would have to be of conversational interaction with all its warts, not simply sanitized sentences extracted from their natural habitat of multiple turns with the overlaps, repetitions and repair.

2.2. The Evolution of Language as a Complex Adaptive System : From Lexicon to Grammar

Diane Larsen-Freeman (1997) showed how Chaos Theory and Complexity Theory can address some questions in Second Language Acquisition (SLA). Lee and Schumann (2003) argue that these theories can also provide an account of the evolution of language. From the perspective of

Complexity Theory and emergentism, languages and the rules of languages may have emerged purely through interaction among hominids, without the need for language genes or grammar modules to evolve.

First, the authors introduce some general principles of complexity and emergentism, and then examine how the principles are compatible with languages and their rules. They then take a macroscopic perspective and explain how hominids equipped with an increasingly large lexicon may have evolved grammar.

2.2.1. Complex Adaptive Systems

Spontaneous pattern emergence is pervasive throughout Nature. There are ample examples that illustrate patterns that emerge spontaneously through self-organization, not dictated by any grand design or preordained rules. For example, the Bénard Convection, is a grid of hexagons that appears when a thin layer of water or oil in a pan is heated to a certain point. Neither the liquid, nor the pan, nor the heat has a hexagonal pattern in its natural state. The pattern emerges spontaneously through self-organization. A flock of geese forms a V-shaped flight pattern. No individual goose has this design in mind. Even though each individual interacts only with its immediate neighbor, this V-shaped pattern emerges. These examples show that patterns emerge from interaction among individuals, and these patterns transcend any individual's goal or design. The spontaneous emergence of patterns is a feature of Complex Adaptive Systems (CASs).

CASs emerge spontaneously, dynamically, unpredictably, and nonlinearly through local interaction among the components of the systems or the agents of the systems. Examples include slime molds, ant colonies, the immune system, the central nervous system, flocks of geese, ecosystems, climate, stock markets, economies, cities, countries, etc.

Lee and Schumann argue that languages are also CASs. Rules of grammar, phonology, and pragmatics are all patterns in languages, and these patterns have emerged spontaneously through linguistic interaction among humans. There is no theoretical need for us to have language genes in our genome or grammar modules in our brains. Languages are not part of our biology. They are not genetically transmitted but are, instead, culturally inherited. They are simply cultural artifacts.

But are languages really CASs? If languages conform to the principles of CASs, we will have strong justification for our claim. The first principle of CASs is that they are all open systems. Open systems are those that receive constant energy input from the outside. Conversely, a closed system is a system that does not get energy input (Prigogine, 1988). For example, a glass of water in a vacuum is a closed system. But a layer of water in a pan in the Bénard Convection is an open system into which heat energy input is constantly fed.

The second principle is that open systems evolve. When there is no energy input, the system maintains an equilibrium state in which no pattern emerges. This is the initial state of the Bénard Convection. When liquid in a heated pan has not yet reached the boiling point, energy flows into this system and it evolves to a near-equilibrium state, but the spontaneous pattern still does not appear. However, when the amount of energy input crosses a threshold, the system plunges into a state far from equilibrium, a state in which dynamic, spontaneous, and unpredictable patterns emerge nonlinearly. This is the state in which the grid of hexagons in the Bénard Convection emerges.

The question, then, is whether languages are open systems and whether they evolve. Our answer is that all living languages are open systems and that they do evolve. Because energy input to language systems comes from linguistic interaction among people, dead languages such as Latin or Sanskrit are closed systems. In the case of a pidgin, where there are no native speakers and interaction is limited, only minimal patterns may emerge. As the number of speakers increases and the pidgin is used in wider social functions, interactions increase. Diverse and unpredictable grammatical patterns dynamically emerge, and the pidgin can evolve.

Here we will list only a few of the other principles governing CASs, but we are confident that languages conform to all the principles of CASs so far identified by scientists (Holland, 1995 ; Johnson, 2001) :

- 1- A CAS contains a large number of agents.
- 2- CASs have multiple levels, and an agent acting at one level can serve as a building block for agents at a higher level.
- 3- Agents act locally and randomly in CASs.
- 4- Bottom-up control is realized through local and random interaction.
- 5- Agents in CASs have pattern-matching abilities.
- 6- CASs have positive feedback systems, and patterns are magnified by this feedback.
- 7- When a pattern is magnified, it can become locked in.

These principles of CAS are compatible with languages (for detailed discussion, see Lee, 2003)

2.2.2. The Emergence of Language

Earlier Lee outlined the biological prerequisites for language evolution. Once these prerequisites were in place, sounds could be concatenated into words, freeing hominid communication from reliance on holistic calls. Bates and Goodman (1997) have shown that the degree of grammatical development in children acquiring their native language is tightly linked to the number of words they know. This link is maintained throughout the major period of grammar growth, which extends from 16 to 30 months. At the end of this period, children have achieved the grammatical complexity that qualifies them as having moved beyond protolanguage to full-fledged language (Bickerton, 1990). At this point, the children have a lexicon of about 600 words.

Now let's imagine a hominid group that has attained all the prerequisites for language presented by Lee at this Roundtable. It develops a basic lexicon that allows it to produce a protolanguage. The speakers continue to develop new words and use them in interaction with one another, with the goal of communicating a wide range of meanings. As the lexicon increases, the conversational interaction generates massive feedback with respect to successful and unsuccessful communication. Feedback of this magnitude typically results in a self-organizing structure as an emergent property (Weber & Deacon, 2000 ; Steels, 1998 ; Batali, 1998 ; De Boer, 2000).

The emergent structure is grammar. It develops so that meanings are expressed in consistent ways, making the language easy to produce, easy to comprehend, and easy for children to learn (Kirby, 1998). It is much easier to mold language to the brain than to change the brain to accommodate language. Thus, the verbal interaction among agents in society leads to the development of a cultural artifact--grammar. This artifact is nonmaterial and invisible ; nevertheless, it exists in the world in the same way as other artifacts (e.g., canoes, buildings, carpets, dances, rituals). Grammars, like other cultural artifacts, vary enormously, but there are constraints on this variation. These constraints are the products of the architecture of the brain, the architecture of the body, and the architecture of conversational interaction among humans.

Language is neither *of* the brain nor *in* the brain. However, the brain processes the external world ; it produces behavior in the world. Language is of the world ; it is « out there. » Therefore, as with material things in the world, the brain can acquire language and use it (i.e., comprehend and generate verbal behavior). Language is not in the brain, but exists instead as a cultural artifact or technology between and among brains.

If language is a CAS, then grammar is not subserved by a dedicated neural module. Instead, it is an emergent property of a large number of human agents attempting to communicate with a large number of items in the form of words. Those interactions generate a powerful and pervasive vetting process in which individuals proffer forms (i.e., strings of words) in an effort to express meaning consistently over time. Successful attempts would be accepted and spread, and unsuccessful attempts would be rejected and forgotten. A polysynthetic grammar might result, or syntactic grammar might emerge. The grammar may be head-last or headfirst, pro-drop or non pro-drop, or adjective before noun, or the opposite. These outcomes would be various attractor states

into which the grammar would fall as a result of the massive positive and negative feedback generated. The grammar would then exist in the environment whenever people spoke and would be mediated by the brain (i.e., learned, comprehended, produced). Grammar would emerge from the feedback generated by communicative interaction. It would not be the result of a special neural linguistic module.

As Diane Larsen-Freeman (1997, p. 148) put it, « Language grows and organizes itself from the bottom up in an organic way, as do other nonlinear systems. While rules can be used to describe such systems, the systems themselves are not the products of rules. » Thus, we cannot ask, for example, *where* in the brain language, tennis, or the waltz reside. All these processes require the interaction of two or more humans and some material artifacts. In the same way, we see grammar (and language, in general) to be a collection of non-material artifacts processed by an individual's brain to communicate.

Let's briefly consider four types of evidence for this position :

- *Pidgins*-In pidginogenic social situations, to use McWhorter's (2001, p. 131) words, language is « crushed to powder. » Generally, pidgins arise when people speaking different native languages are thrown together with very limited access to a target language. This was the typical situation in American plantation societies. Initially, the speakers may have shared only a few dozen words from a superstrate language. Under these conditions, they stretched, altered, and interacted with that shared lexicon to eventually evolve a full-fledged language. In the regard, it is important to remember Véronique's hypothesis that pidginization, early second language acquisition, creolization, and grammaticalization (see below) may all be aspects of the same overall process.
- *Signed Languages*-Signed languages are a genuinely polygenetic phenomenon that has emerged many times in the past (Purlmutter, 2002). Within one or two generations, through interaction, signed languages develop into fully grammatical languages.
- *Grammaticalization*-In the process of language change and the development of pidgin and creole languages there is a general tendency for content words to become grammatical items, then clitics, and finally even inflectional affixes. It is likely that this progression also occurred in language evolution. This process involves a product of one stage of language evolution being fed back into the language system and altering the system itself. Such a feedback mechanism is a major characteristic of complex adaptive systems.
- *Simulations*-There have been several successful computer simulations of the emergence of language as complex adaptive systems (Kirby, 1998 ; Batali, 1998 ; Steels, 1998).

3. Conclusion

Based on the discussion above, the foundational questions that motivated the Roundtable, and the data and arguments in the various presentations, I believe the following considerations should inform evolutionary and neurobiological accounts of language.

How did language evolve ?

- 1- Although many biological changes may have been necessary for language to appear, it is possible that none of these changes involved genetic assimilation or neural instantiation of representations promoting or constraining grammar.
- 2- The structure of language as manifest in grammar may be a complex adaptive system that emerged as a cultural artifact and that exists between and among brains as they mediate its production, comprehension, and the acquisition.

What exactly evolved ?

- 1- As a research heuristic, I assume that the language, as it has evolved and as it was originally processed by the brain, was oral. The brain, with appropriate training, can acquire and mediate language as expressed in decontextualized print, just as it can acquire other knowledge and skills. However, the primordial form of language in our environment of

evolutionary adaptation was most likely spoken conversational interaction. The generativist tradition in linguistics never made a strict distinction in their data between speech and writing ; therefore, analyses are frequently of structures that probably did not face the brains of the initial users of fully fledged language. Understanding how the brain acquires and uses the language of literacy is certainly an important goal, but it is a different goal than that trying to understand what originally emerged as language and how the brain, then and now, processes that language.

- 2- Sign systems characterize communication throughout the biological world. The ability to make symbolic reference may be unique to human language, but it is unlikely to have spontaneously « appeared » within the species as a unique biological endowment. Rather, the ability for symbolic reference both emerged from and remains contingent upon the more primitive abilities of iconic and indexical sign exchange that are characteristic of organisms across the biological spectrum.
- 3- Language as it evolved and as it is used today incorporates multiple semiotic resources (eye gaze, facial expression, body position, gesture, phrasal breaks, repair, and the environmental surround etc.). Evolutionary and neurobiological accounts of language are in defect if they do not incorporate the plenum of the semiotic affordances on which interacting brains rely.
- 4- As iconic and indexical sign systems developed, it is likely that at some point, holistic language was characteristic of hominid communication. Modern language is a mixture of both holistic and analytical forms. Any biological account of language must deal with holistic language both as it might have existed in our environment of evolutionary adaptation and as it is currently processed, along with analytical forms, by the brain
- 5- Given that language likely evolved to serve conversational discourse, the appropriate indexicalization of language is found in transcripts that provide a microanalysis of natural conversation and that include a record of all the semiotic components involved in the interaction.

How does modern language reflect what evolved ?

Notions of « gapping » and « sentences » may only be marginally relevant to evolutionary and neurobiological accounts of language. More pertinent data and constructs may be found in databases of oral language corpora, the analysis of natural conversation, and in data on early SLA, pidginization, creolization, and grammaticalization.

References

- Agha (A.). 1998. « Concept and communication in evolutionary terms ». in : *Semiotica*, 116, pp. 189-215.
- Agajeenian (B.). 2001. Baldwinian organic selection and the origins of language. Los Angeles : University of California [PhD Dissertation].
- Allman (J.-M.). 1999. *Evolving brain*. New York, NY : Scientific American Library.
- Andersen (R.-W.), (ed.). 1984. *Second Languages. A cross-linguistic perspective*. Rowley, Mass : Newbury House Publishers.
- Angioy (A.), Stensmyr (M.), Urru (I.), Puliafito (M.), Collu. (I.), & Hansson (B.). 2003. « Function of the heater : The dead horse arum revisited ». in : *Proceedings of the Royal Society of London* [Available online at : [http://www.biol.sc.edu/~vogt/crab/Angioy-et-al-2003.pdf#search='dead%20horse%20arum\]](http://www.biol.sc.edu/~vogt/crab/Angioy-et-al-2003.pdf#search='dead%20horse%20arum)]
- Arends (J.). 1995. « Introduction ». in : Arends (J.), (ed.). 1995. *The early stages of creolization*. Amsterdam : John Benjamins.
- Baker (P.) & Corne (C.). 1982. *Isle de France Creole : Affinities and origins*. Ann Arbor : Karoma.
- Barbieri (M.). 2001. *The organic codes : An introduction to semantic biology*. Ancona : Pequod.
- Bardovi-Harlig (K.). 1999. « From morpheme studies to temporal semantics : Tense-aspect research in SLA » . in : *Studies in Second Language Acquisition*. 21, pp. 341-382.
- Batali (J.). 1998. « Computational simulations of the emergence of grammar ». in : Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds.). *Approaches to the evolution of language*. Cambridge, MA : Cambridge University Press, pp.405-426.
- Bates (E.) & Goodman (J.-C.). 1997. « On the inseparability of grammar and the lexicon : evidence from acquisition, aphasia, and real-time processing ». in : *Language and Cognitive Processes*, pp. 507-584.
- Becker (A.) & Veenstra (T.). 2003. « The survival of inflectional morphology in French-related Creoles : The role of SLA processes ». in : *Studies in Second Language Acquisition*, pp.283-305.
- Bickerton (D.). 1998. « Catastrophic evolution : the case for a single step from protolanguage to full human language ». in : Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds.). *Approaches to the evolution of language*. Cambridge : Cambridge University Press, pp. 341-358.
- Bickerton (D.). 2001. « Linguistics plays catch-up with evolution ». in : *Journal of Linguistics*, 37, pp. 581-591.
- Briggs (J.) & Peat (F.-D.). 1989. *Turbulent mirror*. New York : Harper & Row Publishers.
- Calvin (W.-H.) & Bickerton (D.). 2000. *Lingua ex machina*. Cambridge, MA : M.I.T. Press.
- Chafe (W.-L.). 1985. « Linguistic differences produced by differences between speaking and writing ». in : Olson (D.-R.), Torrence (N.), & Hildyard (A.), (eds.). *Literacy, language & learning : The nature and consequences of reading and writing*. Cambridge, M. A : Cambridge University press, pp. 105-123.
- Chaudenson (R.). 1981. *Textes créoles anciens (La Réunion et Ile Maurice). Comparaison et essai d'analyse*. Hambourg : Helmut Buske.
- Cinque (G.). 1999. *Adverbs and functional heads : A crosslinguistic perspective*. New York, NY : Oxford University Press.
- Cheney (D.) & Seyfarth (R.-M.). 1993. *How monkeys see the world*. Chicago, IL : Chicago University Press.
- Corballis (M.-C.). 2002. « Did language evolve from manual gestures ? ». in : Wray (A.), (ed.). *The transition to language*. New York, NY : Oxford University Press, pp. 161-179.
- Croft (W.). 1990. *Typology and universals*. Cambridge : Cambridge University Press.
- Deacon (T.-W.). 1997. *The symbolic species*. New York, NY : Norton (W.-W.) & Company.
- Deacon (T.-W.). 2003. « The heirarchic logic of emergence : Untangling the interdependence of evolution and self-organization ». in . Depew (B.-H.-W. & D.-J.), (ed.). *Evolution and learning : The Baldwin effect reconsidered*. Cambridge, Mass. : M.I.T. Press, pp. 273-308.
- Deely (J.). 2001. *Four ages of understanding : The first postmodern survey of philosophy from ancient times to the turn of the twenty-first century*. Toronto : University of Toronto Press.

- De Boer (B.). 2000. « Emergence of sound systems through self-organization ». in : Knight (C.), Studdert-Kennedy (M.) & Hurford (J.-A.), (eds.). *The evolutionary emergence of language : social function and origins of linguistic form*. Cambridge, MA : Cambridge University Press, pp.177-198.
- Degraff (M.). 1992. *Creole grammars and the acquisition of syntax : The case of Haitian*. Pennsylvania : University of Pennsylvania [PhD Thesis].
- Dietrich (R.), Klein (W.) & Noyau (C.). 1995. *The Acquisition of temporality in a second language*. Amsterdam : John Benjamins.
- Donald (M.). 1991. *Origins of the modern mind : Three stages in the evolution of culture and cognition*. Cambridge, Mass. : Harvard University Press.
- Duff (P.). 1993. « Syntax, semantics, and SLA : The convergence of possessive and existential constructions ». in : *Studies in Second Language Acquisition*, 15, 1, pp. 1-34.
- Durreleman (S.). 2000. « The architecture of the clause in Jamaican Creole ». in : *Generative Grammar in Geneva*, 1, pp. 189-240.
- Durreleman (S.). 2003. « The articulation of inflection in Jamaican Creole ». in : *Paper presented at the Summer Conference of the Society for Pidgin and Creole Languages*. Hawaii : University of Hawaii, August 14-17.
- Emmeche (C.). 1999. « The biosemiotics of emergent properties in a pluralist ontology ». in : Taborsky (E.), (ed.). *Semiosis, evolution, energy : Towards a reconceptualization of the sign*. Aachen : Shaker Verlag, pp. 89-108.
- Emmeche (C.). 2000. « Closure, function, emergence, semiosis, and life : The same idea ? Reflections on the concrete and the abstract in theoretical biology ». in : *Annals of the New York Academy of Sciences*, 901, p. 11.
- Fabbro (F.). 1999. *The neurolinguistics of bilingualism : An introduction*. Hove, East Sussex : Psychology Press.
- Favareau (D.). 2000. *State-matching, recognition and relevancy : The role of mirror neurons in establishing mutual intelligibility*. Los Angeles : University of California [Unpublished M.A. Thesis].
- Favareau (D.). 2001. « Beyond self and other : On the neurosemiotic emergence of intersubjectivity ». in : *Sign Systems Studies*, 30, 1, pp.57-100.
- Favareau (D.). 2002. « Constructing representema : On the neurosemiotics of self and vision ». in : *Semiotics, Evolution, Energy and Development Journal*, 2, 4, pp. 3-24.
- Favareau (D.). 2004. *A synthesis of biosemiotics and interaction analysis for the investigation of experience as a natural category*. Los Angeles : University of California [PhD Dissertation].
- Freeman (W.-J.). 2000. *How brains make up their minds*. New York, NY : Columbia University Press.
- Greenspan (S.-I.) & Shanker (S.-G.). 2004. *The first idea*. Cambridge, MA : Da Capo Press.
- Goodman (M.-F.). 1964. *A comparative study of Creole French dialects*. La Haye ; Mouton.
- Goodwin (C.). 1981. *Conversational organization : Interaction between speakers and hearers*. New York : Academic Press.
- Goodwin (C.). 1994. « Professional vision ». in : *American Anthropologist*, 96, 3, pp. 606-633.
- Goodwin (C.). 1995. « Co-constructing meaning in conversations with an aphasic man ». in : *Research on Language and Social Interaction*, 28, 3, pp. 233-260.
- Goodwin (C.). 2000. « Action and embodiment within situated human interaction ». in : *Journal of Pragmatics*, 32, pp. 1489-1522.
- Goodwin (C.). 2003a. « Conversational frameworks for the accomplishment of meaning in aphasia ». in : Goodwin (C.), (ed.). *Conversation and brain damage*. Oxford, New York : Oxford University Press, pp. 90-116.
- Goodwin (C.). 2003b. « Pointing as situated practice ». in : Kita (S.), (ed.). *Pointing : Where language, culture and cognition meet*. Hillsdale, NJ : Lawrence Erlbaum Associates, pp. 217-241.
- Goodwin (C.). in preparation. « Human sociality as mutual orientation in a rich interactive environment : Multimodal utterances and pointing in aphasia ». in : Enfield (N.-J.) & Levinson (S.-C.), (eds.). *The roots of human sociality : Culture, cognition and human interaction*. Oxford, UK : Berg Publishers.

- Goodwin (C.). In Press. « Interactive footing ». in : Holt (E.) & Clift (R.), (eds.). *Voicing : Reported speech and footing in conversation*. Cambridge : Cambridge University Press.
- Gougenheim (G.). 1929/1971. *Étude sur les périphrases verbales de la langue française*. Paris : Nizet.
- Hazaël-Massieux (G.). 1996. *Les créoles. Problèmes de genèse et de description*. Aix-en-Provence : Publications de l'Université de Provence.
- Hoffmeyer (J.). 1996. *Signs of meaning in the universe*. Bloomington : Indiana University Press.
- Hoffmeyer (J.). 2000. « The biology of signification ». in : *Perspectives in Biology and Medicine*, 43, 2, pp. 252-268.
- Holland (J.). 1995. *Hidden Order : How adaptation builds complexity*. Cambridge, MA : Perseus Books.
- Hopper (P.-J.). 1998. « Emergent grammar ». in : Tomasello (M.), (ed.). *The New psychology of language*. Mahwah, N. J. : Lawrence Erlbaum, pp.155-175.
- Jackendoff (R.). 2002. *Foundations of language*. Oxford : Oxford University Press.
- Jennings (W.). 1995. « Saint-Christophe : Site of the first French Creole ». in : Baker (P.), (ed.). *From contact to creole and beyond*. London : University of Westminster Press, pp. 63-80.
- Johanson (D.) & Edgar (B.). 1996. *From Lucy to Language*. New York : Simon and Schuster Publishers.
- Johnson (S.). 2001. *Emergence*. New York, NY : Scribner.
- Kalmar (I.). 1985. « Are there really no primitive languages ? ». in : Olson (D.-R.), Torrance (N.) & Hildyard (A.), (eds.). *Literacy, language and learning*. Cambridge : Cambridge University Press, pp. 148-166.
- Kaminski (J.), Call (J.) & Fischer (J.). 2004. « Word learning in a domestic dog : evidence for fast mapping ». in : *Science*, 304, 11, pp. 1682-1683 [June 2004].
- Kay (R.-F.), Cartmill (M.) & Balow (M.). 1998. « The hypoglossal canal and the origin of human vocal behavior ». in : *Proceedings of the National Academy of Science*, 95, pp. 5417-5419.
- Kelso (J.-A.-S.). 1995. *Dynamic patterns*. Cambridge, MA : M.I.T. Press.
- Kirby (S.). 1998. « Fitness and the selective adaptation of language ». in : Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds.). *Approaches to the evolution of language*. Cambridge, MA : Cambridge University Press, pp.359-383.
- Kite (G.-C.). 2000. « Inflorescence odour of the foul-smelling aroid *Heliconia muscivora* ». in : *Kew Bulletin*, 55, pp. 237-240.
- Klein (W.) & Perdue (C.). 1992. *Utterance structure (Developing Grammars Again)*. Amsterdam : John Benjamins.
- Klein (W.) & Perdue (C.). 1997. « The Basic Variety ». in : *Second Language Research*, 13, 4, pp. 301-347.
- Kull (K.). 1998a. « On semiosis, Umwelt, and semiosphere ». in : *Semiotica*, 120, 3-4, p. 12.
- Kull (K.). 1998b. « Organism as a self-reading text : Anticipation and semiosis ». in : *International Journal of Computing Anticipatory Systems*, 1, pp. 93-104.
- Kull (K.). 2000. « Active motion, communicative aggregations, and the spatial closure of Umwelt ». in : *Annals of the New York Academy of Sciences*, 901, pp. 272-279.
- Lai (C.-S.-L.), Fisher (S.-E.), Hurst (J.-A.), Vargha-Khadem (F.) & Monaco (A.-P.). 2001. « A fork-head-domain gene is mutated in a severe speech and language disorder ». in : *Nature*, 413, pp. 519-523.
- Larsen-Freeman (D.). 1997. « Chaos/complexity science and second language acquisition ». in : *Applied Linguistics*, 18, 2, pp. 141-165.
- Lee (N.). 2003. *Emergence of languages as Complex Adaptive Systems*. Los Angeles : University of California [PhD dissertation].
- Lee (N.) & Schumann (J.-H.). 2003. « The evolution of language and the symbolosphere as complex adaptive systems ». in : *Paper presented at the Conference of the American Association for Applied Linguistics* [Arlington VA, March 22-25].
- Lieberman (P.). 2000. *Human language and our reptilian brain*. Cambridge, MA : Harvard University Press.
- Llinas (R.). 2001. *I of the vortex : From neurons to self*. Cambridge : M.I.T. Press.

- MacNeilage (P.-F.) & Davis (B.-L.). 2000. « Evolution of speech : The relation between ontogeny and phylogeny ». in : Knight (C.), Studdert-Kennedy (M.) & Hurford (J.-A.), (eds.). *The evolutionary emergence of language : social function and origins of linguistic form*. Cambridge, MA : Cambridge University Press, pp. 146-160.
- MacNeilage (P.-F.) & Davis (B.-L.). 2001. « Motor mechanisms in speech ontogeny : phylogenetic, neurobiological and linguistic implications ». in : *Current Opinion in Neurobiology*, 11, pp. 696-700.
- Markos (A.). 2002. *Readers of the book of life : Contextualizing developmental evolutionary biology*. Oxford, New York : Oxford University Press.
- McWhorter (J.). 2001. *The Power of Babel : a natural history of language*. New York, NY : Times Books.
- Meyer (Ch.-F.). 1995. « Coordination Ellipsis in Spoken and Written American English ». in : *Language Sciences*, 17, pp. 241-69.
- Mufwene (S.-S.). 1996. « The founder principle in Creole genesis ». in : *Diachronica*, 13, 1, pp. 115-168.
- Mufwene (S.-S.). 2000. « What do pidgins and creoles tell us about language evolution ? ». in : *Paper presented at the International Conference on the Origin and Evolution of Languages : Approaches, Models, Paradigms* [College de France, Paris, September, 26-27].
- Mufwene (S.-S.). 2001. *The ecology of language evolution*. London : Cambridge University Press.
- Mufwene (S.-S.). 2002. « Competition and selection in language evolution ». in : *Selection*, 3.
- Nair (R.-B.). 2003. *Narrative gravity : Conversation, cognition, culture*. London : Routledge.
- Newmeyer (F.-J.). 2002. « Uniformitarian assumptions and language evolution research ». in : Wray (A.), (ed.). *Transition to Language*. New York, NY : Oxford University Press, pp.359-375.
- O'Grady (W.). 2003. « Language without grammar ». in : *Plenary talk presented to the American Association for Applied Linguistics (AAAL)* [Arlington, VA.].
- Ono (T.) & Thompson (S.-A.). 1995. « What can conversation tell us about syntax ? ». in : Davis (P.W.), (ed.). *Descriptive and theoretical models in alternative linguistics*. Amsterdam : Benjamins, pp. 213-271.
- Parkvall (M.). 1995. « The role of St Kitts in a new scenario of French Creole genesis ». in : Baker (P.), (ed.). *From Contact to Creole and Beyond*. London : University of Westminster Press, pp. 41-62.
- Parmentier (R.-J.). 1994. *Signs in society : studies in semiotic anthropology*. Bloomington : Indiana University Press.
- Peirce (C.-S.). 1931-1935. *Collected papers of Charles Sanders Peirce (Vol. 1-6)*. Cambridge, MA : Harvard University Press.
- Peirce (C.-S.). 1958. *Collected papers of Charles Sanders Peirce (Vol. 7-8)*. Cambridge, MA : Harvard University Press.
- Pfaff (C.), (ed.). 1987. *First and second language acquisition processes*. Rowley Mass : Newbury House Publishers.
- Pinker (S.). 1994. *The language instinct*. New York, NY : Morrow.
- Poldrack (R.-C.) & Packard (M.-G.). 2003. « Competition among multiple memory systems : converging evidence from animal and human brain studies ». in : *Neuropsychologia*, 41, pp. 245-251.
- Pomerantz (A.). 1984. « Agreeing and disagreeing with assessments : Some features of preferred/dispreferred turn shapes ». in : Atkinson (M.) & Heritage (J.), (eds.). *Structures in social action*. New York : Cambridge University Press, pp. 57-101.
- Prigogine (I.). 1988. « Origins of complexity ». in : Fabian (A.-C.), (ed.). *Origins : The Darwin college lectures*. Cambridge : Cambridge University Press, pp. 69-88.
- Purlmutter (D.). 2002. « How many times did language originate ? ». in : *Lecture, UCLA Center for Behavior Evolution, and Culture* [April, 29].
- Reader (S.-M.) & Laland (K.-N.). 2002. « Social intelligence, innovation, and enhanced brain size in primates ». in : *Proceedings of the National Academy of Science*, 99-7, pp. 4436-4441.
- Ross (J.-R.). 1967. *Constraints on Variables in Syntax*. [Doctoral dissertation, MIT., published in 1986 as *Infinite Syntax*. Norwood, N.J. : ALEX.].

- Sacks (H.). 1995. *Lectures on conversation vols. I and II*. Oxford : Basil Blackwell [Edited by Gail Jefferson, with an Introduction by Emanuel A. Schegloff].
- Sacks (H.), Schegloff (E.-A.), & Jefferson (G.). 1974. « A simplest systematics for the organization of turn-taking for conversation ». in : *Language*, 50, pp. 696-735.
- Savage-Rumbaugh (E.-S.), Shanker (S.-G.) & Taylor (T.-J.). 1998. *Apes, language and the human mind*. New York : Oxford University Press.
- Schegloff (E.-A.). 1968. « Sequencing in conversational openings ». in : *American Anthropologist*, 70, pp. 1075-1095.
- Schegloff (E.-A.), Jefferson (G.) & Sacks (H.). 1977. « The preference for self-correction in the organization of repair in conversation ». in : *Language*, 53, pp. 361-382.
- Schegloff (E.-A.). 1995. *Sequence organization*. [inpublished manuscript].
- Sebeok (T.). 1977. « Neglected figures in the history of semiotic inquiry : Jakob von Uexküll ». in : Sebeok (T.), (ed.). *The sign and its masters*. Lanham, MD : University Press of America, pp. 187-207.
- Sebeok (T.). 1990. « Sign science and life science ». in : Deely (J.), (ed.). *Semiotics 1990*. Lanham, MD : University Press of America, pp. 243-252.
- Sebeok (T.). 2001. « Biosemiotics : Its roots, proliferation, and prospects ». in : *Semiotica*, 134, 1, p. 18.
- Steels (L.). 1998. « Synthesizing the origins of language and meaning using coevolution, self-organization and level formation ». in : Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.), (eds.). *Approaches to the evolution of language*. Cambridge, MA : Cambridge University Press, pp. 384-404.
- Stensmyr (M.-C.), Urru (I.), Collu (I.), Celander (M.), Hansson (B.-S.) & Angioy (A.-M.). 2002. « Rotting smell of dead-horse arum florets ». in : *Nature*, 420, pp. 625-626.
- Stjernfelt (F.). 2002. « Tractatus Hoffmeyerensis : Biosemiotics expressed in 22 basic hypotheses ». in : *Sign Systems Studies*, 30.1, pp. 337-345.
- Taborsky (E.). 1998. *Architectonics of semiosis*. New York : St. Martin's Press [1st ed.].
- Tannen (D.). 1987. « Repetition in Conversation : Toward a Poetics of Talk ». in : *Language*, 63, 3, pp. 574-605.
- Thompson (S.) & Hopper (P.-J.). 2001. « Transitivity, clause structure, and argument structure : Evidence from conversation ». in : Bybee (J.-L.) & Hopper (P.-J.), (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam : Benjamins.
- Tomasello (M.). 1998. *The new psychology of language. Vol.1 : Cognitive and functional approaches to Language structure*. Mahwah, NJ : Lawrence Erlbaum associates.
- Tomasello (M.). 1999. *The cultural origins of human cognition*. Cambridge, Mass. : Harvard University Press.
- Tomasello (M.). 2002. *The new psychology of language, vol. 2 : Cognitive and functional approach is to Language structure*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Tomasello (M.). 2003. *Constructing a language*. Cambridge, MA : Harvard University Press.
- Turovski (A.). 2000. « The semiotics of animal freedom ». in : *Sign Systems Studies*, 28, pp. 380-387.
- Uexküll (J. von). [1957] 1934. « A stroll the world of animals and men ». in : Schiller (C.), (ed.). *Instinctive behavior : The development of a modern concept*. New York : International Universities Press, pp. 5-80 [C. Schiller, Trans.].
- Ullman (M.). 2001. « A neurocognitive perspective on language : the declarative/procedural model ». in : *Nature Reviews*, 2, pp. 717-726.
- Valdman (A.). 1978. *Le créole : structure, statut et origine*. Paris : Klincksieck.
- Vargha-Khadem (F.), Watkins (K.-E.), Price (C.-J.), Ashburner (J.), Alcock (K.-J.), Connelly (A.), Frackowiak (R.-S.-J.), Friston (K.-J.), Pembrey (M.-E.), Mishkin (M.), Gadian (D.-G.) & Passingham (R.-E.). 1998. « Neural basis of an inherited speech and language disorder ». in : *Proceedings of the National Academy of Sciences*, 95, 21, pp. 12695-12700.
- Véronique (D.). 1994. « Premières étapes de l'émergence des constructions grammaticales en français, langue étrangère ». in : Giacalone Ramat (A.) & Vedeovelli (M.), (a cura di). *Italiano lingua seconda / lingua straniera. Atti del XXVI Congresso della Società di linguistica Italiana, Siena 5-7 novembre 1992*. Roma : Bulzoni, pp. 139-151.

- Véronique (D.). 1995. « Acquisition des modalités en français langue étrangère et développement des modalités dans les créoles français ». in : Giacalone Ramat (A.) & Crocco Galès (G.), (eds.). *From pragmatics to syntax. Modality in second language acquisition*. Tübingen : Gunter Narr Verlag, pp. 59-82.
- Véronique (D.). 1997. « Le devenir des « petits mots » : *pour* dans quelques créoles français ». in : *Faits de langues*, 9, pp. 61-70.
- Weber (B.) & Deacon (T.). 2000. « Thermodynamic cycles, developmental systems, and emergence ». in : *Cybernetics and Human Knowing*, 7, 1, pp. 21-43.
- Wray (A.). 1998. « Protolanguage as a holistic system for social interaction ». in : *Language and Communication*, 18, pp. 47-67.
- Wray (A.). 2000. « Holistic utterances in protolanguage : the link from primates to humans ». in : Knight (C.), Studdert-Kennedy (M.) & Hurford (J.), (eds.). *The evolutionary emergence of language : social function and the origins of linguistic form*. Cambridge : Cambridge University Press, pp. 285-302.
- Wray (A.). 2002a. *Formulaic language and the lexicon*. Cambridge : Cambridge University Press.
- Wray (A.). 2002b. « Dual processing in protolanguage : performance without competence ». in : Wray (A.), (ed.). *The transition to language*. Oxford : Oxford University Press, pp. 113-137.
- Wray (A.) & Grace (G.-W). forthcoming. « The consequences of talking to strangers : evolutionary corollaries of socio-cultural influences on linguistic form ». in : *Lingua*.



From Calls to Words : How ethology can bridge the divide

By Eric A Salzen

Psychology Department,
Aberdeen University, Aberdeen, Scotland

May 2006

0. Introduction

Speculations on the origins and evolution of speech from animal calls and vocalizations have a long history (cf. Monboddo, 1773 ; Andrew, 1963) but there is an equally long history of their dismissal as « the ravings of an insane mind » (Smith, 1848) or an « amazing thing » to still believe (Calvin & Bickerton, 2001). This rejection of such biological origins for speech and language by linguists and comparative philologists is because there seems to be an unbridgeable divide between the contrasting and conflicting properties of :

- a- **Vocal signals** which are typically considered to be :
 - innate, conative, emotional, involuntary, & high intensity ;

and :

- b- **Speech symbols** which are typically considered to be :
 - learned, cognitive, rational, voluntary, & low intensity.

In addition there are contrasting neural systems for :

- a- **Emotional calls**, which are typically considered to be :
 - limbic, & right hemisphere systems

and those of :

- b- **Rational speech**, which are typically considered to be :
 - cortical, & left hemisphere systems.

Yet as Andrew (1963) wrote « It seems obvious that a vocal language could only arise as a modification of an existing system of vocalization ; the question is what kind of modification ? ». This divide between vocalization and speech, therefore, has to be bridged and bridged in each of the areas of Phylogeny, Ontogeny, and Neurology. The following is an outline of how this can be done using an ethological analysis of the nature of animal signalling, an early study of infant speech development which paralleled some ethological methodology and concepts, and a neuro-ethological analysis of vocalizations in monkeys and man. The principles of ethology and human ethology have been well presented by Eibl-Eibesfeldt (1975 and 1989) but seem not to be well known or to have been employed in discussions of the origin, evolution and development of speech. Thus the present paper brings these concepts and some existing ethologically oriented studies into the discussion. On the grounds of priority of publication I shall use some old studies to show that the answers have been available for a long time and that modern studies, which I shall cite, add detail and confirmation rather than conceptual insight to these old studies which are frequently ignored.

1. Phylogeny of speech and language

1.1. Animal vocalizations. *How can emotional vocalizations become referential words ?*

In an ethological analysis of the nature of social signalling Morris (1956) has shown that vertebrate signals are derived from « intention movements », with their accompanying autonomic support reactions, of behaviours that are aroused but cannot be fully performed because they are blocked (« thwarted ») or are in « conflict » with each other. This analysis is summarised in *Table 1* which I have taken from Morris and modified for an analysis of human emotional displays (Salzen, 1991).

Thwarting situations :

- I. Absence of indispensable stimuli following intense arousal
- II. Simple physical obstruction of aroused activity
- III. Simultaneous arousal of two or more incompatible tendencies

Primary responses to thwarting :

Somatic :

1. **Perseverance** - persistent approach and adjustment (I,II)
2. **Snap Decision** - capricious choice of response (III)
3. **Threshold Intention Movements** - initial element of response (I,II)
4. **Ambivalent Postures** - elements of both responses (III)
5. **Alternating Intention Movements** - successive responses (III)

Autonomic :

1. **Alimentary** - salivation increase or decrease, urination, defaecation
2. **Circulatory** - pallor, flushing, genital tumescence, fainting
3. **Respiratory** - changes in rate and amplitude, panting, gasps, vocalizations
4. **Thermoregulatory** - sweating, raised/sleeked feathers/hair
5. **Lacrimatory** - weeping

Secondary responses to thwarting :

1. **Displacement Activities** - irrelevant behaviour
2. **Redirection Activities** - response to another stimulus
3. **Regressive Activities** - immature responses
4. **Neurotic Inactivity** - loss of responsiveness
5. **Aggressive Behaviour** - intense approach and adjustment
6. **Stereotypic Activities** - repetitive movement patterns
7. **Visceral Dysfunction** - chronic autonomic imbalances

Table 1 : *Thwarted action state signalling*

This analysis of responses to thwarting is a modification of one given by Morris (1956) as the basis for understanding the social signalling displays of vertebrates. These responses will be selected in evolution if they cause the social partner to change behaviour in a way that removes the source of thwarting and enables performance of the thwarted behaviour. It is the basis of the Thwarted Action State Signalling (TASS) theory of emotion (Salzen 1991).

These signals of thwarting and conflict have evolved into quite remarkable complex and specialised displays used in courtship and agonistic encounters. e.g. the peacock display and bird song (cf. Eibl-Eibesfeldt, 1975). These displays are directly comparable with emotional displays in humans and I have published a « Thwarted Action State Signalling » (TASS) theory of emotion which accounts for human emotions in this way (1991). The changes in respiration that accompany these behaviours can give vocalizations - gasps, grunts, groans, screams, sighs, shouts, laughter, cries and calls - and these are shaped by the specific bodily actions so that these sounds may become characteristic of each display pattern (cf. Thiessen & al, 1980 ; see also Blumberg (1992) for the ultrasonic calls of Gerbils).

It is clear from this ethological analysis that a particular display signals the intended action and by its nature it also indicates the underlying motivational and activation state/s. The nature and orientation of the action in relation to the objects in the immediate environment (the physical and/or social situation), adds information that is both referential to the releasing stimulus source and also behaviourally meaningful to the onlooker/listener i.e. they add both referential and motivational information to the signal (cf. Marler & al., 1992 ; see also Leger (1993) for the role of context in animal signalling). Numerous studies have shown such referential meanings and some have even shown « audience » or social partner effects on vocalizations or responses to vocalizations such as predator alarm calls in chickens (Marler & Evans, 1994) as well as monkeys (Snowdon, 1993 ; Gouzoules & al., 1984) and apes (Jolly, 1996).

Many of these effects can be understood if the particular social partners arouse an additional motivational state (e.g. maternal, hierarchical, social security) that conflicts with the elicited response (e.g. escape, freezing. or hiding) and so produces a mixed or compound action-state and response that is distinctive of the particular conflict state. These social signalling processes with their referential capacities towards other animal species will also occur to social

partners and to inanimate objects. In thwarting theory terms the incipient response is the signal, while its meaning with respect to the stimulus object is given by the nature of the incipient response (signal) and the character or identity of the individual (signaller), i.e. the meaning of the signal for the recipient is a function of their particular relationships with the stimulus object and with the signaller and to the particular physical and social situational setting. In this respect vocal signals are no different from visual signals. This still applies in everyday speech where context is helpful for understanding. There is no reason in principle, therefore, why human speech should not have arisen from thwarted-state vocalizations being used in social interactions in relation to objects and hence for social cooperation in the use of objects, a view that is of long standing.

The importance of non-verbal signalling, including emotional expressions, for social interaction in humans is now well recognised and it is still the primary medium for the communication of personal states of thwarting i.e. for what are conventionally called « emotions ». Hence the social communication of our motive states and the efficacy of our social interactions still depends heavily on the non-verbal signalling that we share with our fellow primates. Swadesh (1971, chap. 5) notes that man still uses vocalizations in the form of exclamatory « words » such as « ouch », « wow », « whoops », etc. Otherwise in speech emotional state is well conveyed by intonation, prosody, and voice quality ; see review by Scherer (1992) of vocal affect decoding. One might well ask therefore why the spoken word should ever have arisen. One answer is that it arose from vocalizations being used for social interactions in relation to objects and hence for social cooperation in the use of objects, a view that is of long standing. The thwarting circumstances defined by Morris (1956) may involve activation of responses not only by species partners, but also by other species, and by objects with motivational significance. Thus the same social signalling processes with their referential capacities will operate in reactions to inanimate objects in social settings. There is no reason in principle, therefore, why human speech should not have evolved from vocalizations aroused in the collaborative use of objects. Indeed language reflects this object orientation so that the least literate person has a large lexicon referring to objects and their use but a limited lexicon for feelings and interpersonal relations. When words are used in interpersonal relations for feelings and emotions they are words for objects, events, and situations used in metaphor (Davitz & Mattis, 1964), simile, and analogy (especially in poetry). Skinner (1989) has pointed out that words referring to feelings and states of mind were first used to describe the behaviour or the situations that are now associated with them.

The implication of this social signal origin is that speech depends on thwarted action states for occurrence. Basically people speak when they want something and stop speaking when they achieve it. It may, of course, be only social recognition and attention as in social chit-chat, or for information as in social gossip. But in general speech drops out when behaviour proceeds smoothly - murmurings of sweet nothings are part of the early phase of sexual encounters. The success of speaking in ending thwarting and satisfying needs can make it an acquired self-reinforcing behaviour and become a pleasure in itself so that one comes to like the sound of one's own voice. Failure to achieve end-of-thwarting can equally result in the chronic talking of the boring person or the proverbial nagging spouse, both being in a behavioural bind in that the more they speak to obtain attention the less people listen. Talking is reputedly more frequent and persistent in women perhaps because men are more able both physically and socially to take actions to end their own thwarting than are women who so often depend on inducing men to act for them or cooperate with them. Children speak most in problem solving involving the use of tools where the problem is most difficult. They speak to themselves as if to a social partner and so « children develop a method of guiding themselves that had previously been used in relation to another person » (Vygotsky, quoted by Oatley, 1985, p. 54).

1.2. First Words. How could vocalizations evolve into words ?

Linguists may concede that referential vocalizations may be equivalent to phonemes (C or V sounds) or at best to syllables (i.e. V or CV or VC) but not to words, which are combinations of varied syllables. The ethologist might argue as follows. Ethological studies have identified an evolutionary process of « ritualization » in which thwarting displays assume a constant form and intensity in order to form unambiguous signals (Blest, 1961). Morris (1957) has suggested that because of this move to a « typical intensity » and a « typical form » the strength of the

mood or emotional state underlying the signal shows itself by the number of repetitions of the stereotyped signal. Repetition of elements is seen in bird calls and with variation of the elements produces bird song e.g. chaffinch song (Marler, 1959).

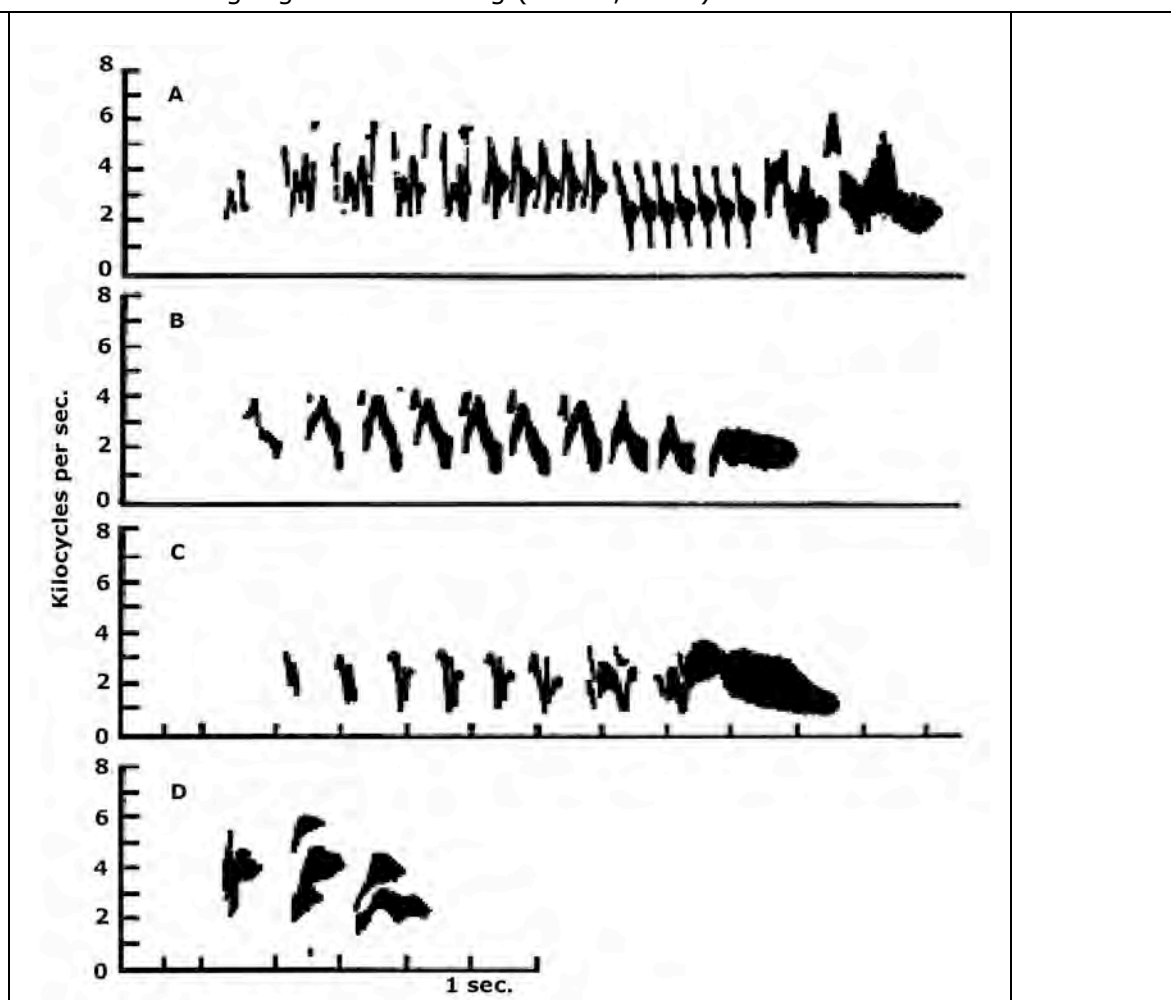


Figure 1. Examples of Chaffinch song (from Marler, 1959)

- A is the song of a wild British bird showing variation of elements.
- B is the song of an isolate reared British bird (from W.H.Thorpe).
- C is the song of an island bird (Tenerife).
- D shows « chink » calls of wild, isolate, (British) and island (Canaries) birds.

This shows how the elements of a repetitive call can become varied through learning giving complex multisyllabic vocalizations ; in this case to local song dialects.

It is evident (*Fig. 1*) that even in bird song variation of the elements can be learned and produces local dialects while such variation is absent with isolate rearing (Marler, 1959). The point here is not that bird song and words are homologous, but that the same biological and behavioural interactions can produce parallel and comparable effects i.e. a group of varied vocal elements in bird song and a group of varied CV-CV elements in human words. Repetition was seen in a gestural language invented spontaneously by some deaf children and reported by Goldin-Meadow & Feldman (1977). The children used gestures to represent objects with their associated action requirement (i.e. equivalent to verb-nouns). Diamond (1959) on the basis of comparative philology suggested that the first speech sounds were of the simple form consonant/vowel or c/v/c or c/v/c/v as in infant speech in English speaking cultures and in many other languages. Swadesh (1971) another comparative philologist has come to the same conclusions (see also Lewis, 1936 ; Ruhlen, 1994). Hence first words can be derived, in ontogeny and phylogeny, from repetition of thwarted-state calls.

Diamond (1959) has claimed that the earliest meanings of which we can find traces in comparative studies of language are actions such as « cut », « break », « crush », and « strike », and that man's first words were uttered to obtain assistance in strenuous activities as in hunting, killing, and perhaps simple tool fashioning. Interestingly a modern comparative

philology study by Foster (1996) suggests a proto-language with similar first words (but no reference to Diamond). See also Ruhlen (1994). These first words represent a combined action-object reference similar perhaps to some monkey vocalizations as already discussed. Diamond also pointed out that the wide variety of movements made by the brachiating arm of Apes and Man would be particularly effective in shaping calls through their physical actions on the chest and breathing. At first these actions, especially arm and facial actions, and any accompanying vocalizations shaped by these actions, would recruit assistance from the social partner, but ultimately voice alone would suffice and the phonations would refer to these actions in relation to present objects in their contexts, presaging the labelling use of words. Some writers on the origin of language (e.g. Revesz, 1956) have rejected biological theories of the kind proposed by Diamond because they imply that speech arose from the crude noises made by animals under strong emotional stress - cries of anger or pain. In situations and states of this kind man indeed continues to use crude vocalizations rather than speech.

There is the question, then, of whether speech began with the use of weaponry in hunting and war or whether it arose in less strenuous and less emotional social activities. Thwarted-state signalling as the basis for words is quite consistent with such social activities and with the view expressed by Oatley (1985, p. 53) that speech developed in relation to actions requiring social assistance with objects in what were evolutionarily new social or group activities for the human primate such as in tool and other artefact preparation and in use by males in hunting or by females in domestic activities such as in the preparation of food and clothing. This view has been restated by Reynolds (1993) who describes observations of tool making as a co-operative activity in modern Australian aboriginals. Pumphrey (1951) argued that language properly appeared when tools were used to make tools ; again a process requiring planning and social co-operation. Pumphrey's case rests on the view that words as symbols derive from degenerate stimuli or signals and that signs of such degenerate stimuli appear in the paintings and carvings made by Cro-Magnon man in the Aurignacian period and are fully advanced by the Solutrean period. The suggestion of speech arising in relation to domestic activities, or in infant care (Joseph, 1996), as with those of food-gathering (Bickerton, 1990) or in maintaining social networks (Dunbar, 1996) would also place women in centre stage and would certainly be consistent with the female superiority for the acquisition and use of speech and with her role in the development of speech in infants. In such activities the thwarting circumstances that give rise to speech would involve moderate arousal and coordinated actions that attempt to resolve the thwarting by variable orienting movements (cf. Primary Somatic Responses to thwarting in Table 1.). At first these actions and any accompanying vocalizations would recruit assistance from the social partner, but ultimately voice alone would suffice and the phonations would refer to actions in relation both to objects and places, presaging the labelling use of speech.

1.3. Lexicon & syntax. How could first words evolve into language ?

It is not difficult to see how a lexicon can develop in speech by variations and additions of the c and v elements. Farb (1974) has given examples of the use of repetition in modern languages to indicate plurality, intensity, size, and adverbs. These include the Indonesian word « igi-igi-igi-igi » for « multitude », the West African Ewe word « gada-gada-gada » for the intensity of a fire, and the Chinese words « tiantian » for « every day » and « kuai » and « kuai-kuaide » for « quick » and « quickly ». Diamond (1959) gives an extreme example with the word « ouatouijipakijouououououou » for « ocean » in Botocudo (a South Pacific island language). Swadesh (1971) came to the same conclusions from his somewhat similar comparative studies of languages, ancient and modern. Both writers have noted how vowels are variable and subject to change while consonants are stable and more consistent across languages (and indeed vowel variation is the main element in regional and local accents). The parallel with bird song is complete.

An earlier suggestion by Revesz (1956) was that the first words were verbs used as imperatives to get other people to do things. The previous section has reviewed speculations as to the nature of the earliest words and their suggestion that they were articulations signifying a demand or request for social actions in relation to an understood object. Thus first words would be, in essence, verb-nouns meaning to the social partner « do the action with the associated object to an understood visible goal » e.g. « stone » for « throw a stone at the thing/person at which I am looking/pointing ». Extra words for Agent, Object, Place etc would result from specific movements to indicate these entities where they were ambiguous and the one-word verb-noun failed to produce an appropriate response.

Thus separate words as nouns appear later specifying the agent and/or the goal (grammatical subject & object). Qualifier words - adjectives, adverbs, prepositions - are supplements and replacements for gestural and intonation cues of object specificity, urgency, intensity, direction, place, and time. Tenses, plurals, and genders are not needed in current contexts with shared knowledge but are needed when referring to agents, actions, and objects that are not present or contextually evident. Pidgin languages still largely lack them or develop them in forms independent of the language of origin. Bickerton (1990) gives an account of pidgin languages and an extensive treatment of the nature of syntax and its possible evolution. The need for maximum specificity is greatest, of course, in written language which can be highly ambiguous without complex syntax.

2. Ontogeny of speech and language.

2.1. First syllables. *How do infant vocalizations develop into speech sounds ?*

Studies of the first vocalizations and pre-speech sounds of infants raise the same issue of how high intensity involuntary infant cries can become low intensity voluntary speech. Modern studies of infant pre-speech and speech sounds (cf. contributions to Papousek & al., 1992 and the text by Vihman, 1996) have added surprisingly little of significance to the detailed and insightful study and review done by Lewis (1936). Lewis is rarely cited except as of historical interest and the essentially « ethological » character of his observations and interpretations is totally unrecognised. From his own study of one infant and from the data of previous « classical » German and French studies Lewis describes how :

- a- **Discomfort cries** give rise to the long and short « a » and « e » vowels ; momentary closure of the lips gives « w » ; the tongue on the hard palate gives « l » ; the tongue on the soft palate gives « ng » ; and vocal chord closure gives « r » in rhythmic crying.
- b- **Comfort sounds**, commonly given by the infant when lying on its back after feeding, result from the tongue falling back on the soft palate with or without air-flow giving the back-consonants « g » (hard), « k », « ch » (Scottish loch), « rr » (French r) and « grrg » (guttural r, absent in European languages). Residual salivation from the recently completed feeding gives gurgling consonants which may be coupled with gastric-wind air-flow (« Vomitiv-Lauten » of Kussmaul, cited by Lewis).
- c- **Incipient sucking** movements before, during, or after feeding periods give the labials « p », « b », and nasal « m » with lip closure, and the nasal « n » and dentals « t », « d » as the tongue presses against the front gums in sucking movements.

Thus Lewis describes beautifully and « ethologically » what thwarted-state theory invokes, i.e. vocalizations shaped by incipient actions arising from activated behaviours occurring in thwarted states of discomfort and pain, and from the declining actions in end-of-thwarting comfort states (e.g. anticipatory and residual feeding/sucking actions in hunger and satiation). This was proposed by Lewis in 1936, before the ethologists of the 1940's analysed the role of intention movements in signalling. More recently MacNeilage (1998) has derived similar first syllables from ingestion-related actions (see also the citations of Stark by Vihman, 1996).

Lewis also described clearly how the infant first discriminates adult speech meaning by its affective intonation and noted its importance for language acquisition, « It is not too much to say that a study of the child's responses to intonation provides one of the chief clues to the nature of his later use and comprehension of language » (p. 43). Lewis cites an earlier source for making this suggestion. Fernald (1992) in her article on « Meaningful melodies in mothers' speech to infants » makes a point of this too. She writes « Speech first becomes meaningful to the infant through prosody rather than through words, and through associations of sound and meaning that are not arbitrary » (p. 277). Her model (Fig. 13.5, p. 271) of changes in the functions of prosody over the first year is essentially what Lewis had described when showing the importance of intonation and melody throughout development - from cries to words, in conveying meaning to the infant before phonetics are responded to, and in mediating the shift to phonetics.

2.2. Babbling. *How do infant speech sounds become voluntary ?*

There is the objection that cries are intense distance-calls and that it is the low-intensity non-cry vocalizations, which are also the by-products of respiratory changes, that are significant for signalling the infant's behavioural state of arousal, comfort, discomfort (Papousek, 1992).

Furthermore, Oller & Eilers (1992) believe that involuntary sounds do not have the potential of arbitrary meaning and that it is the voluntary infant squeals, growls, raspberries that give rise to the « *babbling* » from which voluntary syllabic vocalizations are derived. Lewis in fact suggested that the low-level comfort (and later the nasal discomfort) sounds and incipient sucking actions are the basis of babbling and come to be given for their own sake. He suggested that the process acts through the infant hearing its own sounds which then reinstate their particular associated motivational and sensori-motor states and so reinforce the emitted behaviour (cf. MacTurk & al., 1987) ; this is part of a general feed-back system that operates in the development of competence or coping skills. Thus Lewis may have provided the clue to how babbling bridges the two vocalization systems - involuntary and voluntary.

Babbling becomes involved in first words when adults pick up and use the same sounds made in babbling by speaking to the infant in « motherese » with simplified consonants and « *cvcv* » repetition. Typically the parent intercedes with reinforcing behaviours accompanied by repetition of particular babble sounds to give the double syllable words - *dada*, *mama*, etc. Subsequently when the infant makes these particular babbles again it hears itself making these parentally-reinforced and repeated sounds and so the action becomes self-reinforcing and in effect the infant practices satisfying motor actions. This sensori-motor rehearsal may be a sign of the developing voluntary motor control of the speech apparatus which is essential for the development of an appetite for verbal mimicry or imitation which is a major means of the proliferation of speech and language learning. It is this voluntary control of mimicry or imitation that enables the infant to acquire new words for use in instrumental learning and cognitive association.

2.3. First words. How do infant speech sounds become words ?

Lewis goes on to review and describe the first words of infants the majority being single or duplicated labials or labial-dentals, i.e. front consonants, with the « *a* » vowels in European languages. Lewis (p. 125) also cites Wundt for parallels throughout the world. There are also similar meanings across languages (see *Table 2* taken from Lewis) with words referring to the child itself, people around it, food, bed, and play (see also Clark & Clark, 1977). All are understandable from an origin in anticipatory sucking movements and are also consistent with MacNeilage & Davis's (1998) recent account of proto-word « *cv* » forms. They also show the syllabic repetition essential for development of a lexicon. « Motherese » or nursery- or baby-talk also follows similar patterns in many languages, namely simplified consonants, reduplicated syllables, minimal syntax words, and of course exaggerated pronunciation and intonation. Fernald (1992) cites Ferguson (1964) on elevated pitch and exaggerated intonation in motherese in six languages.

Lewis also discussed the roles of imitation and instrumental learning in acquiring new words. He noted that after having imitated a word the infant may not immediately use it instrumentally but continue with established expressive actions and vocalizations (e.g. « *e-e* » in effortful reach instead of « *ka-ka* » for cake). Lewis says « The new movement must, in fact, be made a more efficient tool for the child than his own accustomed act, before he will replace the latter ». This is where thwarted-action-state theory would suggest that a little frustration or thwarting encourages the use of the newly acquired speech vocalization (i.e. withhold the cake while prompting the word production) and illustrates the role that thwarting still plays in speech production and is the way that Skinnerian instrumental learning can play a part in language learning. Lewis goes on to show how new words are acquired by modification of the first words, by changes in their use, and by refinement of their meaning through imitation and learning (e.g. Lewis, p. 214, *Fig. 2.*).

As speculated in phylogeny so in ontogeny, the infants first words have whole sentence meanings « *do x with y* » etc. These are Diamond's action/object verb-nouns referring an object and to its action or use in respect of the given situation, sometimes referred to as « *holophrases* » (cf. Greenfield & Smith, 1976). Two word utterances follow, separating verb from noun with object naming words. Qualifiers and syntax follow in their course with evidence for imitation of phonetics, generalisation of syntactic elements (« *ed* » past tense, « *es* » plural, etc.), all being used either for instrumental purposes directly or for purposes of competence and mastery of language itself, which indirectly is also for operating on the social environment to facilitate the satisfaction of thwarted motivations.

Typical early front-consonant words spoken by children of various nationalities (European)								
[Adapted from Lewis, 1936, p. 126]								
Meaning		Form (In the notation of the various observers).						
Front consonant		m	n	p	b	t	d	w
1.	Mother	mama			babab			
2.	Nurse	amme	nana		bäbe	tété	dedda	
	Aunt	muhme					deda	
	Grandmother	amma			baba			
3.	Food	mum	nana	pap	bap			wawa
4.	Bed, sleep		nana		bye-bye	teitei	dodo	
5.	Child himself				baby			
6.	Father			papa	baba	tata	dada	
7.	Play :							
	- ball				ba			
	- clock					titta	didda	
	- dog							wauwau
	- pointing					tatta	da	
	- giving						da	
	- thanking					ta	da	
	- going out			pa	ba	tata	ada	

Table 2 : *Early front-consonant words. Adapted from Lewis (1936)*

This table shows the typical earliest words of infants - single or duplicated syllables with labial or labio-dental consonants from a variety of European languages. The meanings are associated with the situations of primary concern of the child - attendant people, food, bed, and play. Lewis describes how front-consonants of these words can be derived from sucking and feeding actions and how these word situational associations and interactions give rise to fixed meanings.

3. Neurology of speech and language.

Phylogeny requires a link between involuntary and voluntary vocalizations. Ontogeny shows the behavioural link between them. Does neurology also involve involuntary and voluntary systems with a connecting link ?

3.1. Motor control. *How does an involuntary vocal neural system come under the control of a voluntary motor control system ?*

Jurgens (1992) has proposed three subsystems for control of vocalizations in primates (Fig. 2) :

- 1- a **lower brain-stem** reticular formation system for motor coordination and possibly patterned routines of laryngeal, articulatory, and associated respiratory functions ;
- 2- a **mid-brain** and an **anterior limbic** sub-cortical and cortical system for initiating vocalizations in response to stimuli in relation to both internal and external states ;
- 3- a **voluntary motor** system - anterior cortex, basal ganglia, thalamus and cerebellum.

Jurgens goes on to note that damage to :

- 1- can give mutism e.g. a patient with a lesion of the peri-aqueductal grey and neighbouring tegmentum ;
- 2- can remove affective intonation e.g. a patient with anterior limbic cortex damage from ant. cerebral artery occlusion could make joy, anger, and pain cries but the intonation was flat and monotonous ;
- 3- can abolish learned utterances in humans - i.e. speech and hummed melodies (but not expletives or exclamations).

Most significantly, and citing Kuypers (1958), Jurgens points out that humans have a direct connection between the primary motor cortex and the nucleus ambiguus which contains the laryngeal motor neurons (Jurgens, 1976). Mammals lack this connection. This may be why humans have voluntary control over the form of their vocalizations while non-human primates appear to lack this ability and are unable to vary the form of a vocalization in instrumental learning. Monkeys can be trained to vary their rate of calling (Sutton & al., 1974) but not the form of their calls (cf. Myers, 1976). Lesions of monkey cortical areas equivalent to the human

cortical speech areas do not affect their vocalizations (Myers, 1976 ; Ojemann & Creutzfeldt, 1987). There is some evidence that the monkey homologue of Wernicke's area is involved in the recognition of sound sequences and the discrimination of the calls of other monkeys from those of the individual itself (cf. Pinker, 1994).

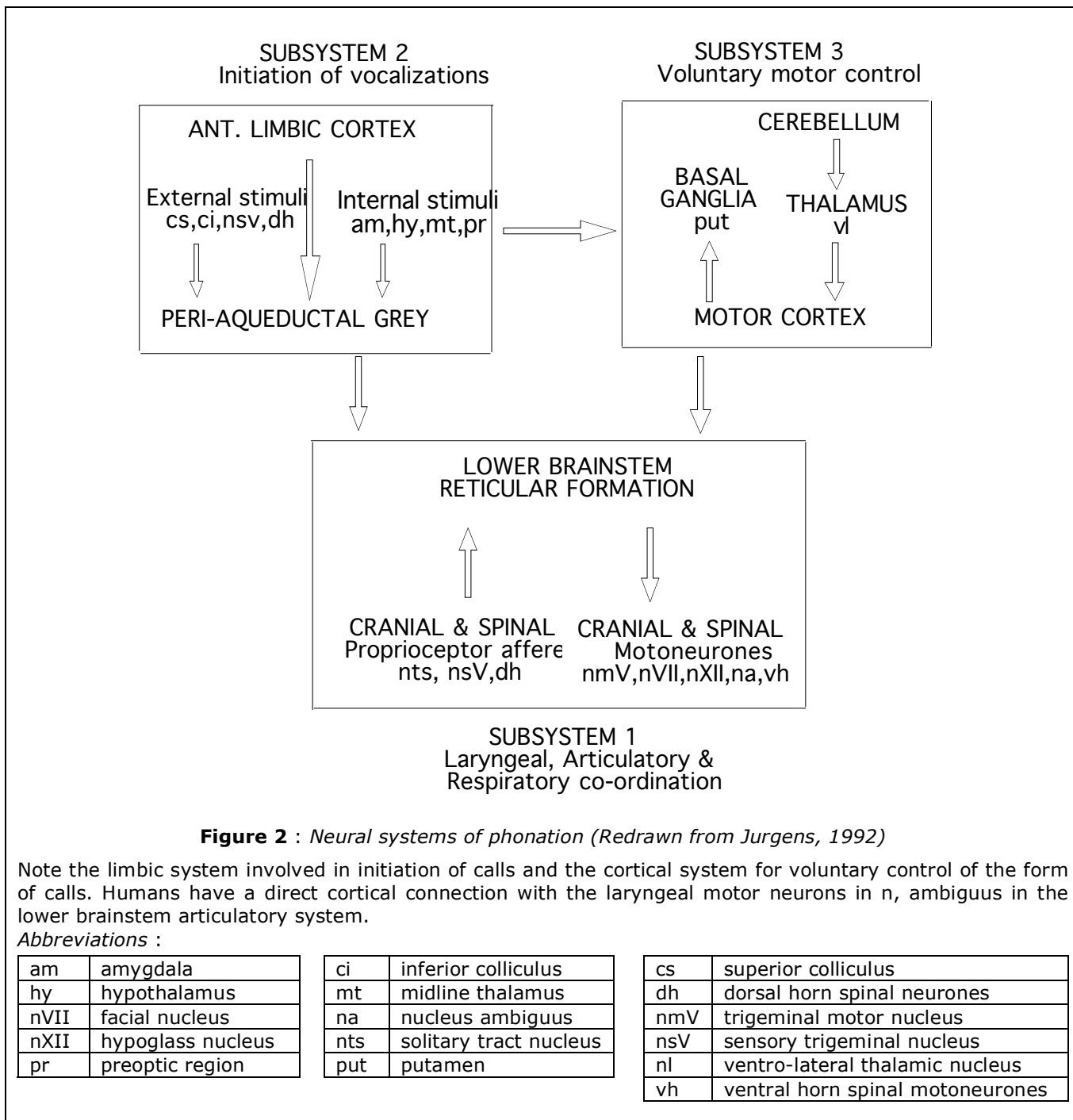


Figure 2 : Neural systems of phonation (Redrawn from Jurgens, 1992)

Note the limbic system involved in initiation of calls and the cortical system for voluntary control of the form of calls. Humans have a direct cortical connection with the laryngeal motor neurons in n, ambiguous in the lower brainstem articulatory system.

Abbreviations :

am	amygdala	ci	inferior colliculus	cs	superior colliculus
hy	hypothalamus	mt	midline thalamus	dh	dorsal horn spinal neurones
nVII	facial nucleus	na	nucleus ambiguus	nmV	trigeminal motor nucleus
nXII	hypogloss nucleus	nts	solitary tract nucleus	nsV	sensory trigeminal nucleus
pr	preoptic region	put	putamen	nl	ventro-lateral thalamic nucleus
				vh	ventral horn spinal motoneurones

MacNeilage (1998) has proposed the distinction between a medial intrinsic motivational language system including the anterior cingulate and supplementary motor areas for vocalization and the « cv » speech form or « frame » and a lateral extrinsic system including Broca's and Wernicke's areas for external exchanges involved in language « content » and learning.

The neurology, therefore, is consistent with the developmental data which distinguish between the emergence of consonant and vowel sounds from incipient involuntary discomfort/comfort vocalizations in low-level thwarting states and the development of word production by imitation and instrumental shaping of voluntary babbling sounds. The neuro-behavioural process by which the voluntary cortical system takes over and makes use of these involuntary sounds (presumably via babbling) is an interesting one and one that is shared with

what must similarly take place in arm-hand control of voluntary grasping and manipulatory movements. The voluntary control of the hand is established by visual feedback in exploratory movements while voluntary control of the voice must be similarly established through auditory feedback in babbling. Interestingly and perhaps significantly both are lateralized to the same (left) hemisphere and this is consistent with the notion that speech evolved in relation to social use and manipulation of objects. But in any event, this move to voluntary lateralized control is the emergent process, both phylogenetically and ontogenetically, in speech vocalization.

3.2. Lateralization. Why are the neural systems for speech and language lateralized ?

Levy (1969) first suggested left hemisphere for analytic serial information processing and output, and right hemisphere for gestalt, holistic, and parallel functioning. An equivalent physiological explanation (cf. Kinsbourne, 1978) is that the left hemisphere executes fine behavioural control through sequences of responses to specific stimuli which are characteristic of learned approach behaviour, while the right hemisphere mediates global unlearned responses to total stimulus arrays which are characteristic of avoidance and withdrawal behaviour (a more restricted suggestion than that of « will » *versus* « automatic » functions made nearly a century earlier by Hughlings Jackson (cf. Springer & Deutsch, 1985). Reviews of data by Davidson (1984) and Fox & Davidson (1984) have also suggested left frontal activity for approach behaviour and for positive affect, and right frontal activity for stopping behaviour and for negative affect. The left-right parietal asymmetry data also suggest a differentiation of verbal versus spatial, and sequential versus simultaneous or holistic perceptual and cognitive processing (Tucker, 1981).

Certainly the role of the left hemisphere in fine motor control both of hand and speech musculature is well established (e.g. Joseph, 1996) but the significance of the sequential character of these actions must be emphasised. Sequential approach movements are the essence of serial learning and this is evident in learned manipulative skills and is the basis of speech sequences and ultimately of syntax as Lashley (1951) appreciated long ago. The study by Curtiss (1980) of the development of speech in a 14-year-old isolate-reared girl (« Genie », see summary by Bickerton, 1990) suggests that the sequential processing abilities of the left hemisphere are necessary for the acquisition of syntax. This girl, following training, mastered one- and two-word utterances and some more complex ones but they were in « telegraphic » style (as in some aphasics), impoverished in syntax words and never achieving normal syntax rules. Tests indicated that her language processing was being done in the right hemisphere even though she was right-handed. There is the suggestion here that sequential or syntactic language development must be linked with the developing processing systems of voluntary control of sequential arm-hand movements in the left hemisphere during reaching and grasping in the neonate (cf. Young, 1977).

3.3. Limbic cortex. Are the neural systems of emotion still involved in speech ?

The ethological explanation of speech implies that the autonomic effects of thwarted action states were, and perhaps still are, involved in the *instigation* of speech but that its *form* is determined by the somatic motor system which comes under voluntary control. Thus one might expect lesions of higher autonomic centres to affect the impulse to vocalize or speak rather than give specific speech defects. Such damage might give absence of vocalization in response to thwarting.

The anterior cingulate cortex is well placed to pass affective output of the sub-cortical limbic system to the effective systems of the frontal cortex (cf. Damasio & Van Hoesen, 1983 ; Kelley & Stinus, 1984). Lesions in this area in monkeys have no clear effects on spontaneous vocalizations apart from some reduction and loss in vocalizations (Myers, *op. cit.*) but they do affect operantly conditioned vocalizing (Sutton & al., 1974). Steklis & Raleigh (1979) reviewed the monkey lesion studies and concluded that the anterior cingulate region is of particular importance for conditioned vocalizations. Subsequently, Trachy et al. (1981) showed that anterior cingulate lesions affected the rate of conditioned responding but not the acoustical structure of the calls while spontaneous vocalizations were unaffected.

Direct electrical stimulation of this area in monkeys gives autonomic effects, vocalizations, and facial expressions reminiscent of hypothalamic stimulation (Myers, 1976). In humans such stimulation gives wakefulness and attention, exploratory type hand movements, mouth « tasting » movements, and compulsive gnawing of objects placed in the hand (Talairach & al.,

1973). Feeling experiences such as anguish, sadness, fear, or euphoria were also reported and Damasio & Van Hoesen (*op. cit.*) suggest that the stimulation incites approach behaviour despite the mixed character of the experienced states. This could represent appetitive intention movements while the affective experience suggests thwarted-action states.

Lesions of this area in humans may result in akinetic mutism and autonomic disturbances. Damasio & Van Hoesen (*op. cit.*) describe a patient with an infarct in this area who showed loss of normal facial expression of affect as well as lack of speech. This patient could speak and repeat speech but had no will or desire to speak. The facial expression was neutral. In time the patient recovered normal speech with some occasional failures and was able to report that during her mute period she had no anxiety and did not talk because she « had nothing to say », that her mind was « empty », nothing « mattered », and there was no « will » to reply to questions. If vocalizations and speech can be understood as part of actions in response to thwarting then it may be that the cingulate's contribution is to increase appetitive approach and intentional actions to which vocalizations are an accompaniment. Joseph (1996), in a comprehensive text, also assigns a role for the human anterior cingulate in the expression of emotion, in emotional learning and vocalization, and in « the initiation of motivationally significant goal-directed behavior ».

If this connection between frustration & conflict, vocalizations, and speech seems somewhat tenuous then some findings from brain imaging studies (Posner & Raichle, 1994 ; Deacon, 1997 ; Bush & al., 1998) may add strength. The anterior cingulate gyrus of the monkey has been shown to contain adjacent areas for « attention », « emotional vocalizations », and « autonomic responses » in a postero-dorsal to antero-ventral succession as shown in *Fig. 3*. (taken from Posner & Raichle, 1994).

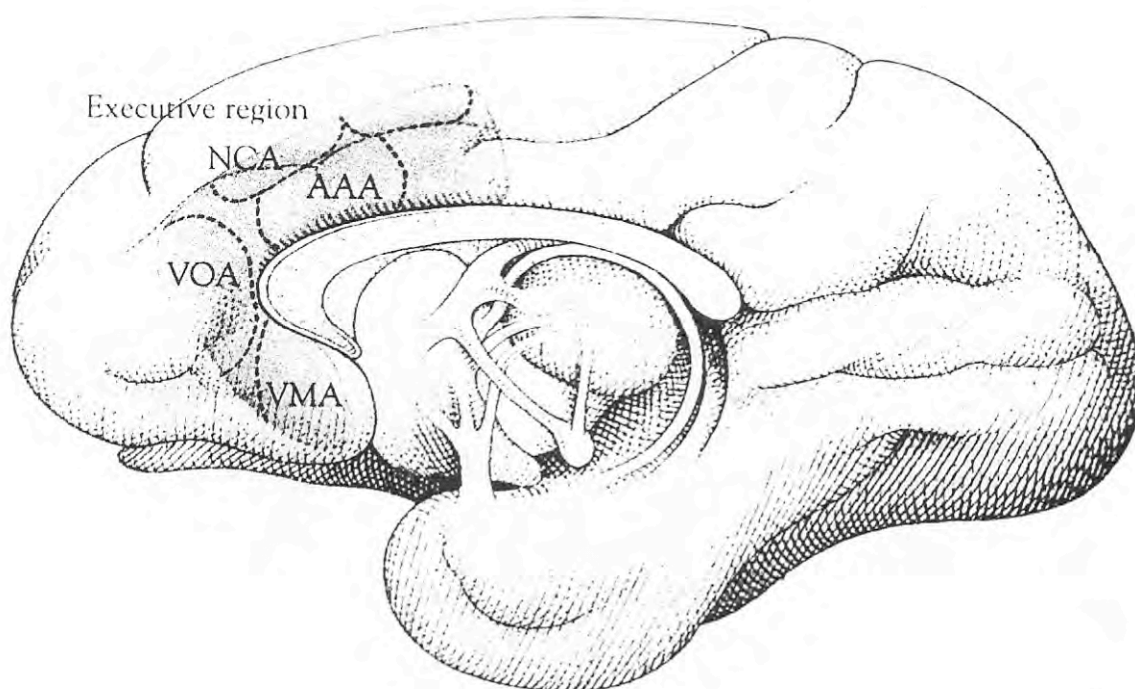


Figure 3 : Taken from Posner (M.-I.) & Raichle (M.-E.). 1994. *Images of the mind*. New York : Scientific American Library, p. 160.

The figure shows regions in the anterior part of the cingulate gyrus of the monkey brain that have been shown to be involved in particular functions : attention (AAA), pain (NCA), emotional vocalizations (VOA), and autonomic responses (VMA).

In humans a variety of imaging studies (PET and MRI) together show that there is a comparable succession, from the postero-dorsal to the antero-ventral regions of the anterior cingulate cortex, of areas that are peculiarly associated with « cognitive conflict », « emotional conflict », « emotional disorders », and « emotional pathologies » respectively (cf. *Fig. 4.*, taken from Whalen & al., 1998).

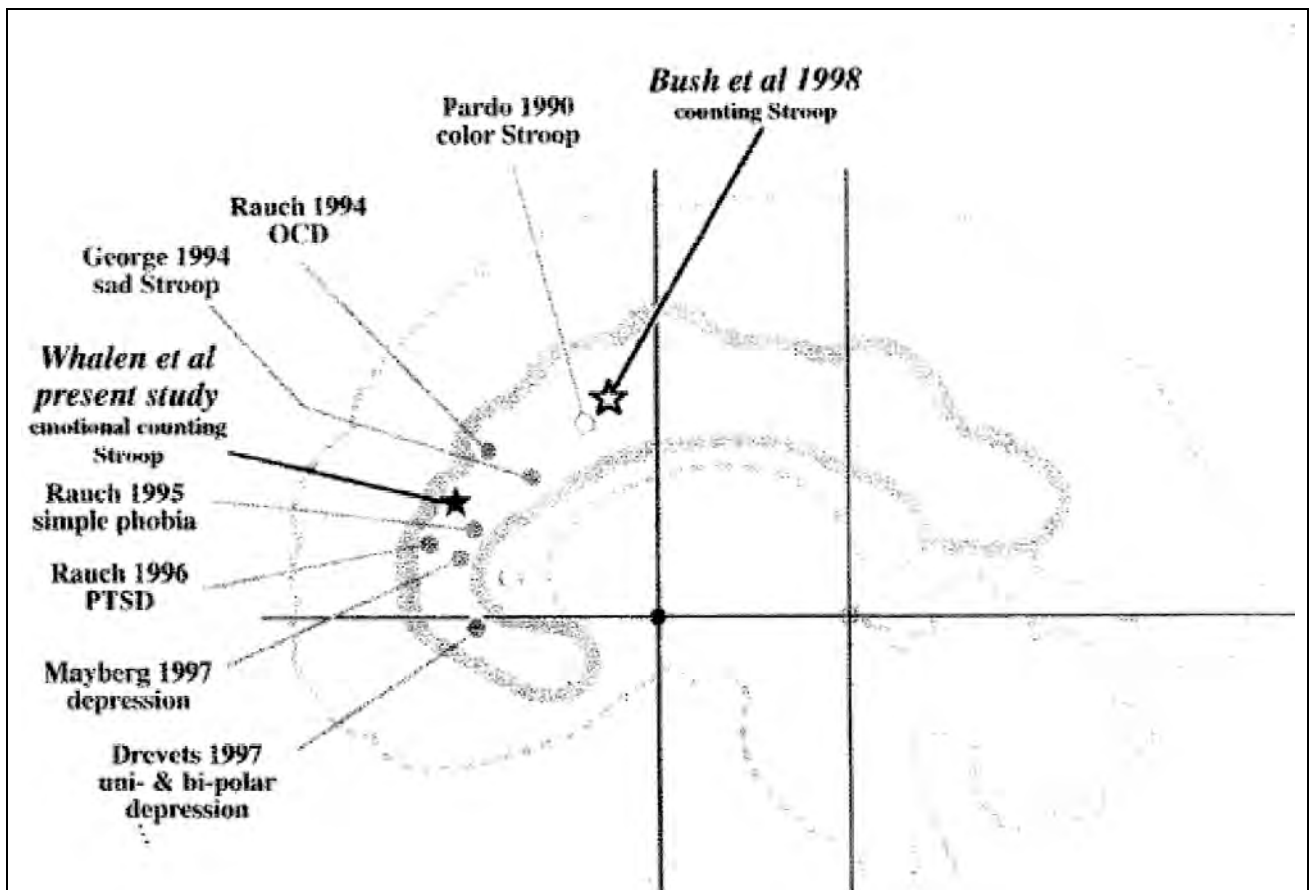


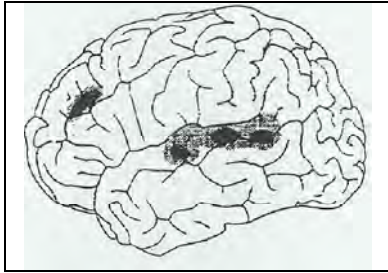
Figure 4 : Figure is taken from Whalen (P.-J.) & al. 1998. *Biological Psychiatry*, 44, pp. 1219-1228.

The figure shows a succession of areas, from the postero-dorsal to the antero-ventral regions of the anterior cingulate cortex, that are peculiarly associated with « cognitive conflict », « emotional conflict », « emotional disorders », and « emotional pathologies » respectively.

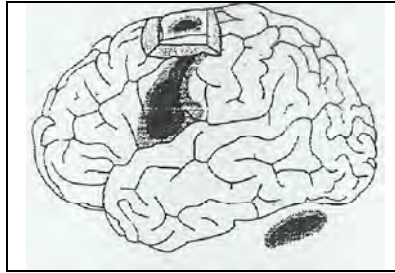
Schematic illustration of a representative sagittal section through the cingulate cortex depicting activation loci for the ecStroop and cStroop as well as selected emotional and cognitive studies. Studies involving emotional tasks or manipulations are represented by solid symbols, and tend to cluster in the rostral anterior cingulate affective division (ACed). Cognitive/motor tasks are represented by open symbols, and are within the more dorsal anterior cingulate cognitive division (ACcd). Highlighted in black ink are the activations from the present study within the ACad during performance of the emotional counting Stroop (ecStroop), and the within-group activation of the ACcd during performance of the cognitive counting Stroop (cStroop). Shaded in gray are selected cognitive and emotional studies referred to by first author, year of publication, and subject matter. For a more thorough presentation of cognitive versus emotional AC activation loci see Figure 6 In Bush et al (1998). CC corpus callosum.

Speech that requires the self-generation or the self-selection of words has been shown to give selective activation of the anterior cingulate in PET imaging (cf. Fig. 5., taken from Deacon, 1997). The same area has been shown to be selectively active in MRI imaging in a task requiring true/false responding i.e. generating a word for a falsehood (Langleben (D.-D.) & al., 2002). In this study the subject was required to give true or false yes/no responses to read questions using a button press (cf. Fig. 6). Reading the test questions arouses conflicting responses when lying is required since this involves the inhibition or blocking of the initial elicited (honest) response tendency and the instigation of the unelicited alternative (false) one. The assumption is that internal verbal responses are generated first and then translated into button presses so that speech processes are being invoked. Deception is essentially a conflict state that involves both unelicited word generation and the emotional arousal that accompanies lying. A comprehensive review of studies showing the sites of activation in the anterior cingulate by cognitive conflict, emotional arousal, and intrinsic word generation may be found in Bush & al. (1998).

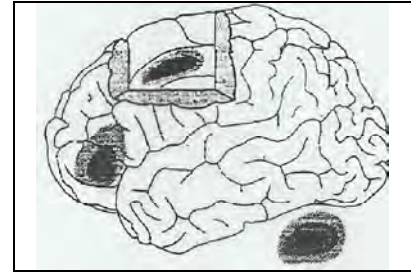
Positron Emission Tomography (PET)



1. Hearing a rapid sequence of words



2. Shadowing : repeating rapidly presented words (-1)



3. Word association : generating a verb for noun (-1&2)

Figure 5 : Taken from Deacon (T.-W.). 1997. *The symbolic species.*

The co-evolution of language and the human brain. London : Allen Lane, pp. 296.

Note an anterior cingulate area that, in addition to those areas mentioned by Deacon, is active only for self-generated words.

Positron emission tomography (PET) studies of language tasks reveal similar patterns to those seen with cerebral blood flow imaging. In these drawings of PET image activation, patterns are superimposed on outline drawings of human brains ; darker areas are more activated by the specific task. The three tasks are hierarchically constructed : 1 = Listening only ; 2 = repeating the input word ; 3 = producing a word associated with the word presented. The images to the right were produced by subtracting the results of those to the left to reveal differences. Notice the critical involvement of the ventral prefrontal cortex in the word association task. The activated area below the back of the brain indicates an additional intense involvement of the contralateral cerebellum (see also fig 9.1). The figure graphically summarizes data from Petersen & al. (1988) ; Posner & al. (1988) ; and reviewed in : Posner & Raichle (1994).

It seems, then, that the « conflict » regions of the anterior cingulate are also activated in self-generated or spontaneous speech and I suggest that inhibited or blocked, and conflicting motivational and cognitive arousal states instigate neocortical attentional and speech systems through this anterior limbic cingulate gyrus and so recruit and maintain the cognitive and social responses that may resolve the frustration and conflict states. Thus the role of this limbic area in speech is consistent with my Thwarted Action State Signalling (TASS) theory of emotion (Salzen, 1991) and the present proposed outline of the origin, evolution and nature of speech.

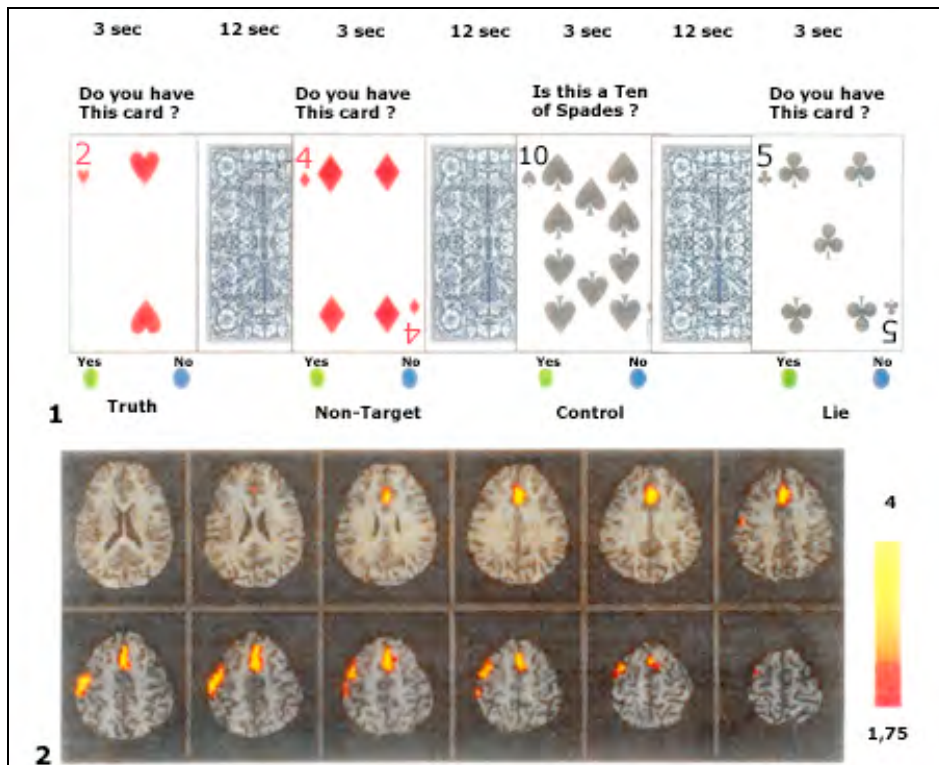


Figure 6 : Taken from Langleben & al. 2002.
NeuroImage, 15, pp. 727-732.

In these functional MRI scans the anterior cingulate is active when the subject is being untruthful (lying) i.e. the subject is required to give a false yes/no response using a button press. Reading the test questions arouses conflicting responses when lying is required since this requires the inhibition or blocking of the initial (honest) response tendency and the instigation of the unelicited alternative (false) one. The assumption is that internal verbal responses are generated first and then translated into button presses so that speech processes are being invoked.

1. A segment from the computerized GKT adapted for event-related [MRI « Each Truth » (2 of Hearts), « Lie » (5 of Clubs), and « Control »(10 of Spades) was presented 16 times, each Non-Target card - twice. Stimulus presentation time was 3s. interstimulus interval - 12s. total number of presentations - 88. Order of presentation was pseudorandom (randomly predetermined).
2. SPM(t) map projected over standart MRS template demonstrating significant increase in BOLD fMRI signal after Lie compared with Truth in the ACC, the medial right SFG, the border of the left prefrontal cortex, the left dorsal premotor cortex, and the left anterior parietal cortex.

4. Conclusion

In conclusion, there seems to be a paradox in the suggestion that speech and language must have arisen from vocal emotional displays yet are used largely for communication about objects and social actions in respect of such objects. Furthermore the neural representations of speech and language are voluntary motor, cortical, and left-hemisphere as in voluntary control of the hand and manipulation, while emotional activities at least for negative affect seem to be characteristic of right-hemisphere processing. Ethological theory suggests that speech and language are the outcome of thwarting and conflict both in phylogeny and ontogeny, and also in present adult function. That function is to communicate the nature of the thwarted actions so as to induce appropriate helping responses in social partners. The facility for voluntary control and use of verbal symbols as referents for the stimulus objects and their associated voluntary manipulation accounts for the precision of this communication and the proliferation and predominance of object reference in language. Voluntary control of both hand and voice involves external perceptual feedback and direct left-hemisphere cortical premotor neuron access to the final motor neurones of the appropriate musculature. Voluntary control of vocalizations develops in conjunction with voluntary control of reaching and manipulatory movements when these movements are thwarted or are unsuccessful. In this sense speech is an extension of the arm and of object manipulation by the recruitment and control of the social partner's manipulations when the individual's own efforts are inadequate for the aroused purpose. Close personal interaction by emotional signalling is still non-verbal as in non-human Primates. Distant social interaction uses object language which becomes essential for social interactions in large groupings and especially for the bureaucratic control of societies.

References

- Andrew (R.-J.). 1963. « The origin and evolution of the calls and facial expressions of the primates ». in : *Behaviour*, 20, pp. 1-109.
- Andrew (R.-J.). 1976. « Use of formants in the grunts of baboons and other nonhuman primates ». in : Harnad (S.-R.), Steklis (H.-D.) & Lancaster (J.), (eds.). *Origins and evolution of language and speech*. New York : New York Academy of Sciences, pp. 673-693.
- Baker (M.-C.). 2001. « Bird song research : the past 100 years ». in : *Bird Behavior*, 14, pp. 3-50.
- Beecher (M.-D.), Petersen (M.-R.), Zoloth (S.-R.), Moody (D.-B.) & Stebbins (W.-C.). 1979. « Perception of conspecific vocalizations by Japanese macaques. Evidence for selective attention and neural lateralization ». in : *Brain, Behavior & Evolution*, 16, pp. 443-460.
- Bickerton (D.). 1990. *Language and species*. Chicago : University Chicago Press.
- Blest (A.-D.). 1961. « The concept of ritualization ». in : Thorpe (W.-H.) & Zangwill (O.-L.), (eds.). *Current problems in animal behaviour*. London : Cambridge University Press, pp.102-124.
- Blumberg (M.-S.). 1992. « Rodent ultrasonic short calls : locomotion, biomechanics, and communication ». in : *Journal of Comparative Psychology*, 106, pp. 360-365.
- Bush (G.), Whalen (P.-J.), Rosen (B.-R.), Jenike (M.-A.), McInerney (S.-C.) & Rauch (S.-L.). 1998. « The counting Stroop : an interference task specialized for functional neuroimaging – validation study with functional MRI ». in : *Human Brain Mapping*, 6, pp. 270-282.
- Calvin (W.-H.) & Bickerton (D.). 2000. *Lingua ex machina. Reconciling Darwin and Chomsky with the human brain*. Cambridge, Mass. : M.I.T. Press.
- Cheney (D.-L.) & Seyfarth (R.-M.). 1982. « How vervet monkeys perceive their grunts : field playback experiments ». in : *Animal Behaviour*, 30, pp. 739-751.
- Cheney (D.-L.) & Seyfarth (R.-M.). 1990. *How monkeys see the world*. Chicago : University Chicago Press.
- Clark (H.-H.) & Clark (E.-V.). 1977. *Psychology and language*. New York : Harcourt Brace Jovanovich.
- Curtiss (S.). 1977. *Genie : a psycholinguistic study of a modern-day « wild child »*. New York : Academic Press.
- Daanje (A.). 1951. « On the locomotory movements in birds and the intention movements derived from them ». in : *Behaviour*, 3, pp. 48-98.
- Damasio (A.-R.) & Van Hoesen. 1983. « Emotional disturbances associated with focal lesions of the limbic frontal lobe ». in : Heilman (K.-M.) & Satz (P.), (eds.). *Neuropsychology of human emotion*. New York : Guilford Press, pp. 85-110.
- Davidson (R.-J.). 1984. « Affect, cognition, and hemispheric specialization ». in : Izard (C.-E.), Kagan (J.) & Zajonc (R.-B.), (eds.). *Emotions, cognition, and behavior, vol. 3*. London : Cambridge University Press, pp. 320-365.
- Davitz (J.-R.) & Mattis (S.). 1964. « The communication of emotional meaning by metaphor ». in : Davitz (J.-R.), (ed.). *The communication of emotional meaning*. New York : McGraw-Hill, pp. 101-112.
- Deacon (T.). 1997. *The symbolic species. The co-evolution of language and the human brain*. London : Allen Lane.
- Diamond (A.-S.). 1959. *The history and origin of language*. London : Methuen.
- Dunbar (R.). 1996. *Grooming, gossip and the evolution of language*. London : Faber & Faber.
- Eibl-Eibesfeldt (I.). 1975. *Ethology. The biology of behavior*. New York : Holt, Rinehart & Winston.
- Eibl-Eibesfeldt (I.). 1989. *Human ethology*. New York : Aldine de Gruyter.
- Evans (C.-S.) & Marler (P.). 1992. « Female appearance as a factor in the responsiveness of male chickens during anti-predator behaviour and courtship ». in : *Animal Behaviour*, 43, pp. 137-145.
- Farb (P.). 1974. *Word play*. London : Jonathan Cape.
- Fernald (A.). 1992. « Meaningful melodies in mother's speech to infants ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 262-282.
- Foster (M.-L.). 1996. « The reconstruction of the evolution of human spoken language ». in : Lock (A.) & Peters (C.R.), (eds.). *Handbook of human symbolic evolution*. Oxford : Blackwell, pp. 747-775.

- Fox (N.-A.) & Davidson (R.-J.). 1984. « Hemispheric substrates of affect : a developmental model ». in : Fox (N.-A.) & Davidson (R.-J.), (eds.). *The psychobiology of affective development*. Hillsdale, New Jersey : Lawrence Erlbaum, pp. 353-381.
- Gardiner (A.-H.). 1932. *The theory of speech and language*. Oxford : Clarendon Press.
- Gibson (K.-R.) & Ingold (T.), (eds.). 1993. *Tools, language and cognition in human evolution*. Cambridge : Cambridge University Press.
- Goldberg (G.). 1985. « Supplementary motor area structure and function : review and hypotheses ». in : *Behavioral & Brain Sciences*, 8, pp. 567-616.
- Goldin-Meadow (S.) & Feldman (H.). 1977. « The development of language-like communication without a language model ». in : *Science*, 197, pp. 401-403.
- Gouzoules (S.), Gouzoules (H.) & Marler (P.). 1984. « Rhesus monkey (*Macaca mulatta*) screams : representational signalling in the recruitment of agonistic aid ». in : *Animal Behaviour*, 32, pp. 182-193.
- Greenfield (P.-M.) & Smith (J.-H.). 1976. *The structure of communication in early language development*. New York : Academic Press.
- Hauser (M.-D.), Evans (C.-S.) & Marler (P.). 1993. « The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations ». in : *Animal Behaviour*, 45, pp. 423-433.
- Jolly (A.). 1996. « Primate communication, lies, and ideas ». in : Lock (A.) & Peters (C.-R.), (eds.). *Handbook of human symbolic evolution*. Oxford : Blackwell, pp. 167-177.
- Joseph (R.). 1996. *Neuropsychiatry, neuropsychology, and clinical neuroscience*. Baltimore : Williams & Wilkins [2nd ed.].
- Jurgens (U.). 1976. « Projections from the cortical larynx area in the squirrel monkey ». in : *Experimental Brain Research*, 25, pp. 401-411.
- Jurgens (U.). 1992. « On the neurobiology of vocal communication ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 31-42.
- Jurgens (U.) & Pratt (R.). 1979. « The cingular vocalization pathway in the Squirrel monkey ». in : *Experimental Brain Research*, 34, pp. 499-510.
- Kelley (A.-E.) & Stinus (L.). 1984. « Neuroanatomical and neurochemical substrates of affective behavior ». in : Fox (N.-A.) & Davidson (R.-J.), (eds.). *The psychology of affective development*. Hillsdale, New Jersey : Lawrence Erlbaum, pp. 1-75.
- Kinsbourne (M.). 1978. « Biological determinants of functional bisymmetry and asymmetry ». in : Kinsbourne (M.), (ed.). *Asymmetrical function of the brain*. New York : Cambridge University Press, pp. 3-13.
- Kuypers (H.-G.). 1958. « Corticobulbar connections to the pons and lower brainstem in man ». in : *Brain*, 81, pp. 364-388.
- Langleben (D.-D.), Schroeder (L.), Maldjian (J.-A.), Gur (R.-C.), McDonald (S.), Ragland (J.-D.), O'Brien (C.-P.) & Childress (A.-R.). 2002. « Brain activity during simulated deception : an event-related functional magnetic resonance study ». in : *NeuroImage*, 15, pp. 727-732.
- Lashley (K.-S.). 1951. « The problem of serial order in behavior ». in : Jeffress (L.-P.), (ed.). *Cerebral mechanisms in behavior, The Hixon Symposium*. New York : Wiley, pp. 112-136.
- Leger (D.-W.). 1993. « Contextual sources of information and responses to animal communication signals ». in : *Psychological Bulletin*, 113, pp. 295-304.
- Levy (J.). 1969. « Possible basis for lateral specialization of the human brain ». in : *Nature*, 224, pp. 614-615.
- Lewis (M.-M.). 1936. *Infant speech*. London : Kegan Paul, Trench, Trubner.
- MacNeilage (P.-F.). 1997. « Acquisition of speech ». in : Hardcastle (W.-J.) & Laver (J.), (eds.). *The handbook of phonetic sciences*. Oxford : Blackwell, pp. 301-332.
- MacNeilage (P.-F.). 1998. « The frame/content theory of evolution of speech production ». in : *Behavioral and Brain Sciences*, 21, pp. 499-547.
- MacNeilage (P.-F.) & Davis (B.-L.). 2000. « On the origin of internal structure of word forms ». in : *Science*, 288, pp. 527-531.
- MacTurk (R.-H.), McCarthy (M.-E.), Vietz (P.-M.) & Yarrow (L.-J.). 1987. « Sequential analysis of mastery behavior in 6- and 12-month-old infants ». in : *Developmental Psychology*, 23, pp. 199-203.
- Marler (P.). 1959. « Developments in the study of animal communication ». in Bell (P.-R.), (ed.). *Darwin's biological work*. Cambridge : Cambridge University Press, pp. 150-206.
- Marler (P.). 1991. « Song-learning behavior : the interface with neuroethology ». in : *Trends in Neuroscience*, 14, pp. 199-205.

- Marler (P.) & Evans (C.-S.). 1994. « The dynamics of vocal communication in birds ». in : Gajdusek (D.-C.) & McKhann (G.-M.), (eds.). *Evolution and neurology of language*. Amsterdam : Elsevier, pp. 81-89.
- Marler (P.) & Evans (C.-S.). 1997. « Animal sounds and human faces : do they have anything in common ? ». in : Russell (J.-A.) & Fernandez-Dols (J.-M.), (eds.). *The psychology of facial expression*. Cambridge : Cambridge University Press, pp. 133-157.
- Marler (P.) & Evans (C.-S.) & Hauser (M.-D.). 1992. « Animal signals : motivational, referential, or both ? ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 66-86.
- Monboddo (J.-B. Lord). 1773. *Of the origin and progress of language*. Edinburgh : Kincaid & Creech [vol. 1].
- Morris (D.). 1956. « The feather postures of birds and the problem of the origin of social signals ». in : *Behaviour*, 9, pp. 75-113.
- Morris (D.). 1957. « « Typical intensity » and its relation to the problem of ritualization ». in : *Behaviour*, 11, pp. 1-12.
- Mueller-Preuss (P.). 1979. « Processing of self-produced vocalizations by single neurons in the auditory cortex of the Squirrel monkey (*Saimiri sciureus*) ». in : *Experimental Brain Research*, suppl. II, pp. 146-151.
- Mueller-Preuss (P.), Newman (J.-D.) & Jurgens (U.). 1980. « Anatomical and physiological evidence for a relationship between the « cingular » vocalization area and the auditory cortex in the squirrel monkey ». in : *Brain Research*, 202, pp. 307-315.
- Myers (R.-E.). 1976. « Comparative neurology of vocalization and speech : proof of a dichotomy ». in : *Annals of the New York Academy of Sciences*, 280, pp. 745-757.
- Oatley (K.). 1985. « Representations of the physical and social world ». in : Oakley (D.-A.), (ed.). *Brain and mind*. London : Methuen, pp. 32-58.
- Ojemann (G.-A.) & Creutzfeldt (O.-D.). 1987. « Language in humans and animals : contribution of brain stimulation and recording ». in : Plum (F.), (ed.). *Handbook of physiology, section I : The nervous system, Vol.V. Higher functions of the brain, Part 2*. Bethesda, Maryland : American Physiological Society, pp. 32-58.
- Oller (D.-K.) & Eilers (R.-E.). 1992. « Development of vocal signaling in human infants : toward a methodology for cross-species vocalization comparisons ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 174-191.
- Papez (J.-W.). 1937. « A proposed mechanism of emotion ». in : *Archives of Neurology and Psychiatry*, 38, pp. 725-744.
- Papousek (H.). 1992. « Early ontogeny of vocal communication in parent-infant interactions ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 230-261.
- Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press.
- Petersen (M.-R.), Beecher (D.), Zoloth (S.-R.), Moody (D.-B.) & Stebbins (W.-C.). 1978. « Neural lateralization of species-specific vocalizations by Japanese macaques (*Macaca fuscata*) ». in : *Science*, 202, pp. 324-327.
- Petersen (M.-R.), Beecher (D.), Zoloth (S.-R.), Green (S.), Marler (P.), Moody (D.-B.) & Stebbins (W.-C.). 1984. « Neural lateralization of vocalizations by Japanese macaques : communicative significance is more important than acoustic structure ». in : *Behavioral Neuroscience*, 98, pp. 779-790.
- Pinker (S.). 1994. *The language instinct*. London : Allen Lane.
- Place (U.-T.). 1998. « Behaviourism and the evolution of language ». in : Man Cheung Chang, (ed.). *Current trends in the history and philosophy of psychology, Vol. 2*. Leicester : British Psychological Society, pp. 55-61.
- Ploog (D.-W.). 1992. « The evolution of vocal communication ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 6-30.
- Posner (M.-I.) & Raichle (M.-E.). 1994. *Images of the mind*. New York : Scientific American Library.
- Pumphrey (R.-J.). 1953. « The origin of language ». in : *Acta Psychologica*, 9, pp. 219-239.

- Revesz (G.). 1956. *The origins and prehistory of language*. London : Longmans.
- Reynolds (P.-C.). 1993. « The complementation theory of language and tool use ». in : Gibson (K.-R.) & Ingold (T.), (eds.). *Tools, language and cognition in human evolution*. Cambridge : Cambridge University Press.
- Ruhlen (M.). 1994. *The origin of language. Tracing the evolution of the mother tongue*. New York : Wiley, pp. 407-428.
- Ruhlen (M.), (ed.). 1994. *On the origin of languages : studies in linguistic taxonomy*. Stanford, Ca. : Stanford University Press.
- Salzen (E.-A.). 1981. « Perception of emotion in faces ». in : Davies (G.), Ellis (H.) & Shepherd (J.), (eds.). *Perceiving and remembering faces*. London : Academic Press, pp. 133-169.
- Salzen (E.-A.). 1991. « On the nature of emotion ». in : *International Journal of Comparative Psychology*, 5, pp. 47-88.
- Scherer (K.-R.). 1992. « Vocal affect expression as symptom, symbol, and appeal ». in : Papousek (H.), Jurgens (U.) & Papousek (M.), (eds.). *Nonverbal vocal communication : comparative and developmental approaches*. Cambridge : Cambridge University Press, pp. 43-60.
- Seyfarth (R.-M.) & Cheney (D.-L.). 1984. « The natural vocalizations of non-human primates ». in : *Trends in Neuroscience*, 7, pp. 66-73.
- Seyfarth (R.-M.), Cheney (D.-L.) & Marler (P.). 1980. « Vervet monkey alarm calls : semantic communication in a free-ranging primate ». in : *Animal Behaviour*, 28, pp. 1070-1094.
- Skinner (B.-F.). 1989. « The origins of cognitive thought ». in : *American psychologist*, 44, pp. 13-18.
- Smith (G.). 1848. *The origin and progress of language*. London : The Religious Tract Society.
- Snowdon (C.-T.). 1993. « A comparative approach to language parallels ». in : Gibson (K.-R.) & Ingold (T.), (eds.). *Tools, language and cognition in human evolution*. Cambridge : Cambridge University Press, pp. 109-128.
- Springer (S.-P.) & Deutsch (G.). 1985. *Left brain, right brain*. New York : Freeman [rev. ed.].
- Steklis (H.-D.) & Raleigh (M.-J.). 1979. « Behavioral and neurobiological aspects of primate vocalization and facial expression ». in : Steklis (H.-D.) & Raleigh (M.-J.), (eds.). *Neurobiology of social communication in primates*. New York : Academic Press, pp. 257-282.
- Sutton (D.). 1979. « Mechanisms underlying vocal control in nonhuman primates ». in : Steklis (H.-D.) & Raleigh (M.-J.), (eds.). *Neurobiology of social communication in primates*. New York : Academic Press, pp. 45-67.
- Sutton (D.), Larson (C.) & Lindeman (R.-C.). 1974. « Neocortical and limbic lesion effects on primate phonation ». in : *Brain Research*, 71, pp. 61-75.
- Swadesh (M.). 1971. *The origin and diversification of language*. Chicago : Aldine Atherton.
- Talairach (J.), Bancaud (J.), Geier (S.), Bardas-Ferrer (M.), Bonis (A.), Szikla (G.) & Rusu (M.). 1973. « The cingulate gyrus and human behaviour ». in : *Electroencephalography and Clinical Neurophysiology*, 34, pp. 45-52.
- Trachy (R.-E.), Sutton (D.) & Lindeman (R.-C.). 1981. « Primate phonation : anterior cingulate lesion effects on response rate and acoustical structure ». in : *American Journal of Primatology*, 1, pp. 43-55.
- Tucker (D.-M.). 1981. « Lateral brain function, emotion, and conceptualization ». in : *Psychological Bulletin*, 89, pp. 19-46.
- Vihman (M.-M.). 1996. *Phonological development. The origins of language in the child*. Cambridge : Blackwell.
- Whalen (P.-J.), Bush (G.), McNally (R.-J.), Wilhelm (S.), McInerney (S.-C.), Jenike (M.-A.) & Rauch (S.-L.). 1998. « The emotional counting Stroop paradigm : a functional magnetic resonance imaging probe of the anterior cingulate affective division ». in : *Biological Psychiatry*, 44, pp. 1219-1228.
- Young (G.). 1977. « Manual specialization in infancy : implications for lateralization of brain function ». in : Segalowitz (S.-J.) & Gruber (F.-A.), (eds.). *Language development and neurological theory*. New York : Academic Press, pp. 289-311.



Evolution of Language as a Gestural System¹

By Michael C. Corballis
Department of Psychology
University of Auckland
Auckland, New Zealand

May 2006

1. Introduction

Over 30 years ago, Gordon W. Hewes (1973) revived the suggestion, often attributed to Condillac (1971 [1746]), that language evolved from manual gestures, and not from animal calls. Although the idea was controversial at the time, and remains so, it continues to be advocated, and appears to have gained increasing acceptance (e.g., Arbib, 2005 ; Armstrong, 1999 ; Armstrong & al., 1995 ; Corballis 2002 ; Givón, 1995 ; Rizzolatti & Arbib, 1998 ; Rubin, 2005). From an evolutionary point of view, the idea makes some sense, since nonhuman primates have little if any cortical control over vocalization, but excellent cortical control over the hands and arms. There has been virtually no success in teaching even our closest relative, the chimpanzee, to speak, but quite good progress toward teaching chimpanzees and other great apes to communicate by a form of sign language (Gardner & Gardner, 1969), or by using visual symbols on a keyboard (Savage-Rumbaugh & al., 1998). The human equivalents of primate vocalizations are probably emotionally-based sounds like laughing, crying, grunting, or shrieking, rather than words. Human speech required extensive anatomical modifications, including changes to the vocal tract and to innervation of the tongue, and the development of cortical control over voicing via the pyramidal tract (Ploog, 2002). Most of the evidence suggests that these changes occurred late in hominid evolution, leading some to argue that language itself emerged suddenly, as a « catastrophic » event (Bickerton, 1995 ; Crow, 2002). Such an unlikely happening is obviated if language was initially based on manual gestures, which can be regarded as « language-ready » from much earlier in evolution, as also proposed by Arbib (2005). The gestural theory allows more time for linguistic complexities to emerge through natural (or sexual) selection.

In any case, language is not tied to speech. Since Hewes wrote, it has become abundantly clear that the signed languages of the deaf have all of the essential properties of true language, and are conducted entirely with movements of the hands and face (Armstrong & al., 1995 ; Emmorey, 2002 ; Neidle & al., 2001). Even in individuals with normal speech, moreover, manual gestures typically accompany speech, and are closely synchronized with it, implying a common source (Goldin-Meadow & McNeill, 1999). In many cases, in fact, gestures carry part of the meaning, especially where some iconic reference is needed, as in describing what a *spiral* is (Goldin-Meadow & McNeill, 2002 ; McNeill, 1992). Hand and mouth are further linked by the fact that, in most people, the left hemisphere is dominant both for manual action and for vocalization, a coupling often claimed as unique to humans (Corballis, 1991, 2003 ; Crow, 2002), even if cerebral asymmetry itself is not (Rogers & Andrews, 2002).

In this article, I develop a version of the gestural theory in which the proposed transition is not abrupt, but there is rather a gradual shift from dominance by manual gesture to dominance by the voice. This view is bolstered by recent developments in speech science, the neurophysiology of action perception, and behavioral genetics, all of which lend new support to the framework developed by Hewes. The first of these developments is the emerging view that speech itself is better understood as a gestural system, rather than as a system for the production and perception of communicative sounds. On this view, then, the emergence of speech did not depend on a switch from a gestural system to an acoustic one, but instead language

¹ *Acknowledgments.* I thank Michael Arbib, Karen Emmorey, and Michael Studdert-Kennedy for comments on earlier drafts, and Tecumseh Fitch, Maurizio Gentilucci, and Russell Gray for contributing to the development of ideas.

evolved from a gestural system in which manual gestures were gradually mingled with, and eventually (mostly) replaced by, articulatory ones. The second development is the discovery of the so-called « mirror system » in humans and other primates. There are reasons to suppose that speech may have emerged out of this system in humans, but in monkeys it appears to be concerned primarily with the production and perception of manual gestures. The third development has to do with the discovery of the role of a gene on chromosome 7, the *FOXP2* gene, in the control of articulation.

2. Language as gesture

Traditionally, speech has been regarded as made up of discrete elements of sound, called phonemes. It has been known for some time, though, that phonemes do not exist as discrete units in the acoustic signal (Joos, 1948), and are not discretely discernible in mechanical recordings of sound, such as a sound spectrograph (Lieberman & al., 1967). One reason for this is that the acoustic signals corresponding to individual phonemes vary widely, depending on the contexts in which they are embedded. This has led to the view that they exist only in the minds of speakers and hearers, and the acoustic signal must undergo complex transformation for individual phonemes to be perceived as such. Yet we can perceive speech at remarkably high rates, up to at least 10–15 phonemes per second, which seems at odds with the idea that some complex, context-dependent transformation is necessary.

These problems have led to the alternative view, known as articulatory phonology (Browman & Goldstein, 1995), that speech is better understood as comprised of articulatory gestures rather than as patterns of sound. Six articulatory organs — namely, the lips, the velum, the larynx, and the blade, body, and root of the tongue — produce these gestures. Each is controlled separately, so that individual speech units are comprised of different combinations of movements. The distribution of action over these articulators means that the elements overlap in time, which makes possible the high rates of production and perception. Unlike phonemes, speech gestures can be discerned by mechanical means, though X-rays, magnetic resonance imaging, and palatography (Studdert-Kennedy, 1998).

This still raises the question, though, of how these gestures are perceived — our ability to understand speech on radio or telephone is incontrovertible evidence that speech can be understood from the acoustic stream alone. The short (but still incomplete) answer is that speech is understood in terms of the articulatory gestures that produce it, rather than in terms of elementary sound units. This is the so-called « motor theory of speech perception » (Lieberman & al., 1967). Although we can understand the radio announcer, there is abundant evidence that watching people speak can aid understanding of what they are saying. It can even distort it, as in the McGurk effect, in which dubbing sounds onto a mouth that is saying something different alters what the hearer actually hears (McGurk & MacDonald, 1976).

McNeill (1992) has pointed out, though, that even today speech-synchronized manual gestures should be considered part of language, so the dominance of speech is not complete. Conversely, facial gestures are increasingly recognized as an important component of the signed languages of the deaf. These gestures tend to focus on the mouth, and are distinct from *mouthing*, where the signer silently produces the spoken word simultaneously with the sign that has the same meaning. Mouth gestures have been studied primarily in European signed languages, and schemes for the phonological composition of mouth movements have been proposed for Swedish (Bergman & Wallin, 2001), English (Sutton-Spence & Day, 2001) and Italian (Ajello & al., 2001) Sign Languages. Facial gestures also play a prominent role in American Sign Language, providing the equivalent of prosody in speech, and are also critical to many other linguistic functions, such as marking different kinds of questions, or indicating adverbial modifications of verbs (Emmorey, 2002). In a recent study Muir and Richardson (2005) found that native signers watching discourse in British Sign Language focused mostly on the face and mouth, and relatively little on the hands or upper body. The face may play a much more prominent role in signed languages than has been hitherto recognized.

In summary, evidence from spoken and signed language suggests that movements of the hands and face feature prominently in both. This suggests that the transition from dominance of the hands to dominance of the face might have been a smooth and continuous one. All that needed to be added was voicing, so that facial gestures could be recovered from sound as well as from vision. The final step in the addition of voicing probably occurred late, perhaps with or even after the emergence of our own species, *Homo sapiens*, as I shall suggest below.

3. The mirror system

Neurons in the region of F5 in the ventral premotor cortex of the monkey typically fire when the animal makes grasping movements with the hand or mouth. A subset of those cells, dubbed « mirror neurons » also fire when the animal observes another individual making the same movements. In the monkey, these responses require the presence of a target, and do not respond to actions that merely mimic an action in the absence of a target, nor do they respond to a target alone (Gallese & al., 1996 ; Rizzolatti & al., 1996). This direct mapping of perceived action onto the production of action seems to provide a platform for the evolution of language (e.g., Arbib, 2005), and to support, albeit indirectly, the motor theory of speech perception. Furthermore, the area of the human brain that corresponds most closely to area F5 in the monkey includes Broca's area, which is one of the main cortical areas underlying the production of speech. This suggests that speech may have arisen from cortical structures that initially had to do with manual action rather than with vocalization (Rizzolatti & Arbib, 1998).

It has also become apparent that mirror neurons are part of a more general mirror system that involves other regions of the brain as well. The superior temporal sulcus (STS) also contains cells that respond to observed biological actions, including grasping actions (Perrett & al., 1989), although few if any respond when the animal itself performs an action. F5 and STS are connected to area PF in the inferior parietal lobule, where there are also neurons that respond both to the execution and perception of actions. These neurons are now known as « PF mirror neurons » (Rizzolatti & al., 2001). Other areas, such as the amygdala and orbito-frontal cortex, may also be part of the mirror system.

A similar system has been inferred in humans, based on evidence from electroencephalography (Muthukumaraswamy & al., 2004), magnetoencephalography (Hari & al., 1998), transcranial magnetic stimulation (Fadiga & al. 1995), and functional magnetic resonance imaging (fMRI) (Iacoboni & al. 1999). Unlike the mirror system in monkeys, the human mirror system appears to be activated by movements that need not be directed toward an object (Rizzolatti & al. 2001), although there is evidence that it is activated more by actions that are object-directed than by those that are not object-directed (Muthukumaraswamy & al., 2004). Activation by non-object-directed action may reflect adaptation of the system for more abstract signaling — as in signed languages. The mirror system in humans appears to involve areas in the frontal, temporal and parietal lobes that are homologous to those in the monkey, although there is some evidence that they tend to be lateralized to the left hemisphere in humans, especially in the frontal lobes (Iacoboni & al., 1999 ; Nishitani & Hari, 1998). It is well established that manual apraxia, especially for actions involving fine motor control, is associated with left-hemisphere damage (Heilman & al., 2000). It is possible that the incorporation of vocalization into the mirror system, perhaps unique to *Homo sapiens*, resulted in lateralization of the manual as well as of the vocal system (Corballis, 2003).

The mirror system leads to what has been termed the « direct-matching hypothesis », according to which we understand actions by mapping the visual representations of observed actions onto the motor representations of the same actions (Rizzolatti & al., 2001). In humans, this system is tuned to the perception of actions that have a reference to people. Evidence from an fMRI study shows, for example, that it is activated when people watch mouth actions, such as biting, lip-smacking, oral movements involved in vocalization, when these are performed by people, but not when they are performed by a monkey or a dog. Actions belonging to the observer's own motor repertoire are mapped onto the observer's motor system, while those that do not belong are not — instead, they are perceived in terms of their visual properties (Buccino & al., 2004). Watching speech movements, and even stills of a mouth making a speech sound, activate the mirror system, including Broca's area (Calvert & Campbell, 2003). This is consistent with the idea that speech may have evolved from visual displays that included movements of the face.

Although most of the evidence on the mirror system has to do with visual input, area F5 of the monkey also contains what might be termed « acoustic mirror neurons ». These respond to the sounds of actions, such as tearing paper or breaking a peanut, as well as to the performance of those actions. That is, even in the monkey, the direct-matching hypothesis is not restricted to visual input (Kohler & al., 2002). There is no evidence for mirror neurons in the monkey that fire to both the production and perception of vocalization. It is likely, though, that vocalization was incorporated into the mirror system in humans, and probably only in humans

or our hominid forebears (Ploog, 2002), providing the mechanism for the motor theory of speech perception. As we shall see, this probably occurred late in hominid evolution, possible within the last 100,000 years.

3.1. The link between hand and mouth

Although there is no evidence that vocalization is involved in the mirror system in the monkey, there is evidence for an intimate connection between hand and mouth. In the monkey, area F5 of the monkey includes also a class of neurons that discharge when the animal grasps an object with either the hand or the mouth (Rizzolatti & al., 1988), and in the parietal lobe there are mirror neurons that respond both when the animal observes and executes a grasping action with the hand or the mouth (Gallese & al., 2002). Gentilucci & al. (2001) infer a similar class of neurons in humans. They showed that when subjects were instructed to open their mouths while grasping objects, the size of the mouth opening increased with the size of the grasped object, and conversely, when they open their hands while grasping objects with their mouths, the size of the hand opening also increased with the size of the object. In the evolution of communication, this mechanism of double command to hand and mouth could have been instrumental in the transfer of a communication system, based on the mirror system, from movements of the hand to movements of the mouth.

In humans, grasping movements of the hand affect the kinematics of speech itself. When a person grasps an object and brings it to the mouth (Gentilucci, Santunione, & al., 2004), or watches another person doing this (Gentilucci, 2003 ; Gentilucci, Santunione, & al., 2004 ; Gentilucci, Stefanini, & al., 2004), the kinematics of speech vary with the size of the object. The larger the object, the wider the mouth opens, and the actual speech sounds are affected accordingly. Again, this implies a mirror system, since action observation induces the same effects as action execution. The effects on voicing and lip kinematics are dependent on the arm movement itself, and not on the nature of the grasped objects. Indeed, the same effects were found when either fruits or geometrical solids were presented, and even when the action was pantomimed with no object present at all.

By using the mirror system, an individual observing an arm action can automatically and covertly execute the same action in order to interpret the meaning of the action. For manual actions functionally related to oro-facial actions the motor command is sent also to the mouth, and reaches the threshold for execution when the mouth is already activated to pronounce the syllable. In a more recent study, Bernardis and Gentilucci (2006) have shown that these interactions are not confined to nonsymbolic movements such as grasping. Spoken words and symbolic gestures influence each other when produced simultaneously, and the voice spectra of responses to gestures and words are also affected by observation of gestures and words produced simultaneously.

It is likely that the connections between hand and mouth reviewed above were established initially in the context of eating, including the acts of grasping and bringing food to the mouth, but adapted later for communication. MacNeilage (1998) has suggested that speech itself originated from repetitive ingestive movements of the mouth. This may well be correct, but it is perhaps only half the story, since it neglects the important role, in primates at least, of hand and arm movements in eating. Eating, moreover, suggest a natural link of hand to mouth that may well have been decisive in the switch from a predominantly manual system to a predominantly vocal one.

4. The *FOXP2* gene

About half of the members of three generations of an extended family in England, known as the KE family, are affected by a disorder of speech and language. The disorder is evident from the affected child's first attempts to speak and persists into adulthood (Vargha-Khadem & al., 1995). The disorder is now known to be due to a point mutation on the *FOXP2* gene (fork-head box P2) on chromosome 7 (Fisher & al., 1998 ; Lai & al., 2001). For normal speech to be acquired, two functional copies of this gene seem to be necessary.

The nature of the deficit in the affected members of the KE family, and therefore the role of the *FOXP2* gene, have been debated. Some have argued that *FOXP2* gene is involved in the development of morphosyntax (Gopnik, 1990), and it has even been identified more broadly as the « grammar gene » (Pinker, 1994). Subsequent investigation suggests, however, that the core deficit is one of articulation, with grammatical impairment a secondary outcome (Watkins,

Dronkers, & al., 2002). The *FOXP2* gene may therefore play a role in the incorporation of vocal articulation into the mirror system.

This is supported by a study in which fMRI was used to record brain activity in both affected and unaffected members of the KE family while they covertly generated verbs in response to nouns (Liégeois & al., 2003). Whereas unaffected members showed the expected activity concentrated in Broca's area in the left hemisphere, affected members showed relative *underactivation* in both Broca's area and its right-hemisphere homologue, as well as in other cortical language areas. They also showed *overactivation* bilaterally in regions not associated with language. However, there was bilateral activation in the posterior superior temporal gyrus ; the left side of this area overlaps Wernicke's area, important in the comprehension of language. This suggests that affected members may have tried to generate words in terms of their sounds, rather than in terms of articulatory patterns. Their deficits were not attributable to any difficulty with verb generation itself, since affected and unaffected members did not differ in their ability to generate verbs overtly, and the patterns of brain activity were similar to those recorded during covert verb generation. Another study based on structural MRI showed morphological abnormalities in the same areas (Watkins, Vargha-Khadem, & al., 2002).

The *FOXP2* gene is highly conserved in mammals, and in humans differs in only three places from that in the mouse. Nevertheless, two of the three changes occurred on the human lineage after the split from the common ancestor with the chimpanzee and bonobo. A recent estimate of the date of the more recent of these mutations suggests that it occurred « since the onset of human population growth, some 10,000 to 100,000 years ago » (Enard & al., 2002, p. 871). If this is so, then it might be argued that the final incorporation of vocalization into the mirror system was critical to the emergence of modern human behavior, often dated to the Upper Paleolithic (Corballis, 2004).

It is unlikely, though, that the *FOXP2* mutation was the only event in the transition to speech, which undoubtedly went through several steps and involved other genes (Marcus & Fisher, 2003). Moreover, the *FOXP2* gene is expressed in the embryonic development of structures other than the brain, including the gut, heart, and lung (Shu & al., 2001). It may have even played a role in the modification of breath control for speech (MacLarnon & Hewitt, 1999). A mutation of the *FOXP2* gene may nevertheless have been the most recent event in the incorporation of vocalization into the mirror system, and thus in the refinement of vocal control to the point that it could carry the primary burden of language.

The idea that the critical mutation of the *FOXP2* gene occurred less than 100,000 years ago is indirectly supported by recent evidence from African click languages. Two of the many groups that make extensive use of click sounds are the Hadzabe and San, who are separated geographically by some 2000 kilometers, and genetic evidence suggests that the most recent common ancestor of these groups goes back to the root of present-day mitochondrial DNA lineages, perhaps as early as 100,000 years ago (Knight & al., 2003). This could mean that clicks were a prevocal way of adding sound to facial gestures, prior to the *FOXP2* mutation. Although Ingman & al. (2000) have estimated that modern humans migrated out of Africa some 52,000 years ago, more recent estimates suggest that the migration was earlier than this, perhaps closer to 80,000 years ago (Oppenheimer, 2003). Those who migrated may have already developed autonomous speech, leaving behind African speakers who retained click sounds. The only known non-African click language is Damin, an extinct Australian aboriginal language. *Homo sapiens* may have arrived in Australia as early as 60,000 years ago (Thorne & al., 1999), not long after the migrations out of Africa. This is not to say that the early Australians and Africans did not have full vocal control of speech ; rather, click languages may be simply a vestige of earlier languages in which vocalization was not yet part of the mirror system giving rise to autonomous speech.

5. When did gesture become language ?

The foregoing suggests that the gradual transformation of language from a predominantly manual system to a vocal one was completed relatively late in hominid evolution, and possibly since the emergences of our own species, *H. sapiens*. By regarding speech itself as a gestural system, the transformation from manual to vocal language can be separated from the question of when gestural communication evolved into true language, with the system of grammatical rules that provide what Chomsky (1988) has termed « discrete infinity » — the capacity to

generate and understand an infinite number of potential messages. This may be in part an extension of the mirror system, since the condition known as « agrammatism » — the loss of grammatical ability — is associated with damage in the region of Broca's area, although the true mirror system probably has more to do with the mapping of input to output than with the elaboration of action.

A critical component of grammatical language, as distinct from animal communication, is recursion (Hauser & al., 2002), which appears to be beyond the capacities of monkeys (Fitch & Hauser, 2004), and probably also of nonhuman apes (Pinker, 1994). The emergence of grammar may well have depended on the increase in brain size that began with the emergence of *Homo erectus* some 2 million years ago, and may have progressed through several stages (Jackendoff, 2002). Therefore, language as a grammatical system may have evolved as a primarily manual system, although no doubt punctuated by vocalizations in the form of grunts.

Language, whether manual or vocal, is expensive in terms of neural circuitry, and occupies a good part of the brain, especially the left hemisphere. The best guess, then, is that gestural communication began to acquire grammatical properties when brain size increased beyond the ape-like proportions seen in the early hominids. This can be dated from about two million years ago, with the emergence of the genus *Homo*. Recent genetic studies have suggested that at least two mutations may have contributed to this increase in brain size. It is of some interest that both mutations involve the inactivation of genes, suggesting that we may have gained some of our humanness by losing genes, rather than by acquiring new ones.

One of these mutations has to do with a gene on chromosome 7 that encodes the enzyme CMP-N-acetylneuraminic acid (CMP-Neu5Ac) hydroxylase (CMAH). An inactivating mutation of this gene has resulted in a deficiency in humans of the mammalian sialic acid N-glycolylneuraminic acid (Neu5Gc). This acid appears to be absent in Neanderthal fossils as well as in humans, but is present in present-day primates. It also seems to have been down-regulated in the chimpanzee brain, and through mammalian evolution, leading to speculation that inactivation of the CMAH gene may have removed a constraint on brain growth in human ancestry (Chou & al., 2002). Chou & al. applied molecular-clock analysis to the CMAH genes in chimpanzees and other great apes, as well as to the pseudogene in humans, which indicated that the mutation occurred some 2.8 million years ago, leading up to the expansion in brain size.

Another inactivating mutation may also have contributed to the increase in brain size has to do with a gene on chromosome 7 that encodes for the myosene heavy chain MYH16, responsible for the heavy masticatory muscles in most primates, including chimpanzees and gorillas, as well as the early hominids. Molecular-clock analysis suggests that this gene was inactivated around 2.4 million years ago, leading to speculation that the diminution of jaw muscles and their supporting bone structure removed a further constraint on brain growth (Stedman & al., 2004). It is a matter of further speculation as to why this seemingly deleterious mutation became fixed in the ancestral human population. It may have had to do with the change from a predominantly vegetable diet to a meat-eating one, or it may have had to do with the increasing use of the hands rather than the jaws to prepare food (Currie, 2004). Perhaps it was also driven in part by the increasing dependence on social intelligence — and language.

These mutations, and the resulting increase in brain size, may have been selected because of a change in environment. With the global shift to cooler climate after 2.5 million years ago, much of southern and eastern Africa probably became more open and sparsely wooded (Foley, 1987). This left the hominids not only more exposed to attack from dangerous predators, such as saber-tooth cats, lions, and hyenas, but also obliged to compete with them in carnivory. The solution was not to compete on the same terms, but to establish what has been termed a « cognitive niche » Whiten (1999, p. 175), relying on social cooperation and intelligent planning for survival. It seems reasonable to suppose that language evolved in this context.

As suggested by Jackendoff (2002), the first steps may have been the emergence of a phrase structure and the formation of grammatical categories. It is likely that recursive grammar evolved relatively late in this progression. The key to this may lie, not in brain size per se, but rather in what is known as « secondary altriciality ». That is, the human brain undergoes most of its growth after birth, and a growing brain is highly plastic. In macaque newborns, the brain at birth has a volume of about 70 % of adult size ; in chimpanzees the figure is about 40 % ; in humans, it is only 25 % (Coqueuniot & al., 2004). Human language, whether signed

or spoken, is dependent on a critical period of growth. The optimal period is probably between two and four, although there is evidence that grammar can be acquired later in childhood (Vargha-Khadem & al., 1997), but probably not after puberty. Recent evidence suggests that early *Homo erectus* showed an ape-like pattern rather than a human-like pattern of post-natal brain growth, suggesting that secondary altriciality may not have emerged until fairly late in the genus *Homo*, although it was probably present in *Homo neanderthalensis* (Coqueugniot & al., 2004 ; Dean & al., 2001), suggesting that it may go back some 500,000 years. This late development may signal the emergence of recursive grammar, and a language system as grammatically complex as modern signed and spoken languages.

If this scenario is correct, and if autonomous speech did not evolved until some time within the past 100,000 years, then grammatical language would have been at least partly manual, perhaps somewhat resembling the present-day signed languages invented by deaf communities.

6. Why speech ?

According to the account presented here, the transition from manual to vocal language was not abrupt. This raises the question, though, of why the transition took place at all. The signed languages of the deaf clearly show that manual languages can be as sophisticated as vocal ones. Further the transition to speech involved the lowering of the larynx, which greatly increased the risk of choking to death. Clearly, the evolutionary pressure toward speech must have been strong. But why ?

There are a number of possible answers. First, a switch to autonomous vocalization would have freed the hands from necessary involvement in communication, allowing increased use of the hands for manufacture and tool use. Indeed vocal language allows people to speak and use tools at the same time, leading perhaps to pedagogy (Corballis, 2002). Indeed, it may explain the so-called « human revolution » (Mellars & Stringer 1989), manifest in the dramatic appearance of more sophisticated tools, bodily ornamentation, art, and perhaps music, dating from some 40,000 years ago in Europe, and maybe earlier in Africa (McBrearty & Brooks, 2000). This may well have come about because of the switch to autonomously vocal language, made possible by the FOXP2 mutation (Corballis, 2004).

Although manual and vocal language can be considered linguistically equivalent, there are other advantages to vocalization. One factor may have been greater energy requirements associated with gesture ; there is anecdotal evidence from those attending courses in sign language that the instructors required regular massages in order to meet the sheer physical demands of sign language expression. The physiological costs of speech, in contrast, are so low as to be nearly unmeasurable (Russell & al., 1998). Further, speech is less attentionally demanding than signed language ; one can attend to speech with one's eyes shut, or when watching something else. Speech also allows communication over longer distances, as well as communication at night or when the speaker is not visible to the listener. The San, a modern hunter-gatherer society, are known to talk late at night, sometimes all through the night, to resolve conflict and share knowledge (Konner, 1982). A recent study also indicates that the short-term memory span is shorter for American Sign Language than for speech (Boutla & al., 2004), suggesting that voicing may have permitted longer and more complex sentences to be transmitted — although the authors of this study claim that the shorter memory span has no impact on the linguistic skill of signers.

A possible scenario for the switch is that there was selective pressure for the face to become more extensively involved in gestural communication as the hands were increasingly engaged in other activities. Our species had been habitually bipedal from some 6 or 7 million years ago, and from some 2 million years ago was developing tools, which would have increasingly involved the hands. The face had long played a role in visual communication, and as outlined above plays an important role in present-day signed languages. Consequently, there may have been pressure for intentional communication to move to the face, including the mouth and tongue. Gesturing may then have retreated into the mouth, so there may have been pressure to add voicing in order to render movements of the tongue more accessible — through sound rather than sight. In this scenario, speech is simply gesture half swallowed, with voicing added. Even so, lip-reading can be a moderately effective way to recover the speech gestures, and as mentioned earlier the McGurk effect illustrates that speech is in part a

visual medium. Adding voicing to the signal could have had the extra benefit of allowing a distinction between voiced and unvoiced phonemes, increasing the range of speech elements.

Changes in the mode of communication can have a dramatic influence on human culture, as illustrated by the invention of writing, and more recently by email and the Internet. These changes were relatively sudden, and cultural rather than biological. The change from manual to vocal communication, in contrast, would have been slow, driven by natural selection and involving biological adaptations, but it may have had no less an impact on human culture — and therefore, perhaps, on human survival.

7. Summary and conclusions

Since Hewes (1973) presented the case for the gestural origins of language, evidence has accumulated to the point that a plausible scenario can be envisaged. We now have evidence that the adaptations for articulate speech were completed late in hominid evolution, possibly even within the past 100,000 years. Since it is unlikely that language itself evolved so late and so suddenly, this provides a good reason to suppose that grammatical language was previously carried by manual and facial gesture, perhaps with increasing vocal accompaniment. I have suggested also that the final achievement of autonomous speech had a dramatic effect on human culture, and was perhaps even instrumental in the human revolution leading to what has been termed « modern » behavior (Corballis, 2004).

The scenario is rendered all the more plausible by the insight that speech itself is a gestural system rather than a vocal one, which is in turn bolstered by the recent discovery of the so-called « mirror system » in the primate brain. Language can thus be conceived as part of the system that directly maps biological action onto perception, present also in primates. Of course language is much more than that, since it involves all of the complexities of grammar and « discrete infinity », and these features probably emerged well before speech became autonomous. The discoveries about FOXP2 provide further potential insight as to how vocalization might have been incorporated into this system, providing the means by which speech became autonomous. We go from there to telephone and radio — although text messaging may be returning us, or our children, to a visuo-manual mode.

These new developments remain somewhat speculative, but will no doubt add further evidence concerning the evolution of human language. Whether that evidence will continue to support the gestural theory remains to be seen, but that theory has come a long way since Hewes' formulation in 1973.

References

- Ajello (R.) & al. 2001. « Linguistic gestures : Mouthing in Italian Sign Languages (LIS) ». in : Sutton-Spence (R.) & Boyes-Braem (P.), (eds.). *The hands are the head of the mouth : The mouth as articulator in sign language*. Hamburg, Germany : Signum-Verlag.
- Arbib (M.-A.). 2005. « From monkey-like action recognition to human language : An evolutionary framework for neurolinguistics ». in : *Behavioral & Brain Sciences*, 28, pp. 105-168.
- Armstrong (D.-F.). 1999. *Original signs : gesture, sign, and the source of language*. Washington : Gallaudet University Press.
- Armstrong (D.-F.) & al. 1995. *Gesture and the nature of language*. Cambridge : Cambridge University Press.
- Bergman (B.) & Wallin (L.). 2001. « A preliminary analysis of visual mouth segments in Swedish Sign Language ». in : Sutton-Spence (R.) & Boyes-Braem (P.), (eds.). *The hands are the head of the mouth : The mouth as articulator in sign language*. Hamburg, Germany : Signum-Verlag.
- Bernardis (P.) & Gentilucci (M.). 2006. « Speech and gesture share the same communication system ». in : *Neuropsychologia*, 44, pp. 178-190.
- Bickerton (D.). 1995. *Language and human behavior*. Seattle, WA : University of Washington Press.
- Boutla (M.) & al. 2004. « Short-term memory span : Insights from sign language ». in : *Nature Neuroscience*, 7, pp. 997-1002.
- Browman (C.-P.) & Goldstein (L.-F.). 1995. « Dynamics and articulatory phonology » in : van Gelder (T.) & Port (R.-F.), (eds.). *Mind as motion*. Cambridge, MA : MIT Press.
- Buccino (G.). 2004. « Neural circuits involved in the recognition of actions performed by non-conspecifics : An fMRI study ». in : *Journal of Cognitive Neuroscience*, 16, pp. 114-126.
- Calvert (G.-A.) & Campbell (R.). 2003. « Reading speech from still and moving faces : The neural substrates of visible speech ». in : *Journal of Cognitive Neuroscience*, 15, pp.57-70.
- Chomsky (N.). 1988. *Language and problems of knowledge : The Managua lectures*. Cambridge, MA : MIT Press.
- Chou (H.-H.) & al. 2002. « Inactivation of CMP-N-acetylneuraminic acid hydroxylase occurred prior to brain expansion during human evolution ». in : *Proceedings of the National Academy of Sciences, USA*, 99, pp. 11736-11741.
- Condillac (E.-B. de). 1971. *An essay on the origin of human knowledge*. in : Nugent (T.), (tr.). Gainesville, FL : Scholars Facsimiles and Reprints [Originally published 1746].
- Coqueugniot (H.) & al. 2004. « Early brain growth in *Homo erectus* and implications for cognitive ability ». in : *Science*, 431, pp. 299-302.
- Corballis (M.-C.). 1991. *The lopsided ape*. New York : Oxford University Press.
- Corballis (M.-C.). 2002. *From hand to mouth : the origins of language*. Princeton, NJ. Princeton University Press.
- Corballis (M.-C.). 2003. « From mouth to hand : Gesture, speech, and the evolution of right-handedness ». in : *Behavioral & Brain Sciences*, 26, pp. 199-260.
- Corballis (M.-C.). 2004. « The origins of modernity : Was autonomous speech the critical factor ? ». in : *Psychological Review*, 111, pp. 543-522.
- Crow (T.-J.). 2002. « Sexual selection, timing, and an X-Y homologous gene : Did *Homo sapiens* speciate on the Y chromosome ? ». in : Crow (T.-J.). *The speciation of modern Homo Sapiens*. Oxford, UK : Oxford University Press.
- Currie (P.). 2004. « Muscling in on hominid evolution ». in : *Nature*, 428, pp. 373-374.
- Dean (C.) & al. 2001. « Growth processes in teeth distinguish modern humans from *Homo erectus* and earlier hominins ». in : *Nature*, 414, pp. 628-631.
- Emmorey (K.). 2002. *Language, cognition, and brain : insights from sign language research*. Hillsdale, NJ : Erlbaum.
- Enard (W.) & al. 2002. « Molecular evolution of FOXP2, a gene involved in speech and language ». in : *Nature*, 418, pp.869-871.
- Fadiga (L.) & al. 1995. « Motor facilitation during action observation — a magnetic stimulation study ». in : *Journal of Neurophysiology*, 73, pp. 2608-2611.
- Fisher (S.) & al. 1998. « Localisation of a gene implicated in a severe speech and language disorder ». in : *Nature Genetics*, 18, pp. 168-170.
- Fitch (W.) & Hauser (W.-T.). 2004. « Computational constraints on syntactic processing in a nonhuman primate ». in : *Science*, 303, pp. 377-380.

- Foley (R.). 1987. *Another unique species : patterns in human evolutionary ecology*. Harlow : Longman Scientific and Technical.
- Gallese (V.) & al. 1996. « Action recognition in the premotor cortex ». in : *Brain*, 119, pp. 593-609.
- Gallese (V.) & al. 2002. « Action representation and the inferior parietal lobule ». in : Prinz (W.) & Hommel (B.), (eds.). *Attention and Performance XIX : III Action perception and imitation*. Oxford, UK : Oxford University Press.
- Gardner (R.-A.) & Gardner (B.-T.). 1969. « Teaching sign language to a chimpanzee ». in : *Science*, 165, pp. 664-672.
- Gentilucci (M.). 2003. « Grasp observation influences speech production ». in : *European Journal of Neuroscience*, 17, pp. 179-184.
- Gentilucci (M.) & al. 2001. « Grasp with hand and mouth : a kinematic study on healthy subjects ». in : *Journal of Neurophysiology*, 86, pp. 1685-1699.
- Gentilucci (M.), Santunione (P.) & al. 2004. « Execution and observation of bringing a fruit to the mouth affect syllable pronunciation ». in : *European Journal of Neuroscience*, 19, pp. 190-202.
- Gentilucci (M.), Stefanini (S.) & al. 2004. « Action observation and speech production : study on children and adults ». in : *Neuropsychologia*, 42, pp. 1554-1567.
- Givón, (T.). 1995. *Functionalism and grammar*. Philadelphia, PA : John Benjamins.
- Goldin-meadow (S.) & McNeill (D.). 1999. « The role of gesture and mimetic representation in making language the province of speech ». in : Corballis (M.-C.) & Lea (S.-E.-G.), (eds.). *The descent of mind*. Oxford, UK : Oxford University Press.
- Gopnik, (M.). 1990. « Feature-blind grammar and dysphasia ». in : *Nature*, 344, p. 715.
- Hari. (R.-N.) & al. 1998. « Activation of human primary motor cortex during action observation : A neuromagnetic study ». in : *Proceedings of the National Academy of Sciences, USA*, 95, pp. 15061-15065.
- Hauser (M.-D.) & al. 2002. « The faculty of language : What is it, who has it, and how did it evolve? » in : *Science*, 298, pp. 1569-1579.
- Heilman (K.) & al. 2000. « Hemispheric asymmetries of limb-kinetic apraxia – A loss of dexterity ». in : *Neurology*, 55, pp. 523-526.
- Hewes (G.-W.). 1973 « Primate communication and the gestural origins of language ». in : *Current Anthropology*, 14, pp.5-24.
- Iacoboni (M.) & al. 1999. « Cortical mechanisms of human imitation ». in : *Science*, 286, pp. 2526-2528.
- Ingman (M.) & al. 2000. « Mitochondrial genome variation and the origin of modern humans ». in : *Nature*, 408, pp. 708-713.
- Jackendoff (R.). 2002. *Foundations of language : Brain, meaning, grammar, evolution*. Oxford, UK : Oxford University Press.
- Joos (M.). 1948. « Acoustic phonetics ». in : *Language Monographs*, n°23. Baltimore, MD : Linguistic Society of America.
- Kohler (E.) & al. 2002. « Hearing sounds, understanding actions : Action representation in mirror neurons ». in : *Science*, 297, pp. 846-848.
- Konner (M.). 1982. *The tangled wing : Biological constraints on the human spirit*. New York : Harper.
- Knight (A.) & al. 2003. « African Y chromosome and mtDNA divergence provides insight into the history of click languages ». in : *Current Biology*, 13, pp. 464-473.
- Lai (C.-S.) & al. 2001. « A novel forkhead-domain gene is mutated in a severe speech and language disorder ». in : *Nature*, 413, pp. 519-523.
- Liberman (A.). 1967. « Perception of the speech code ». in : *Psychological Review*, 74, pp. 431-461.
- Liégeois (F.) & al. 2003. « Language fMRI abnormalities associated with FOXP2 gene mutation ». in : *Nature Neuroscience*, 6, pp. 1230-1237.
- MacLarnon (A.) & Hewitt (G.). 1999. « The evolution of human speech : The role of enhanced breathing control ». in : *American Journal of Physical Anthropology*, 109, pp. 341-363.
- MacNeilage (P.-F.). 1998. « The frame/content theory of evolution of speech ». in : *Behavioral & Brain Sciences*, 21, pp. 499-546.
- Marcus (G.-F.) & Fisher (S.-E.). 2003. « FOXP2 in focus : what can genes tell us about speech and language ? ». in : *Trends in Cognitive Sciences*, 7, pp. 257-262.

- McBrearty (S.) & Brooks (A.-S.). 2000. « The revolution that wasn't : A new interpretation of the origin of modern human behaviour ». in : *Journal of Human Evolution*, 39, pp. 453-563.
- McGurk (H.) & MacDonald (J.). 1976. « Hearing lips and seeing voices ». in : *Nature*, 264, pp. 746-748.
- McNeill (D.). 1992. *Hand and mind : What gestures reveal about thought*. Chicago, IL : University of Chicago Press,
- Mellars (P.-A.) & Stringer (C.-B.), (eds.). 1989. *The human revolution : Behavioral and biological perspectives on the origins of modern humans*. Edinburgh : Edinburgh University Press.
- Muir (L.-J) & Richardson (I.-E.-G.). 2005. « Perception of sign language and its application to visual communications for deaf people ». in : *Journal of Deaf Studies & Deaf Education*, 10, pp. 390-401.
- Muthukumaraswamy (S.-D.) & al. 2004. « Mu rhythm modulation during observation of an object-directed grasp ». in : *Cognitive Brain Research*, 19, pp. 195-201.
- Neidle (C.) & al. 2000. *The syntax of American Sign Language*. Cambridge, MA : The MIT Press.
- Nishitani (N.) & Hari (R.). 2000. « Temporal dynamics of cortical representation for action ». in : *Proceedings of the National Academy of Sciences, USA*, 97, pp. 913-918.
- Oppenheimer (S.). 2003. *Out of Eden : The peopling of the world*. London : Constable.
- Perrett (D.-I.) & al. 1989. « Frameworks of analysis for the neural representation of animate objects and actions ». in : *Journal of Experimental Biology*, 146, pp. 87-113.
- Pinker (S.). 1994. *The language instinct*. New York : Morrow.
- Ploog (D.). 2002. « Is the neural basis of vocalisation different in non-human primates and *Homo sapiens* ? ». in : Crow (T.-J.), (ed.). *The speciation of modern homo sapiens*. Oxford, UK : Oxford University Press.
- Rogers (L.-J.) & Andrew (R.-J.), (eds.). 2002. *Comparative vertebrate lateralization*. New York : Cambridge University Press.
- Rizzolatti (G.) 1996. « Premotor cortex and the recognition of motor actions ». in : *Cognitive Brain Research*, 3, pp.131-141.
- Rizzolatti (G.) & Arbib (M.-A.). 1998. « Language within our grasp ». in : *Trends in Cognitive Sciences*, 21, pp.188-194.
- Rizzolatti (G.) & al. 1988. « Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements ». in : *Experimental Brain Research*, 71, pp. 491-507.
- Rizzolatti (G.) & al. 2001. « Neurophysiological mechanisms underlying the understanding and imitation of action ». in : *Nature Reviews*, 2, pp. 661-670.
- Rubén (R.-J.). 2005. « Sign language : Its history and contribution to the understanding of the biological nature of language ». in : *Acta Oto-Laryngologica*, 125, pp. 464-467.
- Russell (B.-A.) & al. 1998. « Effects of varied vocal intensity on ventilation and energy expenditure in women and men ». in : *Journal of Speech, Language, & Hearing Research*, 41, pp. 239-248.
- Savage-Rumbaugh (S.) & al. 1998. *Apes, language, and the human mind*. New York : Oxford University Press.
- Shu (W.-G.) & al. 2001. « Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors ». in : *Journal of Biological Chemistry*, 276, pp. 27488-27497.
- Stedman (H.-H.) & al. 2004. « Myosin gene mutation correlates with anatomical changes in the human lineage ». in : *Nature*, 428, pp. 415-418.
- Studdert-Kennedy (M.). 1998. « The particulate origins of language generativity : From syllable to gesture ». in : Hurford (J.-R) & al., (eds.). *Approaches to the evolution of language*. Cambridge, UK : Cambridge University Press.
- Sutton-Spence (R.) & Day (L.). 2001. « Mouthings and mouth gestures in British Sign Language ». in : Sutton-Spence (R.) & Boyes-Braem (P.), (eds.). *The hands are the head of the mouth : The mouth as articulator in sign language*. Hamburg, Germany : Signum-Verlag.
- Thorne (A.) & al. 1999. « Australia's oldest human remains : Age of the Lake Mungo human skeleton ». in : *Journal of Human Evolution*, 36, pp. 591-612.
- Vargha-Khadem (F.) & al. 1995. « Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder ». in : *Proceedings of the National Academy of Sciences, USA*, 92, pp. 930-933.

- Vargha-Khadem (F.) & al. 1997. « Onset of speech after left hemispherectomy in a nine-year-old boy ». in : *Brain*, 120, pp. 159–182.
- Watkins (K.-E.), Dronkers (N.-F.) & al. 2002. « Behavioural analysis of an inherited speech and language disorder : Comparison with acquired aphasia ». in : *Brain*, 125, pp. 452-464.
- Watkins (K.-E.), Vargha-Khadem (F.) & al. 2002. « MRI analysis of an inherited speech and language disorder : structural brain abnormalities ». in : *Brain*, 125, pp. 465-478.
- Whiten (A.). 1999. « The evolution of deep social mind in humans ». in : Corballis (M.-C.) & Lea (S.-E.-G.), (eds.). *The descent of mind*. Oxford : Oxford University Press.



Mai 2006

0. Introduction

Cette contribution², que nous envisageons comme la mise en débat d'une hypothèse établie à partir de résultats scientifiques toujours controversés aujourd'hui, part du constat que les animaux (insectes, poissons, mammifères...) ont la capacité de communiquer par le moyen des phéromones³. Or, chez l'Homme, cette communication n'est pas développée, et il semble qu'il ne soit pas en mesure de l'utiliser comme moyen fonctionnel de communication. Or nous considérons ici la communication comme *fonctionnelle*, c'est-à-dire que nous ne l'envisageons pas dans le simple cadre d'une interaction interindividuelle, mais dans le cadre d'une interaction spécifique réciproque, qui organise certaines des interactions sociales essentielles d'un groupe d'une même espèce, pour la survie de l'individu et de son espèce même. Cette conception de la communication implique également des relations d'aide entre individus pour préserver la vie au sein d'une espèce ou d'un groupe social donné. En particulier, les phéromones sont essentielles dans les communications de prédation, de reproduction, et toutes celles qui concernent les actions consistant à prendre soin des petits, mais aussi, par exemple chez les Fourmis, pour l'organisation de la vie sociale⁴. Aussi est-ce également dans ce sens que nous parlerons du langage, de tout langage, comme étant fonctionnel, dans la mesure où c'est en effet par le langage que l'on communique, et que le langage sert à la communication dans le but de la survie de l'individu et de l'espèce⁵.

¹ Cet article reprend partiellement en les développant certains points déjà présentés dans une contribution intitulée « Les métamorphoses du langage ». in : Abdallah-Pretceille (M.), (dir.). 2006. *Les métamorphoses de l'identité*. Paris : Éditions Anthropos-Economica [en cours de publication].

² Nous souhaitons remercier ici tous les participants issus des sciences humaines et des sciences du vivant du groupe de recherche interdisciplinaire en anthropologie de la communication de la MSH Paris-Nord, qui ont permis à cette réflexion d'avancer. Nous remercions également Marina Cavassilas, docteur en sémiotique de l'Université de Limoges pour ses apports de réflexion en sémiotique. Dans le cadre de ces recherches, la position de l'auteure demeure celle de linguiste.

³ Pour des descriptions détaillées sur le rôle des phéromones, voir Leroy (Y.). 1986. *L'univers odorant de l'animal*. Paris : Boubée ; Brossut (R.). 1996. *Phéromones. La communication chimique chez les animaux*. Paris : Belin, CNRS Éditions ; Barbier (M.). 1982. *Les phéromones. Aspects biochimiques et biologiques*. Paris : Masson.

⁴ « La communication chimique étant présente chez les animaux de tous les embranchements, l'analyse de cette fonction donne accès à tous les niveaux de relations interorganismes et permet à elle seule de suivre les étapes évolutives des comportements impliquant des relations interindividuelles. [...] la communication chimique se diversifie considérablement dans les sociétés. Il faut plus de vingt-cinq signaux chimiques différents pour régler l'organisation d'une fourmilière » (Y. Leroy, 1986, p. 11).

⁵ Comme nous l'explicitons tout au long de l'article, nous utilisons ici le terme de « langage » dans un sens beaucoup plus large que celui admis par les éthologues, qui le distinguent en particulier des « communications chimiques ». Nous allons en ce sens totalement à l'encontre de cette distinction, puisque nous parlerons ici de « langage chimique ». C'est pourquoi une partie de l'article s'attache à expliquer et justifier ce rapprochement et que nous redéfinissons ici à chaque étape le sens dans lequel nous employons chacun des termes. Il nous semble en effet que la possibilité d'anticipation, de mensonge et donc de narrativité qui existe chez certains animaux, d'une part, et le fait que nos actes de langage humains puissent eux-mêmes être soumis à des marqueurs chimiques sans que nous nous en rendions compte dans nos interactions, remet en question l'affirmation selon laquelle : « La communication animale qui orchestre les relations entre deux ou plusieurs individus n'est pas affaire de langage. [mais] de physiologie. Par le langage édifié à partir de signes conventionnels et de systèmes de signes, l'Homme a pris ses distances par rapport au réel (I. Meyerson, 1948). Il peut être informé sans réagir d'emblée à l'information reçue. L'animal est rivé de manière obligatoire à l'impact du stimulus perçu, en l'occurrence

L'état actuel des recherches tend à montrer que l'Homme pourrait être probablement émetteur de signaux chimiques proches des phéromones, mais non récepteur, par défaut d'un sens approprié (le sens voméronasal, qui nécessite la présence d'un organe spécifique appelé fosse voméronasale¹) lui permettant de percevoir et décoder les signaux chimiques, très complexes et nombreux, que sont les phéromones. Ce sens en effet « est régressé chez l'Homme dès les premiers mois de vie fœtale » (Leroy, 1986, p. 48). Notre hypothèse part du principe que cette absence, unique chez les Mammifères et relativement peu commune dans le règne animal, pourrait se trouver étroitement liée au développement tout aussi unique d'un langage articulé parlé spécifique à l'être humain. Langage que l'Homme a progressivement développé pour arriver à l'utilisation complexe que l'on sait aujourd'hui. C'est également grâce à ce développement progressif du langage articulé parlé que l'Homme a développé (et remarquons-le, de manière tout à fait récente, à l'échelle de l'histoire de l'humanité) une traçabilité et mémorisation systématique de ses actes de langage, (écriture, imprimerie). Des données récentes en biologie semblent aller dans ce sens en faisant concorder l'apparition de la vision en couleur chez l'un de nos premiers ancêtres – singe, avec la disparition d'une cavité réceptrice². Autrement dit, en termes d'évolution, les animaux compensent et adaptent leur organisme à leur environnement et à leurs besoins. C'est pourquoi, quel que soit le scénario, il importe peu de connaître les raisons du passage d'une espèce avec fosse voméronasale à une espèce *sans*. En revanche, il est certain qu'une perte sensorielle a dû entraîner d'autres développements communicatoires. L'important est de voir que ce passage a visiblement eu lieu à un moment donné, et que l'espèce ayant donné naissance à l'Homme n'en était pas plus pourvue que l'Homme lui-même. Dans ce sens, il est possible que l'une des raisons de l'émergence de la langue articulée parlée, dans le cadre d'une communication qui certainement a toujours été multimodale, a pu se faire en partie au moins *parce que* cette nouvelle espèce n'était plus elle-même dotée de la capacité à communiquer grâce au sens chimique³ voméronasal. C'est-à-dire par un système complexe d'émission et de réception d'un certain type de signaux chimiques, les phéromones.

1. La communication par les phéromones : un sens à part entière

Mais que sont les phéromones ? et à quoi servent-elles ? Elles ont un rôle à la fois sur la croissance, la reproduction, la vie sociale des animaux. C'est pourquoi les hommes ont cherché à en synthétiser un certain nombre afin de pouvoir contrôler, par exemple, l'augmentation du nombre d'agneaux par mise bas ; ou bien afin de perturber la reproduction des insectes dans les vergers. L'odorat, toujours opérationnel chez l'être humain, nous permet de reconnaître quelqu'un par son odeur, sans en être même conscients⁴. Mais le sens voméronasal et le sens de l'odorat sont des sens distincts. Généralement inodores, les phéromones agissent de

au signal de communication. L'animal est en prise directe avec le réel dans toutes ses démarches, y compris dans la communication. En cela, la communication animale ne s'apparente en rien à la communication humaine qu'elle soit langagière ou non verbale ». (Y. Leroy, 1986, p. 11). Notre point est précisément, comme nous le développerons, que puisque nous ne savons pas, à l'heure actuelle, comment l'Homme traite les signaux chimiques qu'il reçoit appelés *phéromones*, en raison de l'absence apparente de cavité voméronasale, nous ne pouvons en déduire qu'ils sont exclus du « langage ». Cela, d'autant plus qu'il n'est pas exclu que les réponses que nous émettons à leur réception soient *verbales*.

¹ « L'organe voméronasal, ou organe de Jacobson, a été découvert par le Danois Ludwog Levin Jacobson en 1809 chez les Mammifères. Mais cet organe existe en fait chez de nombreux vertébrés. « Il est situé à la base du septum nasal. C'est une paire de sacs longs et étroits (...). Mais on sait maintenant que l'organe voméronasal est un organe séparé du système olfactif ». (Leroy, 1986, p. 48).

² « Chez les Primates, l'organe voméronasal existe chez les Prosimiens et les singes du Nouveau Monde mais manque chez les Singes de l'Ancien Monde, chez les Anthroïdes et chez l'Homme (Bailey, 1978) » (Leroy, 1986, p. 50).

³ Lorsque nous parlons ici de communication chimique, nous entendons ne parler *que* des phéromones, selon les distinctions très claires proposées par Y. Leroy dans son ouvrage, certes de 1986, mais qui constitue toujours à notre connaissance la seule étude éthologique de son type sur les phéromones. Yveline Leroy dirigeait le laboratoire d'Etho-écologie des communications animales à l'École Pratique des Hautes Études.

⁴ « Les neurobiologistes connaissent la liaison privilégiée qui existe entre le sens olfactif et les fonctions inconscientes du cerveau, notamment grâce au dispositif anatomique qui permet à l'information olfactive d'atteindre le « cerveau qui sait » (le cerveau cognitif) par l'intermédiaire seulement de deux relais, au lieu de trois habituellement requis pour tous les autres messages de systèmes sensoriels. C'est la raison pour laquelle les messages olfactifs sont envoyés directement dans les zones du cerveau liées aux émotions ». (L. Vincent, 2004, p. 26). À partir de Lledo (P.-M.), Carleton (A.) & Vincent (J.-D.). 2002. « Odors and olfaction ». in : *Journal de la Société de Biologie*, 196, pp. 59-65.

manière totalement indépendante des odeurs (L. Vincent, 2004, p. 41)¹. Si les chercheurs s'accordent aujourd'hui pour reconnaître qu'il existe une communication chimique chez l'homme, les avis restent cependant très partagés quant à l'existence d'un système phéromonal fonctionnel qui serait indépendant du système olfactif (L. Vincent, 2004, pp. 40-41)². L'un des nombreux arguments avancés relativement à leur rôle distinct et indépendant est que « *les messages de chaque organe récepteur ne sont pas relayés aux mêmes endroits du cerveau et [qu'] il y a une différenciation totale entre les voies suivies par les neurones phéromonaux et les neurones olfactifs* »³. C'est pourquoi à ce jour « *même si le système voméronasal ne semble pas fonctionnel chez l'homme [on peut gager] qu'il existe des messages chimiques semblables aux phéromones qui sont reconnus par le système olfactif* »⁴. « *L'existence de phéromones chez l'homme a été et reste controversée, note R. Brossut, mais après une série de travaux récents qui tendent à prouver leur existence et celle d'un système sensoriel spécialisé dans leur réception, le débat est plutôt de savoir si des phéromones humaines peuvent intervenir dans le comportement sans que l'individu en soit conscient, sans que ces informations soient analysées au niveau supérieur par le cortex* »⁵. On voit ici que l'une des questions sous-jacentes à la problématique que nous souhaitons mettre ici en débat est celle de la conscience. Ne pouvant traiter tous les aspects complexes de cette question en un article, nous renvoyons cet aspect de la question à une contribution ultérieure.

2. La fosse voméronasale en débat : l'Homme et les phéromones

En suggérant l'existence d'un système de communication phéromonale chez l'être humain (qui ne passerait pas par un organe voméronasal comme chez les autres Mammifères⁶), on dit également que nous n'aurions aucune maîtrise consciente de son utilisation. Par ailleurs, on ne sait pas non plus quel serait le fonctionnement exact, de ce système ce qui rend impossible de déterminer si nous utilisons réellement les phéromones (à bon escient) au cours de nos activités communicatives et interactions avec notre milieu. Si donc des communications chimiques de type phéromonal existent chez l'Homme qui passent par d'autres voies qu'un organe voméronasal, cela fait malgré tout toujours de l'Homme le seul Mammifère terrestre à être ainsi constitué. La difficulté même que les scientifiques ont à mettre en évidence et isoler ce système et son fonctionnement laisse perplexe ; car on ne peut par conséquent en déduire aucune fonctionnalité vitale pour la communication de l'espèce, comme groupe.

On remarque que là où les autres espèces utilisent pour communiquer des phéromones – souvent complétées par ailleurs par des signaux d'un autre type : vocal, tactile, gestuel etc. – l'Homme utilise, lui, le langage articulé parlé ; dans cette mesure, les *langues* sont devenues pour les êtres humains les principaux instruments de communication intra-spécifique au niveau des individus comme du groupe. Il est assez aisé, dès lors, d'imaginer une corrélation possible entre d'une part, une « perte » de l'utilisation *fonctionnelle* de notre sens « phéromonal » et le

¹ À partir de Meredith (M.). 2001. « Human vomeronasal organ function : a critical review of best and worst cases ». in : *Chemical senses*, 26, pp. 433-445.

² *Ibidem*, et Stern (K.) & McClintock (M.-K.). 1998. « Regulation of ovulation by human pheromones ». in : *Nature*, 392, pp. 177-179.

³ *Ibidem*, pp. 42-43 : « Les neurones partant de l'organe voméronasal ont pour destination les zones hypothalamiques qui sont impliquées dans les fonctions hormonales et reproductives : cela veut dire que les phéromones produisent des signaux dans les structures du cerveau qui contrôlent les comportements sexuels et reproductifs ainsi que l'équilibre hormonal ; les neurones partant de l'épithélium olfactif se projettent, eux, sur le cortex cérébral après un relais dans le bulbe olfactif, sorte d'antenne qui pointe à l'avant du cerveau au-dessus des narines ».

⁴ *Ibidem*, p. 62.

⁵ Brossut (R.). 1996. *Phéromones. La communication chimique chez les animaux*. Paris : Croisée des sciences, Belin, CNRS Éditions, p. 133.

⁶ « Dans les sociétés de Mammifères, il n'est pas rare que les mâles reproducteurs provoquent, par le biais des phéromones, la castration d'autres mâles créant en quelque manière une caste de sous-mâles. Cet état de choses n'est pas sans analogie avec ce qu'on observe dans les sociétés d'Insectes. Mais dans ces dernières, ce sont les femelles reproductrices, les reines, qui, par le jeu de phéromones, inhibent le développement sexuel des autres femelles, celles des castes ouvrières. [...] Chez l'Homme, certes il existe des chimiorécepteurs neurosensoriels, mais l'impact de l'odeur n'agit pas à la manière d'un stimulus déclencheur : il engendre une réponse conceptualisée. Si je sens une odeur (de brûlé), un scénario me vient à l'esprit (une casserole est sur le feu) en fonction duquel je vais décider d'une conduite à tenir (je vais éteindre le gaz) ». (Y. Leroy, 1986, p. 8).

développement, d'autre part, de nos communications linguistiques dites « conscientes », qui passent par le canal des perceptions liées à nos cinq sens. Cette hypothèse reste entièrement spéculative, car nous ne pouvons déterminer – du moins à l'heure actuelle – si l'apparition de la capacité linguistique est concomitante, postérieure ou antérieure à la régression ou non d'une cavité voméronasale chez l'Ancêtre de l'Humain, ni au développement d'un autre type éventuel de communication phéromonale¹. Il faudrait en particulier pouvoir répondre à deux questions préalables ici essentielles pour rendre possible la construction d'un savoir approfondi quelconque sur ce sujet :

- 1- Est-il possible de dater la perte éventuelle des capacités de reconnaissance et distinction des signaux phéromonaux chez l'être humain ou ses ancêtres, avec l'existence d'un véritable « sens » qui leur serait lié ?
- 2- Que peut-on dire aujourd'hui précisément de l'exploration de l'évolution de la cavité voméronasale humaine et de la certitude d'une existence ou non d'un système de communication phéromonale chez l'Homme ? Certains chercheurs disent que celle-ci a été infirmée ; d'autres le contraire, et c'est bien là toujours le problème. À moins que des résultats tout à fait récents permettent de trancher véritablement dans un sens ou dans l'autre.

3. Développement du langage et phéromones

Des articles scientifiques récents vont dans le sens de notre hypothèse en abordant la problématique de la compensation sensorielle (et fonctionnelle, dans le sens que nous avons défini) : la perte des capacités à détecter les signaux phéromonaux serait liée au développement d'autres fonctions venues les remplacer². Un article en particulier se fonde sur une étude du gène TRPC2 (TRP2) qui est essentiel au fonctionnement de l'organe voméronasal chez la souris pour démontrer que l'organe voméronasal était probablement encore fonctionnel chez l'ancêtre commun des singes du Nouveau Monde, et chez les singes et grands singes de l'Ancien Monde, mais n'est plus trouvé qu'à l'état de vestige chez l'ancêtre commun des singes et grands singes de l'Ancien Monde (notre traduction). Les auteurs fondent à partir de ces données l'hypothèse que c'est à ce moment de l'évolution que d'autres modalités, et notamment le développement de la vision en couleur, ont pu remplacer largement les signaux phéromonaux³. Et c'est pourquoi nous faisons l'hypothèse d'un développement progressif du langage articulé parlé pour « remplacer » les communications fonctionnelles nécessaires à la survie, auparavant remplies par les phéromones.

Le langage articulé pourrait être alors une adaptation évolutive interne à l'émergence d'une nouvelle espèce, concomitante à d'autres adaptations (comme la vision en couleur).

¹ « Paraît-il crédible qu'un système aussi élaboré que le système voméronasal, avec organe, récepteurs et voies neuronales spécifiques, et chargé d'accomplir une fonction unique et irremplaçable, ait totalement disparu chez l'homme ? Les scientifiques se disputent sur la question, pour ou contre : les uns décrivent une structure, d'autres ne trouvent pas l'organe, d'autres prouvent un effet phéromonal chez l'homme, d'autres encore soupèsent les « preuves » et les trouvent bien légères. Dans l'ensemble, les données s'accumulent pour montrer qu'il existe une communication chimique entre humains, avec des effets sur le comportement et sur le corps. Reste que ces messages pourraient ne pas passer par un organe voméronasal spécialisé, mais être reconnus par certains récepteurs du système olfactif... ». (Vincent, 2004, p. 43).

² Liman (E.-R.) & Innan (H.). 2003. « Relaxed selective pressure on an essential component of pheromone transduction in primate evolution ». in : *PNAS*, 100, 6, p. 3329. Cambridge, MA : Harvard University [edited by David Pilbeam] ; Gilad (Y.), Man (O.), Pääbo (S.), Lancet (D.). 2003. « Human specific loss of olfactory receptor genes ». in : *PNAS*, 100, 6, pp. 3324-3327 [March] ; Spehr (M.), Schwane (K.), Heilmann (S.), Gisselmann (G.), Hummel (T.) & Hatt (H.). « Dual capacity of a human olfactory receptor ». in : *Current Biology*, 14, 19, R832 ; et Voss (L.-B.). 2004. « Olfaction : Attracting Both Sperm and the Nose Dispatch ». in : *Current Biology*, 14, R919-R920 [November 9].

³ « We presume that, at a certain point in human evolution, ancestral species may have relied more on visual and auditory signals, rather than on chemical signals, for communicating social and reproductive status. Remarkably, it is at the time in evolution when selective pressure on TRPC2 is relaxed that the common ancestor of OW monkeys and apes developed trichromatic color vision through a gene duplication of the green_red opsin gene (32, 33). Indeed, many species of OW monkeys signal sexual and social status via colourful skin pigmentation of the face or genitalia. Thus, it is interesting to speculate that, in the evolution of OW monkeys and apes, an enhanced reliance on vision may have led to a reduced reliance on chemical signalling in mediating social interactions ». Liman (E.-R.) & Innan (H.). op. cit., p. 3329.

À nouvelle espèce, nouveau langage. Par ailleurs, cette hypothèse est tout à fait parallèle à celle défendue par Michael Corballis concernant les origines gestuelles du langage, en termes d'évolution et de chronologie. Le plus probable semble être ici en effet que le passage à une expression parlée soit le résultat d'une série de longues et lentes transformations évolutives parallèles, et à relier les unes aux autres pour mieux les comprendre.

4. Les fonctions du langage

Plusieurs approches des origines du langage proposées jusqu'à nos jours font appel en réalité à d'autres questionnements, renvoyant à la quête de causes ou de conséquences – de quelles transformations le langage est-il né ? quelles transformations le langage a-t-il permis ? à quel moment intervient la conscience ? L'une des principales tentations dès que l'on aborde cette problématique du langage humain, est en effet de faire dériver le propos vers la conscience, la mémoire, la pensée. Nous souhaitons ici clairement éviter toute dérive de ce type pour nous attacher d'un point de vue d'étholinguiste aux seules *fonctions* du langage humain. Nous constatons en effet, simplement, que le langage remplit en lui-même et pour lui-même certaines fonctions qui sont vitales – que l'on peut dire primaires – et qu'il n'est en cela ni un luxe, ni un ersatz, ni un outil développé *pour* la pensée ou la mémoire, mais pour la survie de l'individu et de l'espèce. Partant de cela, notre hypothèse est donc que toutes les *autres* fonctions que remplit à *présent* le langage parlé pour l'homme se seraient développées de façon seconde, en surplus. Mais qu'à ses tous débuts ses raisons d'émergence auraient été liées, sur le plan de la survie de l'individu au sein d'un groupe, et pour son espèce, aux mêmes fonctionnalités que les phéromones, sachant que :

La communication est toujours au service d'un comportement, écrit Yveline Leroy. En ce sens, chaque comportement correspond au déroulement d'une fonction donnée qui comprend, dans la plupart des cas, un enchaînement de séquences, le plus souvent effectuées alternativement par deux individus qui assument ensemble la fonction en cause. [...] L'étude de la communication animale ne peut être pleinement comprise que dans le cadre des fonctions au service desquelles elle est dévolue. C'est la raison pour laquelle [...] la communication est envisagée successivement dans les diverses grandes fonctions vitales : sexualité, soin aux jeunes, vie sociale, prédation, parasitisme, déprédation, zoogamie, endozoochorie.¹

C'est pourquoi, nous proposons une définition fonctionnelle, pragmatique, étholinguistique du langage à ses débuts. Nous posons que la caractéristique de tout langage est d'exister par une motivation fonctionnelle liée à notre survie et d'avoir, d'une façon ou d'une autre, une fonction par rapport à notre survie (nécessité nutritive, reproductive, etc.). Le sens donné à *langage*, dans son acception la plus large, est alors celui d'une tension communicative significative et intéressée émanant d'un individu vers autrui (qui correspond à la dimension illocutoire en termes d'actes de langage), pour sa propre survie et/ou celle de l'espèce. La communication peut se définir alors comme l'interaction qui résulte de la mise en œuvre d'un langage avec un autre individu vivant ou avec le milieu. Dans ce sens le langage est égocentré (dimension illocutoire : la visée émane d'un « je ») alors que la communication est interactive (dimension perlocutoire de l'effet abouti, qui se manifeste par la réciprocité de l'échange) et explique mieux encore dans quelle mesure la langue n'est qu'un langage parmi d'autres, et qui est également susceptible de se superposer à d'autres langages utilisés en simultanéité². Si le langage (un langage) est nécessaire à l'établissement d'une communication, la communication, elle, est ce qui doit résulter d'un acte de langage (au sens large). Ce qui nous ramène à la théorie initiale des actes de langage avec plus de force puisqu'une forme de motivation s'inscrit désormais entre les trois types d'actes, locutoire, illocutoire et perlocutoire, dont le pivot tournant et égocentré semble bien être l'acte illocutoire. Cette distinction entre langue, langage et communication n'est pas superflue, car elle les constitue et les comprend dans un même système : la communication (ou mise en relation d'un individu vivant avec un autre individu vivant – quelle que soit leur espèce) est essentielle à la survie³ ; tout langage émanant d'un individu vers un autre sert à établir une communication.

¹ Y. Leroy, 1986, p. 9.

² Sur la thématique générale d'une prise en compte linguistique de la complémentarité expressive du corps et de la langue, voir par exemple les travaux de Hall (E.-T.), Kerbrat-Orrecchioni (C.), Mondada (L.).

³ Boris Cyrulnik reprend souvent dans ses cours l'exemple des petits enfants en carences affective, linguistique et physiologique dont l'ensemble des carences s'amenuisent jusqu'à disparaître dès qu'ils retrouvent une sécurité affective. Cet exemple montre en effet que l'émergence de la parole dépend de

La linguistique pragmatique prend en effet en compte le langage comme système fonctionnel, à partir de la théorie des actes de langage et du schéma de la communication bien connu, repris par Jakobson. C'est ainsi qu'à la capacité de parler, équivalant au locutoire, correspond la fonction phatique¹ - un individu cherche à entrer en contact avec un autre individu à l'aide d'un langage ; à l'illocutoire correspond une fonction cognitive (d'intentionnalité) ; et au perlocutoire la fonction communicative (de persuasion). Or, bien qu'aucune de ces trois fonctions ne se situe sur le même plan que les deux autres, elles s'articulent cependant *ensemble* pour donner ce qu'on nomme généralement un « acte de parole » (oral) (E. Salzen, 2006).

5. Une écologie des langages et des techniques

Chaque individu développe le langage (dont la langue est un exemple) dont il a besoin par rapport à son environnement direct. Le chat en contact avec l'être humain ronronne, alors qu'il ne le fait pas à l'état sauvage ; les petits enfants apprennent la langue et la culture dans laquelle ils se trouvent élevés - qui ne sont pas nécessairement celles de leurs parents. Et tout être humain migrant doit tout réapprendre, se réadapter dans une autre culture et une autre langue. Ce problème est également celui qui se pose pour les grands singes et apparemment pour toutes les espèces, domestiquées, ou en contact *social* (*versus* de prédation ou de domination stricte) avec l'homme : toutes semblent s'adapter pour communiquer avec lui, et vice-versa. En contexte de sociabilité (dont la domesticité est l'exemple le plus poussé) c'est la relation affective qui détermine les efforts faits des deux côtés². On voit alors en quoi la maîtrise de la parole et de la langue, en tant que compétence communicative avec nos semblables, se trouve effectivement liée à notre survie et notre développement, et donc en quoi le développement de la capacité linguistique chez l'homme peut être dit avant toute autre chose, pragmatique. Les animaux qui ne possèdent pas le langage parlé communiquent par d'autres canaux, d'autres signaux, d'autres langages, qui nous demeurent pour la plupart insaisissables (S. Strum, 1995).

Le travail de terrain d'Anne Robichez³ sur le pointer du doigt chez le petit enfant comme étape nécessaire du développement ultérieur d'un langage parlé vient compléter les travaux de M.-C. Corballis⁴ pour corroborer la théorie générale d'une évolution d'un langage manuel vers le langage parlé. Ce lien, déjà évoqué, entre les zones du cerveau commandant tout à la fois main, outil et langue peut laisser supposer qu'il ne manque aux chimpanzés et aux orang-outans qu'une morphologie adaptée qui leur permettrait d'émettre les signifiants d'une langue parlée pour communiquer par la parole, puisqu'ils partagent à la base les mêmes configurations que l'homme au niveau du cerveau.

Sur le plan morphologique, le petit enfant est potentiellement capable d'apprendre toutes les langues, et donc d'émettre tous les sons qui leur correspondent. Il perd cette capacité en apprenant à émettre certains sons au cours de son développement, et au fur et à mesure que sa morphologie se fixe sur la production de certains sons plutôt que d'autres. Dès le premier âge, c'est par les pleurs et les cris que le petit enfant exprime sa faim, sa soif, sa colère, etc. à quoi les adultes répondent par des paroles. L'enfant apprend alors à parler par imprégnation, c'est-à-dire par le simple fait d'entendre parler autour de lui, et parce que cette parole entendue est inévitablement associée à des messages émotionnels et affectifs de premier ordre (B. Cyrulnik, 1995, pp. 43-45). Ainsi, le langage parlé appartient pleinement au programme naturel (B. Cyrulnik, 1995, pp. 77-79) (i.e. dont l'apprentissage n'a pas à être

l'équilibre émotionnel et affectif de l'enfant, au même titre que ses autres fonctions vitales en dépendent. Par ailleurs, des éléments comparables ont été mis en évidence par les éthologues pour de nombreuses autres espèces : l'évolution « normale » des petits d'animaux nécessite l'existence d'un certain type de communication et de langage, entre le petit et une figure maternelle ou maternelle. L'une des expériences les plus connues étant celle de la « maternité » de Konrad Lorenz avec les oies cendrées.

¹ « De la fonction phatique qui nous intéresse ici, Jakobson dit que, « le message requiert un contact, un canal physique et une connexion psychologique entre le destinataire et le destinataire, contact qui leur permet d'établir et de maintenir la communication ». La fonction phatique est ainsi « accentuation du contact », par laquelle on cherche à « attirer l'attention de l'interlocuteur ou à s'assurer qu'elle ne se relâche pas ». (Marandon, 1989, p. 11).

² « Pour Darwin, [...] c'est la sympathie - ou son contraire - qui préside aux relations entre les individus ». (Y. Leroy, 1986, p. 15).

³ Robichez (A.). 1989. *Le pointer de l'index*. Toulouse [Thèse de médecine].

⁴ Corballis (M.-C.). 2002. *From hand to mouth. The origins of language*. Princeton : Princeton University Press.

forcé) de développement humain comme l'un des éléments naturels *vitaux*, nécessaires à sa survie. En ce sens, la capacité linguistique apparaît dans un cadre normal de développement comme le seul outil auto-généré de façon naturelle par simple imprégnation. Ces résultats concernant l'apprentissage de la langue chez le petit enfant corroborent la définition de la capacité linguistique comme technique.

Contrairement à ce qu'on appelle normalement un *outil*, et qui est une extension exogène du corps humain, la capacité linguistique a pu se développer d'autant plus naturellement – bien que liée à des raisons évolutives – qu'elle représente le seul outil *endogène*, c'est-à-dire généré de manière auto-suffisante *par* le corps humain et *pour* le corps humain et la satisfaction de ses besoins vitaux et sociaux. Selon la perspective pragmatique évolutive ici adoptée, le langage et la spécialisation humaine du langage que sont les langues, a pour principale fonction de permettre à tout homme qui maîtrise celui qui a cours dans un milieu donné, de pourvoir à tous ses besoins sociaux et vitaux, sans recours à aucune autre ressource extérieure. Le langage – tout langage – est bien en cela un outil endogène.

Chez les animaux, les phéromones – qui sont une autre forme de langage, d'un niveau de complexité autre par rapport à ceux que nous manions et maîtrisons en tant qu'humains – remplissent des fonctions identiques en termes de pragmatique et de survie, à savoir communiquer avec ses semblables, en échangeant des messages concernant essentiellement la reproduction et la nutrition (y compris chez les insectes¹), mais également l'organisation de la vie en « sociétés ». À ce stade, nous remarquons que l'écologie, entendue comme une « économie des relations à autrui »², fait loi : un lien existe entre développement du langage et développement des techniques, et l'on retrouve ce lien aussi bien au niveau du cerveau qu'au niveau des pratiques (ou des cultures...), et du développement des pratiques. Le développement de langages et de techniques aurait alors une seule et même fonction *première* (mais de nombreuses autres en sus), qui serait d'assurer la survie de l'individu et de l'espèce à travers l'individu, quel que soit ce langage (qu'il s'agisse de langage parlé ou de langage chimique)³. La langue ou capacité linguistique telle que nous la connaissons aujourd'hui développée chez l'humain, ne serait alors qu'un langage parmi d'autres et qu'un outil parmi d'autres, mais le plus perfectionné et probablement le plus développé de tous, devenu nécessaire à l'être humain pour sa survie.

6. Des techniques et des besoins

Le point de vue fonctionnel et pragmatique que nous adoptons sur le langage conduit à penser l'émergence de la capacité linguistique comme le développement d'une habileté langagière spécifique – c'est-à-dire propre à l'espèce humaine – qui est également une technique.

¹ Voir Y. Leroy, déjà citée, ou Jaisson (P.), Dahbi (A.), Lenoir (A.) & Hefetz (A.). « Comment les fourmis partagent leur odeur ». in : *La Recherche*, 314, pp. 32-34.

² En 1859, Isidore Geoffroy Saint-Hilaire, professeur de sciences à l'université de Paris et président de la société impériale d'acclimatation, définit une nouvelle discipline consacrée à « l'étude des relations des êtres organisés dans la famille et la société, dans l'agrégat et la communauté » qu'il nomme éthologie, mot dont le sens a évolué depuis pour désigner la science qui étudie les comportements animaux. La même année, Haeckel, biologiste allemand, désigne l'*oecologie* (du grec *oikos* « maison » et *logos* « science ») comme « la science des relations de l'organisme avec l'environnement comprenant au sens large toutes les conditions d'existence ». À la même époque, Charles Darwin (1809-1882) s'interroge lui aussi sur les rapports entre les animaux et les plantes. Une dizaine d'années plus tard, lors d'une conférence à Iéna, Haeckel précise sa définition de *oecologie* comme étant une « économie de la nature » et « la recherche de l'ensemble des relations de l'animal à la fois avec son environnement inorganique et organique ; ce qui comprend par dessus tout ses relations amicales ou hostiles avec ceux des animaux ou des plantes avec lesquels il est en contact directement ou indirectement ». Cette définition présente l'écologie comme le principe relationnel économique selon lequel chaque être vivant animal, dont l'homme, gère ses relations avec les autres êtres vivants et non-vivants. Les deux termes écologie et économie ont la même base étymologique *oikos* (maison). Il y a donc une logique indéniable qui mène de l'une à l'autre. Du point de vue de l'humain, ceci permet de fonder l'écologie comme l'économie des relations à autrui. Fracchiolla (B.). 2003. *Écologie et altérité : du discours de valeurs au discours de droits chez les Verts français et les Verdi italiens*. Paris : Université de Paris 3 [thèse de doctorat].

³ Comme nous l'avons déjà dit *supra*, nous avons conscience d'aller ici totalement à l'encontre des théories généralement admises concernant le « langage » chez les éthologues, qui le distinguent des « communications chimiques ». C'est pourquoi nous avons pris le soin de redéfinir à chaque étape le sens dans lequel nous employons chacun des termes.

A. Leroi-Gourhan avait déjà remarqué qu'outil et langage sont liés neurologiquement, et que la possibilité de langage émerge à partir du moment où la préhistoire livre des outils. On sait aujourd'hui que l'hémisphère gauche gère le traitement d'actions fines et précises. Aussi, une première hypothèse lie l'accroissement de la taille du cerveau à l'apparition des outils du langage, et permet d'envisager que le langage oral se soit développé à partir d'un langage manuel (M.C. Corballis : 2002). La langue elle-même – ou encore la capacité linguistique – définie comme un des langages possible, peut dès lors s'envisager comme une technique à part entière qui se serait progressivement développée jusqu'à devenir l'outil langagier principal de l'espèce humaine.

La conservation et l'évolution d'une technique dépendent au sein de chaque espèce de la pertinence de cette technique par rapport au besoin auquel elle répond à un moment donné et qui peut ou non perdurer. Certains grands singes lavent leurs aliments¹, des chimpanzés utilisent une paille pour se nourrir de termites, des céphalopodes « savent », de façon récurrente, ouvrir un bocal contenant une crevette. On a également observé le développement ponctuel puis l'abandon de certaines techniques telles que la chasse et le pillage chez certaines bandes de babouins (S. Strum²), ou le cassage de noix chez certaines communautés de chimpanzés (A. & J. Ducros, F. Jouliau, 1998)³. De façon comparable, l'expérience de D. et R. Fouts⁴ avec la chimpanzée Washoe et son fils adoptif Loulis ont montré que, comme technique, le langage humain pouvait être appris par les grands singes, voire perfectionné, et transmis à leurs petits dans un cadre où, suscité par des humains, l'apprentissage de cette technique était rendu nécessaire à la satisfaction des besoins élémentaires et moins élémentaires des singes. Toujours dans ce sens, D. Lestel remarque que l'orang-outan Chantek semble avoir totalement conscience de l'arbitraire du signe et initie souvent la conversation de son propre fait avec les hommes qui l'entourent⁵. En revanche, si les orang-outans n'utilisent pas le langage verbal dans leur milieu naturel, c'est simplement qu'ils n'en ont pas besoin (J.-A. Rondal, 2000, pp. 83-87). D'une manière comparable, les Hommes ont inventé puis développé toutes leurs techniques. Ainsi, comme conséquence de l'évolution, l'émergence de la capacité linguistique en tant qu'outil correspondrait à la maîtrise d'une technique survenue de manière nécessaire dans le cours de l'évolution ; peut-être provoquée par la nécessité d'aller à la rencontre de l'autre, de se présenter et de faire connaissance, de manière amicale ou combative, lors des déplacements de populations.

7. La fonction communicative de tout langage

L'une des fonctions du langage est de permettre de communiquer avec ses semblables, et éventuellement, avec d'autres animaux⁶. Tout langage a dans cette mesure une dimension fonctionnelle, instrumentale, liée à la base à la survie des individus au sein d'une espèce. Tout

¹ L'exemple de référence est celui des macaques rhésus laveurs de patates douces de l'île de Kôshima au Japon. Nous avons pu observer également cette technique de lavage d'aliment en Indonésie chez les macaques de la Monkey Forest, à Ubud, dans l'île de Bali.

² « Étant donné leur enthousiasme et leur succès en tant que chasseurs, j'avais pensé que cette activité deviendrait un élément permanent de leur comportement [...] mais les qualités de chasseurs des babouins de Pumphouse se sont révélées transitoires. [...] Vers la fin des années soixante-dix, [...] [les] techniques de chasse avaient disparues ». ; [...] « Certaines bandes ont une propension au pillage et d'autres recherchent des solutions différentes, mais il peut y avoir des animaux pilleurs et d'autres non pilleurs au sein de la même bande. [...] Chaque fois que les babouins peuvent se passer de piller les cultures, c'est-à-dire que la nature leur procure de la nourriture, même en petite quantité, ils cessent d'être des nuisibles ». (S. Strum, 1995, pp. 224-225).

³ Ducros (A.), Ducros (J.) & Jouliau (F.). 1998. *La culture est-elle naturelle ? Histoire, épistémologie et applications récentes du concept de culture*. Paris : Édition Errance, Coll. « Hespérides » [avec le concours du groupement de recherche CNRS « Biologie, Société et Culture »] et les travaux menés au cours du laboratoire du SHADYC, EHESS, Vieille Charité, Marseille, dirigé par Frédéric Jouliau.

⁴ D. et R. Fouts sont co-directeurs de « The Chimpanzee and Human Communication Institute » (CHCI) de l'Université Centrale de Washington.

⁵ Voir les travaux de Miles (H.-L.). 1983. « Apes and language : The search for communicative competence ». in : Luce (J. de) & Wilder (H.), (eds.). *Language in primates. Perspectives and implications*. New York : Springer, pp. 43-61, et Mile (H.-L.). « The cognitive foundations for reference in a signing orangutan ». in : Taylor Parker (S.) & Gibson (K.), (eds.). 1990. *Language and intelligence in monkeys and apes. Comparative developmental perspectives*. New York : Cambridge University Press, pp. 511-539.

⁶ Nous distinguons ici, comme le fait le français (mais non l'anglais) et pour éviter toute ambiguïté langue (parlée) et langage (la langue parlée en étant un).

langage est donc spécifique, au sens propre, à l'espèce qui l'a développé. Ainsi en va-t-il également de la langue parlée, articulée, humaine. D'ailleurs, lorsque nous ne pouvons pas ou plus l'utiliser pour diverses raisons (accident cérébral, surdité), nous développons d'autres outils pour communiquer, d'autres aptitudes, voire d'autres langages, qui nous permettent malgré tout de survivre et de continuer à satisfaire nos besoins les plus élémentaires. Même sur le cours d'une vie, l'être humain sait s'adapter physiologiquement et compenser éventuellement par de nouvelles aptitudes celles qui viennent à lui faire défaut. Ces éléments rejoignent les conclusions de l'article précédemment cité sur la disparition d'une fosse voméronasale parallèle au développement d'une vision en couleurs. Par ailleurs, notre observation des animaux et l'étude très partielle – voire partielle – de leurs langages et comportements nous permettent difficilement de formuler des affirmations définitives concernant leurs structurations en systèmes de signes et leur inscription dans le temps. Des chercheurs ont montré que les animaux échangent, et pas seulement dans l'instant ; qu'ils peuvent anticiper et mentir en se servant d'un langage (Voir D. Lestel, 2004). Bref, que la parole articulée n'est certainement pas le seul langage qui existe et que même après des années d'observations et de recherche, les scientifiques les plus aguerris ne viennent eux-mêmes pas plus à bout du langage des grands singes que les grands singes ne peuvent venir à bout du nôtre (S. Strum, 1995). La barrière de l'espèce joue, dans un sens comme dans l'autre¹.

Les origines du langage articulé parlé chez l'homme sont probablement multiples, et son émergence le résultat d'une longue et lente évolution générale. Avant de parler, les grands singes qui ont immédiatement précédé l'homme ne parlaient pas, mais utilisaient d'autres langages. Qu'il s'agisse de gestes précis, du regard, de l'intonation, de la fréquence vocale, ou encore de cris ou de postures/attitudes générales, l'Homme utilise toujours aujourd'hui une communication multimodale et une quantité infinie de signes qui viennent étayer le message purement verbal, dont l'écrit est une trace d'une certaine forme, au même titre qu'un enregistrement vocal : mais ce ne sont là seulement que des traces comparées à l'énonciation elle-même prise dans sa situation de communication².

Les linguistes, avec en particulier les travaux de Catherine Kerbrat-Orrecchioni, se sont intéressés à ces paramètres d'étayage du langage parlé. C'est d'une part sur la base de ces travaux et réflexions en sciences du langage et d'autre part sur la base de travaux en éthologie sur les grands singes et également sur le rôle des phéromones dans la communication entre animaux, que nous nous fondons dans les paragraphes qui suivent.

8. Conclusion

Pour les linguistes, ce qui distinguait la communication animale de la communication humaine était constitué par certaines propriétés des langues naturelles (en particulier la double articulation, l'élasticité du discours ou le débrayage – qui permet à l'homme de parler d'autre chose que lui-même). Les avancées des sciences cognitives, de la biologie, de l'éthologie comme celles de la psychologie animale et de la zoo-sémiotique remettent aujourd'hui en question d'anciennes certitudes. D. Lestel explique qu'une narrativité potentielle émerge dès que se dégage une possibilité de mime, y compris chez les animaux : la narrativité n'est donc pas spécifiquement humaine : « un animal qui joue raconte déjà une histoire comme celui qui trompe un concurrent potentiel » (Lestel, 2004, p. 97). Ainsi voit-on que certains des points qui furent pendant longtemps considérés comme spécifiques du langage articulé humain, ne peuvent plus l'être aujourd'hui. C'est pourquoi l'on parle à présent, sur le plan linguistique, plutôt en termes de gradation que de limite pour dire qu'il existe plutôt des langages ayant différents niveaux de complexité que des langages d'un côté et des « signaux » ou « instruments de communication » qui n'en seraient pas de l'autre³. J.-A. Rondal distingue pour sa part trois niveaux de complexité : « L1 : systèmes de signaux et/ou signes isolés, éventuellement arbitraires et analysables, L2 : systèmes de signes combinables utilisables pour la communication de significations relationnelles, L3 : langages humains modernes » (J.-A. Rondal, 2000, pp. 31-33). Dans cette mesure, l'on peut dire que les langages humains,

¹ Si les orang-outans n'utilisent pas le langage verbal dans leur milieu naturel, c'est simplement qu'ils n'en ont pas besoin (Rondal (J.-A.). 2000, pp. 83-87).

² Et nous ne souhaitons justement pas aborder ici la problématique de la « trace » liée à la construction du savoir et à la complexification exponentielle du langage articulé parlé et des interactions humaines, pour nous en tenir justement à la dimension immédiate de l'interaction (quel que soit son contenu).

³ Voir Courtès (J.), Greimas (A.-J.). 1993. *Sémiotique*. Paris : Hachette, p. 204.

dont le langage verbal constitue aujourd'hui l'aboutissement le plus développé et le plus spécifique, ont évolué en diachronie, au rythme de l'évolution, chacun de manière spécifique. En termes de sémiotiques, les animaux utilisent vraisemblablement des combinaisons d'unités distinctives non motivées (tels les phonèmes) et non essentiellement indicielles lorsqu'elles sont prises séparément, mais qui s'articulent entre elles en de véritables signes (tels les sèmes) pour signifier autre chose¹. Toujours présente, la barrière spécifique oblige au doute sur tous les éléments qui nous échappent. Au regard de tout cela et de nombre de recherches réalisées en éthologie et biologie, l'attitude scientifique la plus prudente à l'égard de la possibilité de langage chez les animaux semble être aujourd'hui à l'opposé de ce qu'elle est et à été depuis toujours. Et nous pensons que l'ensemble des Animaux dispose de langages et que chaque langage existe en lien intrinsèque et unique avec l'espèce qui l'utilise. La véritable question est peut-être alors celle de l'existence de systèmes parallèles dans le règne animal, demeurant distincts car intra-spécifiques, plus que celle de leur différence.

Nous savons que les signes non verbaux participent de façon conséquente à notre communication verbale. Le visage et ses expressions est le support grammatical de la Langue des Signes Française. Comment être certains alors que les attouchements d'antennes de certains insectes, leurs danses de parade, les signes échangés par les animaux en général ne jouent en quelque manière un rôle équivalent et ne correspondent pas à une forme de l'expression faite de traits distinctifs (que nous étudions, voyons, et comprenons sans réellement jamais les comprendre parce que nous ne sommes pas l'autre insecte/animal auquel cela est destiné) ? Comment être sûrs alors également que les phéromones n'entrent pas elles aussi dans la composition et l'articulation d'un langage propre dont émerge un signifié distinct non équivalent avec le signifiant ? Il paraît difficile d'en être sûrs. La notion de « sémiologie symbolique » semblerait ici plus appropriée, si l'on reprend les distinctions formulées par Ch. Peirce². Les phéromones sont des substances chimiques de communication d'une extrême complexité qui entrent en composition (selon une syntaxe spécifique ?) les unes avec les autres pour donner certains signaux de communication.

Chez les Fourmis, la communication chimique atteint un haut degré de complexité. Une même Fourmi, une *Acanthomyops* clavier par exemple, possède des glandes mandibulaires qui contiennent des terpènes dont du citronellal, du néral, du géraniol, une glande de Dufour qui renferme plusieurs hydrocarbures et deux méthylcétone (Blum, 1969, 1974) [...]

Chez les Fourmis du genre *Formica*, on a identifié quarante-six substances volatiles dont surtout des hydrocarbures linéaires, l'undécane étant le principal constituant. À ces substances s'en ajoutent d'autres, moins volatiles, comme le géranylgeraniol et son acétate, l'acétate d'octadécényle, et divers acétates d'alcools aliphatiques. (Leroy, 1986, pp. 66-67).

¹ Nous reprenons ici, pour notre analogie, autour des « langages possibles » les travaux de Hjelmslev. Concernant les langues, il propose la décomposition du signifiant et du signifié en unités plus petites. La notion de signe cède la place à celle de langage. Les deux plans d'un langage – l'expression et le contenu – doivent être étudiés séparément car chacun possède sa spécificité. Les unités minimales constitutives des plans d'un langage sont les *figures* : *phème* est utilisé pour désigner la figure du plan de l'expression et *sème* pour désigner la figure du plan du contenu. Sur le plan du contenu, les figures que sont les sèmes s'assemblent pour former le signifié d'un mot : un signifié est décomposable en sèmes. Sur le plan de l'expression, ces unités ne sont que distinctives, elles s'assemblent entre elles pour former des unités qui, elles, sont significatives : les *formants*. Le formant est le terme sémiotique qui correspond au signifiant de Saussure. En séparant expression et contenu en forme et substance, Hjelmslev ouvre ainsi la voie aux sciences du langage.

² Ch. Peirce distingue l'indice, de l'icône, du symbole. L'« indice » renvoie à l'objet qu'il dénote du fait qu'il entretient un rapport de contiguïté avec lui. Ayant des qualités communes avec l'objet qu'il dénote, il en découle que l'« indice » n'est pas un signe arbitraire. L'« indice » est un signe qui est en continuité et contiguïté avec l'objet. Par exemple, l'empreinte d'un pas sur le sable / est indice du « pas », la /fumée/ est un indice du « feu ». L'« icône » est un signe qui renvoie à l'objet qu'il dénote du fait d'un rapport de similitude à l'objet. L'icône est une image, une ressemblance, un signe analogique qui possède les propriétés de l'objet : il est déterminée par un rapport de ressemblance entre le signifiant et le signifié. L'icône ressemble à la chose sans en être. Le « symbole » renvoie à un objet qu'il dénote en vertu d'une loi, loi qui détermine l'interprétation du symbole par référence à l'objet en question. Le lien unissant le symbole à l'objet est conventionnel et de ce fait totalement arbitraire. Il est établi par référence à un code ou des conventions. Il n'a plus de rapport analogique avec la chose mais simplement conventionnel : arbitraire par rapport à elle, il se déchiffre à l'aide d'un code.

Considérer ces échanges comme relevant d'une communication indicielle est certes la lecture la plus simple et la plus habituelle pour nous. Mais est-il réellement impossible d'imaginer qu'une fourmi capable d'user d'une palette aussi importante de substances en les combinant différemment les une aux autres, « choisisse » d'utiliser tel composant ou signifier telle chose spécifique à l'un de ses congénères ? La question du « choix » de la fourmi demeure. Et avec elle la question de savoir si les unités qui entrent dans la composition de ce « langage » (qui est donc codé par l'émetteur, et que le récepteur doit par conséquent décoder) ne sont pas constituées d'unités distinctives non motivées et elles-mêmes non signifiantes par elles-mêmes (comme cela est le cas pour les phonèmes/lettres utilisés dans la composition des langues), mais qui deviennent signifiantes uniquement par leur combinaison et un agencement complexe des unes avec les autres. « Les phéromones de pistes des fourmis sont souvent des mélanges très complexes de substances ». (Y. Leroy, 1986, p. 70)¹. Si la question du perlocutoire semble aisée relativement aux communications chimiques que sont les phéromones, c'est ici la question de l'illocutoire qui demeure (et demeurera) en suspens, nous renvoyant à nouveau d'abord à la barrière inter-spécifique, mais aussi (et surtout) à la problématique afférente de la conscience, par l'intermédiaire de l'éventualité du choix ; et aussi pourquoi pas à la question de l'inné et de l'acquis². Peut-être que, comme on peut soutenir l'existence d'un langage intra-spécifique, il existerait une forme de conscience intra-spécifique parallèle, qui n'a rien à voir avec celle que nous interrogeons en tant qu'Humains.

Notre objectif était de proposer un point de vue différent sur le langage qui ouvre d'autres questionnements sur la langue « parlée » articulée humaine (nous incluons bien sûr les langues des signes) au profit de la diversité et complexité du vivant et de ses langages. L'objectif était d'observer la langue parlée articulée humaine, notre « langage » comme faisant partie de cette complexité. Nous avons voulu également nous interroger sur les sens donnés à certaines notions, comme celle de « langage », d'« indice », d'« écologie » ou encore de « communication » lorsqu'elles passent les frontières de l'une ou l'autre discipline, et envisager les questionnements qui s'ensuivent. En conclusion, nous constatons que probablement, en tant qu'espèce animale, les humains partagent avec l'ensemble de leurs semblables certaines des caractéristiques liées au(x) langage(s) et à leurs origines et que, sans doute, ces caractéristiques communes sont révélées dans la domestication et les langages développés en commun pour communiquer entre les hommes et les autres espèces (et probablement les mammifères plus que les autres encore). Si les langues sont (devenues) les instruments langagiers privilégiés des Hommes, elles sont cependant loin d'être les seules, et elles ont probablement en commun beaucoup plus d'éléments que nous ne pourrions jamais le savoir avec d'autres langages animaux, si l'on se penche non seulement sur leur(s) origine(s) – qui sont certainement plurielles – mais surtout sur les origines des conditions de leur émergence.

¹ (Leroy, 1986), nous renvoyons à l'ensemble de l'ouvrage, fort riche en éléments de discussion sur le sujet.

² Des expériences faites sur des souris mâles montrent que si on prive les jeunes de l'utilisation de leurs phéromones, ils sont incapables de reconnaître la période d'oestrus chez les femelles. En revanche, la même expérience chez les mâles plus âgés ne semble pas influencer de manière conséquente sur leurs capacités à reconnaître le moment propice à la fécondation. Il semble que cette remarque aille à l'encontre d'une communication phéromonale qui serait purement indicielle. (en termes linguistiques) Si elle l'était, la relation naturelle de cause à effet devrait se répéter normalement, et faire correspondre à l'absence d'indice, une absence de comportement.

Références Bibliographiques

- Barbier (M.). 1982. *Les phéromones. Aspects biochimiques et biologiques*. Paris : Masson.
- Bailey (K.). 1978. « Flehmen in the ring railed Lemur (*Lemur catta*) ». in : *Behaviour*, 65, pp. 309-319.
- Blum (M.-S.). 1969. « Alarm pheromones ». in : *Ann. Rev. Entomol.*, 14, pp. 57-80
- Blum (M.-S.). 1974. « Pheromonal society in Hymenoptera ». in : Birch (M.-C.), (ed.). *Pheromones*, pp. 222-249. Londres : North-Holland.
- Brossut (R.). 1996. *Phéromones. La communication chimique chez les animaux*. Paris : Belin, CNRS éditions, Coll. « Croisée des sciences ».
- Cavalli-Sforza (L.). 1994. *The history and geography of human genes*. Princeton: Princeton University Press.
- Cavé (C.), Gaïtella (I.) & Santi (S.), (eds.). 2001. *Oralité et gestualité. Interactions et comportements multimodaux dans la communication*. Paris : L'Harmattan, pp. 79-96, chap. *Communication animale et interspèces* [Actes du colloque *Orage 2001*, Aix-en Provence, 18-22 juin 2001, Organisé 1- par le Laboratoire Parole et langage (CNRS) et l'Université de Provence ; 2 – par l'Association GeVoix].
- Chapouthier (G.) & Matras (J.-J.) 1988. *Mémoire et cerveau : biologie de l'apprentissage*. Monaco : Éditions du Rocher.
- Corballis (M.-C.). 2000. « Did language evolve from manual gesture ? ». in : *The Evolution of language* [3rd conference, 3-6 avril 2000].
- Corballis (M.-C.). 2002. *From hand to mouth. The origins of language*. Princeton : Princeton University Press.
- Cosnier (J.), Coulon (J.), Berendonner (A.) & Orecchioni (C.). 1982. *Les voies du langage. Communications verbales, gestuelles et animales*. Paris : Dunod.
- Courtès (J.) & Greimas (A.-J.). 1993. *Sémiotique*. Paris : Hachette.
- Cyrułnik (B.). 1995. *De la parole comme d'une molécule, entretiens avec Emile Noël*. Paris : Seuil.
- Cyrułnik (B.) & Soulé (M.). 1998. *L'intelligence avant la parole. Nouvelles approches originales du bébé*. Paris : ESF Éditeur, Coll. « La vie de l'enfant ».
- Ducros (A.), Ducros (J.) & Joulian (F.). 1998. *La culture est-elle naturelle ? Histoire, épistémologie et applications récentes du concept de culture*. Paris : Éditions Errance, Coll. « Hespérides » [avec le concours du groupement de recherche CNRS « Biologie, Société et Culture »].
- Fracchiolla (B.). 2003. *Écologie et altérité : du discours de valeurs au discours de droits chez les Verts français et les Verdi italiens*. Paris : Université de Paris 3 [Thèse de doctorat, dir. Martine Abdallah-Pretceille].
- Fracchiolla (B.). 2006. « Les métamorphoses du langage ». in : Abdallah-Pretceille (M.), (dir.). *Les métamorphoses de l'identité*. Paris : Éditions Anthropos-Economica [en cours de publication].
- Gilad (Y.), Man (O.), Pääbo (S.), Lancet (D.). 2003. « Human specific loss of olfactory receptor genes ». in : *PNAS*, 100, 6, pp. 3324-3327 [March].
- Haeckel (E.). 1866. *Generale morphologie der organismen*. Berlin : Reimer.
- Hall (E.-T.). 1984. *Le langage silencieux*. Paris : Seuil [1^{ère} ed. 1959. *The silent language*. New York].
- Hauser (M.-D.). 2002. *À quoi pensent les animaux ?* Paris : Odile Jacob.
- Hjelmslev (L.). 1968-1971. *Prolégomènes à une théorie du langage*. Paris : Les Éditions de Minuit.
- Gilad (Y.), Man (O.), Pääbo (S.) & Lancet (D.). 2003. « Human specific loss of olfactory receptor genes ». in : *PNAS*, 100, 6, pp. 3324-3327 [March 18].
- Lacks (B.) & Victorri (B.), (dir.). 2002. « L'origine du langage ». in : *Langages*, 140, juin.
- Leroi-Gourhan (A.). 1964. *Le geste et la parole*, t. 1, *Techniques et Langage*. Paris : Albin Michel.
- Leroy (Y.). 1986. *L'univers odorant de l'animal*. Paris : Boubée
- Lestel (D.). 1995. *Paroles de singes. L'impossible dialogue hommes-primates*. Paris : La Découverte.
- Lestel (D.). 2004. *L'animal singulier*. Paris : Seuil.
- Lledo (P.-M.), Carleton (A.) & Vincent (J.-D.). 2002. « Odors and olfaction ». in : *Journal de la Société de Biologie*, 196, pp. 59-65.

- Liman (E.-R.) & Innan (H.). 2003. « Relaxed selective pressure on an essential component of pheromone transduction in primate evolution ». in : *PNAS*, 100, 6, p. 3329. Cambridge, MA : Harvard University [edited by David Pilbeam].
- Meyerson (I.). 1948. « Discontinuité et cheminement dans l'histoire de l'esprit ». in : *Journal de psychologie normale et pathologique*, juillet-septembre. Paris : P.U.F.
- Meredith (M.). 2001. « Human vomeronasal organ function : a critical review of best and worst cases ». in : *Chemical senses*, 26, pp. 433-445.
- Miles (H.-L.). 1983 « Apes and language : The search for communicative competence ». in : Luce (J. de) & Wilder (H.), (eds.). *Language in primates. Perspectives and implications*. New York : Springer, pp. 43-61.
- Mile (H.-L.). « The cognitive foundations for reference in a signing orangutan ». in : Taylor Parker (S.) & Gibson (K.), (eds.). 1990. *Language and intelligence in monkeys and apes. Comparative developmental perspectives*. New York : Cambridge University Press, pp. 511-539.
- Pierce (C.-S.). 1978. *Écrits sur le signe*. Paris : Seuil [Textes rassemblés, traduits et commentés par Gérard Deledalle].
- Robichez (A.). 1989. *Le pointer de l'index*. Toulouse [Thèse de médecine].
- Rondal (J.-A.). 2000. *Le langage : de l'animal aux origines du langage humain*. Sprimont (Belgique) : Mardaga.
- Salzen (E.). 2006. « From Calls to Words : How ethology can bridge the divide ». in : *Marges Linguistiques*, 11, mai 2006. Saint-Chamas : M.L.M.S. Éditeur.
- Spehr (M.), Schwane (K.), Heilmann (S.), Gisselmann (G.), Hummel (T.) & Hatt (H.). « Dual capacity of a human olfactory receptor ». in : *Current Biology*, 14, 19, R832.
- Stern (K.) & McClintock (M.-K.). 1998. « Regulation of ovulation by human pheromones ». in : *Nature*, 392, pp. 177-179.
- Strum (S.-C.). 1995. *Voyage chez les babouins*. Paris : Seuil.
- Tyack (P.). 1993. « Animal language research needs a broader comparative and evolutionary framework ». in : Roitbalt (H.), Herman (L.) & Nachtigall (P.), (eds.). *Language and communication : Comparative perspectives*. Hillsdale, NJ : Erlbaum.
- Ulhin (T.-P.). 2000. « The role of hand in the evolution of language ». in : *Psychology*.
- Victorri (B.). 2002. « *Homo narrans* : le rôle de la narration dans l'émergence du langage ». in : *Langages*, 146, Juin.
- Vincent (L.). 2004. *Comment devient-on amoureux*. Paris : Odile Jacob.
- Vossh (L.-B.). 2004. « Olfaction : Attracting Both Sperm and the Nose Dispatch ». in : *Current Biology*, 14, R919-R920 [November].
- Watzlawick (P.). 1975. *Changements*. Paris : Seuil [trad. fr.].



May 2006

1. Introduction²

I focus in this essay on only some aspects of language evolution, viz., language speciation and language death, although I also discuss some structural changes, more specifically phonological and syntactic ones. I argue, contrary to the dominant practice in the 20th century, that a biological approach to evolution is applicable to languages, although I must clarify at the outset that languages should be analogized to species rather than to organisms. Moreover, the application works best when linguistic species are analogized with viral rather than with animal species. I argue eventually that the approach is analogical only to the extent that it is inspired by scholarship on biological evolution. Otherwise, linguistic species can be posited in their own right (Mufwene, 2001). They share properties with other species, biological and otherwise, while they also differ from the latter in interesting ways that are specific to their ontogenetic, architectural peculiarities. Thus genetic linguistics can contribute to theories of evolution, adding hypotheses that are specific to, and/or inspired by, languages as species.

The invocation of population genetics in the title of this essay underscores two working assumptions of mine since Mufwene (2001), viz., (1) the agency of the evolution of a language lies in the individual communicative acts of its speakers, which is similar to the fact that a biological population or species is cumulatively affected by the experiences or activities of the individuals it consists of³; (2) the communicative activities that produce language evolution are largely determined by the socio-economic ecologies in which speakers evolve, which is similar to saying that the ecology rolls the dice in evolution. However, I should now explain why I return to an approach that has been unpopular until recently and justify some of the terminology that I now prefer to use in my work, such as *evolution*, *competition*, and *selection*.

2. Why this return to a biological approach ?

By the time Charles Darwin published *On the origin of species* (1859), genetic linguists had already been addressing some of the questions that I address here. According to Maher (1983, p. xxv), August Schleicher felt so vindicated by this book that he responded to it with an elaborate discussion of analogies that he perceived between linguistic and biological evolution (Schleicher, 1863, 1869). He covers notions such as « variation », « selection », « life of languages » (as new ones emerge and some of the current ones die), « gradual development » from the most « primitive » to more advanced and complex stages, « genus »/« family », and « species » (not in the way I use it in the present discussion). He also notes that confusion on whether a language variety should be identified as a language or as a dialect is parallel to what Darwin reports about identifying a class of biological organisms as a species or a subspecies.

¹ This paper was first published in 2005 in : *Gene, Sprachen, und ihre Evolution*, Schriftenreihe der Universität Regensburg, volume 29, pp. 30-52, ed. by Günter Hauska, Universitätsverlag Regensburg. We thank the author, the editor, and the publisher for authorizing *Marges Linguistiques* to reprint this text, which has been edited only minimally. May 2006, M.L.M.S. Publisher. Saint-Chamas (France).

² I am grateful to David Hull for feedback on an earlier version of this essay. We disagree on whether viruses can be grouped into species, on a par with animal species. I saw no reason to change my position, as I also assume that the evolutionary peculiarities of various species are largely determined by their respective ontogenetic properties and their particular modes of gene or feature transmission. I am solely responsible for the shortcomings of the positions submitted here.

³ Paul (1891, chapters 1 and 2) says something that amounts to this position.

Overall, he concludes, in the spirit of his self-professed uniformitarianism :

The rules now, which Darwin lays down with regard to the species of animals and [of] plants, are equally applicable to the organisms of languages, that is to say, as far as the main features are concerned (cited from Koerner, ed. 1983, p. 30).

Schleicher saw in Darwin's principle of natural selection an explanation for the changes that produce the language speciation patterns represented by the *Stammbaum* « genetic tree » that he was developing, which apparently inspired the biologists' cladogram. He also assumed, like Darwin (Radick, 2002), that different human populations had not reached the same stages of evolution – some being rather primitive and others particularly advanced – and that their languages reflected this putative variation. He thus posited an evolutionary trajectory of languages of the world from the isolating morphosyntactic type to the fusional morphology type. The former type, which he considered as the most primitive, includes, for example, Chinese, Thai, and the Kwa languages of West Africa, whereas the second type, which he claimed to have reached the most advanced level of evolution, includes Western European languages (Schleicher, 1863, 1869).

This position, which was apparently widely accepted by the late 19th century – even by Charles Darwin (Hull 2002, p. 13 ; Radick, 2002, p. 7, p. 13), led French philologists such as Adam (1882, 1883) and Vinson (1882, 1888) to treat creoles and pidgins, which tend to have a predominantly isolating morphosyntax, as languages in their pristine, natural, or primitive state, i.e., as degenerations from the « refined » structures of the European languages they have evolved from. (See DeGraff (2003) for an elaborate discussion.) Although creoles and pidgins have continued to be associated with non-ordinary, contact-based developments and as special cases that putatively do not fit naturally in the domain of genetic linguistics (however, cf. Mufwene, 2003), the Schleicherian evolutionary model has generally been either discredited or ignored over the past century. Except for Bickerton (1984, 1990), who, through his language bioprogram hypothesis, has suggested that creoles and pidgins give us an idea of the protolanguage in human phylogeny, no language typologist has ever suggested any particular evolutionary ranking of structural types¹.

It is not clear why the whole biological approach was abandoned in the 20th century, except that the comparison of languages with organisms was not particularly insightful, for reasons which I articulate below². Hoenigswald (1990, p. 11) also notes that a « language is no corporeal something with physical boundaries, or molecular permanence, or genome ». According to him, this ontological difference from a biological organism or species would make it difficult to account for language evolution on the biological model. He adds, « Try as we may, we can only think of it as the totality of the very traits that are subject to change » (p. 11). Hull (2002, p. 18) is right on the mark in observing that « most of [the] effort in [genetic] linguistics was expended to working out the *histories* of various languages. Much less attention was paid to the *processes* by which languages change ».

¹ Assuming mistakenly that creoles have evolved from pidgins, Bickerton has argued since his *Roots of language* (1981), that pidgins have no syntax or grammar, having been made by adults, who presumably no longer had (full) access to Universal Grammar (UG) or the biological endowment for language. Accordingly, children, guided predominantly by this language organ, would have imposed a UG-based grammar onto their parents' structure-less pidgins, which would account for world-wide structural similarities among them. The same UG must account for the emergence of protolanguage, for which he has produced no structural evidence, among our hominid ancestors. He assumes that its protosyntax must have been similar to that of child language, which shares features with that of creoles. (Anybody who pays attention to cross-creole structural variation would speak of creole grammars, in the plural, instead !) Every one of these claims is highly disputable, supported by tenuous evidence, but it would be digressive to discuss this issue here. See, e.g., Mufwene (1991, 1992a), Lieberman (1991, 2002), McNeill (1992), and Pinker (1992) for informative discussions of Bickerton's thesis and Mufwene (2001) and Chaudenson (2001) for alternative interpretations of the facts that constitute the sociohistorical ecologies of the evolution of creoles. DeGraff (1999a, 1999b, 2005) is equally informative, from a language acquisition perspective. Suffice it to note here that Bickerton (1984, p. 141, pp. 157-158) also suggested that creoles instantiate languages in a primitive developmental stage.

² Linguists such as Franz Bopp and August Schleicher seem, however, to have used the term *organism* with the meaning of *species*, a term that also occurs in the work of Schleicher. They speak of interindividual, now known also as interidiolectal, variation within a language, to which Darwin's evolutionary principle of natural selection could also be invoked to apply (Schleicher, 1863). One important 19th-century linguist whose views foreshadowed some of those expressed in this article is Paul (1891).

Indeed, the comparative method, on which genetic linguistics has critically relied, is used to determine the extent to which languages share linguistic materials (typically, words, sounds, and morphemes) and can be claimed to be genetically related. However, it cannot be used alone to determine whether the shared forms and structures have been inherited from a common ancestor or borrowed from a common, influential non-ancestral language they all came in contact with. Nor can it be used alone to exclude the possibility of mutual influence among the relevant languages if they have been spoken in the same geographical area or in adjacent ones. In other words, used alone, the comparative method cannot distinguish between « genetic relatedness » (based on materials inherited from a common ancestor), « areal diffusion » (due to borrowings from the same language), and « convergence » or *Sprachbund* (due to mutual influence in a contact area).

Areal linguistics (see, e.g., Heine & Kuteva, 2005) underscores the importance of factoring the history of population contacts in genetic linguistics. Moreover, as shown in Mufwene (2003a), there has been very little cross-pollination between genetic linguistics and genetic creolistics (which focuses on the development of creoles, under the conditions of population contacts in plantation settlement colonies). It is not clear why linguists, including some of the most influential creolists, have ignored the fact that genetic creolistics has dealt with ecological details of recent cases of language speciation. Yet, these bear on genetic linguistics, especially in regard to the seemingly gratuitous assumption of asexual parentage (without mating/contact), i.e., the iterated splitting of languages into new varieties without external causation (see below).

On the other hand, genetic linguistics has stuck loyally to Schleicher's *Stammbaum*, despite its significant share of shortcomings. For instance, the approach has focused more on the outcomes of changes than on the mechanisms of changes themselves. Among other things, it does not account for « substrate » influence, from the other languages that the relevant language has been in contact with. Thus the role of substrate influence in the speciation of Latin into the Romance languages has received only lip service. The role of Celtic languages in the evolution of British English dialects other than Irish (either on the Germanic languages that the Jutes, Angles, and Saxons brought from continental Europe or since the development of Old English) is only now starting to receive some attention (e.g., Tristram, 1999, 2000 ; Vennemann, 2001 ; Filppula & al., 2002).

The monoparental, *Stammbaum* approach also makes it impossible to account for hybridization in language evolution as a natural process. I argue below that the approach has been a major obstacle to learning from the development of creoles lessons that should apply to all cases of language speciation such as the development of the Romance languages or, on a larger scale, the diversification of Indo-European and of Bantu languages.

In Mufwene (2001), I submit the following reasons to explain why communal languages are more like species than like organisms, which I claim to be the counterparts of idiolects. Languages are indeed populations of idiolects¹ and are as abstract as species, which are mere categories (Hey, 2001) projected from structural or ontogenetic properties shared by classes of individuals or organisms². As the universe of our experience and knowledge is continuous, the boundaries of both species and languages are naturally fuzzy and operationally arbitrary, imposed by particular ideologies or other practical or theoretical considerations, especially within the same genetic family. It must be difficult to draw the boundary between adjacent vernacular varieties of Italian and French or between those of Dutch and German, a problem aggravated by the arbitrariness of political boundaries.

Unlike organisms, languages qua species are heterogeneous beyond the non-monolithic architecture of idiolects, a property that follows from the fact that idiolects vary among themselves and are alike by the family resemblance principle³. They are thus like biological

¹ For non-linguists, idiolects are the distinct ways in which individuals, as opposed to groups, speak their language, each one with his/her own peculiarities.

² The categories can of course also be posited a priori, based, for instance, on assumed common ancestry (O'Hara, 1994), with structural or ontogenetic features being invoked only to justify the grouping or to clarify their internal division into sub-categories.

³ Even Chomsky (2000) now adheres to this particular view of a language, explaining that lumping idiolects into the same language is a matter of likeness, not of sameness or identity. On the non-monolithic nature of language architecture – i.e., how the different structures coexist, with the functions

populations in which each organism preserves its genotypic and phenotypic individuality, despite the many features it shares with its conspecifics. Languages evolve in non-uniform ways, with some idiolects, sociolects (social varieties), or dialects (regional varieties) being more engaged in some changes than others. They are thus like biological species, whose members are differentially affected by, or react in different ways to, changes in their ecologies. As noted in the Introduction, changes affecting languages are a cumulation of changes affecting individual speakers' idiolects, which raises the following interesting question : Under what particular conditions do ontogenetic features of individuals or organisms spread to become those of (part of) a population ? We should note, along with McMahon (1993, p. 248), that there are many idiosyncratic innovations that are short-lived and do not so spread¹.

Languages are like populations/species in the ways they die and are born, viz., in a protracted way (Mufwene, 2004a). A language is dead when it has no more speakers left, just like a species is dead when no more specimen is left that can instantiate it². Cases of genocides precluded, languages do not typically suffer sudden deaths, because their speakers die or suffer linguistic dysfunction at different times – which entails gradualism. Moreover, they can die in one geographical or social setting but thrive in another, as is evidenced by Old World languages that have died in different places of the New World but continue to be spoken in their homelands or other parts of the New World. Like species, languages are not born in the same way as organisms. They have no moment of conception ; nor do they have a gestation or incubation period. They are identified as new ones post factum, after a particular variety (of idiolects or of organisms) has been identified that is considered (significantly) different from an earlier population.

Other characteristics can be invoked, but the present list should suffice to support the position that as communal extrapolations, languages should not be thought of as organisms. This position will enable us to capitalize on the interaction of both variation and ecology to account for evolution. I assume that a communal language exists only as a social construct, suggested by the ability of speakers in a particular setting to communicate successfully when they use similar sequences of spoken or signed gestures.

The bottom line is that every individual speaks or signs in a way that is internally systematic. Communication with other individuals triggers mutual accommodations, which make the individual systems similar both in the ways individual sounds or signs are produced and in how they combine together into more complex interpretable units. What is really required in such settings is ability on the part of each participant to understand and be understood by others.

Nobody is required to speak or sign in exactly the same way, and sameness hardly happens among idiolects. However, doing things in similar ways must facilitate mutual intelligibility. Languages as communal systems are construed on the basis of similarities in the ways people speak or sign in order to facilitate mutual intelligibility.

Perhaps it is no accident that folk definitions of languages do not include notions such as systems (consisting of units and principles) but rather boil down to « the way a particular group of people speaks ». Thus *German* means the way the people called *Germans* speak and (*Ki*)*Swahili* means the way the people identified as (*Wa*)*Swahili* speak. The fact that world languages such as English and French are spoken in different parts of our planet now does not contradict my observation, because these languages can also be defined etymologically as, respectively, « (originally) language of the Angles or English people » and « (originally) language of the Franks or French people ».

of some of them overlapping or sometimes conflicting with each other – see Mufwene (1992b). As explained in Mufwene (2001), the piecemeal way in which a language is « acquired », with features selected incrementally from diverse inputs/idiolects, accounts for this state of affairs.

¹ Similar ideas can be seen in Paul (1891, chapters 1 and 2).

² As Thomason (2001) points out, the common characterization of language death that I just repeated may be oversimplified, because the lone speaker would not have anybody else to speak it with and his/her knowledge would probably have fallen into attrition by then. One sense in which the characterization may be correct is when a language is thought of as knowledge rather than as practice. Then arises the question of the death of a language by transformation. What proportion of its original features must it preserve in the last speaker or in the last group of speakers in order to be considered the same and alive ?

All these considerations explain why it is necessary to focus on the agency of individuals and dynamics of their interactions. Accommodations made during individual communicative acts not only bear on the development of their communal properties but also drive evolution within their populations. Interesting questions include the following : Does a population interact collectively with its ecology ? Or, instead, is what is associated with a population a cumulation of the effects of individual interactions of its members with the relevant ecology ? What kind of equation can be posited to account for such a cumulation of effects ? How artificial is it to assume that the ecology acts on a population ? If the ecology operates on individuals, what precisely is the ecology of the individual ? If it is normal to assume an ecology for a population too, is the ecology of a population the same as the ecology of an individual ? What's the significance of these epistemological questions ? I must first explain some concepts that are fundamental in this intellectual exercise.

3. Competition, selection, ecology, evolution, and other related notions

I should clarify at this point that I discuss « language evolution », not « the evolution of language ». The latter deals only with the higher abstraction of what various languages share, with the ability to encode, store, and communicate information through a spoken or signed system, and with how this capacity evolved in mankind. *Language evolution* can apply to specific languages in a way similar to *language change*, to which it is closely related semantically. Since the latter term is more established in historical linguistics, the following question arises : why don't I stick to and use it ?

The term *evolution* covers more than the traditional term *change*. In addition to traditional concerns with structural and pragmatic changes, it also covers language speciation and language birth and death, processes to which the term *change* has not applied in linguistics. Also, although language speciation has been a concern of genetic linguistics, a branch of historical linguistics, seldom has the topic been related to those of language birth and death, which are so germane to it. Thus, the development of creoles and pidgins has typically been treated as anomalous or unusual (see, e.g., Hock & Joseph, 1996), if not as an aberration, despite the fact that the topic is, at least from a phylogenesis perspective, closely related to historical dialectology.

The traditional practice is made more bizarre by invocations of substrate influence in both genetic creolistics and genetic linguistics, especially in Romance linguistics, where the term *substratum* originated. It is as necessary to invoke substrate influence from the Celtic languages to account for the speciation of Latin into the Romance languages as it is to invoke African substrate influence to account for the evolution of French and Portuguese into various creoles, the role of other ecological factors discussed in Chaudenson (2001, 2003) and Mufwene (2001) notwithstanding. Both cases are clear instantiations of Pyrrhic victory – where the prevailing language is so clearly affected by the displaced ones – which makes the distinction between creole and non-creole languages more sociological than linguistic (Mufwene, 2001 ; DeGraff, 2003). While the term *change* has been used in cases of substrate influence, it has not been applied to those of language speciation as consequences of divergent structural changes. The term *evolution* offers the advantage both of its applicability to this case and of aligning research on language evolution with that of species evolution, which covers a similar range of developments¹. At least in the way I approach the subject matter, looking for causation in the communicative activities of individual speakers (see below), the term *evolution* brings along a refreshing perspective, viz., the alternative of interpreting the relevant evolutionary processes as adaptive responses to changing communicative ecologies, both external (including other speakers) and internal.

Contact plays an important role in my approach to language evolution, the most significant part being contact between individuals rather than between populations (Mufwene, 2001, chapter 2). In language evolution, the coexistence of two populations in the same geographical area is not a sufficient condition for language contact. They must interact with each other. At the level of linguistic communication, this is made possible by interacting individuals, who can

¹ Those who are not impressed by the distinction I make between *evolution of language* and *language evolution* should note that in biology the term *evolution* is also ambiguous between the two senses that I highlight here. The advantages that the term offers over the traditional term *change* in broadening the scope of genetic linguistics with topics which are germane to language speciation outweigh the inconvenience of an ambiguity that is easily resolved by the context of the discussion.

spread features from the other language among monolingual members of their respective languages. Moreover, as pointed out by Weinreich (1953), the real locus of language contact is the minds of individual speakers.

Given the piecemeal way in which language acquisition proceeds (see below), speakers' minds are the arenas where selection resolves the competition that takes place among features received from various speakers and, in the case of multilingual speakers, also between features of the target system and linguistic systems that they have already worked out. The phenomenon known in contact linguistics as *interference* represents those cases when features of, say, a speaker's mother tongue are substituted for features of the target language – for instance when the English *th* in *think* is pronounced as *s* or *f* (thus *sink* or *fink*) by a non-native speaker. From the point of view of the development of linguistic or communicative competence, the total amount of linguistic knowledge that speakers have, including the variants that compete for the same structural or communicative functions, is comparable to a « gene pool » in biology, both at the individual and at the population levels¹. In the case of language, the term *feature pool* (Mufwene, 2001) seems to be an apt analog. All of these observations make more sense once we re-examine the concepts of « language acquisition » and « transmission », as I do below.

Language transmission is a convenient misnomer in linguistics, because no speaker ever transmits a ready-made system to any learner, although caretakers simplify the learning task to the child by seriously limiting the number of topics for communication and the complexity of utterances. Both the range of topics and structural complexity (in terms of utterance length and embedding) increase as the child matures in age and in cognitive capacity, including processing skills. Otherwise, the burden of developing communicative competence rests on the learner, who, undoubtedly aided by *Universal Grammar*, must discover the units and principles of the targeted language or dialect and must (re)construct (an approximation of) its system².

Language acquisition is thus a (re)construction process comparable to gene recombination in biology (Mufwene, 2001, chapter 2). It can be called a *restructuring* process, though the term has mistakenly been associated with structural divergences that make creoles different from the European languages they have developed from. The main differences from gene recombination in biology lie in the fact that, in the best known cases, gene recombination takes place at the conception of an organism, while its genotype is being formed, with no agency on the part of the gene carrier. In those species of organisms such as viruses whose genotypes can change during their life time, one can argue that gene recombination takes place more than one time.

Languages are species whose phenotypes – the linguistic features on which typological classifications are based – correspond to no genotypes³. This peculiarity, which may appear strange to a biologist, is a consequence of the piecemeal way in which speakers develop

¹ One clarification is needed here. Although the number of variants at the population level is a set-theory union of variants at the level of individuals, the strengths of the variants at the population level is a function of different kinds of dynamics of interaction and influences among individual speakers. This is also the level where it becomes more obvious that selection does not necessarily eliminate disadvantageous variants. As observed by Kretzschmar & Tamasi (2003) in American dialectology, populations have long memories of variation. Even the most marginal features can be resilient.

² This otherwise useful statement is problematic. The contents of UG – that part of human predisposition to cognition that putatively specializes in the development of linguistic and communicative competence – has not been explicitly articulated, at least not in terms of how what it is claimed to do cannot be accomplished by a general learning module of the mind. It is not evident either that, while he/she develops competence in a particular language, the learner really purports to develop a system, as systematic as he/she wants to be. One can also argue that a system simply emerges to the analyst – or is projected by him/her – out of the routines that an individual develops while purporting to communicate with particular sequences of spoken or signed gestures. Fortunately, these issues do not bear (significantly) on the main theses of this essay.

³ Clearly languages and idiolects cannot have genotypes, because they are not biological systems. What is meant here is that they have nothing similar to a genotype in a biological organism or species. I also deliberately ignore here the fact that there is no isomorphism between phenotypes and genotypes, as well as the fact that ecology plays an important role in determining the phenotypes of a population. It is nonetheless true that only some genotypes can be associated with particular phenotypes, for instance the kinds of noses, hairs, lips, ears, limb morphology, and complexion that are found only among humans but not among other mammals.

competence in a language. From the point of view of population genetics, interesting questions arise regarding, for instance, why dialects are not more different from each other than they are ; why they wind up with similar, though not identical, combinations of features ; and how some, but not all of, the individual changes they undergo can amount to communal changes. These questions are similar to those regarding how, with genotypes that are different from each other's, organisms of a biological species do not wind up being more different internally than they actually are and why their phenotypes are so similar.

Needless to say biology will not always inspire linguists who address such issues, nor will insights have to flow only one way in such comparisons between biological and language evolutions. Linguistics can contribute research avenues for biology too, if ontogenetic properties of languages and their peculiar mode of transmission can provide alternative and independent ways of addressing them. One way of accounting for the piecemeal way in which a language is « acquired » is the fact that its features are copied (typically with modification) in ways that are closer to horizontal transmission in biology, as in epidemiology, than to vertical, generational transmission¹.

Although interactions between caretakers and children suggest that language is « vertically transmitted », there is much more peer influence from the time the child interacts with other children, despite the fact that every child takes something from the older people they interact with. The importance of social interaction in language development should actually question the wisdom of capitalizing on age differences, rather than on experience, in the way language is « acquired »/« transmitted ». More remains to be thought out on this topic.

Another important feature of language development is also polyploidy, which makes it possible for the learner to incorporate in their emergent dialects features originating in diverse dialectal sources, including variants. These include alternative terms for the same concept (e.g., *pail* vs. *bucket*), alternative pronunciations for the same word ([fi'næns] vs. [faynæns] *finance*), or alternative grammatical strategies (e.g., *there's/there are many things to do* vs. *it's many things to do*)². Identifying and articulating the principles which regulate learners' selection in the development of their own dialects (i.e., determination of preferred ways of saying things) is a challenge that linguists must face. Individual learners make their selections here in ways that do not seem so similar to the selection that takes place during the formation of a genotype, for instance, what particular genes the process of recombination makes dominant (determining the carrier's phenotypes) or recessive and under what particular conditions. To begin with, no agency of the gene inheritor is involved in the biology case. If ecology plays a role at the level of the formation of genotype, it is not in the same obvious way it seems to do during the formation of dialects. In other words, it is not clear what factors or what particular selection algorithm a biologist would invoke to account for why particular genes become dominant or recessive in the genotype of a particular organism or for why a particular individual winds up with a particular combination of phenotypical traits³.

¹ I speak of « copying » tongue in cheek here. As Sperber (1996, p. 141) observes, « (1) De façon générale, les représentations ne sont pas copiées, elles sont transformées dans le processus de la transmission. (2) Les représentations se transforment par l'effet d'un processus cognitif constructif ». That is, (1) « Generally speaking, representations are not copied, they are transformed during the process of transmission. (2) The representations are transformed by a constructive cognitive process ». David Hull (2004) remarks that selection as an evolutionary mechanism is restricted to vertical, not to horizontal, transmission. I wonder whether the difference does not lie in the complexity of the process, rather than in whether or not selection applies. After all, it is the outcome that shows whether a gene or feature has been preserved in a population, not how it has been preserved or eliminated.

² The process is very similar, if not identical with, what Sperber (1996, p. 147) identifies as « synthesis ». The bearer of, for instance, an idea or a story in the case of culture, or of a feature or even an idiolect in the case of language, receives input from several individuals, some of whom influence him/her more than others. The inputs contribute selectively and differentially to the formation of the new version of the original idea or story, or to the formation of the new speaker's feature or idiolect.

³ We should remember here that selection does not necessarily operate in terms of total exclusion ; it operates in terms of dominance. As explained by Sober (1984), it works more like in a golf tournament, in which scores are added up for every player. While the winner is the one with the highest score, the other players are still part of the overall competition and each occupies the position determined by their respective scores. The competition of linguistic features proceeds more or less the same way, making allowance for those features that are not dominant to coexist with the dominant ones in particular

Linguists can invoke factors such as the statistical frequency of a feature, semantic transparency, regularity, salience, and social status of the model speakers.

For instance, a learner of English who is focused on regularities could easily and incorrectly substitute *falled* for *fell* as the past tense of the verb *fall*. However, although this kind of error often occurs, especially in child language, it usually does not crystallize into an idiolectal idiosyncrasy. The reason is simply that it cannot compete with the high statistical frequency of *fell*, which makes it less competitive. Interestingly, this is contrary to those cases where some speakers produce *dove* instead of *dived* as the past tense of *dive*, although the basic principle at work is the same. They do this by analogy to *drove* as the past tense of *drive*, because the latter also has a high statistical frequency.

Likewise, because the comparative construction with *more* is more common, many speakers will say *more fit* instead of *fitter*, which should be expected to be more typical of monosyllabic adjectives such as *short* > *shorter*. On the other hand, one also hears *funner* instead of *more fun*, simply because it is more consistent with *fitter* and *shorter*, despite the fact that its status as an adjective is less clearcut. There are thus various factors that affect what particular linguistic habits (features) a speaker internalizes in his/her idiolect. This is how selection works, especially at the population level, often not resolving the competition in exclusive terms, only in terms of dominance and differentially from one speaker to another.

Before proceeding, it may not be redundant to clarify that, like in biology, the terms *competition* and *selection* entail no agency on the part of genes or features. The term *competition* refers to the condition of inequality that obtains among variants in a feature pool, with some factors of their internal or external ecologies (dis)favoring some of them for dominance. Contrary to what is suggested by Dawkins' (1967) memetics, it is speakers who make the selections. They are in fact the ones who impose the inequality condition, based on factors, sometimes non-structural, that guide their preferences.

However, the way in which linguistic competence develops also suggests that in the linguistic species selection applies at the level of features (units or combinatoric principles). Although some ethnographic considerations suggest that selection also applies at the level of languages, when speakers target primarily features of a particular language over those of others, what we know about language mixing and the development of creoles suggests otherwise. Languages are selected indirectly through the fact that their features (sounds, words, combinatoric rules, and particular ways of packaging meanings) wind up constituting the majority of those selected from the combined feature pool of the language varieties in contact. Although clearly favored, the indirectly selected language (variety) also bears the influence of (some of) the disfavored varieties and is therefore modified into a new variety. This is what I identified above as Pyrrhic victory. It appears to explain how, under the influence of the Celtic languages over which it prevailed in especially southwestern Europe, Latin evolved into the Romance languages. The same evolution by selection, under the influence of substrate languages, is also true of creoles in former European plantation settlement colonies of the New World and the Indian Ocean. As the relevant socioeconomic history of the relevant plantation settlement colonies shows, a European vernacular was typically selected over African and other European vernaculars, but the emergent colonial variety also bears the influence of these other languages.

Thus, in partial support of Dawkins' (1976) hypothesis that it is genes rather than species which are involved in the selection process, it is indeed still features¹ which are the units of

idiolects and to be used every now and then. Because non-dominant features in some speakers' idiolects can be dominant in other idiolects, languages can have longer memories than an oversimplified approach to language evolution, especially one conceiving of languages as organisms, may lead us to expect. The lives of linguistic features are thus to some extent comparable to those of genes, subject to mating/interaction practices.

¹ Following Dawkins (1976), Hull (1988) opts for the term *replicator* to identify units of selection, with basically the same meaning as what I express here with *linguistic feature*, viz., units and principles that the learner of a language, in the present case, would endeavor to replicate. Croft (2000) follows both, alternating between *replicator* and *lingueme* « linguistic meme ». I am sticking to the traditional term in sociolinguistics, because the others represent no improvements over it. They merely identify features as the elements that can be « replicated », albeit imperfectly – which is just a difference of perspective. Actually, the term *lingueme* conveys as much vagueness and inaccuracy as *meme*, especially because a

selection, although being associated with a particular language becomes one of the factors that favor those features¹. In contact settings that produced creoles, only one target language, often misnamed the *superstratum* or *lexifier*, lends most of its vocabulary and grammar to the emergent vernacular.² Since the target has typically been structurally heterogeneous, availing competing variants, the selection of particular options, rather than others, into the emergent creoles shows again that selection applies to smaller units and combinatoric principles that collectively make up a language ; it applies only implicitly or indirectly to languages themselves.

Selection is constrained by the specific external and internal ecologies of linguistic interaction. The external, socio-economic ecology imposes a particular ranking of variants (dialects and structural features) to which the selections made by speakers are sensitive. It is not enough of course to know whether a feature is standard or nonstandard. Another important factor is whether a particular feature will promote or discourage the acceptance of a learner/speaker by the particular socio-economic group that he/she wishes to fit in. The literature on covert prestige shows that in informal settings, where vernacular varieties are spoken, nonstandard features are normally preferred to their standard counterparts. This behavior accounts very well for the resilience of nonstandard vernacular features, as stigmatized as they are by the elites of various populations. The personality of the speaker/learner is also another important factor, which partly determines which group he/she wants to be associated with, if given a choice. Similarity to epidemiology is not difficult to establish here, as social practices bear on how viruses spread in a population, although different species are subject to different constraints specific to them in the selection of their features.

The internal ecology, which is no less important, consists of all the other variants that a particular one coexists with. The factors that contrast them with each other determine, relative to the linguistic background of the learner (part of the external ecology), which of them is likely to become dominant. For instance, in the case of creoles, the periphrastic comparative (with *more* in English) has been favored not only because the variant exists in the European language but also, and quite significantly, because the substrate languages typically have only a similar periphrastic option. This partial congruence of structural features favored the generalization of the periphrastic comparison in the new, plantation colony variety of the European vernacular. It is also important to remember in this case that the population of linguistic variants consists of those attested in the 17th and 18th-century nonstandard varieties of European colonial languages, excluding some of the variants that we may be more familiar with today. These considerations prepare us not to be too easily impressed by the fact that structures of creoles are quite different from those of the standard varieties of the same languages today.

language is not a body of utterances or texts (*pace* Croft) but rather the « system » that produces them.

¹ Although Lewontin (1970, p. 14) makes allowance for populations to be selected, his primary position is that « many [population] adaptations turn out to be explicable by simple selection at the individual level » (p. 13). This takes us back to idiolects and the piecemeal pattern of their emergence, through the cumulative selection of individual features and their (re)integration into a new « system ». This process justifies the position I defend here. Another way of explaining it is that particular units, such as words (which also implies sounds and morphemes), and particular combinatoric rules (such as in syntax) are associated with specific languages. In practice, however, it is individual units and combinatoric rules which are perceived and can be selected in or out of the emergent idiolects. In multilingual or multidialectal contexts, units and rules from different systems (languages or dialects) are often mixed and can lead to the emergence of significantly different language varieties, such as the Romance languages (in relation to Latin and the Celtic substrate languages) or creoles (in relation to western European and African languages in European plantation settlement colonies of the New World and Indian Ocean). As should also be obvious from the literature on code-switching or mixing, what language or dialect a speaker claims to be speaking is sometimes determined more by the speaker's intention than by the actual text of his/her discourse.

² The terms are convenient misnomers because, in the case of *lexifier*, it is inconceivable that one would naturally target only the vocabulary of a language and ignore the grammar associated with the words, thus hoping to get it from some other source (Chaudenson 2001, 2003). The terms *superstratum* and *substratum* are also inaccurate in the case of creoles because they are based on the social stratification of the populations in contact but not on the time their languages arrived in the settings where creoles developed (Goodman, 1993). Technically, in parts of the New World and the Indian Ocean where creoles developed, the relevant African languages arrived later than the European languages that the slaves appropriated as vernaculars and modified.

Thus, we should not claim uncritically and too hastily that the origins of their present nonstandard peculiarities must be African. Rather, they should prompt us to compare creoles with nonstandard varieties of the same languages, which should give us a closer sense of the extent to which they have diverged.

Then we must also recall that the target for those who made the creoles consisted of several nonstandard varieties competing with each other, which should make it normal to notice that features of creoles have diverse « super- and substrate » origins and therefore can differ maximally from non-creole varieties that did not develop under similar ecological conditions¹. Even the European target itself was structurally heterogeneous, including not only features of diverse dialects but also xenolectal features from especially the European indentured servants who did not speak it natively. In the case of non-creole colonial vernaculars (such as White American English varieties), whose beginnings have been associated with koinéization (Montgomery, 1995 ; Mufwene, 2003b), the situation is thus comparable to one where members of different subspecies can mate with each other, across the subspecies boundaries. In biological terms, it is like having a colony where different subspecies of dogs can interbreed across their normal social boundaries and produce new, hybrid canine sub-species. In the case of creoles, the situation would amount to one where the dogs can also interbreed with members of related species such as foxes, jackals, and wolves, always bearing in mind that the modes of « transmission » in animal biology and in language are different, especially the fact that features of one particular language are likely to be heavily favored by the particular socio-economic setting of the contact².

However indirect the process may be, community-wide targeting or selection of a particular language or dialect over (an)other alternative(s) can have particular ethnographic consequences. Languages that are less often selected for communication are endangered and may eventually die, as explained above. There are various reasons why speakers in a multilingual or multilectal community would prefer a particular language or dialect, or would often/usually not choose to speak another language or dialect. They are all practical, associated with the *hic et nunc* conditions of interaction, and need not be discussed here. They can also vary from one speaker to another in the same population. In the vast majority of cases around the world, the decisions are individual and taken at different times (however repeatedly) and in different tokens or kinds of settings, though members of the population do influence each other.

Similarities between language and biology arise here too in how competition and selection operating on features/genes and on organisms can lead to evolution at the species level. Future research should shed more light on details of similarities, how far they go and where they end. Note, for instance, that when such selections eventually lead cumulatively to situations where members of the population can no longer speak their language or dialect (fluently), loss of competence does not proceed uniformly within the overall population, confirming Mayr's (2001, p. 86) observation that « Since all changes take place in populations of genetically unique individuals evolution is by necessity a gradual and continuous process³ ».

¹ This phenomenon is a consequence of both polyploidy and the fact that populations meet not like armies or sport teams engaging in contact at the same time and in an organized way, but rather as individuals engaging in particularistic contacts at different times and in different settings, all of them having the potential to influence the development of each relevant speaker's idiolect in some respect. How competition and selection of features apply within a population, under specific constraints determined by the relevant ecologies, accounts for the specific kinds of mix and heterogeneity that distinguish one language or dialect from another. Although the transmission mechanisms by which a viral species can display such genetic mixing and heterogeneity are not the same (see Part 4), the similarities in outcomes are too close to escape notice.

² I assume successful communication to be the closest counterpart of biological interbreeding. Differences between these notions are quite consistent with the fact that idiolects have no genotypes and develop gradually. An individual can learn only a limited number of linguistic features from a particular communicative event. As utterances share many features, there are many communicative events from which little or nothing new is learned, even during the early stages of the development of idiolects traditionally identified as « language acquisition ».

³ This is of course in the context of species, with the process of evolution premised on variation, inheritance, and differential reproduction, although layman's language allows individuals to evolve. As explained in Mufwene (2001, chapter 1), inheritance need not be interpreted in too strictly biological terms of gene transmission. It can be extended to « information copying », as in culture and language,

4. Ecology and language evolution

An important dimension that for too long has received little attention in historical linguistics is the motivation for change, i.e., the (chain of) factors that trigger the process. Weinreich & al. (1968, p. 102) formulate it under the name *actuation problem* :

What factors can account for the actuation of change ? Why do changes in a structural feature take place in a particular language at a given time, but not in other languages with the same feature, or in the same language at other times ?

McMahon (1993, p. 248) articulates it in the following fairly complementary way :

The real actuation question is why some of these innovations [by individual speakers] die out and others catch on, spreading through the community, or why certain instances of variation become changes while others don't¹.

Having focused on the subject matter for over twenty years, Labov (2001, p. 466) reformulates the « problem » as follows :

Why here and now ? The beginnings of change [applied to a population] are as mysterious as ever. Why *not* here and *not* now ? Endings are equally difficult to understand. The obverse of the actuation problem is *continuation*. If change has already begun and is not coming to an end, it is continuing. What was the force that was missing a hundred years ago, that fuels the engine of the Northern Cities Shift² today and keeps it moving ?

He thus widens the scope of the question, showing in the rest of his book that a variety of social ecological factors account especially for the ways changes are driven in particular language communities, often affecting only segments of the relevant populations but not others.

Recasting the actuation question/problem the population genetics way, I submit that the interaction of the external ecology of a language with its internal ecology should shed light not only on causes of linguistic changes but also on how these phenomena spread. Here too, there are similarities between, more specifically, linguistic and viral species in the ways changes occur, as they are affected by the social practices of their hosts. Critical to this approach is recognizing that changes often consist of only modifications of patterns of variation within a language, not necessarily of the introduction of new variants (McMahon, 1993, p. 248) or loss of some others. They can consist, at the population level, of shifts in the statistical frequencies of the variants, with some becoming dominant that used to be recessive. My basic assumption is that variation within a population is likely to remain stable unless something happens in its external ecology that disturbs the « balance of power » between the competing variants.

In the case of linguistic and viral species, factors such as migration to a new habitat, contact with another population, and other changes in the composition of the population bear on actuation. Migration also implies split of the proto-population, which, if proceeding randomly, can affect variation both in the migrating, colonial sub-population and in the metropolitan one (the individuals left behind in the motherland). Depending on how it was (self-)selected, the colonial sub-population may carry less than the totality of variants that obtained in the proto-population. The « balance of power » between the variants can shift in the colony, so that (some of the) features that continue to be marginal in the metropolitan population can become dominant in the colony. In the case of language, such a shift in the constituency of variants can eventually produce a new language variety, or it can simply cause the colonial and metropolitan sub-populations to evolve in divergent directions. Although this is not the full explanation, it contributes to accounting for why, for instance, English on the Falkland Islands is not the exact match of any dialect in England, despite the fact the Britons settled permanently in this archipelago in the late 19th century, much later than the English colonized North America, in the early 17th century. Other reasons for the linguistic divergence emerge from the discussion that follows.

pace Fracchia & Lewontin (1999).

¹ It would be misguided to assume that the actuation question does not apply to language or dialect choice, as indirect as this process is. It is important to examine what particular changes in the ethnographic settings of communication would have speakers prefer to use one but not (an)other language(s) or dialect(s).

² The term applies to chain-style changes undergone by vowels in urban American English dialects.

The above observations should remind us that even if the English had not come in contact with other European populations in North America and in Australia, their language would have speciated anyway. Shifts in the constituency of the variants or in their statistical frequencies would have changed the « balance of power » and (dis)avored different variants. Traditional accounts that have simply invoked geographical separation by rivers, mountains, and oceans to account for language speciation had barely addressed the question, especially if they assumed a common ancestor for all the new varieties¹. Likewise, changes in the patterns and dynamics of social interaction – which migrations help produce – affect the vitality of a particular (strain of a) virus in two or more allopatric populations that have the same ancestor. Socio-economic speciation in language and epidemiology is in some ways analogous to geographic diversification.

Language contact can also disrupt the extant pattern of variation, provided the host and immigrant populations interact regularly with each other, at least through some of their members. We are quite aware of the most drastic effects of recent European migrations to settlement colonies of the Americas and Australia on indigenous languages. Owing to specific population structures² – for instance who holds political and economic power and who makes concessions to whom – significant proportions of indigenous languages have been driven to extinction. An important reason lies in the changing socio-economic ecologies which have prompted the indigenous populations to assimilate the European populations' practices, including the adoption of the colonists' languages as vernaculars. In other cases where European vernaculars were appropriated by subordinated non-indigenous populations that have not been integrated by the Europeans, typically Africans in plantation settlement colonies of the New World and Indian Ocean, new language varieties have emerged that have been substantially influenced by African languages, which were also driven to extinction³. These are the new, colonial vernaculars that have been disfranchised from the Indo-European language family by the name *creoles*.

One somewhat oversimplified but still plausible account of such developments is that the other languages that the European languages came in contact with affected the patterns of variation within the European languages, owing to various cases of (partial) structural congruence between them and the other features that the non-European languages sometimes introduced into the systems. (See Mufwene (2001) for a more accurate and nuanced account).

¹ Paul (1891) often invokes shift in the frequency of variants to account for language change. It is curious why this factor has figured so little in 20th-century qualitative historical linguistics.

² I borrow this term from Goodnight & Wade (1999) and Wimsatt (1999), from whom the meaning is more implicit than explicitly articulated. It is clearly used in reference to the internal social organization of a population, as suggested by another term used by Goodnight & Wade : *population subdivision*. I use it below for a variety of social factors, including how a population is variously sub-categorized (often in overlapping terms, such as by gender, age, race, education, profession, and socio-economic class), how it is stratified, and who interacts or socializes with whom. My working assumption is that linguistic features spread along speakers' patterns of socialization. Needless to say that I am rearticulating, from a population perspective, parameters that sociolinguists have traditionally invoked to account for language variation. My primary intention is to highlight not only ecological similarities between the linguistic and viral species but also methodological ones between linguistics and biology (too often ignored in traditional, 20th-century genetic linguistics) that must be used to advantage in order to better understand certain general mechanisms of evolution.

³ The reasons for the extinction of the African languages in the plantation colonies are not identical to those for the extinction of the indigenous languages, although the institution of the new socio-economic world order by European settlers was a major stochastic event in both cases. As explained in Mufwene (2004a, 2004b), the particular way in which the plantation societies developed played a central role in an evolution that could otherwise be described as « chaotic » (in the context of chaos theory). They started from homesteads on which the Africans were minorities, integrated, and their children acquired colonial varieties of European languages as their primary, if not exclusive, vernaculars. These creole, locally-born slaves would become cultural, and especially linguistic, models for the *bozal*, incoming African-born slaves of the larger plantation societies, in which African languages were underrated. Societal multilingualism among the slaves just compounded the problem, leading to the abandonment of African languages in the same way that many of them are now losing ground to urban vernaculars in African cities. On the other hand, Native American languages would be endangered much later by European languages, when the new American societies would assimilate them and, in the case of North America, lured more and more of them out of the reservations, where living conditions worsened and the maintenance of their traditions became more difficult.

The same processes operated in less drastic ways to produce new colonial varieties of European languages in settings where only these came in contact with each other under varying population structure conditions. White American English varieties are generally byproducts of contacts among various metropolitan English dialects and other European languages¹. *Mutatis mutandis*, the same can be said of colonial varieties of other European languages. The whole evolution is similar to bringing populations infected by various strains of a particular virus in contact with each other, creating ecologies in which new strains can emerge, with each ecology producing its own dominant kind. In both cases, population structure, in terms of who interacts with whom, is an important ecological factor, to which I return below.

Both contact and migration influence the ways competition and selection affect variants differentially in the colonial and the metropolitan (sub-)populations. Eventually they produce new and divergent language varieties. Population structure – interpreted now as who lives where and under what conditions he/she can interact with members of another group or population – has a great deal to do with how evolution starts and proceeds. Changes in the population structure also affect whether or not a particular change ends. In the way that Labov (2001) discusses actuation, changes occur even when the structure of a population does not seem to have changed. Assuming several changes to have started from below, in the way societies are economically or ethnically stratified, he considers social mobility and non-conformity as critical factors in the causation of changes.

Social mobility is actually another way the structure of a population changes, when individuals leave one socio-economic class for another and thus join new networks of interaction. « Non-conformity », which means refusal to abide by all the norms of the new class, suggests the expectation on the part of the current members that one's language remains the same all the time, thus that the current pattern of variation is there to stay². Other factors of course determine whether the non-conformists will have followers and which of their « innovations » will spread and be recognized communally as change. In population genetics, they are more or less like individuals who introduce new viruses in a population and who can spread them largely thanks to how the host population interacts with them.

All the above observations make more sense once one factors in the following : (1) populations are constructed from individuals behaving singly most of the time, (2) they are not homogeneous, and (3) their members typically interact within small, overlapping networks. At least, from a geographical perspective, they are more like what ecologists call *metapopulations*, i.e., habitat patches interconnected by dispersing individuals (Hanski, 1996). They are basically convenient groupings of individuals who interact with each other, within and across accepted social boundaries (racial, ethnic, gender, age, professional, economic class, etc.), as individuals, not as team members³.

¹ In the case of the United States, segregation by nationality among the European colonists during the colonial period must have reduced, by retardation, influence from continental European languages. Gradual language shifts from continental European languages must have taken place after the American English koinés had « crystallized » primarily from the contact of metropolitan English dialects, in the same way as on the Falkland Islands. Continental European linguistic influence must have been kept in check in the same way that influence from today's waves of immigrants' xenolectal speech is. By the Founder Principle, children acquire the extant variety natively and their parents die with their xenolectal features identified as such and largely not accepted by the host populations.

² Labov (2001, p. 514) captures something similar with his « Golden Age Principle », according to which « At some time in the past, language was in a state of perfection », consistent with the common reaction among purists that one's language is degenerating or decaying.

³ Saussure (1916) may have been a little mistaken in analogizing a language, interpreted as a system, to a chess game, in which players follow communal conventions. Regarding languages, the putative conventions are more like emergent patterns (i.e., partial regularities emerging to the analyst) than something whose existence is independent of the speaker's knowledge. Those partial regularities do not entail that a person who acquires a language naturalistically learns particular rules. Linguistic regularities are like beauty, in the eyes of the beholder. As McCawley (1976) points out, it is not evident that the generality of the rules by which speakers putatively operate coincides with those that the linguist seeks or claims to capture. This observation follows from the fact that a speaker or a naturalistic language learner is not a linguist. He/She does not process the materials from which his/her competence gradually develops in the same way as the linguist who gathers a (significant) corpus of (varying) data, analyzes the corpus, and seeks to capture the broadest generalization(s) applicable to the data that interest him/her.

An important question is : How do communal patterns emerge from individual speakers' behaviors ?¹ Changes typically apply in non-uniform fashions, affecting idiolects differentially and applying only to some sub-groupings of the larger population. Thus, Labov (2001) reports that, overall, African Americans have not participated in the Northern Cities Shift in the United States. In this particular case, population structure provides a useful ecological explanation, bringing to bear the fact that, as observed by Labov (see also Wolfram, 2000), Americans are sensitive to race distinctions. In fact, most Americans tend to socialize along race/ethnic lines and maintain a strong sense of ethnic identity. While most Whites would not like to be associated with their stereotypes of African American linguistic peculiarities, most African Americans, at least those associated with African American English, would not like the opposite evolution either. According to Wolfram (2000), young rural African Americans prefer to identify with urban African American vernacular English (see also Cukor-Avila & Bailey, 1996), considering the conservative variety of their parents too similar to White English². White and African American speakers would thus prefer to follow evolutionary trends within their respective ethnolects, ignoring or only following with interest those occurring in the other(s). The analogy with virology is clear, as viruses tend to be transmitted through interacting hosts.

We are thus also in a position where we can justifiably speculate on divergence phenomena. Could a population's determination not to be assimilated by, or confused with, another population account for cases where two varieties such as European American and African American English varieties evolve in divergent directions, even in the Southern states where over two centuries of intimate cohabitation preceded the last century of segregated life ? Isn't this what the literature on the divergence of White and African American English varieties is ultimately telling us, thus indirectly underscoring the significance of population structure as an ecological factor in evolution ?

5. Conclusions

It is evident that studies of language evolution stand to gain a great deal by adopting a population genetics approach, one that acknowledges that populations are groupings of individuals, that the agency of many diachronic processes lies in the activities of individual members, that such processes need not be uniform across populations, and that the activities are largely determined by the ecologies in which the individual agents evolve. I submit that languages or dialects are species of idiolects, based on family resemblance. They share with viral species several properties that are relevant to understanding evolution ; therefore the biological evolution model that linguists should consider for inspiration appears to lie in virology and/or epidemiology but not in animal biology.

Languages are like viruses essentially in that they are both parasitic species whose lives depend on the activities of their hosts, thus also on the latter's patterns of social interaction.

However, it is equally important not to overlook some important differences between them, as these can account for their respective evolutionary peculiarities. Unlike viruses, idiolects as linguistic organisms do not have anything corresponding to a biological genotype, simply because, if one can at all claim that they are begotten through the communicative practices of their speakers, their structures develop piecemeal. They develop incrementally as individuals learn to communicate and grow to produce increasingly complex utterances.³ On

¹ This question is connected to the disputable assumption among linguists that native speakers of the same variety use the same grammar (identically represented in all their minds) to process and produce utterances (Mufwene, 1992b).

² This social attitude accounts for the resilience of African American English, as stigmatized as it is, because it also functions as a marker of identity. It also accounts for why the Ocracoke brogue, spoken by White islanders on the coast of North Carolina, is endangered by the vernaculars of Whites in-migrants from the neighboring mainland (Wolfram & Schilling-Estes 1995), whereas Gullah, the creole vernacular spoken by African Americans in coastal South Carolina, has so far been endangered by the exodus of its speakers, not by the White mainlanders' « invasion » (Mufwene, 1997). White mainlanders have mixed with White Islanders but not with coastal African Americans. In coastal South Carolina and Georgia, the residential communities remain racially segregated, just like many parts of American cities.

³ By puberty, the cognitive infrastructure that enables this linguistic development is assumed to have reached an acceptable level of maturity, and the « linguistic system » – as inferred by the linguist – is considered full-fledged, although the vocabulary will continue to grow and more stylistic variation will emerge.

the other hand, a virus starts its life with a fully-structured genotype (by gene recombination at conception) and has an incubation stage. This peculiarity is worth bearing in mind, although, like an idiolect (even in its mature stage), it can change its ontogenetic structure several times over, in a Lamarckian style, and it can be modified when it spreads to other hosts.

On the other hand, a virus may have started its life outside its current host, whereas an idiolect may not – only its features can. An idiolect's host is also its maker. Unlike the genes of a biological organism, the features of an idiolect cannot be transmitted – literally, in the form of a copy – from one host to another. They are typically copied with modification, under competition with similar inputs, by the learners, and they are recombined into new « systems » with some inter-idiolectal variation. Thus, no two speakers produce the same sound in physically and acoustically identical ways and no two speakers have exactly the same range of denotational and connotational meanings, as well as pragmatic constraints, associated with the same words. Cases of misunderstandings between speakers of the same dialect also suggest that no two idiolectal « systems » are identical, though speakers more familiar with each other have an easier time understanding each other. These observations follow from the fact that every speaker differs physiologically from other speakers and is thus equipped with a different configuration of speech organs¹. Moreover, from a social ecological perspective, no two speakers have had identical experiences of social interaction, hence of being exposed to identical samples of speech or signing, which have influenced their language « acquisition » processes.

Still, despite these differences, there are similarities striking enough to justify the approach advocated in this essay. Polyploidy suggests that gene selection and recombination apply to viruses in more or less the same ways feature selection and recombination apply to idiolects, gradually in the latter case, but subject to ecological constraints in both cases. Idiolects are more likely to be influenced by those of speakers that their hosts/makers have interacted the most frequently with. Speakers accommodate each other, minimizing chances of being misunderstood. Likewise, viruses in a population are more likely to share a lot of genetic materials when their hosts socialize with each other than when they do not.

However, even under such conditions, it is still informative to understand under what more specific conditions a particular gene or feature becomes dominant or recessive, while the host has actually received, or been exposed to, all of them. How does selection work and what particular factors constrain it? Linguists can invoke markedness, or some optimality constraints, to explain why some variants are (dis)avored at the level of both idiolects and languages or dialects. While a wide range of ecological factors can account for selection of particular genes for dominance within a biological population, it is not clear to me what factors can be invoked to account for a similar process at the level of (the conception of) organisms. Comparisons such as undertaken in this essay suggest that linguistics and biology can very well inspire each other in addressing evolutionary issues². Questions regarding the specific ways in which selection proceeds at different levels are, I hope, among those that future interdisciplinary research in biological and language evolution will explore.

¹ The physiology of a speaker is part of the immediate external ecology of language « acquisition ». The variation alluded to here has to do with factors such as the following : specific shape and size of the oral cavity, length and width of the throat, size and length of the nasal cavity, alignment of teeth, and width and thickness of the tongue.

² This is a position consistent with part of the conclusions of Hull (2000, p. 27), when he states, « Of course the flow of information may just as well go in the other direction. Biologists may find the solutions that linguists have suggested useful in biology ».

References

- Adam (L.). 1882. *Les classifications, l'objet, la méthode, les conclusions de la linguistique*. Paris : Université de Paris VII.
- Adam (L.). 1883. *Les idiomes négro-aryens et malayo aryens : essai d'hybridologie linguistique*. Paris : Maisonneuve.
- Bickerton (D.). 1981. *Roots of language*. Ann Arbor : Karoma.
- Bickerton (D.). 1984. « The language bioprogram hypothesis and second language acquisition ». in : Rutherford (W.-E.), (ed.). *Language universals and second language acquisition*. Amsterdam : John Benjamins, pp. 141-161.
- Bickerton (D.). 1990. *Language and species*. Chicago : University of Chicago Press.
- Bopp (F.). 1833. *Vergleichende Grammatik des Sanskrit, Zend, Griechischen, Lateinischen, Gothischen und Deutschen*. Berlin.
- Chaudenson (R.). 2001. *Creolization of language and culture*. London : Routledge.
- Chaudenson (R.). 2003. *La créolisation : théorie, applications, implications*. Paris : L'Harmattan.
- Chomsky (N.). 2000. *New horizons in the study of language and mind*. Cambridge : Cambridge University Press.
- Croft (W.). 2000. *Explaining language change : An evolutionary approach*. London : Longman.
- Cukor-Avila (P.) & Bailey (G.). 1996. « The spread of urban AAVE : A case study ». in : Arnold (J.), Blake (R.), Davidson (B.), Schwenker (S.) & Solomon (J.). *Sociolinguistic variation : Data, theory, and analysis*. Stanford, CA : CSLI Publications, pp. 469-485.
- Darwin (Ch.). 1859. *On the origin of species*. London : John Murray.
- Dawkins (R.). 1976. *The selfish gene*. Oxford : Oxford University Press.
- DeGraff (M.). 1999a. « Creolization, language change, and language acquisition : A prolegomenon ». in : DeGraff (M.), (ed.). *Language creation and language change : Creolization, diachrony, and development*. Cambridge, MA : M.I.T. Press, pp. 1-46.
- DeGraff (M.). 1999a. « Creolization, language change, and language acquisition : An epilogue ». in : DeGraff (M.), (ed.). *Language creation and language change : Creolization, diachrony, and development*. Cambridge, MA : M.I.T. Press, pp. 473-543.
- DeGraff (M.). 2003. « Against creole exceptionalism. Discussion note ». in : *Language*, 79, pp. 391-410.
- DeGraff (M.). 2005. « On creole genesis and language acquisition : Some methodological and theoretical preliminaries ». in : *Journal of Pidgin and Creole Languages*.
- Filppula (M.), Klemola (J.) & Pitkänen (H.), (eds.). 2002. *The Celtic Roots of English*. Joensuu.
- Fracchia (J.) & Lewontin (R.-C.). 1999. « Does culture evolve ? ». in : *History and Theory : Studies in the Philosophy of History*, 38, 4, pp. 52-78.
- Goodman (M.). 1993. « African substratum : Some cautionary words ». in : Mufwene (S.-S.), (ed.). *Africanisms in Afro-American language varieties*. Athens : University of Georgia Press, pp. 64-73.
- Goodnight (Ch.-J.) & Wade (M.-J.). 1999. « The ongoing synthesis : A reply to Coyne, Barton, and Turelli ». in : *Evolution*, 54, pp. 317-324.
- Hanski (I.). 1996. « Metapopulation ecology ». in : Rhodes (O.-E.), Chesser (R.-K.) & Smith (M.-H.), (eds). *Population dynamics in ecological space and time*. Chicago : University of Chicago Press, pp. 13-43
- Heine (B.) & Kuteva (T.). 2005. *Language contact and grammatical change*. Cambridge : Cambridge University Press.
- Hermann (P.). 1891. *Principles of the history of language*. London : Longmans, Green, and Co. [translated from the second edition of the original by Strong (H.-A.)].
- Hey (J.). 2001. *Genes, categories, and species : The evolutionary and cognitive causes of the species problem*. Oxford : Oxford University Press.
- Hock (H.-H.) & Brian (D.-J.). 1996. *Language history, language change, and language relationship*. Berlin : Mouton de Gruyter.
- Hull (D.-L.). 1988. *Science as a process : An evolutionary account of the social and conceptual development of science*. Chicago : University of Chicago Press.
- Hull (D.-L.). 2002. « Species, languages and the comparative method ». in : *Selection*, 3,1, pp. 17-28.

- Koerner (K.), (ed.). 1983. *Linguistics and evolutionary theory : Three essays by August Schleicher, Ernst Haeckel, and Wilhelm Bleek*. Amsterdam : John Benjamins [with an introduction by J. Peter Maher].
- Kretzschmar (W.) & Tamasi (S.). 2003. « Distributional foundations for a theory of language change ». in : *World Englishes*, 22, pp. 377-401.
- Labov (W.). 2001. *Principles of linguistic change : Social factors*. Malden, MA : Blackwell.
- Lewontin (R.-C.). 1970. « The units of selection ». in : *Annual Review of Ecology and Systematics*, 1, pp. 1-18.
- Lieberman (Ph.). 1991. « Review of Bickerton 1990 ». in : *American Anthropologist*, 93, pp. 998-1000.
- Lieberman (Ph.). 2002. « On the nature of evolution and the neural bases of human language ». in : *Yearbook of Physical Anthropology*, 45, pp. 36-62.
- McCawley (J.-D.). 1976. « Some ideas not to live by ». in : *Die neueren Sprachen*, 75, pp. 151-65.
- Mayr (E.). 2001. *What evolution is*. New York : Basic Books.
- McMahon (A.). 1994. *Understanding language change*. Cambridge : Cambridge University Press.
- McNeill (D.). 1992. « Review of Bickerton 1990 ». in : *International Journal of Primatology*, 13, pp. 575-583.
- Montgomery (M.). 1995. « The koineization of colonial American English ». in : Philipponneau (C.), (ed.). *Sociolinguistic studies and language planning : Proceedings of the 16th Annual Meeting of the Atlantic Provinces Linguistic Association*. Moncton, NB : Centre de Recherche en Linguistique Appliquée, pp. 309-331.
- Mufwene (S.-S.). 1991. « Language genesis and human evolution. Review article on Bickerton 1990 ». in : *Diachronica*, 8, pp. 239-254.
- Mufwene (S.-S.). 1992a. « The reviewer responds ». in : *Diachronica*, 9, pp. 149-151.
- Mufwene (S.-S.). 1992b. « Why grammars are not monolithic ». in : Brentari (D.), Larson (G.) & MacLeod (L.-A), (eds.). *The joy of grammar : A festschrift in honor of James D. McCawley*. Amsterdam : John Benjamins, pp. 225-250.
- Mufwene (S.-S.). 1997. « The ecology of Gullah's survival ». in : *American Speech*, 72, pp. 69-83.
- Mufwene (S.-S.). 2001. *The ecology of language evolution*. Cambridge : Cambridge University Press.
- Mufwene (S.-S.). 2003a. « Genetic linguistics and genetic creolistics. Short note ». in : *Journal of Pidgin and Creole Languages*, 18, pp. 273-288.
- Mufwene (S.-S.). 2003b. « The shared ancestry of African-American and American White southern Englishes : Some speculations dictated by history ». in : Nagle (S.-J.) & Sanders (S.-L.), (eds.). *English in the southern United States*. Cambridge : Cambridge University Press, pp. 64-81
- Mufwene (S.-S.). 2004a. « Language birth and death ». in : *Annual Review of Anthropology*, 33, pp. 201-222.
- Mufwene (S.-S.). 2004b. « Multilingualism in linguistic history : Creolization and indigenization ». in : Bhatia (T.) & Richie (W.), (eds.). *Handbook of bilingualism*. Malden, MA : Blackwell, pp. 460-488.
- Radick (G.). 2002. « Darwin on language and selection ». in : *Selection*, 3, 1, pp. 7-16.
- Pinker (S.). 1992. « Review of Bickerton 1990 ». in : *Language*, 68, pp. 375-382.
- Saussure (F. de). 1916. *Cours de linguistique générale*. Paris : Payot [ed. by Charles Bally & Albert Sechehaye in collaboration with Albert Riedlinger ; trans. Wade Baskin. 1966. *Course in general linguistics*. New York : McGraw-Hill Co].
- Schleicher (A.). 1863. *Die darwinsche Theorie und die Sprachwissenschaft : Offenes Sendschreiben an Herrn Dr. Ernst Häckel*. Weimar : Böhlau.
- Schleicher (A.). 1869. *Darwinism tested by the science of language*. London : J.-C. Hotten [translated by Alexander V.-W. Bickers, reprinted in : Koerner (ed.). 1983].
- Sober (E.). 1984. *The nature of selection : Evolutionary theory in theory and focus*. Chicago : University of Chicago Press.
- Sperber (D.). 1996. *La contagion des idées*. Paris : Odile Jacob.
- Thomason (S.-G.). 2001. *Language contact : An introduction*. Washington, DC : Georgetown University Press.
- Tristram (H.-L.-C.). 1999. *How Celtic is Standard English ?* St. Petersburg : Nauka.

- Tristram (H.-L.-C.), (ed.). 2000. *The Celtic Englishes II*. Heidelberg : Carl Winter.
- Vennemann (Th.). 2001. « Atlantis Semitica : Structural contact features in Celtic and English ». in : Brinton (L.), (ed.). *Historical Linguistics 1999 : Selected papers from the 14th International Conference on Historical Linguistics, Vancouver, 9-13 August 1999*. Amsterdam : John Benjamins, pp. 351-369.
- Vinson (J.). 1882. « Créole ». in : *Dictionnaire des sciences anthropologiques et ethnologiques*. Paris.
- Vinson (J.). 1888. « La linguistique ». in : *La grande encyclopédie*, 22, pp. 286-296. Paris.
- Weinreich (U.). 1953. *Languages in Contact : Findings and problems*. New York : Linguistic Circle of New York.
- Weinreich (U.), Labov (W.) & Herzog (M.-I.). 1968. « Empirical foundations for a theory of language change ». in : Lehman (W.-P.) & Malkiel (Y.), (eds.). *Directions for historical linguistics*. Austin : University of Texas Press, pp. 97-195.
- Wimsatt (W.-C.). 1999. « Genes, memes and cultural heredity ». in : *Biology and Philosophy*, 14, pp. 279-310.
- Wolfram (W.). 2000. « On the construction of vernacular dialect norms ». in : Okrent (A.) & Boyle (J.), (eds.). *Papers from the 36th Meeting of the Chicago Linguistic Society*. Vol. 2 : *The Panels*. Chicago : Chicago Linguistic Society, pp. 335-358.
- Wolfram (W.) & Schilling Estes (N.). 1995. « Moribund dialects and the endangerment canon : The case of the Ocracoke Brogue ». in : *Language*, 71, pp. 696-721.



May 2006

1. Presentation

More than 5000 languages are spoken today. A few are used by hundreds of millions of people, but the great majority have a very restricted distribution. Languages with only a hundred speakers or fewer are in danger of imminent extinction; many have disappeared already.

It doesn't take a linguist to know that some languages are more closely related than others. Spanish and Italian, my mother tongue, are obvious examples. I can get by in Spanish- or Portuguese-speaking countries without much difficulty. However, words that are identical or similar but have different meanings cause trouble. For example, *burro* means « butter » in Italian but « donkey » in Spanish; *equipaggio* means « crew » in Italian and *equipaje* means « luggage » in Spanish; *salire* means « to go up » in Italian, and *salir*, « to go out » in Spanish. We call these words « false friends, » and fortunately there are not many of them. Italian, French, Spanish, Romanian, and so on, derive from a common source-Latin. Likewise, Germanic languages include Swedish, German, Dutch, Flemish, and English. The Slavic languages of eastern Europe are also quite similar. The resemblance between Sanskrit, a classical language in India, and some ancient European languages was well-known as early as the eighteenth century.

The study of Sanskrit provided the first linguistic clues to the relationships among what would become known as the Indo-European family of languages. Since then, many other linguistic families have been recognized (some linguists prefer to call these families phyla). Like plant and animal taxonomists, linguists have reconstructed trees illustrating linguistic relationships that they call « genetic » - equivalent to the word's use in biology. But linguists have had trouble reconstructing relationships above the family level - we have yet to agree on a single tree linking all the existing families. In fact, many linguists believe that the question of the unity or diversity of modern languages can never be answered. The difficulty rests with the rapidity of linguistic evolution.

Figure 1 shows the geographic distribution of language families recently proposed by Merritt Ruhlen. The least well-known languages, those of the Australian Aborigines and New Guineans, are more difficult to classify. But opinions about other families also differ, and bitter feuding has characterized much of historical linguistics over the past century. One of the most contentious issues among American linguists has been the classification of Native American languages.

At the beginning of the century, having noted the resemblance of many Amerindian languages, the linguist Edward Sapir and the anthropologist Karl Kroeber claimed that only a few Native American language families existed. Their hypothesis met considerable resistance from the majority of American linguists who strongly opposed such a unification. A new cycle of contention began after Joseph Greenberg of Stanford University published a book in 1987 called *Language in the Americas*, showing that the languages spoken by pre-Columbian Americans could be grouped into just three families: Eskimo-Aleut, Na-Dene (comprising languages that are spoken mostly in the Pacific Northwest, but also including Navajo and Apache), and Amerindian, which comprises most languages of North and South America.

¹ This paper was first published by Penguin Press in 2001 in a book entitled: *Genes, Peoples, and Languages* (chap. 5, pp. 133-172). We thank the author, the editor, and the publisher for authorizing *Marges Linguistiques* to reprint this text. May 2006, M.L.M.S. Publisher. Saint-Chamas (France).

Greenberg's proposal agreed with the classifications of American biologists Christie G. Turner and Stephen Zegura, who used measurements of modern and fossilized teeth, and blood groups and proteins, respectively. What's more, these three linguistic families seem to correspond with three major migrations suggested by archeological data. Amerindians appear to have come first, followed by Na-Dene speakers and finally Eskimos. The first group occupied all of America, while the second and third groups have remained near the Arctic, where they originated. We also found that Native Americans can be divided genetically into the three distinct groups Greenberg recognized on linguistic grounds. It must be said, however, that Amerindians are genetically extremely variable and that linguistic sub-groupings within the Amerindian family level do not correspond terribly well with the genetic results. Southern Na-Dene (Apaches and Navajos) are genetically similar to Northern Na-Dene, but the southern populations have absorbed genes from their Amerindian neighbors.

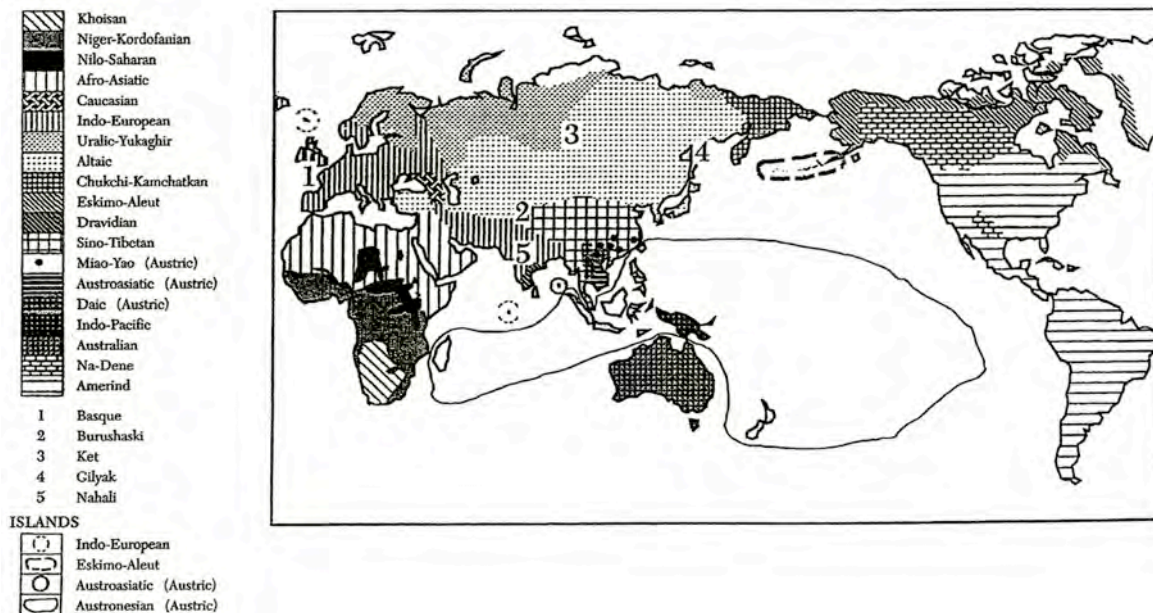


Figure 1 : Geographic distribution of the seventeen linguistic families, and location of several linguistic isolates (drawn from maps in : Ruhlen (1987), vol. 1).

The Amerindians appear to have reached the Americas in a migration much older and more complex than those of later Na-Dene and Eskimo-Aleut speakers, and there may even have been more than one Amerindian migration. Genetic data indicate that Amerindians arrived at least 30 000 years ago, but this date may only represent an average of the most important migrations. In addition, it may be biased upwards if it is true that the first Amerindian migrations involved very few individuals, as some new data on Y chromosomes suggest. A strong founders' effect tends to increase the length of branches in neighbor-joining trees, discussed later, and thus also to exaggerate the time of first settlement calculated genetically.

The publication of *Language in the Americas* unleashed a new war between American linguists and anthropologists supporting Greenberg's thesis. A large group of linguists held a meeting and declared it impossible to recognize fewer than sixty or so taxonomie groups of Native American languages. Taxonomists can be divided into « lumpers » and « splitters » : these synthetic and analytic tendencies probably reflect a fundamental dichotomy in the human spirit. However, in the case of Amerindian classification, methodological differences can explain much of the dispute, as Greenberg discusses in detail. I am not a linguist, but I find Greenberg's arguments most convincing. Furthermore, Greenberg has been through this before. Many years ago, he proposed a now widely accepted classification of African languages into only four families : Afroasiatic, including all Semitic languages and most languages in Ethiopia and North Africa ; Nilo-Saharan, comprising languages spoken along the upper Nile and the southern Sahara ; Niger-Kordofanian, which includes most central, southern, and West African languages, especially the Bantu ones ; and Khoisan languages, spoken by Khoi-Khoi and San populations in southern Africa. Greenberg's classification endured a barrage of criticism when it was first proposed, and it is now widely accepted. A similar change of attitude toward the Amerindian classification will probably come with time.

An examination of some of the objections that Greenberg's colleagues have raised against his classifications can help us understand both the objective difficulties that afflict studies of linguistic evolution, and also the subjective ones, which are typified by Greenberg's attackers. Languages change very rapidly and it is terribly hard to establish clear connections between distant languages. Significant phonological and semantic changes to all languages occur over time. The magnitude of these changes complicates the reconstruction and evaluation of linguistic commonalities. Grammar also evolves, although usually sufficiently slowly to aid the recognition of more ancient linguistic connections. Under the pressure of phonetic and semantic change, a language rapidly becomes incomprehensible. Modern languages derived from Latin would not be understandable to a Roman of two thousand years ago—a thousand-year separation is often enough to render a language incomprehensible to its first speakers. After a separation of five to ten thousand years, the rate of recognizably similar words can drop to ten percent or less. Fortunately, certain words and certain parts of speech exhibit a slower rate of change and give us a better chance of discerning more distant linguistic relationships.

As for the problems caused by misguided methods, some anti-Greenberg linguists believe it is impossible to posit a quantitative relationship between any two languages. By disallowing reliable measurements, and by limiting the relationship between two languages only to « related or not related, » the American linguists opposing Greenberg have ruled out the possibility of hierarchical classifications, an essential prerequisite to taxonomy.

Interestingly, this position completely contradicts the view of linguists who use sophisticated methods to measure linguistic similarity, based on the fraction of words from a standard list that have a detectable common origin. This approach was developed by an American linguist, Morris Swadesh. He suggested that the probability that a word will lose its original meaning stays constant over time. When the fraction of related words retained after a known period of time has been calculated (by observing the change from Latin to the modern Romance languages, for example), we can construct a « calibration » curve, allowing us to read the time elapsed since two living languages shared a common ancestor. This method, which has been called « glottochronology, » uses a « linguistic clock » very similar to 'the « molecular clock » we discussed previously for genetics. In biology, we have the advantage of using many proteins or DNA sequences to obtain many independent estimates of the date of separation between two species. Unfortunately, there is no such variety and wealth of data in linguistics to strengthen our conclusions. Glottochronology is a less rigorous method than those used in biology, and is especially difficult to apply to distant comparisons when the fraction of cognate words more than ten thousand years ago is very small. The word lists cannot grow, because only a limited number of words change slowly ; moreover, every word has its own rate of change, a fact neglected by glottochronology, which assumes a constant rate of change.

Other linguists insist that the resemblance between similar words in different languages be examined in light of « classical sound correspondences », which are very strict rules of phonological change. If these rules are not followed exactly, they say, two words cannot be considered « cognates, » that is, one cannot determine if they share a common origin. Greenberg has answered them with an impressive list of exceptions to these rules within the Indo-European language family and others. He concluded that it would be impossible to establish the Indo-European language family if these rules were applied too rigorously. Fortunately, the Indo-European family was proposed and accepted before the theory of sound correspondences adopted its most rigid format.

Finally, some linguists believe that the parent language that gave rise to a family or, in general, a cluster of languages must be reconstructed in order to demonstrate a phylogenetic relationship among families. Here too, biology provides an analogy - when the « consensus » DNA sequence is calculated from the sequence of two modern species. The consensus sequence is the best guess of the ancestral sequence that would require the fewest changes to generate the diversity observed in a particular sample. But the search for consensus is less rigorous in linguistics because linguistic variation is much wider than biological variation, since only four nucleotides make up DNA. In biology, some proteins are so important to an organism that little or no change could be tolerated. Thus, many sequences of proteins change extremely slowly, and it is possible to prove their relationship without reconstructing ancestral sequences of millions or even billions of years ago. Knowledge of a proto-language maybe

helpful in comparative analyses, but imposing this exercise on all linguistic classifications is a serious limitation when so few proto-languages have been generated. Furthermore, reconstructions have a low probability of being entirely reliable. Greenberg's method avoids this impasse. It may be more subjective than is desirable, but it can go much further than other methods.

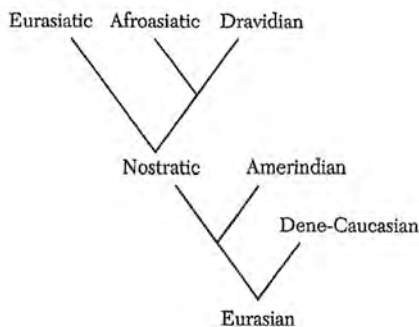
The classification of families by Merritt Ruhlen (a student of Greenberg's) appears to me to be satisfactory for comparing genetic and linguistic evolutions, as we will do in the next section. Defining a family does not appear to be an entirely objective task, but the distinctions between families, subfamilies, and superfamilies are mostly a matter of convenience and are unnecessary for certain purposes. What matters is the possibility of establishing a simple, logical, and hierarchical relationship. Unfortunately, most modern classifications stop at the level of families, of which there are as many as seventeen in Ruhlen's unifying system. There are some superfamilies, but, as already noted, modern linguistic methods have not yet generated a complete tree growing from a single source.

It is of interest to consider some of the proposed superfamilies, even if they are rather controversial. According to Ruhlen, Austric is a superfamily consolidating four families : Miao-Yao (spoken in pockets of southern China, and northern Vietnam, Laos, and Thailand) ; Austroasiatic, comprising Munda languages spoken in northern India and Mon-Khmer (spoken mostly throughout Southeast Asia) ; Daic (spoken widely in southern China and much of Southeast Asia), and Austronesian. There are some 1000 Austronesian languages, spoken by about one hundred eighty million people, including Taiwanese aborigines and Malayo-Polynesians. The latter group ranges from Taiwan to Polynesia, parts of Melanesia, the Philippines, Indonesia, Malaysia, and as far west as Madagascar. The oldest Austronesian languages are spoken by Taiwanese aborigines. Whether or not this superfamily is accepted, it ties together a very wide geographic region, including Southeast Asia, both insular and non-insular areas, and a large number of islands in the two oceans separated by Southeast Asia.

Superfamilies extending over Europe and Asia are of particular interest, Two linguistic groupings in this region are current today and are closely related : Nostratic and Eurasiatic. Rejected at first by most linguists, they are slowly being recognized. The Nostratic superfamily, as originally described by various Russian scientists, includes the Indo-European, Uralic (spoken across the Uralic mountains), Altaic (widely spoken in Central Asia), Afroasiatic, which comprises many North African and also Semitic languages, Dravidian (currently spoken almost only in southern India), and South Caucasian families. One Russian linguist, Vitaly Shevoroshkin, showed that the Nostratic superfamily has strong similarities to the Amerindian group, as defined by Greenberg. The Eurasiatic superfamily proposed by Greenberg is similar to Nostratic, but it is different in the extension given to some families like Altaic and comprises smaller families like Eskimo and Chukchi, in addition to Japanese. Thus, Eurasiatic extends further east than Nostratic, but not as far southwest, as it does not include Afroasiatic and Dravidian, which, according to Greenberg, have an older origin.

One could continue building on this tree by adding an earlier branch, leading to the Nostratic/Amerindian group on one side, and to a new, older superfamily on the other, Dene-Caucasian. Sapir initiated this new grouping, but Sergei Starostin proposed it officially only a few years ago. The Dene-Caucasian superfamily includes essentially three families : North Caucasian, Na-Dene, and Sino-Tibetan. The last family of languages is spoken by almost a billion individuals (in China, India, Nepal, Burma, and also Southeast Asia, plus some isolates in Europe and West Asia), and is therefore the most populous one.

Thus, one could draw this approximate diagram :



This hierarchy of superfamilies includes almost all of Europe, northern Africa, most of Asia, and all of America. It lacks only three African families, Khoisan, Niger-Kordofanian and Nilo-Saharan, as well as Australian (170 languages), and Indo-Pacific, a group of 700 languages, spoken mostly in New Guinea but also in neighboring islands and in the Andaman Islands near Malaysia.

There is, however, a small group of languages called isolates, which most linguists are unable to classify in any of the better - established families. The best known is Basque. Still spoken by approximately 12 000 French and perhaps a million and a half Spanish, this language is likely a relic of a pre-Neolithic period and is possibly related to the language spoken by Cro-Magnons, the first modern humans in Europe. But it has certainly changed enough that modern Basques and Cro-Magnons would not be able to communicate if they had the chance to meet. In fact they probably wouldn't even recognize their languages as related. Several linguists suggest a relationship between Basque and modern languages of the northern Caucasus. It is thus possible that one or more pre-Indo-European languages were spoken in Paleolithic Europe. Other linguists see even more encompassing resemblances among Basque, Caucasian, Sino-Tibetan, and Na-Dene languages. The last are spoken in the northwestern region of North America. Another claim is that Burushaski, an isolate spoken in a high valley in the Himalayas, is related to Basque and Caucasian. And, according to other linguists, Sumerian, Etruscan, and other linguistic « fossils » belong to the same ancient family, Dene-Caucasian. If the Nostratic/Amerindian group were added to the Dene-Caucasian group to form a hypothetical Eurasian superfamily, it may have extended across all Europe and Asia (except the southeast), and spread to the Americas. This giant superfamily later differentiated into several branches, and local twigs of this tree flourished and extended to regions far and wide.

These are interesting and hopeful hypotheses that need to be explored further. If we want to stay on absolutely firm ground, the situation is worse than simply lacking a reliable tree to link all modern languages : it is not even certain that all languages share a common origin. Most linguists consider both problems insoluble. It's a bit like trying to determine if all life on the planet had a single origin. (Many biologists believe in a single origin, since there is only one form of the twenty aminoacids found in proteins.) Greenberg noticed that there is at least one word all the linguistic families share : the root *tik*. It means finger, or the number one (a semantic shift, which requires no explanation). In other languages we find other semantic changes of this root, which also appear acceptable, « hand » and « arm, » for example, or « point, indicate. » In French, *doigt* (pronounced « dwa ») and in Italian, *dito* (meaning « finger ») come from the Latin root *digit*.

Extending this example, American linguist John D. Bengtson has proposed, with Ruhlen, about thirty other roots that are nearly as universal. But it will take a long time for other linguists to examine and accept these newest results. As might be expected, there are very few roots that are common to most languages. Most of them designate parts of the body, or are personal pronouns or small numbers (one, two, and three). It is not surprising that words conserved since the beginning of linguistic diversification are among the first words we learn : eyes, nose, mouth, and so on. But there are others that were certainly very important in Paleolithic life and have been preserved in many languages ; « lice » is one example.

2. Comparison of Linguistic Families with the Genetic Tree

Even without a comprehensive linguistic tree, we can still compare our genetic tree to existing linguistic trees. There are some impressive similarities.

In figure 2, the language families have been drawn next to the populations that speak the respective languages. We see that a family corresponds to one or more branches in the genetic tree. Sometimes a language family is represented in the joint genetic-linguistic tree by only one branch, because the populations speaking these languages were grouped together in the genetic analysis. In effect they show great similarity, either genetic or ethnographic, and live near each other. An example is the Bantu subfamily, belonging to Niger-Kordofanian, which is genetically homogeneous and distinct from other African groups. Although the word Bantu designates a language group, it is also useful as a biological category. Other genetic groupings also have been corroborated by linguistic information ; for example, southern Indians speak Dravidian languages and the Na-Dene speak Native American ones. Shared language families often point to a common genetic and ethnic background.

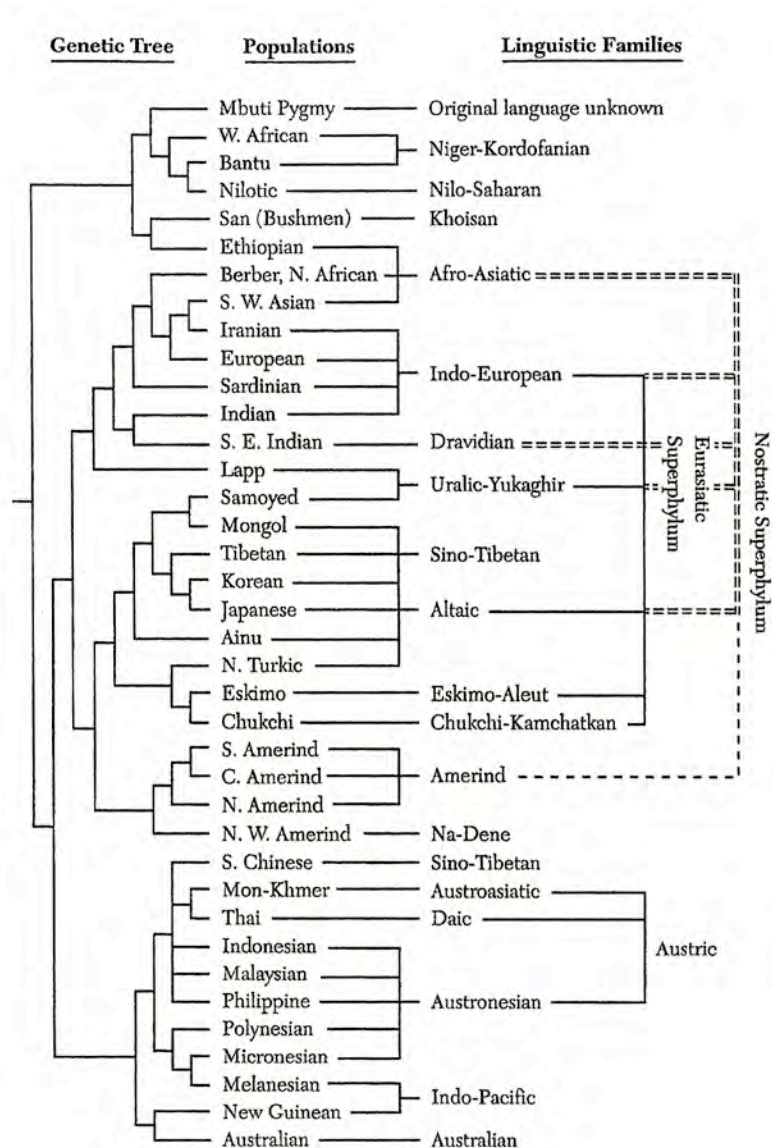


Figure 2 : *The comparison of genetic and linguistic trees* (Cavalli-Sforza et al. 1988, pp. 6002-6006).

The genetic tree shown in figure 2 is composed of 38 populations, some of which are broadly grouped (e.g., Europeans, Melanesians). There are only 16 language families (we had no genetic data on Caucasian populations when we drew this tree). Therefore some populations in the genetic tree must and indeed usually do belong together to the same linguistic family. We can immediately note that populations that are adjacent in the genetic tree usually speak languages of the same family. Because of this we can use the genetic tree to help us date the approximate origin of a linguistic family. Most language families appear to have developed during a brief period, between 6000 and 25 000 years ago.

There are, however, exceptions to the tendency of populations speaking related languages to be close in the genetic tree. Ethiopians, for example, are part of the African genetic branch, but they generally speak languages from the Afroasiatic family, which are widespread in North Africa and the Middle East, where the people are generally Caucasoid. Ethiopians are in effect a bit more African than Caucasoid genetically and more Caucasoid than most Africans linguistically. The Saami (Lapps) illustrate another exception to this rule : they are genetically Caucasoid but speak a Uralic language whose other representatives live mostly in northeastern European Russia and northwestern Siberia near the Urals. The Uralic people of Asia are generally Mongoloid, and the Lapps are a mix of Caucasoid (probably from Scandinavia) and Mongoloid (of Siberian origin), with a prevalence of the former. Even without looking at their genes, we can see this in the color of their skin and hair, and in the color and shape of their eyes, which vary in type from Mongoloid to Caucasoid.

There is a simple explanation for these disagreements between genetic and linguistic classifications. As we have explained, these two populations are the products of relatively recent genetic admixture : of Europeans and Siberians for the Lapps, and of Africans and Arabs for the Ethiopians. In the genetic tree, the populations are placed with the groups that contributed the greater proportion of genes. Extended admixture may also put them in a more isolated and somewhat intermediate position in the tree. The genetic effects of population admixture are much simpler and more predictable than linguistic change. Genes of a mixed population occur in proportions corresponding to those of its ancestral parental populations. But a genetically mixed population tends to preserve only one of the two original languages. Sometimes, the language of a mixed population will not change at all ; more often, however, we find a few words or, sometimes, sounds borrowed from the other language. Grammar is more resistant to change than vocabulary. As for the origin of the mixtures that generated Ethiopians and Lapps, we know of intimate contacts between Arabs and Africans in Ethiopia during the first millennium B.C. An Arab-Ethiopian kingdom established a capital first in Arabia (in the Saba region) and later in northern Ethiopia (at Aksum). But earlier contacts may have occurred that took place too early for history to record. We also know that the Lapps have been in their present territory for at least 2000 years. In both these cases, the shortage of written records limits our ability to know just how far back contact originated. In each case, the degree of genetic mixing established can vary and depends on the amount and mode of contact between two populations.

It is easy to calculate that the genome of a people who received a constant genetic flow of 5 percent per generation from its neighbors would keep only 70 percent of its original genome after three centuries. This is the average value of admixture for African Americans, who have retained 70 percent of their original genome and received 30 percent from mostly white settlers. If this flux continues with the same speed, African Americans will have little more than 10 percent of their original genes after 1000 years of habitation in America. In the cases of Lapps and Ethiopians, their parent populations may have been in reciprocal contact for a long time (perhaps several thousand years), and degrees of admixture are greater than those observed in African Americans, who were in contact with whites for a much shorter period, and under conditions of strong social inferiority.

We can find still other interesting exceptions to exact correspondence of the genetic and linguistic trees. Genetically, Tibetans belong to the group of northern Mongols, but they speak a Sino-Tibetan language like the Chinese. The Chinese represented in our tree, however, originate from southern China and are genetically more similar to southern Mongols. History comes to our rescue in this case as well. According to Chinese historians, Tibetans are related to the northern Chinese. Starting from the third century B.C.

In northern China, pastoral nomads headed south to Tibet. Some remained nomadic shepherds, but most maintained their original language after migrating. The unification of China began around the third century B.C. under the short-lived Qin (pronounced Chin) dynasty, and was completed under the Han Dynasty, which reigned for the next four centuries. These two northern dynasties can be credited with the spread of their language beyond northern China into the entire country. During the following 2000 years, this language naturally differentiated into several others. Nevertheless, a large number of ethnic minorities (55 are officially recognized, forming about 10 percent of the total Chinese population), especially in southern China, have preserved their original language and genes, showing their different origins. The large majority of Chinese (close to 90 percent) call themselves Han, but the original genetic difference between north and south is still clearly visible among them. Therefore, it is not surprising that Tibetans have conserved their northern Chinese genes, although they live in the south, while southern Han look genetically more like Southeast Asians. But they all speak Sino-Tibetan languages, of northern origin.

The Nostratic and Eurasiatic superfamilies are represented at the extreme right of the group of language families in the tree in figure 2. With few exceptions, they correspond with the deeper genetic branches that we have called North Eurasian (not labeled) uniting the Caucasoids, northern Mongols, and Native Americans. This branch begins near the second fission, where non-Africans split into Southeast Asians (including Australians and New Guineans), and North Eurasians. The most important exceptions to this correspondence are the two Sino-Tibetan speaking populations, Tibetans and Chinese, whose genetic and linguistic associations noted in the figure disagree. They both speak Sino-Tibetan languages

but Tibetans associate genetically with North Eurasians, and Chinese with South Asians. As we have just explained, however, the disagreement is superficial, because the Chinese shown in our tree are from the south, and have only adopted their current language in the last 2000 years, while they are genetically more closely related to the speakers of Austric languages, that is, people living in Southeast Asia with whom they are grouped in the genetic tree.

Figure 2 shows one other apparent discrepancy between the genetic and linguistic trees : Melanesians (from the Pacific islands closer to New Guinea) are genetically close to Southeast Asians and are linguistically assigned to the Indo-Pacific family. This is not entirely accurate, because the languages spoken in Melanesia (as well as in coastal parts of New Guinea) are in part Austronesian, and part non-Austronesian. The latter are mostly Judo-Pacific, a very heterogeneous family. In practice, Melanesia's situation is not a disagreement but an exceptionally complicated situation generated by several superimposed migrations that can be considerably clarified by a more detailed analysis.

The genetic tree in figure 2 still has some flaws, which are discussed below, but, because of the complexity of the relationship and the frequency of mixed origins of populations, a much more detailed representation is needed. The tree we published in 1988 shows that after the first split (separating Africans from non-Africans), the second branch separates Eurasians and Americans from Oceanians (Australian Aborigines and New Guineans) and southeast Asians. By 1994, new genetic information showed that the second branching split Oceanians from all the rest ; Southeast Asians actually split from Eurasians in the third branch, Admixture between Southeast Asia and the rest of Asia is probably the reason for this uncertainty, which has yet to be resolved because of the lack of adequate data from Southeast Asia.

Other difficulties have also intervened, because a popular tree reconstruction method developed by Naruya Saitou and Masatoshi Nei (1987), called neighbor joining, produces a genetic tree somewhat different from that of figure 2. In neighbor-joining trees, Europeans are attached to a short branch near the center of the tree between Africans and Asians. The most likely explanation is that Europeans received some of their genes from Asia and some (a smaller portion) from Africa. Using the neighbor-joining method, mixing between populations shortens the mixed branch and moves it toward the tree's origin. I believe that the conflicts between these different tree-building methods are related mostly to admixture between populations.

Mixing between North Asians and Africans contributed to the European makeup. Several European genetic traits are intermediate between the two parent populations. Africa probably contributed genes to Europe via a number of different routes, including the Middle East. This region is reachable directly from both Africa and Asia without crossing the sea, and may have been a starting point for the occupation of Europe by modern humans 40 000 years ago. Neolithic humans came from the Middle East 10 000 years ago. But we cannot exclude the possibility that modern humans also entered Europe from northwestern Asia. This may explain the relationship which several linguists claim exists between Basque and the Dene-Caucasian languages spoken in extreme western Europe. Survival of the evidence of old relationships may be due to the refuge offered by geography, especially mountains, to relic languages and populations.

A genetic tree built using a reasonably large number of micro-satellite markers showed that Basques were more similar to the Hunza (who speak the Burushaski language) than to four other Pakistani populations investigated. This observation, a result of the magnificent research effort done in Pakistan by Dr. Qasim Mehdi and his colleagues, will need confirmation with a greater number of individuals and markers. In the meantime, it provides the first genetic indication of a relationship between these two peoples, for which some remote linguistic resemblance was independently postulated.

3. Why Is There a Similarity between Biological and Linguistic Evolution ?

There are important similarities between the evolution of genes and languages. In either case, a change, which first appears in a single individual, can subsequently spread throughout the entire population. For genes, these changes are called mutations ; they are passed from one generation to the next and can - over many generations - increase in frequency and even eventually completely replace the ancestral type. The genome is conserved and well protected from outside influences. Genetic mutations are rare, and transmission from one individual to another occurs only from parent to child, while linguistic changes are much more frequent and

can pass also between unrelated individuals. As a result, languages change more quickly than genes. In effect, if a word can resist change for 1000 years, a gene can remain substantially unchanged for millions, and even billions of years. Despite these differences, there are two reasons for anticipating important similarities in the evolution of the two systems.

Let me start by emphasizing that there is no reason to think that genes influence the ability to speak one language over another. If there are any such differences, they must be small indeed. Modern humans possess the capacity to learn any language, and the first language learned is a function of the time and place of birth. All modern languages share a similar level of structural complexity-ethnic groups that live at a primitive economic level do not speak a more primitive language than wealthier groups. If there is any interaction between genes and languages, it is often languages that influence genes, since linguistic differences between populations lessen the chance of genetic exchange between them.

Linguistic evolution is a special type of cultural evolution, as we shall discuss more generally in the following chapter. How is it possible for these two very different systems to follow parallel evolutionary trajectories, or to « coevolve »? The explanation is quite simple. Two isolated populations differentiate both genetically and linguistically. Isolation, which could result from geographic, ecological, or social barriers, reduces the likelihood of marriages between populations, and as a result, reciprocally isolated populations will evolve independently and gradually become different. Genetic differentiation of reciprocally isolated populations occurs slowly but regularly over time. We can expect the same thing to happen with languages: isolation diminishes cultural exchange, and the two languages will drift apart. Even if glottochronological estimates of the time of separation are not always as exact as we would like, in general languages do diverge increasingly with time. In principle, therefore, the linguistic tree and the genetic tree of human populations should agree, since they reflect the same history of populations splitting and evolving independently.

Nevertheless, there are several major sources of divergence between genetic and linguistic trees. One language can be replaced by another in a relatively short time. In Europe, for example, Hungarian is spoken in the geographic center of many Indo-European branches: Slavic, Germanic, and Romance; but it belongs to the Finno-Ugric branch of Uralic. The other languages of the same family are spoken in the northeast of Europe and in the west of Siberia. At the end of the ninth century A.D., the nomadic Magyars left their land in Russia, crossed the Carpathians and invaded Hungary; which had already been occupied by the Avars. The conquest resulted in a Magyar monarchy, which imposed its language on the local Romance-speaking population. The number of conquerors was large but did not constitute the majority of the population - perhaps less than 30 percent of the total. The genetic effect of this conquest was therefore modest, and further diluted by subsequent exchanges with neighboring countries. Today, barely 10 percent of the genes in Hungary can be attributed to Uralic conquerors.

Elsewhere, the Barbarian conquests following the collapse of Rome faced greater difficulties in replacing or modifying the language of the conquered, who were always more numerous than their invaders. The earlier inhabitants also usually possessed a higher level of socioeconomic organization and were able to retain their cultural identity. The Lombards probably originated in Sweden or northern Germany, but they began to conquer Italy in the middle of the sixth century A.D. About 35 000 warriors, coming from Austria or Hungary, rapidly occupied most of Italy, except for the extreme south, and established a powerful state that lasted until the eighth century, but had no significant effect on the local language. The same is true of the Franks, a German population that played an important role in French political history without affecting their language. But in England after the fall of the Roman Empire, the Anglo-Saxons-Roman mercenaries of Germanic origin - did succeed in imposing their language, after establishing political control around the sixth century. The British Isles have witnessed dramatic linguistic change in a very short time. The indigenous population spoke pre-Indo-European languages unknown to us today. In the last millennium B.C., Celtic languages were spread throughout most of Europe, from a center probably located between Austria and Switzerland. At the time of Rome's conquest, Celtic languages were spoken throughout most of the British Isles. The Romans imposed Latin, which was followed by the adoption of Anglo-Saxon. Norman invaders ultimately brought many French words into English after 1066.

Another important replacement occurred in Turkey at the end of the eleventh century, when Turks began attacking the Byzantine Empire. They finally conquered Constantinople (modern Istanbul) in 1453. The replacement of Greek with Turkish was especially significant

because this language belongs to a different family-Altaic. Again the genetic effects of invasion were modest in Turkey. Their armies had few soldiers and even if they sometimes traveled with their families, the invading populations would be small relative to the subject populations that had a long civilization and history of economic development. After many generations of protection by the Roman Empire, however, the old settlers had become complacent and lost their ability to resist the dangerous invaders.

In general, the survival of a language like Basque or Burushaski is more likely to happen in *refugia* (isolated places-like mountainous regions-resistant to invaders). A strong social identity also helps to retain a population's language.

Examples of language replacement are not limited to Europe, but because Europe's written history is quite long, the most recent replacements are documented to a unique extent. The Aryan invasions of Iran, Pakistan, and India brought Indo-European languages to Dravidian-speaking areas. The great geographic discoveries of the Malayo-Polynesians, extraordinarily skillful navigators, led to the diffusion of their Austronesian languages to parts of New Guinea, Melanesia and Micronesia, and Polynesia. To the west, Austronesian languages spread as far as Madagascar near the African coast. The Polynesian migrations had lesser genetic consequences in Melanesia and New Guinea, which were already occupied. The genetic-linguistic mosaic of Melanesia is very complex, reflecting a 5000 - year history of migration and admixture of different people. But when the latest Austronesian migrants-passing Melanesia and central Polynesia-reached eastern Polynesia, starting about 3000 years ago, they still appeared nearly Mongoloid, because they had not had time to mix with the Melanesians.

Exploration enthusiasts will be happy to learn that, from a genetic viewpoint, it is still impossible to exclude the possibility that South Americans contributed in some ways to eastern Polynesia, as Thor Heyerdahl implied in his voyages with the *Kon-Tiki*. The genetic difference between Mongoloids and Amerindians is not sufficient to say exactly if and how South Americans may have contributed to Polynesia. Recently discovered genetic Amerindian markers could undoubtedly provide a clearer answer to these questions.

The total substitution of one language for another occurs more easily under the pressure of a strong political organization of newcomers, as witnessed in the Americas. Otherwise, the separate languages spoken in nearby countries can remain relatively unaffected for thousands of years, even when the genes of neighbors experience a partial, and sometimes major, substitution. It is difficult to quantify the extent of substitution that has occurred in the Basque genome through admixture with neighboring populations, but it must be considerable. However, given the length of time during which Basques were exposed to gene flow from neighbors, especially farmers who arrived in the area some 5000 years ago, the gene flow per unit of time was small. There were perhaps only one or two mixed marriages per thousand each generation. By contrast, near complete genetic substitution without language replacement probably did occur in the Hadza and Sandawe. These two populations from Tanzania speak Khoisan languages, but their genes are unlike South African Khoisan. Both groups are quite small and must have lived among Bantus for quite a long time. Bantus probably arrived in the general area about 2000 years ago. A population isolated among other, different ones, undergoing a gene flow of 5 percent per generation over 1000 years, could result in the replacement of 87 percent of the population's original genes, and 98 percent over 2000 years. The Hadza and Sandawe were hunter-gatherers and were thus separated from Bantu farmers by socioeconomic factors sufficient to preserve their own languages but insufficient to prevent genetic exchange with their neighbors. It must be admitted, however, that it is difficult to exclude the opposite hypothesis : that both the Hadza and Sandawe have basically maintained their original non-Khoisan genetic background, and have changed their language to a Khoisan one because of original contact with Khoisan speakers. Eventually, however, this contact disappeared after Khoisans retreated south. There are many such examples of language replacement with little gene replacement in the very recent expansions of Europeans that followed the introduction of transoceanic travel, resulting in the adoption of the invaders' language. The reverse also happened : Finns speak a Uralic language but they have very little, perhaps 10 percent, of Uralic genes (still to be confirmed with more powerful markers than those on which this estimate is based). It is possible they originally spoke a language of the Balto-Slavic subfamily, when they settled in Finland, a very vast country that must have been inhabited by a low density of Uralic-speaking hunter-gatherers or nomadic pastoralists, probably kin of the Saami still living in the northern part of Finland.

As discussed earlier, there is genetic evidence that the original farmers who settled in Finland perhaps 2000 years ago were a very small population, perhaps 1000 or so. This is inferred from strong evidence of genetic drift, especially for certain genetic diseases. The new settlers probably joined a good number of native inhabitants, and peaceful contact with them helped the immigrants settle and spread. The process was facilitated by learning the natives' language, and eventually adopting it. Most likely there was little genetic exchange between the two.

In summary, the replacement of languages is not the only force that disturbs the parallelism between genetic and linguistic evolution. Genetic change due to gene flow from neighbors into a small group can be another one. Deeper analysis and, especially, historical information can frequently help to distinguish between the various explanations. It is remarkable that, despite the opportunity for genetic and linguistic replacement, we can still find sufficient coherence in the modern linguistic and genetic jumble to reconstruct a common tree for the two evolutionary tracks. But the rate of disappearance of traditional languages, a serious loss which is difficult to fight, is so high that this investigation may become impossible in a few generations.

4. Interpreting the Great Human Expansions on the Basis of Genetic and Linguistic Data, Particularly in Asia

We have already observed that, judging from the genetic tree, most linguistic families appear to date anywhere from 6000 to 25 000 years ago. Some families are older ; based on their time of colonization, the Indo-Pacific languages of New Guinea and those of Australian Aborigines may be older than 40 000 years. In this case the definition of families is aided by the great geographic isolation of these two regions, an island and a continent.

Khoisan languages must also be very old, given the uniqueness of some of their characteristics (e.g., the presence of click sounds), but it is difficult to know precisely how old. I would not be surprised if the ancestors of these peoples were responsible for the first expansion from Africa to Asia. There is some support for this perhaps surprising, certainly titillating, hypothesis, other than the possible antiquity of Khoisan languages. According to some anthropologists, the Khoisan people who now live in southern Africa once lived further north, in East Africa or perhaps even northeastern Africa. If this was the case between 50 000 and 80 000 years ago, then they were in the best position to expand to Asia. It is true that of all Africans, East Africans are most similar to Asians. This is encouraging, but it might also be due to more recent migrations between East Africa and Arabia, which certainly happened much later in both directions. It is also worth noting that Khoisans bear some physical resemblance to East Asians, notably in their long-drawn eyes and large round heads. They also show a remarkable genetic similarity to western Asians (but not to East Asians, in spite of their superficial resemblance for facial traits). At the time of writing, the first results of a major breakthrough in the study of Y chromosomes promise to give us some much anticipated answers to these problems.

Let us consider, in parallel, both genetic and linguistic data in other parts of the world, as they might illuminate the oldest expansions. There clearly must have been major expansions and migrations in Asia, soon after the arrival of modern humans. The principal components of Asia clearly indicate some regions that may be centers of major demographic developments. The first five PCs give us, in order : (1) a region of northwest Iran, south of the Caspian Sea, bordering Iraq on the western side and Turkmenistan on the north-eastern side ; (2) Southeast Asia ; (3) the region around the Sea of Japan, including Japan, Korea, Manchuria, and northeastern China ; (4) northern India ; (5) Central Asia. One also notices a major genetic gradient of the first PC extending between east and west. It must have been generated by several migrations in eastern and western directions, of which there are many indications in history and in immediate prehistory. Obviously, there must have been many similar ones before. Also the Y chromosome analysis indicates that there must have been several different demographic developments in Asia, leading to the settlement of the three continents : Oceania, Europe, and America, in this order. It gives some idea of main migratory paths, but the data are not yet to the point of giving the precision one would like to reach.

Recent research by Li Jin et al. (1999) comparing the Y chromosome and also a narrow but highly variable region of chromosome 21 shows that there must have been more than one major migration from Africa to Asia. The Y chromosome is particularly informative on older migrations, as shown by recent work of Peter Underhill and Peter Oefner, who use a

new technique of discovering genetic variation called DHPLC. At the time of writing they have found 165 genetic variants of the Y chromosome, which can be divided into ten major groups (called haplogroups, and numbered from I to X). The first three are the oldest and originated in Africa, but the third born, III, migrated from Africa to Asia. The other seven, from IV to X, originated in Asia and several of them are found also in Oceania, Europe, or America. They also correspond to centers of expansion indicated by the PCs : the VI and IX haplogroups correspond to the south Caspian. These two haplogroups may well include the whole Middle East, and may be a composite of several centers of expansion at different times, including earlier expansions from and to North Africa, and the development of agriculture. The VII and VIII correspond to the Southeast Asian center ; the IV to the expansion from the Sea of Japan ; the V to northern India. The expansion from Central Asia (the fifth PC) probably corresponds to a later branch of the V haplogroup. One can also assign, at least tentatively at this stage, a linguistic family or superfamily to each expansion center : Greenberg's Eurasiatic to the south Caspian region, Austric to the second PC center in Southeast Asia, Dene-Caucasian to the eastern center of the third PC.

The center of origin of Dravidian languages is likely to be somewhere in the western half of India. It could be also in the south Caspian (the first PC center), or in the northern Indian center indicated by the fourth PC. This language family is found in northern India only in scattered pockets, and in one population (Brahui) in southern Pakistan. It was spoken earlier further west-certainly in Elam (southwestern Iran), and possibly in the Indus valley (eastern Pakistan). The major group of residual Dravidian languages is spoken, as is well-known, in the south of India. It may seem strange to place the origin of a language family in an area in which almost no language of the family is represented today. But it is a reasonable assumption that this family was removed from its place of origin by the arrival in Pakistan and northern India of Indo-European speakers, 3500 to 4000 years ago. The effectiveness and cruelty of the Indo-Europeans' war against earlier settlers of India is told in vivid images in battles described in the *Mahabharata*.

The fifth PC indicates a center of expansion further north, approximately in the Altai region, and a natural suggestion is that this is the center of origin of Altaic languages. Two major expansions appear to have originated from this area, both late : by Mongols (third century B.C.) and by speakers of Turkic languages (beginning with the eleventh century AD). There surely could have been many other expansions before. In a period preceding the Mongolian and Turkic expansions, the three millennia before the Christian era, the area was in part controlled by Indo-European speakers (Tocharians).

Interestingly, the first suggestion of a reverse migration from East Asia was a hypothesis by Michel Hammer on the basis of a peculiar Y chromosome mutant he discovered in Japan and also in Africa. We found other mutants supportive of Hammer's observation. The origin of the specific Hammer mutant was not necessarily in Japan, but the migratory chain seems reasonable and independent from others. The haplotypes involved may be the Asian branch of the HI Y-chromosome haplogroup, and the IV haplogroup. It is possible that this expansion may have brought the oldest Eurasian linguistic family (called Dene-Caucasian by Starostin) across Asia to reach Europe, perhaps 40 000 years ago. This family is represented in *refugia* by linguistic isolates like Basque, a Caucasian family, Burushaski in northern Pakistan, Ket (on the Yenisei River, in central Siberia, possibly ancestral to the Na-Dene family in north-western America) and some other extinct, less clearly related languages (like Sumerian and Etruscan). Two major families also belonging to the Dene-Caucasian superfamily, Sino-Tibetan and Na-Dene, survived over more extensive regions. They were nearer to the postulated region of origin, and must have survived in larger numbers than the other branches of the family.

Dene-Caucasian may have been the Eurasian superfamily par excellence until perhaps 20 000 years ago, which some linguists have given as the time of origin of Nostratic. Nostratic may have been a late branch of Dene-Caucasian. It started growing perhaps 10 000 to 20 000 years ago and gave rise to the families which largely replaced Dene-Caucasian in Eurasia : Indo-European, Uralic, Altaic. Greenberg's Eurasiatic family lengthens the northern Asian range of Nostratic to the east. It excludes Afroasiatic, whose geographic origin is obscure, but is earlier than Eurasiatic and originated in Africa, according to Greenberg.

5. The Indo-European Family

Indo-European remains the most studied language family. Attempts to determine its place of origin have given incredibly variable results. Many locations have been proposed, ranging from Germany as far as the northeastern Caucasus and from the Baltic States to Suez. Some hypotheses are even wilder. Not long ago, one of the most popular theories was proposed by the archeologist Marjia Gimbutas, who postulated an origin above the Black Sea and associated the earliest speakers of Indo-European with the Kurgan culture of the Asian steppes. But when Gimbutas published her hypothesis, the Kurgan dates were poorly known. She assumed 3000 to 3500 years B.C., a date which was rejected as too old by English archeologists. Gimbutas's dates appear to have been vindicated by new excavations, which have also shown that horses were probably domesticated and mounted at that time and that war chariots were built in this area.

In 1987, Colin Renfrew proposed that Indo-European languages were conducted north by the Neolithic farmers of the Middle East. I mentioned his influential book previously, which corroborated our hypothesis that Neolithic agriculture spread by a demic and not a purely cultural process. It is tempting to champion the correspondence between the spread of Indo-European languages and the diffusion of agriculture, which geography brings clearly to light. However, in my discussions with Albert Ammerman, my archeological collaborator in the initial research on the spread of farming and farmers, we avoided linguistic correlations because archeology cannot tell us about them in the absence of a written record. Nevertheless, on the basis of theoretical anthropological considerations, archeologist Renfrew came courageously to the conclusion that Indo-European was spread by Middle Eastern farmers.

I learned about Renfrew's hypothesis before he published it, on the occasion of a visit to Cambridge. Further connection between the spread of agriculture and language came to mind when I learned from linguistic literature that the language written in a cuneiform script around 5000 years ago in the region of Elam (southwestern Iran) was Dravidian. Both Renfrew and I independently suggested that Dravidian may have originated in the Middle East and been spread by mideastern farmers east toward Pakistan and India. But in the last section of this book I tried to shift the origin of Dravidian away from the Fertile Crescent further east, either to the south Caspian, eastern Iran, or northern India. It seems very reasonable to assume that agricultural developments helped spread the languages spoken by the first farmers. This must have happened repeatedly, and we will see other examples. But agriculture developed no earlier than 10 000 years ago, and therefore the relevant linguistic families are late ones. If Greenberg is right in stating that Dravidian as well as Afroasiatic is older than Eurasiatic, then the center of origin of Dravidian is not necessarily connected with the Middle East, and maybe further to the east.

Another interesting question related to the difficult problem of centers of origin of linguistic families arises with Renfrew's hypothesis that the Indo-European languages originated in Turkey, and then spread into Europe with Neolithic peasants. Obviously, all immigrants bring their language with them, and have no reason to learn a new one if they fail to encounter anyone in their new territory. It is worth pointing out that the inhabitants of Europe before the agricultural expansion (often called Mesolithics) usually had a very low population density. Since they were hunter-gatherers, they may have preferred living in areas that were geologically different from agriculturally suitable land. These earlier inhabitants and the new settlers, the Neolithic farmers, did not therefore have much contact, especially at the beginning of the agricultural expansion, when the density of farmers was lower than at later stages.

Renfrew's hypothesis, if correct, provides a date for the dissemination of Indo-European languages equivalent to that of the initial spread of farming, around 9500 to 10 000 years ago. This date may seem to be problematic, since old linguistic estimates (although very approximate) suggested an earlier date (6000 years ago). Moreover, this latter date would fit more comfortably Gimbutas's hypothesis of a Kurgan origin 5000 to 5500 years ago (kurgans are tombs in the form of mounds, which have been quarries of art objects in southern Russia).

As we shall see, however, there is no necessary contradiction between Gimbutas and Renfrew. On the contrary ; Alberto Piazza and I believe that their proposals reinforce each other. If we accept this idea, it may be useful to refer to the original Indo-European spoken in Turkey 10 000 years ago as the primary Indo-European, pre-*proto*-Indo-European, and to that spoken 4000 to 5000 years later in the Kurgan region as secondary or *proto*-Indo-European.

It is clear that, genetically speaking, peoples of the Kurgan steppe descended at least in part from people of the Middle Eastern Neolithic who immigrated there from Turkey. To arrive north of the Black Sea farmers from Turkey may have expanded west of the Black Sea, through Romania, and/or along the eastern coast of the Black Sea. Shortly after their arrival, these Neolithic farmers domesticated the horse, which was not as abundant elsewhere, and developed a predominantly pastoral economy. This allowed them to survive and even prosper in an environment ill suited to an exclusively agrarian life. This adaptation took time, but with the first development of bronze (around 5000 years ago), they were on the brink of an expansion. They had food, a means of transport, and powerful new weapons. Actually, the Kurgan region extended fairly widely, and generated many expansions after this first one, over the next 3000 or 4000 years. The very first area of origin may have been between the rivers Volga and Don, but there were many expansions, both eastward to Central Asia and westward toward Europe. Kurgans are found over much of the steppe in both western and eastern directions.

The eastward expansion may have been first. It led east and south through Central Asia toward Iran, Afghanistan, Pakistan, and India, generating the « Indo-Iranian branch » of Indo-European. These languages later completely replaced almost all the Dravidian languages previously spoken from Iran to Pakistan and in northern India, but not those in southern India. Most inhabitants of India are Caucasoid, even if their skin is darker than that of northern Europeans. Populations in the south that speak Dravidian languages are genetically slightly different from, and darker than, northern Indians. At least three ethnic layers are superimposed in this part of the world. The oldest and the most limited (the pre-Dravidians, or Australoid) have unfortunately not been studied in detail. They are said to resemble Australian Aborigines in some respects ; the similarity can only be superficial, but these people are likely to be more or less direct descendants from first African immigrants. As for the Dravidians, they were most probably the first Neolithic farmers of India, but it is unclear where they actually came from - perhaps from the Middle East, as originally hypothesized both by Renfrew and myself, or perhaps from northern Iran or northern India, as outlined above. Unfortunately, not much is known about the development of Indian agriculture.

The expansions in the opposite direction, westward, toward central and northern Europe, generated, one after the other, the Celtic, Italic, and Germanic branches of Indo-European languages. Northward expansions may have originated the Balto-Slavic expansion, which was perhaps the last. The southward expansion was less successful, as the area was already heavily inhabited, but from the second millennium B.C. there were various Indo-European speaking peoples and dynasties in Turkey and the Middle East - the Hittites and the Mitanni - whose probable origin was from the Kurgan.

That Gimbutas's and Renfrew's ideas seem more reasonable in combination than either alone was confirmed by a recent study of the Indo-European language tree. We conducted this research using material published in 1992 by two linguists, Isidore Dyen and Paul D. Black, and the statistician Joseph B. Kruskal, to make the first complete, quantitative analysis of the similarities between Indo-European languages. The data published are the frequency of common origin of two hundred words in some six dozen Indo-European languages. All possible pairs of languages were compared and the similarity of each pair evaluated by calculating the percentage of words that showed common origin by classical linguistic criteria. For instance, « mother » in English and *mère* in French have common origin, while « head » and *tête* do not. The words were those of a standard list used for « glottochronology » a method for dating the separation of languages. Their statistical method is multidimensional scaling, a sophisticated type of principal components analysis. Using their data we applied two modern methods of tree reconstruction developed for genetic studies, and obtained easily reproducible trees that, interestingly, correspond reasonably well to August Schleicher's original one. The biggest difference is the position of the root, which as always is the most difficult to assess.

The most important groups of Indo-European languages are the following subfamilies : Germanic (which includes English and the Scandinavian languages), Italic (issuing Latin, among other languages spoken in Italy, in the first millennium B.C.), Balto-Slavic, Celtic, Greek, Indian, and Iranian. Most linguists consider Indo-Iranian a single branch, although Dyen and his colleagues say they are distinct. In our tree, several languages have an early, separate origin : Albanian, Armenian, and, later, though somewhat less clearly, Greek. Extinct languages like Hittite and Tocharian could not be considered in our analysis. The same tree was obtained by two other major methods of reconstruction and is shown in figure 3.

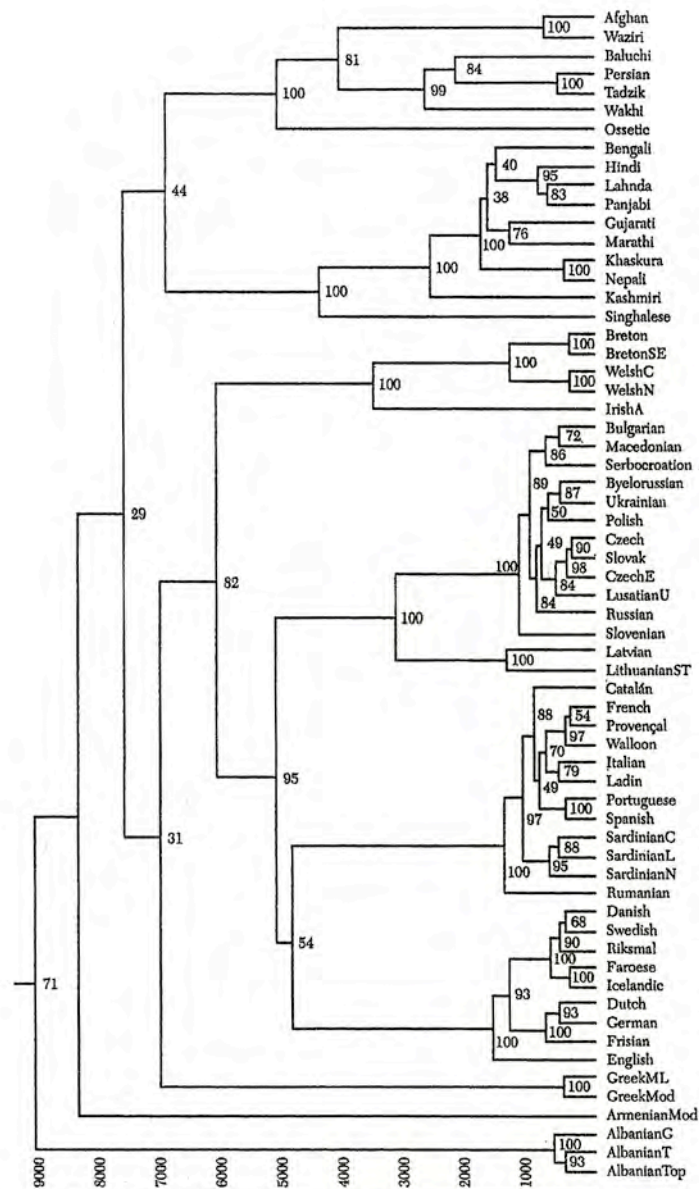


Figure 3 : Tree of 63 Indo-European languages. Numbers near branches indicate the reliability in percent of the specific branch, calculated by the method of the bootstrap.

The scale on the bottom indicates years.

(From an unpublished manuscript by Piazza, Minch, and Cavalli-Sforza, based on data from Dyen, Kruskal, and Black 1992).

It is reasonable to think that isolated languages like Albanian and Armenian (and with less evidence Greek) originated with the first wave of Neolithic farmers from Turkey. Their greater age with respect to other families is responsible for their early position in the tree. They are also geographically closest to Turkey.

Our analysis groups the Indian and Iranian languages into one Indo-Iranian group, in the classic tradition of Indo-European studies, but slightly contradicting the results of Kruskal and his colleagues.

The later branches are likely to be derived from the second wave of Indo-European migrations from the Kurgan area, central European ones from its western part, and the Indo-Iranian branch from the eastern part. The order of branching in the tree is interesting : that of Celtic, Balto-Slavic, Italic, and Germanic subfamilies corresponds reasonably well to their geographic distance from the center of origin. Here the first branch of the tree is the Celtic, which is still spoken in the extreme western areas of Europe, and therefore is the most distant one from the Kurgan area. The next split generated the Italic-Germanic and the Balto-Slavic branch. The Italic branch settled in the southwest of Europe ; it was unable to

completely replace the earlier pre-Indo-European language Basque but was successful in replacing Etruscan in the Italian peninsula. The Germanic branch settled in the northwest, north, and center of Europe, being unable to completely replace the earlier Indo-European subfamily which had also survived Celtic. The Baltic and Slavic branches settled in the northeast and the south-east, respectively, competing with the earlier settlers who spoke Uralic languages.

Recently, Tandy Warnow from Philadelphia and others have proposed a completely independent tree analysis of Indo-European history. Their results are not fully published and therefore difficult to analyze. They use a small number of word roots, believed to be especially reliable, and a very small number of languages that includes extinct ones. Their conclusions differ from ours mainly because they propose a very late branching of Celtic. This is difficult to reconcile with the early diffusion of Celtic languages to almost all of Europe, and their suppression by latecomers, speaking Germanic and Italian languages, which confined these languages to the extreme geographic periphery of northwestern Europe. The small number of words employed and the lack of a statistical test of robustness of conclusions may be the main problem of this otherwise - very interesting analysis.

6. The Bantu Expansion

Many other expansions have brought new languages to new lands. The demic expansions of peoples that we are familiar with were almost all accompanied by their original languages. Among the prehistoric expansions studied both genetically and linguistically, the Bantu expansion is of considerable interest. Despite contacts and genetic flow from tribes speaking other languages, like Nio-Saharan in East Africa, and the Khoisan in South Africa, the Bantus have largely maintained their genetic distinctiveness, which makes them somewhat different from the other West Africans from which they originated. Starting from Nigeria and Cameroon, they headed south, toward the Atlantic coast. The Bantus of the first expansion were still using Neolithic tools 3000 years ago, but later Bantu expansion was aided by the introduction of iron. Only around A.D. 1000 the Bantus reached the Great Lakes region of Uganda and Kenya, and from there they expanded south, both near the coast of the Indian Ocean and further inland. From this time on, archeologists find Bantus relied heavily on iron.

The western and eastern currents moving into the south central continent met eventually. The Bantus were apparently still a few hundred kilometers from the Cape of Good Hope when the Dutch landed there around 1650. Both archeology and linguistics also show that the Bantus arrived earlier in Namibia along the West Coast.

7. Darwin's Prophecy

The origin of scientific linguistics can be dated to 1786. In that year, the English judge Sir William Jones advanced the theory that Sanskrit, Greek, and Latin, and possibly Celtic and Gothic (the ancestor of Germanic languages), appeared to have a common origin, at a famous conference at the Bengal Asiatic Society of Calcutta, which he had just founded and served as president. The similarity of Sanskrit and European languages had already been noticed by the Florentine merchant Filippo Sassetti and by the Jesuit priest Coerdoux. The latter sent notes from Pondicherry to the Academy of Inscriptions in Paris showing that Sanskrit, Greek, and Latin must have a common origin, but his conclusion did not have the impact of Jones's conference. In 1863 the German linguist August Schleicher published a tree showing the origins of Indo-Europeans very much like one we would draw today using modern methods. The ties between biology and linguistics were evident at once. Schleicher was certainly influenced by Charles Darwin's use of trees to explain the theory of organismal evolution. In *On the Origin of Species*, Darwin clearly stated that if we knew the tree of biological descent of the human groups, we could extract the tree relating languages. This effort was not attempted until 1988. I am ashamed to say I was not aware of Darwin's prophecy at the time. I was reminded by a friend who is a historian of our science and had read our 1988 paper, with Alberta Piazza, Paolo Menozzi, and Joanna Mountain, in which we correlated together on a global scale data of genetics, archeology, and linguistics. Here are Darwin's words :

The natural system [of classification] is genealogical in its arrangement, like a pedigree. It may be worthwhile to illustrate this view of classification by taking the case of languages. If we possessed a perfect pedigree of mankind, a genealogical arrangement of the races of man would afford the best classification of the various languages now spoken throughout the world ; and if all extinct languages, and all intermediate and slowly changing dialects, were to be included, such an arrangement would be the only possible one.

The correlation between genes and languages cannot be perfect, since the rapid conquest of large territories may favor replacement of indigenous languages with unrelated ones, as happened in much of the Americas. But these events do not appear to have occurred frequently enough to erase all trace of a correlation. We see equally that in the case of prolonged genetic exchanges with different neighbors, genes can be replaced. Nevertheless, despite the two sources of confusion, the correlation between genes and languages remains positive and statistically significant.

Even at a microgeographic level, the regions subject to detailed study have usually shown strong correlations between geography, genetics, linguistics, and other cultural aspects like surnames. Often the genetic-linguistic mosaic we observe clearly shows the effects of numerous expansions - some are known historically - and of their superimpositions and interactions. Perturbations do occur, but they do not manage in most cases to obscure the clarity of the correlation between genes, peoples, and languages.

8. A Hypothetical Model of Linguistic Evolutionary History Based Mostly on Genetic and Archeological Knowledge

Linguistic evolution is a subject of unusual interest : In this chapter, we have restricted ourselves to explaining the similarities between genes and languages. But linguistic evolution is also very important to understand as an example of a more general phenomenon, cultural evolution.

Following Darwin's suggestion, one can attempt to use our knowledge of genetic evolution to make hypotheses about the earlier part of the linguistic tree. Figure 4 shows the linguistic history reconstructed when a number of cues taken from the genetic tree are correlated to linguistic information. Merritt Ruhien drew the tree, using our 1988 genetic tree as a guideline. But he also took into account new linguistic superfamilies that had been daringly proposed in the interim. I have made few changes to his tree, adding some probable dates. When our genetic data is entirely satisfactory ; the tree will probably not be as simple as this one, but it seems likely that its major features will not change.

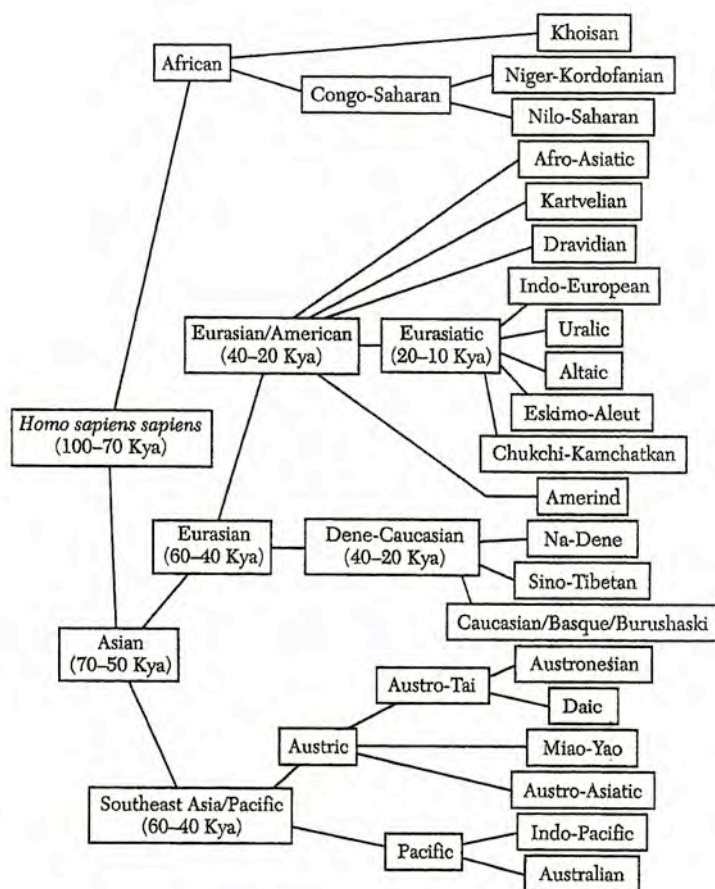


Figure 4 : The tree of origin of human languages, drawn by Merritt Ruhlen in 1994, using Darwin's suggestion to base it on the genetic tree (given in figure 2), somewhat modified, with the likely range of initial divergence dates (Kya = 1000 years ago).

The oldest linguistic families must be African : of the four existing today, Khoisan is considered to be the oldest ; Afro-Asiatic the most recent ; Niger-Kordofanian and Nilo-Saharan probably had a common origin (the Congo-Saharan superfamily suggested by some linguists), and must have arisen at an intermediate time. Khoisan may be the most direct descendant of the people who left Africa, as explained before.

For those Africans who remained, the first linguistic split is likely to have been between the branch leading to modern Khoisan on one side and a branch leading to the ancestral languages of Congo-Saharan on the other. Afroasiatic is likely to have originated much later, probably in northeast Africa, or otherwise in the Middle East, or in Arabia.

The Niger-Kordofanian family consists of two branches. The very small one is the Kordofan, named after a large mountain range in western Sudan, and the other is called Niger-Congo. Perhaps there was a westward spread from East Africa, first to the Kordofan, then to West Africa, or the expansion may have been in the opposite direction. In West Africa a major population burst coupled with the introduction of agriculture may have occurred 4000 to 6000 years ago. I hope archeologists will note the genetic cues from principal component analysis of genes that suggest this may have begun between Mali and Burkina Faso (former Upper Volta). Agriculture then spread from Nigeria and Cameroon to central and southern Africa with the Bantu expansion, which took 3000 years to reach as far south as it would.

The agricultural expansion originated in West Africa found a forest population of hunter-gatherers, Pygmies. Most of my work in Africa was in, the forest areas of the tropical belt, where Pygmies have survived the arrival of farmers. They are very few and, wherever the forest has disappeared, their descendants are hardly recognizable. Unfortunately, the original Pygmy languages have all disappeared, replaced by languages of neighboring farmers. The only possible traces of these languages are in their names for forest animals and plants. Working in an area where foraging (hunting-gathering) barely survives side by side with agriculture has given me a chance to observe the transition from foraging to farming, which is still going on but will soon be over. I believe that archeologists interested in the Neolithic transition in Europe and elsewhere should spend time in this area, while it is not too late, in order to see a living model of what must have happened in similar circumstances in other parts of the world. All the farming people in proximity are dark-skinned, while those who live in the forest are less so.

A group of very dark-skinned Africans (sometimes called « elongated » because of their very elegant, tall and thin bodies), many of whom could make a good living as fashion models, inhabit East Africa and other neighboring regions. Linguistically, they are Nilo-Saharan, a name that indicates their dominant living area and their origin. The Nilo-Saharan started domesticating cattle in the Sahara perhaps 8000 years ago, but many of them had to leave when it became a desert. Even today they are still mostly shepherds.

The earliest confirmed date of modern humans in Asia is in China, 67 000 years ago ; and the first settlement of New Guinea and Oceania is believed to have been 50 000 to 60 000 years ago at the earliest, or at most somewhat later, 40 000 years ago. Did modern humans reach East Asia by land or perhaps by boat, along the southern coast? Probably both. There exist archeological estimates of the rate at which the recolonization of northern Europe took place at the end of the last glaciation (around 13 000 years ago) : it varied from 0,5 to 2 kilometers per year, not too different from that of farmers. But the major limiting factor may have been largely that of ice withdrawal, rather than that of human movement. It is possible that the coastal route allowed a faster movement. How long did it take to travel by the hypothetical coastal route from East Africa, a likely point of departure, to Southeast Asia (a necessary point of arrival, from which some may have continued north to East Asia along the coast of the Pacific, and others south to New Guinea and Australia) ? We may venture a minimum guess of 10 000 years. Perhaps modern adventurers might try to repeat a small part of the journey in conditions similar to those of our ancestors. This could be very informative, even if today's coast and the availability of seafood must be different from what it was then. Assuming that it took that much time to go from East Africa to Southeast Asia, the average displacement that may have taken place would have been on the order of 50 to 60 kilometers per generation (2 km/year). This is about twice the rate of advance of farmers many thousands of years later. We are speaking of a mode of life of which there do not seem to exist living or historical examples (other than perhaps in Borneo). I would describe it as nomadic fishing. But the model of demographic expansion would not be very different from that of Neolithic farmers,

in the sense that there must be both migration and reasonably active reproduction for an expansion to take place. Of course, over the generations, some extended families or small groups may have settled or chosen to leave the coast and go into the interior, as others continued their almost random wandering along the coast. The settlement of Southeast Asia, New Guinea, and Australia led to the development of the Indo-Pacific and Australian linguistic families. Andamanese and other Negritos of Southeast Asia are the closest living descendants of Africans who were among the first settlers of Southeast Asia and Oceania.

China and Japan may have been settled before Australia, and may have been the first areas of development of the Dene-Sino-Caucasian family, which must have spread west through central Asia to Europe. The Na-Dene branch went to Siberia, from which it later (about 10 000 years ago) migrated to North America, after the Amerindians first colonized it (some 15 000 to 30 000 years ago).

A late expansion centered in Southeast Asia was that of the Austrie superfamily, which connects genetically Taiwanese aborigines and the southern Mongols of Southeast Asia. The second PC of Asia indicates a possible genetic expansion centered in Southeast Asia. This may have happened both very early, and also again as late as the local development of agriculture. Principal components cannot distinguish two expansions having the same area of origin which took place at different times.

The multiplicity of migrations and expansions between Europe and Asia, in both directions, is well substantiated. A recent discovery of a late west-to-east migration was that of the Judo-Europeans who went as far as western China as recently as 4000 to 1000 years ago, but the languages spoken by them, Tocharian, are extinct. Among the last people to expand from a center located in the east were Mongolians who engaged China, forcing the emperors to build the Great Wall more than 2200 years ago. Attila the Hun went as far west as Italy. Their relatives speaking Turkic languages started expanding from Central Asia about eight or nine centuries ago, eventually reaching Turkey and the Balkans.

We have mentioned that the roughly continuous genetic gradient from Europe to East Asia is the result of all these migrations. Many Central Asians are herders and nomadic. Languages, especially those forming the Eurasiatic family, introduce apparent discontinuities in the gradient, as communities inevitably speak a single, common language. Some extend widely and may, move, creating a complex human geography. Political changes and military events may force the replacement of languages in a relatively short time. Here the correlation of languages and genes cannot be perfect but it is still evident to some extent, in spite of the turbulence of the Eurasian history of the last four or five millennia.

Genetic research can certainly help the understanding of linguistic evolution, and vice versa.

References

- Bengtson (J.-D.). [with Merritt Ruhlen] 1994. « Global Etymologies ». in : Ruhlen (M.). *On the Origin of Languages : Studies in Linguistic Taxonomy*. Stanford, CA : Stanford University Press.
- Cavalli-Sforza (L.-L.), Piazza (A.), Menozzi (P.), Mountain (J.). 1988. « Reconstruction of Human Evolution : Bringing Together Genetic, Archaeological and Linguistic Data ». in : *Proceedings of the National Academy of Science*, 85, pp. 6002-6006.
- Darwin (Ch.). 1859. *On the origin of species by means of natural selection [or, The preservation of favoured races in the struggle for life]*. London : John Murray.
- Dyen (I.), Kruskal (J.-B.) & Black (P.). 1992. « An Indoeuropean classification : A lexicostatistical experiment ». in : *Transactions of the American Philosophical Society*, 82,5. Philadelphia : American Philosophical Society.
- Gimbutas (M.). 1982. *The Goddesses and Gods of Old Europe 3500-3500 BC*. London : Thames and Hudson [updated edition].
- Greenberg (J.-H.). 1987. *Language in the Americas*. Stanford : Stanford University Press.
- Hammer (M.-F.). 1995. « A recent common ancestry for human Y chromosomes ». in : *Nature*, 378, pp. 376-378.
- Hammer (M.-F.) & Zegura (S.-L.). 1996. « The role of the Y chromosome in human evolutionary studies ». in : *Evol. Anthropol*, 5, pp. 116-134.
- Hammer (M.-F.), Spurdle (A.-B.), Karafet (T.), Bonner (M.-R.), Wood (E.-T.), Novelletto (A.), Malaspina (P.), Mitchell (R.-J.), Horai (S.), Jenkins (T.) & al. 1997. « The geographic distribution of human Y chromosome variation ». in : *Genetics*, 145, pp. 787-805.
- Jin (L.), Underhill (P.-A.), Doctor (V.), Davis (R.-W.), Shen (P.), Cavalli-Sforza (L.-L.) & Oefner (P.-J.). 1999. « Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations ». in : *Proceedings of the National Academy of Sciences*, 96, pp. 3796-3800.
- William (J.). 1786. *The Sanscrit Language*. London.
- Naruya (S.) & Masatoshi (N.). 1987. « The Neighbor-joining Method : A New Method for Reconstructing Phylogenetic Trees ». in : *Molecular Biology Evolution*, 4, 4, pp. 406-425
- Qasim (S. Mehdi) & al. 2002. « Y-Chromosomal DNA Variation in Pakistan ». in : *American Journal of Human Genetics*, 70, 5, [May], pp. 1107-1124.
- Renfrew (C.). 1987 *Archaeology & Language : The Puzzle of the Indo-European Origins*. London : Jonathan Cape.
- Ruhlen (M.). 1987. *A Guide to World's Languages - Volume I : Classification*. Stanford : Stanford University Press.
- Ruhlen (M.). 1994. *The origin of language. Tracing the evolution of the mother tongue*. New York : Wiley [trad. Fr. 1997. *L'origine des langues*. Paris : Belin].
- Ruhlen (M.), (ed.). 1994. *On the origin of languages : studies in linguistic taxonomy*. Stanford, Ca. : Stanford University Press.
- Sasseti (F.). [1578-1588] 1970. *Lettere [Letters]*. [edited by Vanni Bramanti].
- Schleicher (A.). 1863. *Die Darwinsche Theorie und die Sprachwissenschaft*. Weimar : H. Böhlau.
- Shevoroshkin (V.). 1992. *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum: Bochumer Universitätsverlag [ACTA First International Symposium on Language and Prehistory. Ann Arbor, 8-12. November, 1988].
- Starostin (S.). 1990. « A statistical evaluation of the time-depth and subgrouping of the Nostratic macrofamily ». in : *Symposium on Molecules to Culture*. Cold Spring Harbor, NY : Cold Spring Harbor Laboratory Press.
- Underhill (P.-A.), Jin (L.), Oefner (P.-J.), Cavalli-Sforza (L.-L.). 1996. « A Pre-Columbian Y Chromosome-Specific Transition and Its Implications for Human Evolutionary History. in : *Proceedings of the National Academy of Sciences*, 93, pp. 196-200.
- Warnow (T.). 1996. « Reconstructing the evolutionary history of natural languages ». in : *Proceedings of the seventh annual ACM-SIAM symposium on Discrete algorithms*, pp. 314-322. Atlanta : Georgia (United States).



**La convergence pluridisciplinaire
dans la recherche de l'origine des langues :
l'exemple américain de la « Nouvelle Synthèse »
- Aspects historiques, théoriques et étude critique**

Par Laurent Métoz
Laboratoire HTL, CNRS UMR 7597,
ENS LSH, Lyon, France

Mai 2006

1. Introduction

« Nouvelle Synthèse », nom donné à la convergence pluridisciplinaire des travaux de L.-L. Cavalli Sforza (généticien), M. Ruhlen (linguiste) et C. Renfrew (archéologue). Ces scientifiques se donnent comme objectif ultime de comprendre l'origine de l'Homme par l'étude de sa diversité génétique, linguistique et culturelle et de prouver la monogenèse des langues du monde et de l'espèce humaine.

Une telle convergence ne résulte pas d'une création *ex nihilo* mais, au contraire, d'un lent processus qui, depuis l'étude des langues amérindiennes par M. Swadesh (1954, 1956, 1962) et J. Greenberg (1979, 1986, 1987), n'a jamais cessé de chercher à lier linguistique, archéologie et génétique. D'ailleurs, M. Ruhlen (1987, 1994), pilier linguistique de la « Nouvelle Synthèse », doit beaucoup, nous le verrons, aux travaux de ces derniers. Ainsi, en nous focalisant sur cette succession des méthodes de classification linguistique, une réflexion centrée autour de trois questions essentielles peut être menée : comment ces méthodologies ont-elles émergé *i.e.* quels en ont été les facteurs déclencheurs ? Quelle est leur valeur tant théorique que méthodologique ? Comment cette succession de méthodes de classification linguistique a-t-elle abouti à l'émergence de la « Nouvelle Synthèse » ?

Pour répondre à ces questions, il est impératif de décrire les méthodologies employées par M. Swadesh, J. Greenberg et M. Ruhlen et d'analyser les liens méthodologico-théoriques qui les unissent. Cette première phase permettra de faire émerger les futurs fondements de la « Nouvelle Synthèse ». Il s'agira ensuite, par l'étude des arguments avancés par la génétique des populations, d'en montrer certaines limites.

2. Historique des fondements de la « Nouvelle Synthèse »

2.1. La place de la glottochronologie dans l'émergence de la « Nouvelle Synthèse »

Au début des années 1950, M. Swadesh, par analogie avec les découvertes contemporaines en archéologie (notamment le carbone 14, découvert en 1945 par F.-W. Libby), développe une méthode permettant une datation des divergences linguistiques : la glottochronologie. Les fondements théoriques sur lesquels elle repose sont doubles : l'existence d'un vocabulaire de base commun à l'ensemble des langues du monde d'une part, un taux constant d'évolution des langues d'autre part.

Les objectifs de M. Swadesh sont clairs : il s'agit non seulement d'adopter une méthodologie qui, sur la base d'une étude lexicale des langues, va lui permettre d'introduire dans les recherches glottochronologiques des langues pour lesquelles les documents scripturaux sont rares voire inexistantes mais aussi de démontrer, mathématiquement, la preuve des relations génétiques qui peuvent exister entre les différentes langues du monde (ce qui lui permettra de remonter au plus près de l'origine des langues et du langage humain).

2.1.1. Fondements théoriques de la glottochronologie

M. Swadesh, même si sa méthodologie est foncièrement lexicale, n'omet nullement l'importance de la phonétique dans l'étude des langues. Il en constate cependant la faiblesse méthodologique : « Chaque fois que les linguistes ont essayé de comparer des langues très lointainement parentes, ils se sont heurtés à des faits non concluants » (Swadesh (M.). [1967] 1986, p. 29). Il semble gêné par la loi de régularité phonétique (développée par les « néogrammairiens » vers la fin du XIX^e siècle) qui, même si elle constitue un postulat méthodologique fort de la démarche comparatiste traditionnelle, reste proche du dogme. Cependant, le

phénomène d'érosion phonétique, dont il ne cesse de souligner l'impact diachronique, constitue un obstacle majeur à ladite loi. Ceci l'amène à conclure que l'efficacité de la méthode comparatiste traditionnelle est loin d'être établie d'où un rôle minoré de la phonétique au sein de la glottochronologie.

Dans le comparatisme classique, le fait que le lexique, partie de la langue la plus instable, puisse être soumis à un renouvellement rapide dû aux nombreuses contraintes qui pèsent sur lui est un autre postulat méthodologique fort : plus nous nous éloignons de la période commune plus des facteurs de divergence lexicale ont pu jouer. L'ultime frontière au-delà de laquelle il n'est plus possible, pour ces linguistes, d'étudier l'apparement des langues est de six mille ans. M. Swadesh, ne l'admettant pas, va, dès lors, développer deux concepts qui lui permettront d'aller à l'encontre de cette vision classique à savoir, premièrement, l'existence d'un fond lexical commun à toutes les langues du monde (environ cent notions) qui, deuxièmement, s'avérerait stable temporellement.

2.1.1.1. Le vocabulaire de base

L'existence d'un vocabulaire de base universel a d'emblée été soumise à critique. En effet, en partant du principe selon lequel il existe toujours plus de langues que de cultures, M. Swadesh parvient à réduire la diversité culturelle mondiale engendrant ainsi une universalisation de certaines notions linguistiques. L'élargissement global des dimensions sémantiques d'une notion de base, ce qu'il appelle la constellation lexicale (cf. figure 1) fait qu'il devient plus aisé de chercher et de retrouver des items lexicaux communs à toutes les langues du monde. De plus, dans ce type de recherches, fondées sur un postulat universaliste, certaines notions du vocabulaire de base s'avèrent inappropriées selon l'aire linguistique considérée (e.g. notions relatives aux systèmes de parenté, cf. R. Blust, 2000).

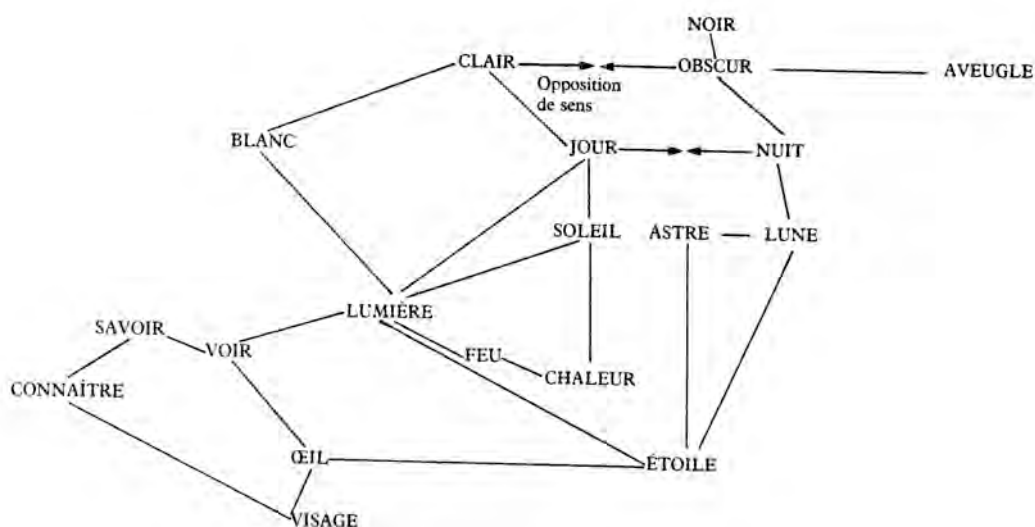


Figure 1 : Exemple de constellation lexicale (Swadesh, 1967, 1986, p. 134).

Bien que fortement soumis à critique, cette définition du vocabulaire de base, avec sa dimension universaliste (D. Hymes 1960a, b), sera reprise sans aucune réévaluation dans les travaux ultérieurs de J. Greenberg et M. Ruhlen. L'intégration dans une recherche pluridisciplinaire comme la « Nouvelle Synthèse » de concepts linguistiques erronés ne peut constituer qu'un obstacle à toute crédibilité tant théorique que méthodologique.

2.1.1.2. Le taux constant d'évolution

M. Swadesh, en créant une analogie entre datation au carbone 14 et rythme constant d'évolution des langues, pose les premières pierres de la « Nouvelle Synthèse ». Postuler une évolution à vitesse constante des langues revient à appliquer le gradualisme darwinien à la linguistique. Ce présupposé théorique permettra, trente ans plus tard, la mise en parallèle, déjà supposée par Ch. Darwin, des évolutions linguistiques et humaines. Lorsque L.-L. Cavalli-Sforza considère comme intellectuellement très proches l'horloge linguistique de M. Swadesh et l'horloge moléculaire, cela n'est nullement anodin. Mais ce taux constant sera remis en question par plusieurs de ses contemporains comme par exemple K. Bergsland et H. Vogt

(1962), D. Hymes (1960a, 1960b, 1964), H. Hoijer (1956) ou encore I. Fodor (1961, 1982) et, plus récemment, par R. Dixon (qui a tenté, en 1997, d'appliquer la théorie des équilibres ponctués de S.-J. Gould et N. Eldredge aux langues) et J. Matisoff (qui, en 2000, a développé le concept d'*Accretion Tolerance Quotient*).

2.1.2. Les étapes de la démarche glottochronologique : description et critique

2.1.2.1. Constitution d'une liste de vocabulaire de base

La construction d'une liste de vocabulaire de cent notions (pour suivre les préceptes de M. Swadesh) commune à l'ensemble des langues sélectionnées pour l'étude constitue la première étape de la glottochronologie. Les conditions d'intégration d'une notion dans le vocabulaire de base sont triples : une notion doit être commune (*i.e.* d'usage courant mais les items les plus couramment employés dans une langue correspondent-ils réellement au vocabulaire de base ?), universelle (*i.e.* existante dans toutes les sociétés connues) et historiquement persistante (*i.e.* résistante à l'emprunt). Ce dernier critère est le plus complexe à établir car il nécessite une compréhension fine des différents phénomènes linguistiques ayant pu avoir un impact sur les langues étudiées (le tabou...). Outre ces difficultés spécifiques, d'autres viennent s'y agréger *e.g.* notions inappropriées suivant l'aire linguistique considérée, limitation artificielle de la liste (Matisoff (J.). 2000) ou encore la création de listes de vocabulaire spécifique suivant les familles linguistiques considérées (seule condition pour que la glottochronologie, uniquement fondée sur des données indo-européennes, puisse être universelle, *cf.* Blust (R.). 2000 ; Matisoff (J.). 2000).

2.1.2.2. Remplissage des listes

Cette étape permettra de calculer le taux de résistance de chaque langue. M. Swadesh a énoncé trois critères permettant de remplir la liste de vocabulaire pour chacun des états d'une langue donnée. Premièrement, essayer de trouver un équivalent simple pour chaque notion, en ne tenant pas compte de formes spécialisées ou liées, et le moins commun des deux équivalents. Deuxièmement, employer un seul mot ou élément plutôt qu'une phrase, même si le sens est plus large que celui de la notion de la liste. Troisièmement, là où il est impossible de trouver un seul élément, omettre la forme. Malheureusement, l'application stricte de ces critères s'avère difficile : leur multiplication constitue un obstacle majeur dans une méthode où des langues sans tradition écrite ou pour lesquelles peu de documents sont disponibles peuvent être introduites *e.g.* rareté des documents, nature des documents survivants, subjectivité de l'intervention humaine dans le choix des items à incorporer dans la liste.

2.1.2.3. Comparaison des listes

À ce stade, il s'agit de déterminer si les deux items lexicaux correspondant à une seule notion de la liste sont, ou non, cognats *i.e.* génétiquement apparentés. Cette étape n'est pas exempte de difficultés méthodologiques puisque le phénomène d'érosion phonétique peut conduire à un non-apparement d'items lexicaux historiquement cognats. Les travaux de J. Matisoff (2000) ont montré que deux langues distinctes peuvent posséder un pourcentage élevé de vocabulaire commun sans pour autant être génétiquement apparentées. Les causes d'un tel phénomène sont multiples : une relative jeunesse de la culture de l'écriture (impliquant une évolution normative inachevée), une aire linguistique complexe où les possibilités d'emprunt sont nombreuses (langues tibéto-birmanes...). D. Hymes (1960a, 1960b, 1964) a, lui, montré que le glissement sémantique des items lexicaux, qui s'opère de manière aléatoire, peut générer une interprétation erronée en termes d'évaluation des cognats. Ces erreurs potentielles peuvent entraîner une surestimation des divergences temporelles des langues étudiées (en comptant moins de cognats qu'en réalité) ou, au contraire, une sous-estimation (en comptant plus de cognats qu'en réalité).

2.1.2.4. Calcul des dates glottochronologiques

Ici, le linguiste est à même d'appliquer les formules glottochronologiques lui permettant de calculer les dates de divergence des idiomes linguistiques étudiés. Deux formules (fonctions logarithmiques) rendent ce calcul possible : la première représente la résistance de la liste par millénaire pour une langue donnée. Le temps écoulé entre deux états de cette langue sera calculé en employant la formule : $t = (\log c) / (\log r)$ où t représente le temps écoulé exprimé en millénaire, c est le vocabulaire commun aux deux états de langues exprimé en pourcentage et r indique le taux de rétention (de résistance propre à la liste exprimé également en pourcentage).

S'il s'agit de calculer le temps pendant lequel deux langues ont divergé l'une de l'autre, une deuxième formule peut être dérivée de la première : $d = (\log c) / (2 \log r)$ où c correspond au vocabulaire commun des deux langues et d au temps pendant lequel la divergence a eu lieu.

La comparaison des listes pour les deux états de chaque langue aboutit à l'établissement d'un pourcentage d'items lexicaux retenus pendant la période de temps écoulé entre les deux états. Les pourcentages pour toutes les langues doivent être ramenés à une base commune et arbitraire comme le millénaire selon une fonction logarithmique (S. Gudschinsky en 1964 tentera d'effacer cette dimension temporelle mais cette tentative s'avérera infructueuse). Cependant, les langues font partie intégrante des groupes sociaux, ce qui implique que la division linguistique ne peut jamais s'opérer de manière radicale et abrupte. Si les langues ne divergent pas d'un coup d'un seul, alors il faut faire nôtre l'idée que les dates obtenues par la glottochronologie qui, elle, postule une divergence soudaine des langues, soient plus courtes qu'en réalité (cf. schéma ci-après). Ceci constitue d'ailleurs le paradoxe central de la glottochronologie : elle se propose de dater les séparations d'idiomes linguistiques mais que faut-il entendre par « séparation » ? Ce terme renvoie à un processus lent invalidant d'emblée l'idée d'une rupture brutale dans la filiation linguistique (Cosériu (E.). 1965). Ce type de raisonnement, résurgence de la théorie caduque de l'arbre généalogique schleicherien fait, qu'en réalité, il n'y a pas de date à calculer puisque les langues ne se séparent jamais brusquement.

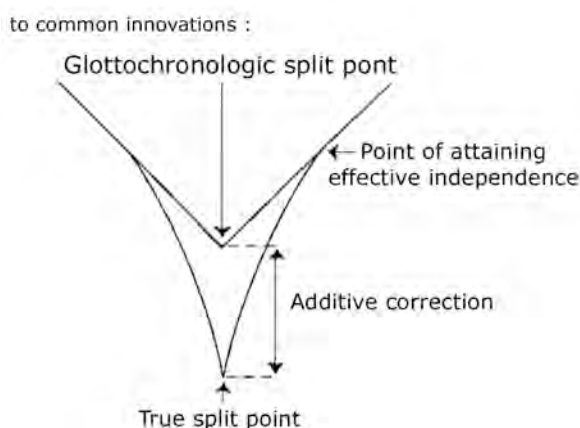


Figure 2 : L'affinement des datations glottochronologiques (Hymes (D.). 1960, p. 20).

2.1.2.5. En résumé

Bien qu'effleurée ici, cette méthodologie, aussi largement contestable et contestée qu'elle puisse être a toutefois permis d'éclaircir d'une lueur timide quelques pans de l'histoire humaine. Mais globalement, elle soulève plus d'interrogations qu'elle n'apporte de réponses.

Malgré ces critiques, notons toutefois que par l'analogie théorique entre la datation au carbone 14 et le concept d'évolution constante des langues, les premières pierres de la « Nouvelle Synthèse » sont posées : en appliquant le concept darwinien de gradualisme aux langues, nous nous orientons vers une possible démonstration de la convergence des évolutions humaines et linguistiques, comme le présentait Darwin et comme tentera de le démontrer M. Ruhlen (1987, 1994) et surtout L.-L. Cavalli-Sforza (1988).

2.2. Joseph Greenberg : comparaison multilatérale et génétique

2.2.1. Fondements de la méthode de comparaison multilatérale

Au lieu de comparer des langues appartenant à diverses familles linguistiques, il s'agit, dans cette méthode, de comparer des langues génétiquement proches dans le but d'extraire les caractéristiques de chaque famille (« with recurrent significant resemblances », Greenberg (J.). 1957, p. 40) pour ensuite comparer chaque famille entre elles. Au final, ce qui est comparé, ce ne sont pas les langues mais les caractéristiques générales de la famille à laquelle elles appartiennent. Les groupements linguistiques ainsi réalisés ne se font donc pas sur l'héritage mais sur la ressemblance : « [...] what is involved in classification is not the registering of resemblance, but a noting of the comparative degree of resemblance » (Greenberg (J.). 1987, p. 5). Ceci nous permet de percevoir aisément la conceptualisation de l'existence de macro-familles. En appliquant à chaque étape de la comparaison c'est-à-dire à n'importe quel niveau dans l'arborescence (langues, familles de langues et superfamille) la même méthode, J. Greenberg parvient à définir des macro-familles et notamment l'Amérinde.

Cette méthode repose sur un autre postulat : plus il y a de langues intégrées lors d'une étude comparative, plus le linguiste est à même de remonter à la racine de l'arbre des langues. En effet, puisque les langues ne sont pas considérées en elles-mêmes mais comme des indicateurs de la tendance générale de la famille linguistique à laquelle elles appartiennent (caractéristiques phonétiques, morphologiques, grammaticales), plus le nombre de langues est élevé, plus ces informations seront précises et multiples.

2.2.2. Language in the Americas (1987)

Ce livre vise à démontrer, par application de la méthode de comparaison multilatérale, que les langues amérindiennes (*i.e.* les langues d'Amérique du Nord, d'Amérique du Sud et de la Méso-Amérique) peuvent être regroupées en trois grands phyla que sont la macro-famille Amérinde, l'Eskimo-aléoute et le Na-Dene.

2.2.2.1. Principales critiques

Les linguistes ayant étudié les travaux de J. Greenberg se sont majoritairement focalisés sur le *Dictionnaire Étymologique Amérindien* (désormais D.E.A.) publié en 1987 en annexe de *Language in the Americas* (1987, pp. 181-270). Ce dictionnaire répertorie deux cent quatre-vingt-une « étymologies » (dans le sens greenbergien du terme, nous y reviendrons) preuves de l'existence de la famille Amérinde. Son étude a fait apparaître, parmi d'autres, les biais suivants : ambiguïté du terme « étymologie », qualité des données phonétiques et problèmes de traduction des termes amérindiens en langue anglaise.

2.2.2.1.1. J. Greenberg et le terme « étymologie »

Chez J. Greenberg, le terme « étymologie » n'est pas à envisager comme une possible explication des glissements sémantiques des items lexicaux (qui aurait permis d'opérer une analyse historique des langues) mais bien comme un terme générique regroupant un ensemble d'items semblant posséder un sème commun et une structure phonétique proche. Le terme « étymologie » ne regroupe, en réalité, que des items lexicaux suffisamment similaires pour être génétiquement apparentés (ces « étymologies » correspondent aux constellations lexicales de Swadesh). Voici plusieurs exemples : « excrement / night / grass, ask / wish / seek / pleasure, child / copulate / son / girl / meat / be greasy / fat / deer » (Campbell (L.). 1988, p. 600) auxquels nous pouvons ajouter l'étymologie 85 du D.E.A. signifiant « dirty » où nous trouvons : dirty / black / blue / green / night / verdure / dark in color. La manière dont le D.E.A. est construit laisse supposer que la recherche des items phonétiques s'est faite préalablement à la mise en constellation sémantique facilitant, par conséquent, la création d'une macro-famille. En effet, ces extensions sémantiques augmentent artificiellement les chances d'apparementements des items phonétiques *ipso facto* des langues : en multipliant les connexions sémantiques, J. Greenberg parvient à élargir le champ de recherche des items phonétiques en réduisant la diversité lexicale des langues amérindiennes. Nous percevons immédiatement le problème soulevé par cette définition idiosyncrasique de l'étymologie : en fondant les regroupements linguistiques non sur l'héritage mais sur la ressemblance des structures phonétiques, cette méthode ne peut en rien traduire la phylogénie des langues amérindiennes, ce qui fait dire à I. Goddard : « [...] there are essentially no hypothesis of linguistic history in L[anguage] I[n] A[mericas], and nothing can be learned from this book about the history of Indian languages » (Goddard (I.). 1990, p. 557). Nous sommes ici dans une classification statique (phénéti-que) plus que dynamique (cladistique).

2.2.2.1.2. La qualité des données phonétiques : extrapolation méthodologique

Même si, au sein de cette méthode, la référence à la phonétique est explicite, nombreux ont été les linguistes à démontrer, qu'en réalité, la manière dont elle est employée ne peut en aucun cas servir de base à une classification historique des langues amérindiennes (Adelaar (W.). 1989 ; Goddard (I.). 1990) : outre l'absence de table de correspondances, nombre de données phonétiques à partir desquelles J. Greenberg construit sa classification des langues amérindiennes semblent être erronées (Poser (W.). 1992 ; Kimball (G.). 1992 ; Goddard (I.). 1990 ; Ringe (D.). 1996). W. Poser constate que, pour les langues Salinan et Yurumangui, respectivement 63 % et 65 % des données phonétiques de ces deux langues sont erronées. G. Kimball, pour les familles Gulf et Muskogean, constate que 69 % des items phonétiques de ces deux familles nécessitent une correction. Le poids du facteur chance constitue un élément majeur de l'émergence des groupements linguistiques.

Il est à noter que les préceptes méthodologiques de la comparaison multilatérale constituent un pas supplémentaire dans la biologisation des classifications linguistiques (biologisation à l'origine de la « Nouvelle Synthèse »). En effet, par le rejet de la phonétique et par l'élaboration d'une comparaison fondée non plus sur l'héritage mais sur la ressemblance, la classification linguistique perd de sa dimension historique pour s'orienter vers une dimension phénétique (n'oublions pas le concept greenbergien de « valid genetic unit » comme unité systématique de la classification linguistique, conceptuellement identique au taxon génétique, qui ne fait que renforcer l'analogie croissante entre classification génétique et linguistique).

2.2.2.1.3. La traduction des langues amérindiennes

La traduction anglophone de termes amérindiens contribuerait à faire augmenter le poids du facteur chance dans l'élaboration de groupements linguistiques, c'est en tout cas ce que laisse suggérer l'étude de « On the Best Method of Studying the North American Languages » article publié en 1891 par J.-H. Trumbull, un linguiste américain. La nature même des langues amérindiennes semble pouvoir expliquer à elle seule les difficultés de traduction : la plupart d'entre elles sont polysynthétiques *i.e.* elles offrent à leurs locuteurs la possibilité de construire des items lexicaux si complexes qu'ils peuvent incorporer une quantité de sens souvent équivalente à celle contenue dans toute une phrase d'une autre langue *e.g.* « nup-pakh-nuh-tö-pe-pe-nauwut-chut-chuh-quô-ka-neh-cha-e-nin-nu-mun-nô-nok » signifiant en anglais « our well-skilled looking-glass makers ». Extrapolons ces remarques aux étymologies 126 *Go*₁, 127 *Go*₂, 128 *Go*₃ et 129 *Go*₄ du D.E.A.. Selon J. Greenberg, seules trois acceptions apparaissent dans ces quatre étymologies à savoir « go », « come » et « road ». « Come » représente un cas sur dix-neuf, « road » cinq cas sur dix-neuf et « go » treize sur dix-neuf or Trumbull nous indique que « no indian languages [...] has any verb which exactly corresponds to the English verb « to go » [...] ». » (Trumbull (J.-H.). 1891, p. 62). Il va plus loin dans son article en démontrant l'existence d'environ quatorze acceptions amérindiennes différentes au verbe anglais « to go ».

Le biais de la comparaison multilatérale devient dès lors évident : en ne nous informant pas de l'intégralité des acceptions amérindiennes pour un terme donné du D.E.A., J. Greenberg se réserve la possibilité de chercher les items phonétiques qu'il nous présente dans l'intégralité des acceptions du terme (que lui seul connaît). Par conséquent, le nombre réel d'items phonétiques peut être augmenté de manière drastique et il lui devient de plus en plus facile de trouver des formes phonétiques proches permettant de démontrer l'existence d'une parenté génétique entre les langues amérindiennes. D'après notre exemple, s'il choisit ses items phonétiques dans l'ensemble des quatorze acceptions du verbe « to go » établies par J.-H. Trumbull et fait croire au lecteur qu'il n'y en existe que trois, le biais de la démarche devient évident.

2.2.3. Les implications anthropologiques des travaux linguistiques de Joseph Greenberg

J. Greenberg, dans son chapitre « Conclusions : An Overview » de *Language in the Americas* (Greenberg (J.). 1987, pp. 331-337), propose une extrapolation de ses résultats linguistiques à des hypothèses anthropologiques sur le peuplement des Amériques. En effet, la découverte de la macro-famille Amérinde (qui regroupe toutes les langues du continent à l'exception des familles Na-Dene et Eskimo-Aléoute), lui permet d'émettre l'hypothèse selon laquelle l'existence de trois familles linguistiquement distinctes sur le sol américain doit résulter d'un schéma de propagation des populations spécifique. Il va dès lors tenter, grâce à sa formation anthropologique, de résoudre la question épineuse du peuplement des Amériques.

Pour cela, J. Greenberg postule que chaque *phylum* linguistique présent sur le continent américain est le fruit d'une vague de migration humaine. Les répartitions géographiques des populations humaines et des familles linguistiques vont être de précieux indices dans la compréhension de la colonisation du continent. J. Greenberg conclut que les populations parlant des langues de la macro-famille Amérinde sont les plus anciennes du continent *i.e.* environ 12 000 ans (datation proche des travaux de l'anthropologue A. Hrdlicka, défenseur d'une colonisation récente du continent américain et mouvement majoritaire aux E.A.U. (*cf.* Métoz (L.). 2005, pp. 354-356). L'espace géographique recouvert par ces populations lui fait dire qu'elles n'ont rencontré aucune résistance et qu'elles devaient donc être les toutes premières. Seule l'étude de la répartition géographique et, dans une moindre mesure, de leur diversification linguistique permet de différencier l'ancienneté des *phyla*. Dès lors, et en toute logique, la deuxième vague de colonisation du continent s'est heurtée aux populations humaines de la première migration limitant ainsi leur progression à l'intérieur du continent, c'est la famille Na-Dene. Le même raisonnement est appliqué à la dernière famille, la famille eskimo-aléoute.

Quoi qu'il en soit, nous percevons la faiblesse argumentative des hypothèses proposées par J. Greenberg. Sans pour autant les remettre en cause, il semble évident que cette justification d'une triple colonisation du continent américain doit s'appuyer sur des preuves plus tangibles. C'est d'ailleurs ce qui a été fait avec la publication en 1986 d'un article fondamental dans la construction de la « Nouvelle Synthèse ».

2.2.4. Introduction de la génétique comme moyen de vérification des hypothèses linguistiques

En 1986, soit un an avant la publication de *Language in the Americas*, est publié un article de J. Greenberg, C. Turner et S. Zegura « The settlement of the Americas : A Comparison of the Linguistic, Dental, and Genetic Evidence » dans lequel il nous est proposé deux expériences visant à prouver, par une démonstration extralinguistique, le bien fondé des regroupements linguistiques de J. Greenberg. L'argumentation de cet article repose sur deux éléments distincts : premièrement, l'existence d'une corrélation entre la morphologie dentaire et langues amérindiennes, deuxièmement, une corrélation entre gènes et langues amérindiennes. Cet article est fondamental dans la création de la « Nouvelle Synthèse » puisqu'il marque officiellement l'entrée de la génétique dans la linguistique historique. Malheureusement, les preuves génétiques, à l'instar des preuves dentaires, sont si faibles que les auteurs nous demandent de ne pas en tenir compte dans la démonstration : « This is why we regard the genetic evidence as supplementary rather than primary » (Greenberg (J.) & al. 1986, p. 487). Que les preuves génétiques soient pertinentes ou pas, la question n'est pas là, l'important est ailleurs à savoir dans la possibilité théorique et méthodologique, déjà effleurée avec la glottochronologie de M. Swadesh, d'une mise en parallèle des théories génétiques et linguistiques.

2.2.5. En résumé

La méthode de comparaison multilatérale créée par J. Greenberg, qui a permis la découverte de la macro-famille Amérinde, n'est pas exempte de critique : qualité des données phonétiques, ambiguïté du terme « étymologie », poids prépondérant du facteur chance avec, notamment, les difficultés posées par la traduction des idiomes amérindiens... Cependant, malgré l'invalidation scientifique de la macro-famille Amérinde, les travaux de J. Greenberg marquent une nouvelle étape dans la construction de la « Nouvelle Synthèse » puisque la génétique s'invite désormais dans la recherche de l'origine des langues.

2.3. M. Ruhlen, la comparaison multilatérale à son paroxysme

Élève de J. Greenberg lui-même élève de M. Swadesh, M. Ruhlen publie en 1994 *The Origin of Language. Tracing the Evolution of the Mother tongue* (trad. fr. 1997. *L'Origine des Langues*). Il entend, dans ses travaux, appliquer la méthode de comparaison multilatérale à l'ensemble des langues du monde et plus seulement aux langues amérindiennes (vœu pieux énoncé par J. Greenberg à la fin de *Language in the Americas*). En l'appliquant à mille trois cent vingt-deux langues et proto-langues issues de trente-deux familles linguistiques, M. Ruhlen est parvenu à mettre en évidence vingt-sept racines mondiales preuves, pour lui, d'une monogénèse des langues du monde.

2.3.1. Principales critiques

2.3.1.1. Les connexions sémantiques inter-racines

À l'instar des travaux de M. Swadesh et de J. Greenberg, les relations sémantiques entre les items lexicaux ont été largement étudiées. L'étude des connexions sémantiques des racines mondiales de M. Ruhlen (recherches du groupe *Origine de l'Homme, du Langage et des Langues*, CNRS, 2003) nous apprend qu'il en existe non vingt-sept mais quatre *i.e.* quatre racines correspondant à autant de macro-classes étymologiques :

- la première est constituée des racines 1, 9, 11, 12, 15 *i.e.* « enfant », « chien », « homme », « parent féminin plus âgé », « femme » ;
- la deuxième n'est constituée qu'autour d'une seule racine, la 14 signifiant « rester sur place » ;
- la troisième n'est construite qu'autour d'une seule racine, la 18 signifiant « deux » ;
- la quatrième, la plus importante de toutes, est constituée de vingt racines *i.e.* les racines 2-8, 10, 13, 16, 17, 19-27 *i.e.* « genou », « cendres », « nez », « tenir (en main) », « bras », « os », « trou », « qui ? », « sucer », « penser », « quoi ? », « voler », « bras », « vulve », « jambe », « doigt », « terre », « cheveu », « eau ».

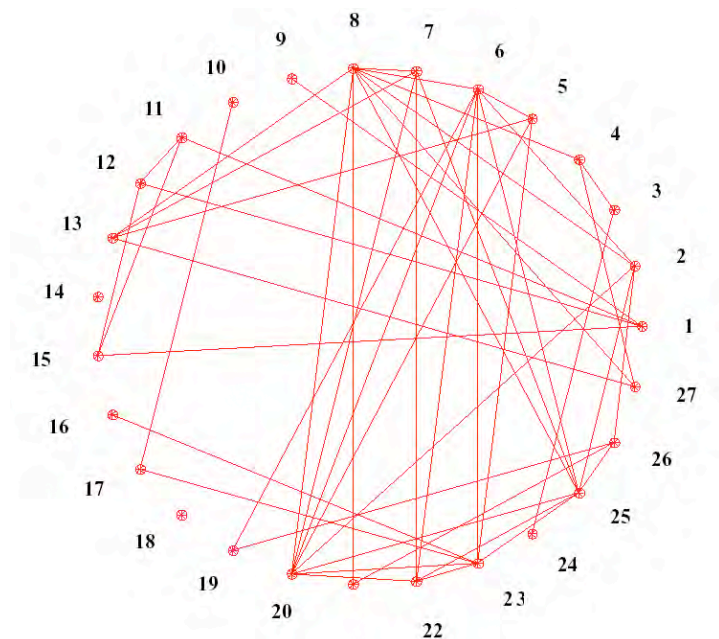


Figure 3 : Connexions sémantiques des racines mondiales de M. Ruhlen (Boe (L.-J.) & al., 2003).

2.3.1.2. Les données phonétiques

Outre les connexions sémantiques inter-racines, les correspondances phonétiques semblent de nature à remettre en cause le bien fondé des résultats des travaux de M. Ruhlen. En parcourant son ouvrage, les éléments linguistiques qui vont permettre de classer les langues sont aisément repérables (n'oublions pas qu'il propose aux lecteurs d'opérer seuls la classification des langues du monde, mécanisme connu d'objectivation de la méthode). Ces derniers sont, outre les glissements sémantiques, les correspondances phonétiques. La liste ci-dessous référence toutes les équivalences suivant l'ordre d'apparition dans le corps de l'ouvrage :

[ʃ]≡[θ] / [p], [t], [k]≡[b], [d], [g] / [l]≡[y] / [p]≡[ph]≡[f]≡[b]≡[β]≡[β] / [t]≡[th]≡[θ]≡[d]≡[ð]≡[d'] / [ph], [th], [kh]≡[p], [t], [k] / [ʌ]≡[ɰ]≡[l] / [l]≡[r] / [l]≡[n] / [u]≡[w] / [t]≡[d] / [k]≡[q] / [p]≡[b]

Nous retrouvons dans cette liste des équivalences attestées e.g. « Loi de Grimm », phénomènes phonétiques fréquents dans les langues du monde... or une analyse plus fine des données figurant en annexe de son livre fait apparaître des équivalences beaucoup plus larges. Ainsi, M. Ruhlen fournit au lecteur six équivalences attestées au phonème [k] mais l'analyse des annexes en fait émerger vingt-six (Métoz (L.). 2002). Au final, ce phonème et ses équivalents représentent 77.2 % de toutes les occurrences de consonnes dans les familles les mieux représentées dans le livre de M. Ruhlen à savoir les familles indo-européenne, afro-asiatique et Niger-Congo (Métoz (L.). 2002, p. 105). Il est clair qu'une extrapolation de ces résultats à l'ensemble des données phonétiques fournies par M. Ruhlen dans son ouvrage conduit irrémédiablement à penser qu'un tel réductionnisme phonétique ne peut conduire qu'à une chance infinitésimale de ne pas retrouver des items phonétiques communs à toutes les familles de langues.

2.3.2. En résumé

Compte tenu des connexions sémantiques inter-racines qui ramènent la diversité sémantique des racines mondiales à quatre macro-classes étymologiques et compte tenu également du réductionnisme phonétique, il apparaît clairement que l'application de la méthode de comparaison multilatérale à l'ensemble des langues du monde et des racines mondiales qu'elle permet de faire émerger est scientifiquement inacceptable. Aussi inacceptable soit-elle, c'est cette classification linguistique qui sera utilisée par le généticien L.-L. Cavalli-Sforza dans la mise en parallèle des arbres linguistiques et génétiques.

3. La « Nouvelle Synthèse »

Ce bref rappel historique des conditions d'émergence des fondements théoriques et méthodologiques permet de mettre en évidence les mécanismes de biologisation des classifications linguistiques. Une fois ces derniers acquis, nous sommes à même d'étudier la mise en parallèle des arbres linguistiques et génétiques.

3.1. La relation gènes/langues

Tout l'intérêt scientifique, et médiatique, dans la mise en évidence de cette relation de congruence est la possibilité de mettre en parallèle les évolutions génétiques et linguistiques.

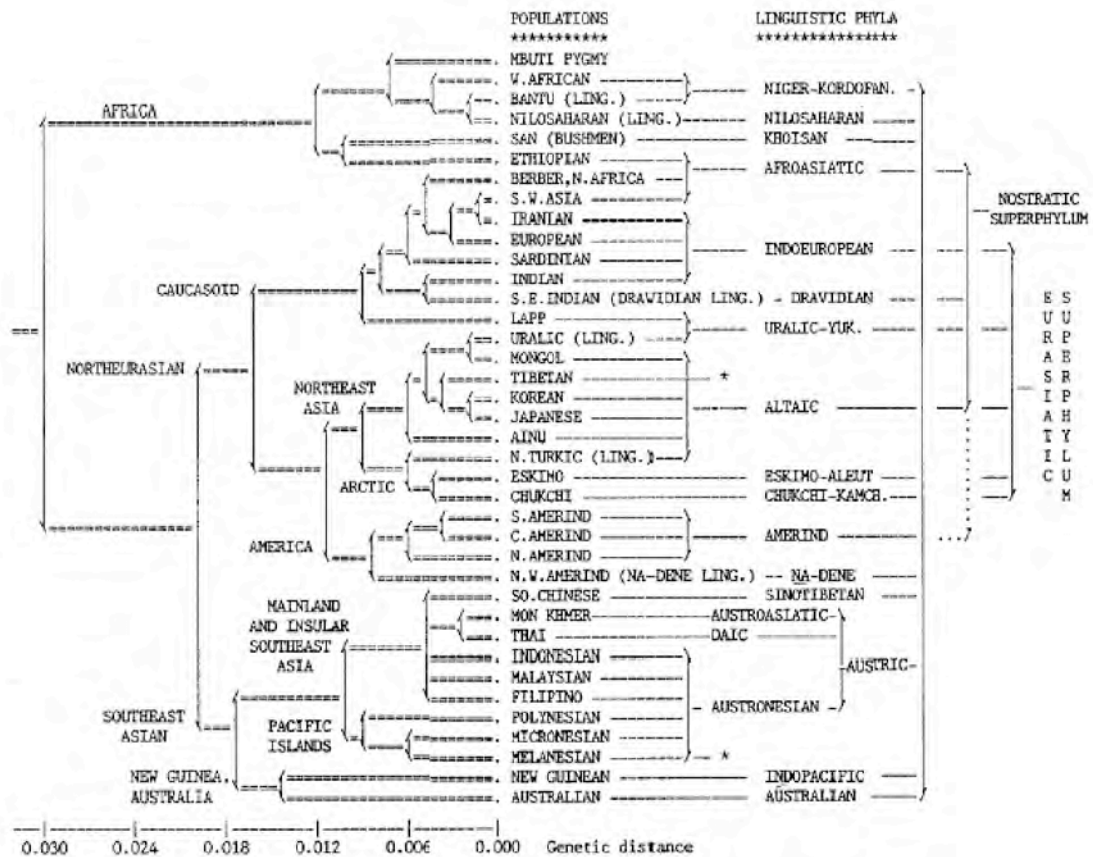


Figure 4 : L'arbre des gènes et des langues (Cavalli-Sforza (L.-L.) & al., 1988).

De prime abord, cette mise en parallèle de structures arborescentes pourrait nous faire penser que l'origine des langues a enfin cédé à la science grâce à la génétique. En réalité, il n'en est rien puisque ces arbres présentent de grandes incohérences rendant impossibles les conclusions évoquées par les fondateurs de la « Nouvelle Synthèse » à savoir une commune évolution des populations génétiques et linguistiques. Voyons ces incohérences.

3.1.1. Les erreurs de correspondance

La première, facilement identifiable puisqu'il suffit de lire la légende, est que les populations humaines biologiques suivantes : Bantous, Nilo-sahariens, Dravidiens, Ouraliens, Türks du Nord, Na-Dene sont, en réalité, des populations regroupées d'après leur classification linguistique. Autrement dit, aucune analyse génétique des dites populations ne cautionne la validité scientifique du regroupement. Dans ce cas, il est normal qu'une population soit semblable à elle-même. Ce phénomène affecte six populations sur trente-huit soit approximativement quinze pour cent des populations humaines totales.

Deuxièmement, des populations biologiques, bien que très proches sur la structure arborescente, parlent des langues de famille différentes *e.g.* les Sans et les Éthiopiens, les Asiatiques du Sud Est et les Iraniens, les Indiens et les Dravidiens, les Eskimos et les Tchouktchis, les Néo-Guinéens et les Australiens et enfin, les Mon-Khmers et les Thaïs. Parallèlement à ce phénomène, des populations qui peuvent être génétiquement très éloignées parlent des langues de la même famille. C'est le cas pour les Éthiopiens et les asiatiques du Sud-Est, les Saamis et les Ouraliens, les Aïnous et les Türks du Nord.

Certains phénomènes quantitativement moins importants doivent également être soulignés : le schéma que présente M. Ruhlen (1997, p. 167) nous apprend que les Mélanésiens, qui parlent des langues de la famille Austronésienne, peuvent, pour une partie d'entre eux, parler des langues de la famille Indo-Pacifique. Il est possible d'apercevoir tout en haut de la figure une population, les Pygmées Mbuti qui, eux, n'ont tout simplement pas de langue.

En résumé, la congruence entre l'arbre génétique et l'arbre linguistique s'estompe dès que nous rentrons dans les détails. De fait, les deux hypothèses, l'arbre généalogique de tous les gènes humains et l'arbre généalogique de toutes les langues humaines, discutables chacune dans son champ, se confortent l'une l'autre dans ce qui pourrait ressembler à un cercle vicieux d'argumentation interdisciplinaire ou à une supercherie pseudo-scientifique. Ce phénomène de circularité du raisonnement est très présent au sein de la « Nouvelle Synthèse » et A. Leroi-Gourhan, dès 1945, condamne ce type de pratique :

C'est pourquoi la linguistique ou l'anthropologie, qui donnent dans leurs aspects comparatifs de belles séries continues, de robustes ensembles, dès qu'elles se surimposent au cadre histoire, avec des dates, des lieux et des noms de peuples, deviennent hasardeuses et discontinues. (Leroi-Gourhan (A.). 1945, 1973, 2002, p. 306).

3.1.2. Nature des arbres

Nous avons déjà souligné combien les arbres obtenus par l'application de la comparaison multilatérale relevaient plus de la phénétique que de la cladistique (c'est-à-dire plus du statique que du dynamique). Le phénogramme linguistique (représentation obtenue après application de la méthode phénétique) ne permet pas, puisqu'il est impossible de l'enraciner, de postuler la moindre hypothèse quant à un éventuel évolutionnisme linguistique (telle langue diverge de telle autre). De plus, l'étude des critiques émises envers la nature phénétique des arbres linguistiques montre qu'elle est généralisable aux arbres génétiques produits par L.-L. Cavalli-Sforza puisque, dans les deux disciplines, les arbres sont construits non sur la recherche d'un héritage commun mais au contraire sur une présence ou une absence de traits – génétiques ou linguistiques. Par conséquent, la mise en concordance des deux arbres ne peut pas se faire dans des conditions pleinement satisfaisantes (*cf.* pour la nature des arbres génétiques Bateman (R.). 1990 ; Ben Hamed (M.). 2004 ; Métoz (L.). 2005 pour la nature des arbres linguistiques).

Ainsi, comme nous l'avions déjà conclu pour les arbres linguistiques construits à l'aide de la méthode de comparaison multilatérale, les arbres génétiques, mis en symétrie aux arbres linguistiques, ne sont aucunement de nature à rendre compte de l'évolution historique des gènes. D'apparence visuelle extrêmement plaisante, cette image pieuse pour tous les défenseurs de la « Nouvelle Synthèse » ne résiste pas à une lecture fine.

3.1.3. Généalogie des gènes et des individus

L.-L. Cavalli-Sforza (Cavalli-Sforza (L.-L.) & al., 2000), à partir d'un échantillon de mille sept chromosomes Y européens, parvient à identifier dix familles de mutations chez plus de quatre-vingt-quinze pour cent des échantillons. Ces familles, obtenues grâce aux différentes mutations du chromosome Y, donnent la possibilité d'opérer une classification des Européens. Cette classification lui permet, grâce à l'apport de l'horloge moléculaire, d'affirmer que la majorité des chromosomes Y proviendrait du paléolithique.

Nous percevons ici la manière avec laquelle la généalogie des gènes se lie à celle des individus et ouvre la voie, grâce à des modèles rendant compte de l'évolution des populations, à des modèles de migrations humaines. L'hypothèse d'une langue-mère des langues européennes se trouve ainsi confirmée par l'étude du chromosome Y. La reconstitution de l'histoire des migrations humaines en Europe, en partie due au développement de l'agriculture (preuves archéologiques), sous-entend une histoire des migrations linguistiques. Cette idée est d'ailleurs étayée par les travaux de L.-L. Cavalli-Sforza qui débouchent sur une similitude (quasi) parfaite entre les arbres linguistiques de J. Greenberg et M. Ruhlen et l'arbre génétique de la population mondiale. Depuis, lier migrations humaines et migrations linguistiques est un phénomène plausible grâce, en partie, à la médiation de l'archéologie.

Pourtant, tout généticien des populations sait qu'il ne faut pas confondre ces deux types de généalogies. Pourquoi ? Simplement parce que le gène en tant qu'unité évolutive fait partie d'un système dynamique : l'Homme. La reconstruction de l'histoire d'un gène ne résume pas pour autant l'histoire du système. Établir une histoire des migrations humaines à partir d'un seul gène est scientifiquement contestable. En effet, l'homme est constitué de quelque trente mille gènes, c'est pourquoi il semble difficile à partir d'un seul gène de retracer l'histoire du système. De plus, tous les gènes ne possèdent pas la même vitesse d'évolution *i.e.* certains gènes évoluent plus vite que d'autres et d'autres sont soumis à des facteurs de sélection qui ne sont pas les mêmes selon le gène. Il y a des questions de recombinaisons qui sont une

histoire propre à un gène donné, et qui ne seront donc pas généralisables au système. Tout cela fait qu'il est possible de reconstruire l'histoire d'un gène, et que deux gènes peuvent avoir des histoires différentes, *a fortiori* différentes de celle de l'individu.

3.2. La construction argumentative du proto-monde

3.2.1. Vers une montée en puissance

3.2.1.1. La dimension éphémère

Malgré l'invalidation scientifique de la « Nouvelle Synthèse », force est de reconnaître l'impact du livre de M. Ruhlen sur le grand public. Devant cet engouement inexplicable, R. Nicolai, dans son livre *La traversée de l'empirique* (2000), se propose d'étudier les mécanismes d'émergence du livre de M. Ruhlen ainsi que la manière dont la structure argumentative de l'ouvrage « fait apparaître une frontière fragile entre la croyance et le savoir » (Nicolai (R.). 2000, p.33). Cet ouvrage constitue, selon lui, la cristallisation d'un mécanisme d'ampleur : la réception par le grand public des théories scientifiques et les moyens de mise en œuvre pour sa propagation.

Il constate que le renouveau des thématiques de recherche autour de l'origine des langues trouve ses premiers fondements dans une série de publications dont la parution aurait pu passer inaperçue si, justement, elles n'avaient pas été à l'origine de ce regain d'intérêts : *L'Express* (août 1991, spécial été, n° 2093, pp. 15-21), *Le journal du C.N.R.S.* (1996), « Les langues du monde » (paru dans *Pour la Science*, octobre 1991) et enfin *La Recherche* (février 1998, n° 306). Même si nous restons dans le domaine français, ces parutions font écho aux travaux comparatistes que nous avons étudiés dans la première partie de cet article. Ces travaux ont largement favorisé la renaissance du phœnix et ceci pour plusieurs raisons : ils sont à la fois à l'origine du renouveau de la thématique monogénétique et source d'une remise en cause fondamentale des méthodologies traditionnelles, ils développent également l'idée d'un proto-monde, bref ils nourrissent les fantasmes des lecteurs.

3.2.1.2. La *Linguist List* et la dimension épiphénoménale

Les débats contemporains autour du proto-monde ont été résolus, aux E.A.U, vers le milieu des années 1990. Ainsi, dès les premières publications de J. Greenberg ou de M. Ruhlen, de nombreux linguistes ont commencé, grâce à la *Linguist List*, à opérer une critique des méthodes linguistiques mises en œuvre dans la démonstration du Proto-Monde. C'est au cours d'un de ces débats qu'apparut la thématique qui nous intéresse ici à savoir la manière dont la presse (populaire et/ou de vulgarisation) rendait compte de ces débats. Ces problématiques, bien que constituant deux pôles de réflexion différents, représentent des axes de réflexion à ne pas négliger.

Les débats qui ont lieu à cette époque sur ce forum ont eu un rôle fondamental dans la réactualisation de la thématique de l'origine monogénétique des langues du monde. En effet, la *Linguist List*, contrairement aux quelques publications citées précédemment, s'adresse à des professionnels ou du moins à des internautes sensibles à ces thématiques. Faire pénétrer le débat dans la communauté scientifique est un moyen de créer une émulation tant théorique que méthodologique parmi les chercheurs. De plus, compte tenu de l'impact de ces recherches sur le grand public, ces thématiques s'avèrent être très valorisantes pour les chercheurs tant économiquement que médiatiquement (Ruhlen (M.), Cavalli-Sforza (L.-L.) et Renfrew (C.)).

Ainsi, les diverses opinions publiées sur ce forum convergent vers une interrogation unique : comment des journaux scientifiques de vulgarisation peuvent-ils médiatiser, et ainsi faire connaître au grand public des théories erronées voire fausses ? La réponse semble être relativement simple : l'exploitation économique du fantasme des origines pousse les groupes de presse à publier des articles ou des ouvrages qui, non seulement, répondent aux attentes des lecteurs mais qui représentent également un rendement financier non négligeable. Ainsi, les lecteurs, tout comme les groupes de presse (ou d'édition), ont ce qu'ils désirent. Notons que ces remarques sont largement extrapolables au domaine audiovisuel.

3.2.1.3. M. Ruhlen et la dimension du durable

C'est dans ce contexte particulier que s'inscrit la publication anglophone du livre de M. Ruhlen *The Origin of Language. Tracing the Evolution of the Mothertongue* (1994). Il est important de préciser qu'il ne cache nullement l'objet de son livre qu'il considère comme un ouvrage de vulgarisation.

Mais qu'est-ce qui, dans ce livre, est vulgarisé ? Rien moins que les débats qui secouent depuis 1986-1987 (*i.e.* publication de *Language in the Americas* et de « The settlement of the Americas ») la linguistique américaine. Nous avons ici la bible des dernières recherches en linguistique et en génétique des populations, l'archéologie y est également traitée mais avec moins de vigueur que les deux autres disciplines. En réalité, ce qui est vulgarisé dans cet ouvrage, ce sont des éléments de connaissances qui ont atteint le statut de travaux connus et reconnus. Mais il semblerait que ses objectifs soient plus complexes qu'une simple vulgarisation synthétique d'un débat en cours depuis vingt ans car M. Ruhlen, en même temps qu'il diffuse l'information concernant un état de connaissance, rend compte de la dénonciation d'une cabale freinant la diffusion de théories dissidentes auprès des mandarins de la méthode comparatiste traditionnelle. Selon lui, ce frein résulte du rejet de la potentielle existence d'une relation génétique entre les grandes familles de langues. Ce rejet semble être ancré dans les conceptions linguistiques de la majorité des linguistes non qu'ils remettent en cause une éventuelle monogénèse des langues du monde mais ils jugent que la méthode de comparaison multilatérale n'est pas un outil suffisamment fiable pour parvenir à une telle conclusion. Il semblerait, selon M. Ruhlen, que les linguistes soient hermétiques à toute remise en cause de leurs croyances, croyances à la limite du dogme. Cependant, loin de tout dogmatisme, nous avons pu rendre compte avec l'étude critique des travaux de M. Swadesh, de J. Greenberg et de M. Ruhlen que la méthodologie était biaisée et que cette remise en cause se basait sur des éléments de preuves irréfutables.

3.2.2. Les mécanismes d'objectivation de la méthode

Deux mécanismes sont à l'œuvre, dans *L'Origine des Langues* (1997), pour démontrer le bien fondé de la « Nouvelle Synthèse » et notamment celui de la méthode de comparaison multilatérale : M. Ruhlen va, d'une part, faire participer le lecteur à la construction des arbres linguistiques, d'autre part, opérer une décrédibilisation des linguistes.

La méthode proposée dans cet ouvrage est, pour le lecteur néophyte, très habile et difficilement critiquable puisqu'il lui demande de reconstruire lui-même l'arbre linguistique. Le lecteur, tout au long de la démonstration, est impliqué dans le processus, ce qui rend l'objectivation de la méthode et des données plus efficace. Il est toutefois évident que l'adresse au lecteur n'est pas uniquement destinée à montrer la simplicité de la démarche de comparaison multilatérale. Ce que M. Ruhlen cherche à accomplir, c'est de faire entrer le lecteur dans le jeu c'est-à-dire de le prendre comme témoin du bien-fondé de sa démarche. De plus, en considérant le lecteur comme une personne libre de tout préjugé linguistique sur l'origine des langues (puisque en réalité, il n'y connaît rien), l'adoption des théories du proto-monde est grandement facilitée. Cette caractérisation du lecteur opère une critique, implicite mais violente et démagogique, du linguiste professionnel, celui qui ne remet pas en cause les dogmes de la méthode comparatiste indo-européenne et qui est trop corrompu théoriquement pour accepter l'idée d'une monogénèse des langues du monde :

Je m'aperçus que la plupart des linguistes (en fait, pratiquement tous) n'ont pas conscience d'une telle preuve dans la littérature linguistique [preuve d'une monogénèse des langues]. En effet, la plupart des détenteurs d'un doctorat en linguistique croient non seulement qu'elle n'existe pas, mais surtout qu'elle ne peut pas exister, pour des raisons imaginaires que nous examinerons dans cet ouvrage. (Ruhlen (M.). 1994, pp. 9-10).

L'étude d'un corpus, basé sur les opinions des lecteurs sur le site d'une librairie sur ce livre, montre bien que la stratégie argumentative fonctionne parfaitement, voici quelques exemples : « la démarche présentée ici est rigoureuse, et très similaire à celle de Charles Darwin », « un livre incontournable », « This is one of the best guidebook for anyone who wonders where are we from »...

De prime abord, nous pouvons nous demander pourquoi une telle analyse de la structure argumentative de l'ouvrage de M. Ruhlen. La réponse est, en réalité, fort simple. Compte tenu de l'atmosphère actuelle et de la réactualisation de la question de l'origine des langues, il semblait intéressant de chercher à comprendre comment se diffusait le savoir *via* les ouvrages de vulgarisation et comment ces derniers étaient perçus par les lecteurs. En analysant la structure argumentative de l'ouvrage de M. Ruhlen, nous nous sommes rendus compte que tout est fait pour rendre, par des formules de rhétorique, l'esprit du lecteur ouvert à des théories scientifiques très plaisantes pour lui mais peut-être trop simplistes. En faisant croire à des lecteurs, souvent peu conscients des problématiques linguistiques sous-jacentes à l'hypothèse du Proto-

monde, que cette théorie est exacte, c'est faire rentrer la linguistique non pas dans la science mais dans la croyance. Mais bien au-delà de l'anecdote, l'objectif de cette analyse est de faire réfléchir à la manière dont la *Société du Spectacle*, qui est la nôtre aujourd'hui, médiatise certaines théories scientifiques par les revues de vulgarisation scientifique et comment la rhétorique employée parvient à faire pénétrer dans les esprits des théories fausses et aliénantes.

4. Conclusion

L'histoire de la linguistique historique américaine de 1950 à nos jours permet de comprendre le développement intrinsèque de la « Nouvelle Synthèse », d'opérer une synthèse des critiques émises envers tous ses constituants. Ce type d'entreprise épistémologique permet, parallèlement à une meilleure compréhension des phénomènes qui sous-tendent les mécanismes d'émergence des problématiques en sciences humaines, une meilleure concentration des efforts de recherche sur des pôles, des points particuliers n'ayant pas encore été soumis à critique.

En réalité, c'est par une imperceptible mais implacable biologisation des classifications linguistiques ainsi qu'un rôle croissant accordé à la génétique que s'opère l'émergence de la « Nouvelle Synthèse » *i.e.* analogie des modèles, analogie terminologique, mise en symétrie des arbres linguistiques et génétiques... Précisons que l'ultime pierre de l'édifice sera posée par C. Renfrew. Le problème inhérent à la biologisation des classifications linguistiques, c'est que ces dernières ne sont plus en mesure de donner aux langues, donc à l'humain, une dimension temporelle. Ni la génétique ni la linguistique ne permettent, dans ce cas, d'avoir accès à l'histoire de l'Homme. Pour se faire, il faut trouver le chaînon manquant *i.e.* la discipline qui fera le lien entre les gènes et les langues. Cette science, c'est l'archéologie de C. Renfrew (pour une critique des travaux de C. Renfrew au sein de la « Nouvelle Synthèse », *cf.* Métoz (L.). 2005, pp. 529-566). Quoi qu'il en soit, ce phénomène de biologisation vide la linguistique historique de sa substantifique moelle en réduisant la classification linguistique à une expérience de génétique des populations.

Compte tenu de la diversité des disciplines présentes au sein de la « Nouvelle Synthèse », il ne serait nullement maladroit de s'interroger sur la dimension temporelle d'une telle mise en convergence des disciplines. En fait, son élaboration a été rapide puisqu'il n'a fallu que trois ans pour que l'ensemble de l'édifice soit achevé : publication de « *The Settlement of America* » (Greenberg (J.) & al.) en 1986, publication de *Language in the Americas* (J. Greenberg), de *A Guide to World's Languages* (M. Ruhlen) et de *Archaeology and Language. The Puzzle of Indo-European Origins* (C. Renfrew) en 1987 et enfin publication en 1988 de « *Reconstruction of Human Evolution : Bringing Together Genetic, Archaeological and Linguistic Data* » (Cavalli-Sforza (L.-L.) & al.). Nous avons volontairement omis l'ouvrage *The Origin of Language. Tracing the Evolution of the Monthertongue* de M. Ruhlen (1994) puisque ce dernier n'est qu'une résurgence médiatique de la « Nouvelle Synthèse » mais, paradoxalement, le seul moyen d'appréhender la méthode de comparaison multilatérale « à la » M. Ruhlen.

Historiquement, l'existence d'une coopération entre ces différents domaines scientifiques ne constitue pas une révolution conceptuelle, même si de nombreux travaux, comme ceux de Leroi-Gourhan par exemple, en avaient déjà montré les limites. Cependant, là où la « Nouvelle Synthèse » marque une différence nette, c'est dans la systématisation des références extra-disciplinaires. Que ce soit les publications de M. Ruhlen, de L.-L. Cavalli-Sforza ou de C. Renfrew, toutes font intervenir une argumentation extra-disciplinaire basée sur les travaux des deux autres. Cette systématisation amène inévitablement une circularité du raisonnement dans laquelle la critique est un sacrilège (or, nous l'avons vu, des pans entiers de la « Nouvelle Synthèse » ne résistent pas au raisonnement scientifique). La non remise en cause de ces conclusions concourt à faire de la « Nouvelle Synthèse » non pas un cadre théorique et méthodologique strict permettant une convergence objective des résultats mais plutôt le creuset d'une réflexion circulaire frisant la supercherie.

En outre, le rôle majeur accordé actuellement à la génétique dans la mise en évidence d'une monogénèse des langues (et des gènes) doit être fortement réévalué car au sein de la « Nouvelle Synthèse », la génétique se fait le médium d'une supercherie scientifique. La génétique, dont il ne fait aucun doute qu'elle puisse être un support solide de l'argumentation scientifique sur l'origine des hommes et des langues, doit se faire plus humble dans des recherches pluridisciplinaires faisant intervenir en majorité des sciences humaines.

Les recherches pluridisciplinaires telles que la « Nouvelle Synthèse » ne sont pas inintéressantes pour l'avancée de la science. Malheureusement, ce type de recherches doit opérer une re-définition de ses ambitions : la science se donne-t-elle pour but d'entrer dans la *Société du Spectacle* (dont l'image la plus pieuse reste, pour ce qui nous intéresse, la mise en parallèle des arbres génétiques et linguistiques) au risque de produire des raisonnements circulaires vides de sens voire stériles non seulement pour la recherche mais également pour le grand public se rapprochant ainsi plus de la croyance que de la science ou se donne-t-elle pour but de participer à une meilleure connaissance fondamentale et universelle des origines des langues et des hommes ? La marchandisation de la science ne doit pas être un prétexte à l'aliénation du grand public.

Quoi qu'il en soit, il ne fait aucun doute que la « Nouvelle Synthèse » peut éclairer les ténèbres qui entourent nos origines mais il faut pour cela qu'elle accepte un postulat méthodologique fort : la possibilité que les travaux des différentes disciplines puissent se contredire voire s'infirmer. Ce n'est pas dans la convergence basique et superficielle des modèles mais dans la divergence des conclusions que se construisent les théories les plus prometteuses et les plus solides.

Références bibliographiques

- Adelaar (W.). 1989. « A Review of Languages in the Americas by Joseph Greenberg ». in : *Lingua*, 78, pp. 249-255.
- Bateman (R.), Goddard (I.), O'Grady (R.), Funk (V.-A.), Mooi (R.), Kress (J.) & Cannell (P.). 1990. « Speaking of Forked Tongues : The Feasibility of Reconciling Human Phylogeny and the History of Language ». in : *Current Anthropology*, 31, 1, pp. 1-24.
- Ben Hamed (M.). 2004. *Évaluation de la Congruence entre Évolutions Génétique, Morphologique et Linguistique : Modèles et Méthodes*. Paris : Université Paris 6 [Thèse de doctorat].
- Campbell (L.). 1988. « Review of Languages in the Americas, by Joseph Greenberg ». in : *Language*, 64, pp. 591-615.
- Cavalli-Sforza (L.-L.), Piazza (A.), Menozzi (P.), Mountain (J.). 1988. « Reconstruction of Human Evolution : Bringing Together Genetic, Archaeological and Linguistic Data ». in : *Proceedings of the National Academy of Science*, 85, pp. 6002-6006.
- Coseriu (E.). 1965. « Critique de la glottochronologie ». Paris : Klincksieck, 1, pp. 87-96. [Straka (G.), (ed.). *Actes du Xe congrès international de linguistique et de philologie romane*. Strasbourg, 1962].
- Dixon (R.). 1997. *Rise and Fall of Languages*. Cambridge : Cambridge University Press.
- Eldredge (N.), Gould (S.-J.). 1972. « Punctuated equilibria : an alternative to phyletic gradualism ». in : Schopf (T.-J.-M.), (ed.). *Models in Paleobiology*. San Francisco : Freeman Cooper, pp. 82-115.
- Fodor (I.). 1961. « The validity of glottochronology on the basis of the Slavic Languages ». in : *Studia Slavica*, 7, pp. 295-346.
- Fodor (I.). 1982. *A Fallacy of Contemporary Linguistics. J.-H. Greenberg's Classification of the African Languages and his « Comparative method »*. Hamburg : Helmut Buske Verlag.
- Goddard (I.). 1990. « Review of Language in the Americas by Joseph Greenberg ». in : *Linguistics*, 28, pp. 556-558.
- Greenberg (J.). 1979. « The Classification of American Indian Languages ». in : Cooley (R.-E.), Barnes (M.-R.) & Dunn (J.-A.), (eds). *Papers of the 1978 Mid-America Linguistics Conference at Oklahoma*. Norman : University of Oklahoma, pp. 7-22.
- Greenberg (J.), Turner (C.) & Zegura (S.). 1986. « The Settlement of the Americas : A Comparison of the Linguistic, Dental and Genetical Evidences ». in : *Current Anthropology*, 27, 5, pp. 477-497.
- Greenberg (J.). 1987. *Language in the Americas*. Stanford : Stanford University Press.
- Gudschinsky (S.). 1964. « The ABC's of Lexicostatistics (Glottochronology) ». in : Hymes (D.), (ed.). *Language in Culture and Society : A Reader in Linguistics and Anthropology*. New York : Harper & Row, pp. 612-623.
- Hoijer (H.). 1956. « Lexicostatistics : a critique ». in : *Language*, 32, pp. 49-60.
- Hymes (D.). 1960a. « Lexicostatistics so far ». in : *Current anthropology*, 1, 1, pp. 3-44.
- Hymes (D.). 1960b. « More on Lexicostatistics ». in : *Current Anthropology*, 1, 4, pp. 338-345.
- Kimball (G.). 1992. « A critique of Muskogean, « Gulf » and Yukian Material in Language in the Americas ». in : *International Journal of American Linguistics*, 58, 2, pp. 447-501.
- Leroi-Gourhan (A.). 2002. *Milieu et technique*. Paris : Albin Michel [1^{ère} ed. 1945].
- Métoz (L.). 2005. *De la glottochronologie à la « Nouvelle Synthèse »*. *Histoire de la linguistique américaine de 1950 à nos jours : concepts et méthodes*. Lyon : l'École Normale Supérieure [Thèse de doctorat de l'École Normale Supérieure Lettres et Sciences Humaines, 3 volumes, 808 pages].
- Métoz (L.). 2002a. « Mise à l'épreuve des « racines universelles » de M. Ruhlen : analyse méthodologique des équivalences phonétiques et sémantiques dans les reconstructions linguistiques ». in : *Institut de la Communication Parlée*. Grenoble : Université Stendhal, Grenoble 3.
- Métoz (L.) et al. 2002b. « Des formes phonétiques aux proto-formes de la langue originelle : analyse méthodologique et évaluation des limites ». in : *Actes des XXIV^{èmes} Journées d'Étude sur la Parole. LORIA et ATILF*. Nancy, pp. 37-40.
- Métoz (L.). 2001. « L'hypothèse des « racines mondiales » de Merritt Ruhlen : analyse méthodologique et évaluation statistique ». in : *Institut de la Communication Parlée*. Grenoble : Université Stendhal, Grenoble 3.
- Nicolai (R.). 2000. *La traversée de l'empirique. Essai d'épistémologie sur la construction des représentations de l'évolution des langues*. Gap : Ophrys.

- Poser (W.). 1992. « The Salinan and Yurumangui Data in Language in the Americas ». in : *International Journal of American Linguistics*, 58, 2, pp. 202-229.
- Renfrew (C.). 1987. *Archaeology and Language. The Puzzle of Indo-Europeans Origins*. Londres : Jonathan Cape Ltd.
- Ringe (D.). 1992. « On calculating the factor of change in language comparison ». in : *Transactions of the American Philosophical Society*, 82, 1.
- Ringe (D.). 1996. « The Mathematics of « Amerind » ». in : *Diachronica*, 13, 1, pp. 135-154.
- Ruhlen (M.). 1987. *A Guide to World's Languages - Volume I : Classification*. Stanford : Stanford University Press.
- Ruhlen (M.). 1997. *L'origine des langues*. Paris : Belin [trad. fr. de *The Origin of Language. Tracing the Evolution of the Mothertongue*, 1994].
- Swadesh (M.). 1952. « Lexico-statistic dating of prehistoric ethnic contacts ». in : *Proceedings of the Philological Society*, 96, pp. 453-463.
- Swadesh (M.) et. al. 1954. « Symposium : Time Depths of American Linguistic Groupings ». in : *American Anthropologist*, 56, 3, pp. 361-377.
- Swadesh (M.). 1956. « Problems of Long-Range Comparison in Penutian ». in : *Language*, 32, 1, pp. 17-41.
- Swadesh (M.). 1986. *Le langage et la vie humaine*. Paris : Payot [1^{ère} ed. 1967].
- Trumbull (J.-H.). 1891. « On the Best Method of Studying the North American Languages ». in : *Transactions of the American Philological Association*, 1, pp. 55-7.



De l'origine du langage à l'émergence du milieu sémiotique

Par François Rastier

Directeur de recherche au Cnrs,
Paris, France

Mai 2006

« La question de l'origine des langues n'a pas l'importance qu'on lui donne ; elle ne se pose même pas (Question de la source du Rhône : puérile !) ». Ferdinand de Saussure (1997, p. 11).

« C'est une idée très fautive que de croire que le problème de l'origine du langage soit un autre problème que celui de ses transformations ». Ferdinand de Saussure (2002, p. 159).

1. Retour de l'origine

La question de l'origine du langage fait à présent l'objet de programmes de recherche opulents, tant au plan national qu'au plan international¹. Pour situer cette question, il n'est pas inutile de rappeler son histoire.

À partir de la Renaissance, les réflexions sur les origines du langage se sont multipliées, redoublant ou contestant les explications religieuses. Il s'agit d'en finir avec les mythes adamiques. On cherche d'abord les langues les plus anciennes et parées d'une autorité : pour Jan van Gorp, alias Goropius, ce serait le flamand (*Origines Antwerpianae*, 1579) ; pour Georg Stiernhelm (1598-1672), le suédois ; pour Leibnitz (*Brevis designatio de Originibus gentium ductis potissimus ex indicio linguarum*, 1710 p. 4), les langues japhétiques, parmi lesquelles la langue scythe et le celtique dont dérive l'allemand.

Au XVIII^e siècle, la théorie des stades, toujours active dans l'évolutionnisme contemporain, se met en place, avec Vico, Rousseau, Condillac, Herder. Le langage chez les trois premiers auteurs serait issu d'un langage naturel inné qui est le langage d'action. À ces gestes se joindront les onomatopées, puis les premières interjections, prédicats synthétiques, enfin développés en phrases dans un stade suivant. L'onomatopée, partie vocale du langage d'action, se trouve ainsi parée d'un rôle charnière entre communication gestuelle et communication verbale (elle le conserve encore chez Darwin dans *La descendance de l'homme et la sélection sexuelle*, 1871).

Ces théories de l'origine ne font aucune place significative à l'histoire des langues ; ainsi Maupertuis, dans ses *Réflexions philosophiques sur l'origine des langues et de la signification des mots* (1748) écarte-t-il explicitement la comparaison des langues. C'est alors l'étude *a priori* des principes de la connaissance (cf. Bergougnieux, 2005, p. 21), c'est-à-dire une forme classique du cognitivisme qui l'emporte.

Au siècle suivant, alors que s'est formée la linguistique historique et comparée, Pictet crée la *paléontologie linguistique* pour retrouver la langue des anciens Aryas² ; mais comme son nom l'indique assez, elle passe bien vite de l'histoire conjecturale des reconstructions aux spéculations sur l'origine du langage, comme en témoigne la paléontologie linguistique française, développée par des auteurs comme Abel Hovelacque et Honoré Chavée. Le modèle des sciences de la vie était alors prégnant, comme il le redevient aujourd'hui. Le darwinisme linguistique de Schleicher considère les langues comme des espèces vivantes et leur histoire sur le mode de l'évolution biologique. La typologie des langues est alors étroitement liée à la typologie des races, et les affirmations de Haeckel, fondateur du racisme « scientifique », sont au centre des débats qui portent sur les regroupements des langues selon les races (cf. Hovelacque, 1876, p. 371).

¹ Voir par exemple le programme *The Origin of Men, Language and Languages*, de la Fondation européenne pour la science.

² Les langues aryennes sont aussi dites japhétiques, et en 1958 Bakhtine parlait encore de Marr comme d'un « remarquable érudit », fondateur de la « paléontologie linguistique ».

Par ailleurs, selon la thèse de Haeckel que l'ontogenèse récapitule la phylogenèse, le développement des enfants devient une référence majeure. Ainsi, dans la seconde édition (1876) de *De l'intelligence* (1870), Taine ajoute-t-il une section annexe intitulée : « De l'acquisition du langage chez les enfants et dans l'espèce humaine »¹.

Le XX^e siècle ne marque aucune rupture significative. Si Jespersen publie en 1922 *Language, its Nature, Development and Origin*, il réfute les quatre principales théories de l'origine, dans l'onomatopée, l'interjection, la résonance et les cris accompagnant l'effort physique, il maintient des thèmes fréquents au XVIII^e siècle, comme celui d'une langue « chantonnée », celui des unités synthétiques (rôle que jouait l'interjection), enfin celui, classiquement empiriste, de la prééminence du lexique concret.

Dans la seconde moitié du siècle, la découverte de nouveaux fossiles d'hominidés et l'essor de la génétique vont coïncider avec un nouvel essor des linguistiques universelles pour appuyer des programmes de naturalisation dans lesquels la question de l'origine du langage pourra devenir un enjeu majeur pour la réduction néo-darwinienne des cultures.

On cherchera à redéfinir la linguistique et son objet en fonction du développement des connaissances en génétique et en paléo-anthropologie. Le marrisme et le cognitivisme d'inspiration chomskienne se sont diversement partagé la question de l'origine du langage en imaginant des parallèles entre génétique et linguistique. La découverte de l'ADN a naturellement nourri un grand nombre de spéculations sur le « code génétique », que sa désignation même, par une métaphore exorbitante, invitait à comparer à un langage, dès lors que l'on réduit le langage à un code. Ainsi Roman Jakobson (dans son débat en 1968 avec François Jacob et Claude Lévi-Strauss) a-t-il comparé le code génétique au code linguistique, en supposant que le langage se fondait dans sa phylogénèse sur un isomorphisme structural avec le code génétique. Cette analogie est reprise à présent par des théoriciens néo-marristes : « La variété infinie de tout le vivant peut se ramener aux très longs « messages » génétiques [...] possédant les traits frappants d'une ressemblance structurale avec le code linguistique » (2005, p. 140). Or Marr ramenait la diversité des langues historiques à quatre syllabes primitives dénuées de sens que Gamkrelize (2005) compare aux quatre molécules de l'ADN. Cette position n'est pas isolée et Ángel López-García (2005) entend démontrer que le code linguistique émerge du code génétique par « duplication ». Le génome ayant pris la place de la Providence comme puissance explicative, cette analogie entre les deux « codes » inverse la détermination mystique qui faisait de la structure du langage divin le modèle de toute chose. Elle retrouve à sa manière la théorie des correspondances de la *magie naturelle*, florissante à la fin du XVI^e, et dont on reconnaît des traits dans le *New Age* actuel².

1.1. Le langage est-il une faculté ?

Expliquer une action par une faculté supposée reste une facilité récurrente : on explique le récit par la faculté narrative, comme naguère le mythe par la faculté mythopoétique (cf. Hocart, 2005, p. 150). Voulant trouver des explications causales, la pensée scolastique rendait compte systématiquement de l'acte par la puissance : par exemple la pensée trouvait sa cause dans l'intelligence définie comme faculté de penser. Reformulant la métapsychologie antique, issue pour l'essentiel du *De Anima* d'Aristote, la philosophie médiévale a certes distingué des facultés, mais pour réaffirmer l'unité de l'âme ; il en va ainsi de Guillaume d'Auvergne à Descartes et à Bossuet. Ce n'est qu'au XVIII^e siècle que l'unité de l'âme commence à être mise en doute : en la ramenant aux principes de la nature, on transpose les facultés psychiques en fonctions organiques. Or, comme dès lors que la fonction crée l'organe, les fonctions mentales doivent être rapportées à des localisations cérébrales. C'est là le principe de la théorie de Gall, fondateur de la phrénologie. Avec les programmes cognitifs de naturalisation de l'esprit, ces théories ont repris de la vigueur.

Naguère, Jerry Fodor, ouvrant par un éloge de Gall son ouvrage *The Modularity of Mind*³, décrivait un esprit morcelé en modules correspondant à autant de *devices* anatomiques. Si

¹ Nous verrons que cette thèse, infirmée par le développement de la biologie, se trouve maintenue par des linguistes néo-darwiniens.

² Nous rencontrerons plus loin une autre correspondance, entre l'ontogenèse et la phylogenèse, l'enfant et les premiers temps de l'humanité, qui demeure récurrente chez les darwiniens classiques (Haeckel), comme chez les néo-darwiniens d'aujourd'hui (Bickerton, Hombert).

³ Tr. fr. *La modularité de l'esprit*. Paris : Éditions de Minuit, 1986.

donc l'homme est doué de langage, cette fonction se concrétise dans un organe du langage, expression d'un équipement génétique particulier. Avec le retour des linguistiques universelles, des théories influentes ont ainsi fait du langage une fonction organique et de la linguistique une « neuroscience cognitive » (Jackendoff). On revient de la linguistique comparée à la spéculation sur l'origine, de la critique historique à une métaphysique « positive », naturalisante et néo-darwinienne. Ses thèses sont claires. La faculté de langage résulte de l'évolution biologique. L'évolution étant réduite — on ne sait pourquoi — à la sélection naturelle, sans considération pour d'autres facteurs comme la dérive génétique, le langage résulte d'une adaptation.

1.2. Le langage a-t-il une fonction ?

Certains animaux n'ont pas de langage inné : c'est le cas par exemple de l'homme ou du perroquet gris du Gabon. Ces deux espèces fort sociales sont toutefois douées pour l'imitation, deux qualités complémentaires pour ces animaux supérieurs intarissables dont l'intelligence pragmatique se traduit par un sens de l'à-propos et de la pertinence que l'on juge remarquables¹.

Le fait que l'homme n'ait pas de langage inné ne suffit pas à infirmer l'hypothèse d'un organe du langage : on suppose simplement que cet organe sert à apprendre (*Language acquisition device*, ou LAD)².

Pour justifier ce substrat organique, on suppose que la fonction crée l'organe, ou, en termes néo-darwiniens, qu'elle procède d'un avantage adaptatif. Il faut donc savoir à quoi sert le langage.

Rousseau avait une réponse galante et sensible, car il faisait des passions le moteur de l'histoire humaine : le langage servirait d'abord à faire sa cour, né au bord des fontaines où bergers et bergères menaient leur troupeau³. Les néo-darwiniens d'aujourd'hui ont transposé cette thèse sur le plan de la reproduction : une fonction biologique contribue par définition au succès reproductif de l'organisme qui en est pourvu (*cf.* Sperber et Origgi, 2005, p. 285). Il reste bien entendu à rendre compte de l'avantage reproductif que constitue l'exercice de la fonction de langage.

Les passions sont devenues des instincts, mais la légitimation fonctionnelle demeure. On peut lire ainsi de petits romans anthropologiques⁴ qui expliquent l'origine du langage en termes d'avantages adaptatifs : le langage servirait à assurer la paix dans de meurtrières hordes ancestrales (Victorri), ou encore à médire des absents tout en favorisant l'unité du groupe (*grooming and gossip hypothesis* de Robin Dunbar), à assurer un prestige social (Jean-Louis Dessales), etc. Quand on connaît un peu leurs auteurs, on s'aperçoit que ces apologues plaisants sont d'excellents tests projectifs. Ils reposent sur l'adage qui fait de l'avantage adaptatif un avatar de la Providence divine ; et surtout ils concrétisent des hypothèses dont l'avantage, somme toute inestimable, est de ne pouvoir être ni infirmées, ni confirmées.

À la différence de la mystérieuse émergence qui faisait du code linguistique un analogue dérivé du code génétique, la voie naturalisante adoptée par le cognitivisme entend réduire le linguistique au biologique par deux voies principales : en faisant de la Grammaire Universelle (désormais GU) une « composante hypothétique du patrimoine génétique », selon les mots mêmes de Chomsky (1984) ; en postulant un organe du langage ou Language Acquisition Device.

Examinons les arguments qui fondent cette seconde voie de naturalisation du langage. La démonstration emprunte deux directions : (i) La recherche d'un gène du langage qui serait altéré dans des familles souffrant d'aphasie. (ii) La recherche de l'organe du langage (iii) La postulation d'un protolangage de *Homo Erectus* (Bickerton) concrétisant la GU.

- a. *Le gène du langage.* Si l'on part du principe que le langage est une fonction biologique, on peut chercher son substrat anatomique dans un organe, et, au-delà, dans les gènes qui commandent le développement de cet organe. Ainsi on a vu le gène du langage dans FOXP2, situé sur le chromosome 7, supposé responsable de l'aphasie héréditaire

¹ Le parallèle avec l'homme s'arrête là, car le perroquet gris du Gabon n'est sans doute pas saussurien, et l'on peut supposer qu'il traite les symboles comme de simples signaux, alors que nous faisons l'inverse. Alex, perroquet gris du Gabon (*Psittacus erithracus*), éduqué depuis trente ans par Irene Pepperberg (MIT), maîtriserait quatre-vingt-dix mots, et compterait jusqu'à six.

² *Device* est un terme de mécanique qui signifie appareil ou engin. L'usage de ce terme introduit une téléologie mécanique dans la téléonomie du vivant.

³ *Essai sur l'origine des langues* (1817). Paris : Belin, ch. X, p. 525. Il s'agit ici de l'origine du langage dans les pays du sud. Dans ceux du nord, le langage sert à appeler du secours (p. 527).

⁴ Je reprends cette expression à André Pichot.

qui frappe plusieurs membres d'une famille anglaise (la célèbre famille désignée par les initiales KE). Il s'agit d'un gène de transcription qui régit la formation d'une protéine composée de 751 acides aminés¹.

Or, le gène FOXP2 a été associé au chant du diamant mandarin et du canari, et pourrait tout simplement avoir une incidence sur la motricité vocale : ainsi les mutations provoquées de ce gène affectent les performances vocales des souriceaux. On doit donc chercher ailleurs.

b. *L'organe du langage.* Selon l'hypothèse, propagée par le néo-darwinisme, mais largement reprise hors de ses rangs, rien n'apparaît que par avantage évolutif ; donc le langage présente des avantages qui sont la cause de son apparition. Et comme toute fonction repose sur une faculté et donc sur un organe — le problème de l'origine du langage devient celui de l'origine de l'organe du langage.

Il n'est pas nécessaire qu'une faculté soit liée à un organe, et comme le soulignait Rousseau « l'art de communiquer les idées dépend moins des organes qui servent à cette communication, que d'une faculté propre à l'homme, qui lui fait employer ses organes à cet usage » (1817, p. 504).

Prenons l'exemple des calculateurs prodiges. On ignore si l'extraction des racines carrées de très grands nombres résulte d'une fonction adaptative aux bénéfices indéterminés. Du moins, l'imagerie cérébrale a permis de montrer que les calculateurs prodiges stockent les résultats intermédiaires dans l'aire de la prosopagnosie (reconnaissance des visages) : en somme, c'est moins le substrat anatomique que l'utilisation qu'on en fait qui compte.

Pourquoi faudrait-il malgré tout que les organes soient dédiés à des fonctions prédéfinies ? C'est faux pour le cerveau comme pour la main : le cerveau sert-il à élaborer des grammaires universelles ou des théories néo-darwiniennes ? La main à les écrire ? Il se pourrait bien que le langage humain soit issu de la rencontre contingente d'un appareil phonatoire (assez ordinaire chez les primates), d'un cortex préfrontal exceptionnellement développé et capable d'imaginer des objets en leur absence, enfin d'interactions sociales complexes.

En outre, si la faculté de langage est une faculté d'apprendre les langues (*language acquisition*), son exercice présuppose l'existence des langues ; en ce cas, la nature humaine présuppose la culture, ce qui va à l'encontre de l'hypothèse initiale de la naturalisation.

c. *Le protolangage.* Alors que Haeckel estimait que le pithécantrophe était muet, l'*Homo Erectus*, son moderne successeur, devient parlant grâce à Bickerton (1990, 1996), qui lui confère un protolangage. Dans l'hypothèse d'un déterminisme génétique, le langage chez *Sapiens* doit être issu d'une mutation d'*Erectus*.

Essentiellement composé de mots ostensifs, le protolangage est dépourvu de syntaxe. Au plan linguistique, l'histoire de l'hominisation, le passage d'*Erectus* à *Sapiens*, consistera essentiellement dans le développement de la syntaxe.

De nombreux auteurs, à la suite de Derek Bickerton, font dériver les langues d'un protolangage axé sur la communication pragmatique liée à la situation (le « ici » et le « maintenant »)². Mais un tel langage n'aurait rien de fondamentalement différent d'un langage animal : Bickerton s'appuie d'ailleurs explicitement sur le « langage » de Kanzi, bonobo inlassablement conditionné qui parvint à répondre à des questions de sa mère humaine adoptive, Sue Savage-Rumbaugh, en frappant des icônes sur un clavier.

Ce langage aurait été fait d'un lexique enrichi progressivement, mais dépourvu de syntaxe : on retrouve là un retour de la théorie de la langue adamique, faite uniquement de termes. La séparation entre lexique et syntaxe procède de la distinction platonicienne entre *onomazein* et *legein*, le nommer et le dire. La syntaxe, dans l'école chomskienne, vient après, car elle couronne l'histoire de l'hominisation.

¹ Or cette molécule diffère de trois acides aminés seulement de celle que l'on trouve chez les mammifères autres que les primates.

² Le singe, l'enfant et le « sauvage » voisinent encore bizarrement, comme au temps des anthropologies philosophiques des Lumières. Cependant, rien ne permet d'affirmer que les langues auraient formé d'abord leur lexique puis leur syntaxe : même si l'on en fait un moyen de catégorisation, les catégories syntaxiques n'ont pas moins d'importance que les catégories lexicales, car elles déterminent ou reflètent les grandes classifications *a priori* qui structurent le lexique, voire l'expérience des sujets parlants (ainsi animé vs inanimé, actif vs passif, etc.).

On s'étonne à peine que Bickerton décrive, de façon révélatrice, le protolangage comme proche de celui de Kanzi, bonobo célèbre, mais aussi de celui des jeunes enfants, et des pidgins. Victorri le décrit didactiquement comme un « langage de Tarzan », laissant transparaître le substrat mythique de cet idiome¹.

Cependant, rien ne permet d'affirmer que les langues auraient formé d'abord leur lexique puis leur syntaxe : même si l'on en fait un moyen de catégorisation, les catégories syntaxiques n'ont pas moins d'importance que les catégories lexicales, car elles déterminent ou reflètent les grandes classifications *a priori* qui structurent le lexique, voire l'expérience des sujets parlants (ainsi animé vs inanimé, actif vs passif, etc.).

Si le protolangage est fait de mots isolés — à l'exemple des signes qu'emploient les chimpanzés conditionnés — comment passer à la prédication ? La solution qu'imagine Sperber est particulièrement révélatrice : à la suite d'une mutation, certains auraient développé des inférences pragmatiques à partir de ces mots isolés, par exemple auraient inféré *il boit* à partir de *eau* (cf. Sperber & Origi, 2005, p. 252). On retrouve une démarche analogue à celle de Destutt de Tracy, quand il explique, dans ses *Éléments d'idéologie* (1805-1817), comment les premiers mots, des interjections, se sont développés en prédications, *Ouf !* devenant par exemple *J'étouffe !*

L'émergence du langage à partir du protolangage serait accessible de trois manières : en observant : (i) le développement de l'enfant ; (ii) l'élaboration de langages de signes par de petites communautés de sourds isolés dont le langage gestuel spontané témoignerait du caractère naturel de la Grammaire Universelle ; (iii) en observant l'évolution des créoles à partir des pidgins.

a. *Le développement de l'enfant.* Selon Haeckel, l'ontogenèse récapitulait la phylogenèse. Cette thèse simpliste a été abandonnée depuis un siècle par les sciences de la vie. Elle garde cependant toute sa séduction mythique : la préhistoire n'est-elle pas l'enfance de l'humanité ? Aussi quand des enfants juxtaposent leurs deux premiers mots, ils recréeraient le protolangage. Cette hypothèse est gratuite : *Erectus* n'était pas dans la situation de mal parler ou de ne pas encore maîtriser le langage de *Sapiens*². N'importe, des spécialistes du développement persistent à soutenir que « les cinquante premiers mots émis par le bébé ainsi que les premières combinaisons qu'il réalise sont porteurs du même sens dans toutes les langues du monde » (Jisa, 2005, p. 256). Voilà donc cinquante primitives universelles qui auront échappé jusqu'ici à la linguistique comparée.

Ce type d'affirmation suffit à Hombert pour affirmer que « la collaboration entre les tenants de cette discipline [les spécialistes de l'acquisition du langage chez l'enfant] et les spécialistes de l'apparition du langage met à mal l'ancienne dichotomie ontogenèse/phylogenèse » (2005, p. 464). Comment les millions d'années de la phylogenèse et les mois d'apprentissage se répliqueraient-ils, sauf à sortir, non seulement de l'histoire, mais de toute échelle temporelle ?

b. *Les sourds.* Que le langage humain soit dérivé d'un langage gestuel, c'est un thème que l'on retrouve fréquemment au XVIII^e siècle, de Condillac à Destutt de Tracy ; cette hypothèse était encore soutenue par Marr. De nos jours, la communication spontanée des sourds est supposée révéler le passage du protolangage au langage.

Les groupes privés d'accès au langage parlé élaborent des langages gestuels. Depuis 1977, des sourds regroupés dans une école de Managua sont étudiés par des cognitivistes nord-américains qui étudient l'évolution de leur langage, dans des conditions d'isolation. Ils ont affirmé que les productions de la première génération étaient un protolangage, développé en langage dans les « générations » suivantes.

Même si l'isolement imposé interdit aux cobayes l'intégration dans la société nicaraguayenne, les jeunes sourds n'entrent évidemment pas dans l'école à leur naissance. En outre, le fait qu'ils utilisent certains signes de l'*American Sign Language* (cf. Cuxac, 2005, p. 206-207) laisse deviner qu'ils n'ont pas été sans relation avec le monde extérieur, ne serait-ce qu'avec les enquêteurs américains. N'importe, Steven Pinker saluait la découverte du protolangage chez les sourds-muets de Managua sous le titre quelque peu emphatique « A linguistic

¹ L'illustration de son étude reproduit opportunément une photo où Tarzan converse avec Cheeta (2005, p. 222).

² Par un étrange renversement de situation, *Erectus* se trouve dans la situation de l'enfant.

Big Bang » (*New York Times*, 24 octobre 1999) en le considérant comme un langage *ex nihilo* (*created out of thin air*). On retrouve le mythe de l'isolat : plusieurs fois dans l'histoire, des souverains curieux, du pharaon Psammétique I^{er} à Frédéric II de Hohenstaufen, jusqu'à Akbar le Grand, ont fait enfermer des nouveaux-nés pour savoir s'ils parleraient spontanément et en quelle langue. Là, on a simplement isolé des sourds dans leur institution¹.

On sait cependant que les sourds-muets créent ou adaptent des signes gestuels à proportion de leur degré d'intégration à la société (des répertoires personnels allant jusqu'à 1 500 signes ont été dénombrés, *cf.* Yau, 1992). La notion d'isolement reste illusoire. Les sourds-muets vivent dans un monde d'échanges linguistiques et gestuels déjà codés. Ils entrent dans un monde sémiotique riche qu'ils enrichissent encore. Par exemple, comme le remarque Descola (1992), les Achuar sourds-muets connaissent la mythologie des Achuar : ils ont constitué des langues signées dont la capacité expressive est comparable à celle du jivaro et permettant de communiquer aussi bien sur le *hic* et *nunc* que sur ce qui n'est pas là². La création des langues de signes apparaît alors comme un cas particulier, certes éminent, de la sémiogenèse.

- c. *Les créoles*. Selon Bickerton, certains contacts entre populations de langues différentes recréent spontanément le protolangage dans les pidgins, qui évoluent en langues dans les créoles. Comme il attribue le protolangage à *Homo Erectus*, prédécesseur de *Sapiens*, ne reprend-il pas obliquement la thèse de l'archaïsme des « nègres » ? Du moins, inexplicablement, ne considère-t-il que les créoles de plantation et non les autres. Le protolangage et les pidgins sont explicitement comparés, mais ce serait persifler que de noter l'analogie entre le « langage Tarzan » ou protolangage et cet idiome ineffable de l'imaginaire colonial, le *petit-nègre*. À défaut de fondement scientifique, il n'est guère douteux que le protolangage ait un riche substrat imaginaire, comme d'ailleurs l'ensemble de la thématique de l'origine du langage³.

Avec les enfants isolés, les sourds, les créoles, on retrouve implicitement les trois références majeures de l'anthropologie positiviste du XIX^e siècle : l'enfant, le déficient, le sauvage. D'où la fascination continue à l'égard de ceux qui conjoignent deux ou trois de ces traits, comme jadis Victor, le Sauvage de l'Aveyron, ou Caspar Hauser, voire les enfants sourds de l'Abbé de l'Épée, qui voulait faire du langage gestuel la langue universelle de l'avenir.

1.3. Du protolangage à la langue universelle de l'avenir

On stigmatise rituellement la « curieuse censure » du règlement intérieur de la Société Linguistique de Paris qui excluait les communications sur l'origine du langage. Relisons l'article 2 des statuts : « La société n'admet aucune communication concernant, soit l'origine du langage, soit la création d'une langue universelle » (1866, révisé en 1876). Ces deux aspects sont liés, car la langue originelle est de fait universelle, comme le sera la langue parfaite de l'avenir. Cet article reflète ainsi une réflexion épistémologique assurée : la linguistique est une science descriptive et historique, qui ne se mêle pas d'imaginer des langues, que ce soient celles de premiers ou des derniers hommes. Une science définit son objet en récusant les faux problèmes dont elle se prive, et cette privation fondatrice la distingue décisivement de la métaphysique.

L'origine, comme l'avenir, échappe à l'histoire, qui, se confondant avec le monde humain, n'a ni commencement ni fin. Les théories de l'origine comme celle de la fin de l'histoire sont également métaphysiques, car elles adoptent nécessairement un point de vue externe au monde humain. Ce point de vue transcendant ne peut être scientifique, sauf à considérer que les langues ne soient pas des formations historiques, et à renvoyer le temps humain de l'histoire au temps biologique de l'évolution.

On n'a pas assez remarqué que les théoriciens de l'origine du langage sont souvent aussi des théoriciens de la langue universelle : l'en-deça et l'au-delà des langues communiquent

¹ On étudie à présent une tribu de bédouins du Neguev qui compte trois générations de sourds : des chercheurs comme Carol Paden et Wendy Sandler interprètent également leur cas dans le paradigme de l'origine.

² Dans ces petits groupes dispersés et semi-nomades, et qui ne comptent pas toujours des sourds, on peut écarter l'hypothèse que ces langues aient été transmises de père en fils sur de longues durées.

³ L'iconographie de Hombert, éd. (2005) est particulièrement révélatrice : chimpanzés et autres singes (pp. 55, 75, 85, 93, 99, 114, 241) ; communications homme-singe (pp. 77, 78, 218, 222) ; gens du tiers-monde, généralement nus (pp. 119, 153, 213, 224, 253, 309, 326, 329, 378, 413, 421, 454, 457) ; dont femmes dépoitraillées, parfois allaitantes (pp. 63, 122, 169, 180, 332, 339, 403) ; nette prévalence, comme jadis, des africains.

aisément. L'uchronie du langage naissant et l'utopie de la langue unique se complètent dans l'idée d'un langage sans langues que Sériot rapproche justement du langage angélique.

Marr étudia la langue originelle japhétique (c'est-à-dire aryenne) et la glottogonie (origine du langage), mais il fut aussi un théoricien de la langue universelle future. Jespersen a promu la supériorité de l'anglais comme langue universelle : c'est précisément ce que Ogden mettra en œuvre en créant le *Basic English*, langue internationale universelle. Dan Sperber a théorisé aussi bien sur l'origine que sur l'évolution future du langage et la disparition de l'écriture (cf. 2001).

L'archaïsme de l'origine et le futurisme du progrès vont dans le même sens, celui d'une sortie de l'histoire et d'une entrée dans le mythe : la pré-humanité et la post-humanité y trouvent des traits communs (cf. l'auteur, 2004). Du moins le protolangage et le *Basic English* se ressemblent-ils : d'une part les exemples de protolangage que donne Bickerton sont naturellement en anglais, mais surtout il s'agit de deux langues à vocabulaire restreint et que leur syntaxe élémentaire réduit à communiquer dans le *hic et nunc* – comme les systèmes de communication animale. Plus qu'à une inobservable langue en voie d'apparition, le protolangage semble une langue restreinte, en voie de disparition : Christos Clairis a ainsi décrit l'appauvrissement syntaxique et morphologique de langues fuégiennes en cours d'extinction (2005, ch. 4).

Au moment même où naissaient les sciences historiques, Hegel prédisait la fin de l'Histoire et la voyait dans une animalité retrouvée. La sortie du temps historique, le fait que les analyses néo-darwiniennes juxtaposent des fossiles, des singes, des nourrissons et des sourds, sans jamais se référer à l'histoire des langues attestées, tout cela confirme que le programme de naturalisation se situe dans un temps sans échelle ni métrique, une *uchronie* – notion proposée par Thomas More dans son *Utopie*, en même temps qu'une langue universelle.

Mythique, le thème de l'origine peut être exploité en ce sens : une « discipline » qui met à égalité Humboldt et Burgess (auteur des dialogues de *La guerre du feu*),¹ ou qui recrée le langage de Tarzan répond certes aux attentes du grand public et remplit une fonction idéologico-politique². Si traditionnellement le matérialisme éliminateur a une fonction polémique, on en vient à souhaiter qu'il ne s'agisse pas de substituer un dogmatisme à un autre, tout en demeurant dans le mythe où Lucy aurait simplement assumé le rôle d'Ève.

Du moins les comparaisons récurrentes (Ève africaine) et le monogénéisme maintenu favorisent une lecture adamique de l'histoire humaine, alors que les hypothèses polygénétiques et diffusionnistes (Weidenreich, Caspari et Wolpoff) restent à peu près absentes de la littérature néo-darwinienne.

La linguistique a pour objet les langues dans leur diversité, alors que le langage est demeuré une abstraction philosophique. Pourquoi le langage supplanterait-il les langues dans la réflexion des linguistes ? À l'époque de la mondialisation, il n'est pas impossible que les langues dans leur immense diversité soient abandonnées³ : le protolangage unique pour toute l'humanité et la langue universelle de l'avenir procèdent du même imaginaire simplicateur. Par ailleurs, il reste plus facile de pratiquer des prises de sang en Afrique que de décrire les langues africaines, de remplacer l'anthropologie culturelle par l'anthropologie physique, ce qui est déjà une forme pratique de naturalisation.

On a oublié que dans les années 1930 des sociétés savantes allemandes procédaient ainsi, alors que Gottfried Benn s'écriait : « Jusqu'à il y a peu, l'homme était un être de raison [...], aujourd'hui il est un être métaphysique, dépendant et environné d'origine et de nature. Autrefois son interprétation de l'histoire était le progrès dans un sens civilisateur, aujourd'hui elle est la liaison avec l'antérieur comme continuité mythique et raciale »⁴.

¹ Dans la somme dirigée par Hombert, ils comptent du moins le même nombre de mentions. Ayant été amené, avec une collègue, à créer des langues préhistoriques pour un film, j'ai eu toutes les peines du monde à convaincre les journalistes scientifiques que ce travail « poétique » n'avait rien de scientifique (Ainsi parlait Rahan – Entretien, avec la collaboration de Rossitza Milenkova-Kyheng, *Sciences humaines* (2005), n° 158, p. 14).

² Par exemple, alors que personne dans le champ du débat ne soutient depuis deux siècles l'origine divine du langage, pourquoi souligner que la recherche sur l'origine s'oppose aux croyances religieuses ?

³ David Rothkopf, 1997. « English is linking the World ». in *Foreign Policy*.

⁴ In Betz, Albrecht, Die « Worthelfer der Gewalt » in der Dritten Walpurgisnacht, in Gilbert Krebs & Gérard Stieg, éd. Dir. *Karl Kraus et son temps*, Publications de l'Institut d'Allemand, Asnières, 1989, p. 159.

2. Hypothèses

2.1. Peut-on parler d'une « nouvelle synthèse » ?

La préhistoire a beaucoup changé, naguère, quand trois disciplines restées jusque-là isolées, la linguistique, la génétique des populations et l'archéologie, se sont si bien rencontrées que plusieurs auteurs, non sans enthousiasme, ont évoqué une *nouvelle synthèse*¹.

Darwin écrivait en 1859 : « Si nous possédions un arbre généalogique parfait de l'humanité, un arrangement généalogique des langues humaines fournirait la meilleure classification des diverses langues parlées aujourd'hui à travers le monde »². Des rencontres récentes ont renouvelé l'espoir de répondre à ses vœux.

- a. Dans le domaine de la *typologie linguistique*, la convergence des travaux de Greenberg et Ruhlen sur l'eurasiatique et des chercheurs russes sur le nostratique a d'abord permis³ de ramener, selon une hypothèse forte, le nombre des familles linguistiques à 19, puis à 12. En outre, un renouveau de la glottochronologie a réactivé la thèse du *monogénéisme linguistique*, formulée au début du XX^e siècle par Trombetti⁴, Sapir, puis son élève Swadesh. Les travaux discutés de Ruhlen tendent à l'accréditer.
- b. D'autre part, on a relevé des corrélations entre génétique des populations et typologie linguistique. Greenberg, selon des critères linguistiques (1987) a regroupé les deux cents familles amérindiennes en trois seulement, obtenant un résultat comparable aux classements génétiques de Cavalli-Sforza (1988, publiés en 1994). Cette corrélation a été confirmée également, au niveau anatomique, par l'étude des dentitions des indiens par Turner (1989 ; cf. Ruhlen, p. 94). Dans le domaine européen, après les travaux de Ammermann et Cavalli-Sforza, en 1978, sur le dégradé génétique entre Anatolie et Europe de l'Ouest, l'hypothèse de leur corrélation avec les langues indo-européennes a été proposée séparément, en 1987, par Colin Renfrew et Aron Dolgopolski. L'étude de ces corrélations entre gènes et langues a été étendue en Europe par Robert Sokal et ses collègues, en tenant compte de facteurs tant historiques que géographiques⁵.

Enfin, dans le domaine africain, les travaux de Langaney sur les langues des groupes khoisan et pygmée ont aussi confirmé l'hypothèse initiale. D'importantes nuances doivent être apportées : les marqueurs génétiques peuvent être hétérogènes, et par exemple, les Damas (Afrique du Sud) ont des marqueurs féminins (ADN mitochondrial) khoisan et des marqueurs génétiques masculins (chromosome Y) nigéro-congolais : or, ils parlent une langue khoisan. En revanche, les Hereros ont des marqueurs masculins khoisan, des marqueurs féminins nigéro-congolais et ils parlent une langue nigéro-congolaise. Enfin, on ne trouve pas de spécificité génétique des Basques, alors que leur langue fait exception.

Tous ces travaux sont débattus : même un accord sur l'hypothèse d'un arbre évolutif unique ne pourrait pas faire oublier les caractères historiques et interculturels de l'évolution. Bien que les scénarios monogénétiques soient traditionnellement privilégiés, car ils se superposent mieux aux croyances religieuses qu'ils entendent périmer, les hypothèses diffusionnistes et aréales ne sont pas pour autant infirmées (cf. notamment Wolpoff et Caspari, 1997).

Par ailleurs, des corrélations locales entre types linguistiques et types génétiques ne permettent pas par elles-mêmes d'établir une détermination du génétique au linguistique : dans des conditions de relatif isolement, comme le « patrimoine sémiotique » se transmet parallèlement au patrimoine génétique⁶, les deux types de spécificités peuvent voisiner sans relation causale et de manière contingente. La corrélation entre gènes et langues plaide moins pour la nature que pour la culture, car on se marie de préférence dans les groupes qui parlent

¹ L'expression fait allusion à la synthèse du darwinisme et de la génétique dans les années 1940, qui a assuré l'essor du néo-darwinisme contemporain.

² *L'Origine des espèces*, d'après Ruhlen, 1997, p. 176.

³ Cette convergence est un des effets de la *glasnost* au milieu des années 80.

⁴ Trombetti, *L'unità d'origine del linguaggio*, Bologne, 1905. Saussure n'écarte pas ses hypothèses, mais les estime invérifiables.

⁵ Ces développements de la linguistique historique et comparée marquent la fin du splendide isolement de la grande famille indo-européenne — l'allemand et l'hébreu sont apparentés — sinon du mythe indo-européen auquel elle avait donné naissance ou qui du moins s'appuyait sur elle. C'est aussi la fin du mythe guerrier : les indo-européens ne sont plus le peuple nomade et guerrier des kourganes (selon l'hypothèse de Gordon Childe reprise par Marija Gimbutas), mais des agriculteurs anatoliens que l'on dit paisibles.

⁶ Cf. l'auteur, 1991, épilogue.

nature que pour la culture, car on se marie de préférence dans les groupes qui parlent la même langue – ne serait-ce que pour conclure le mariage, si bien que dans l'espèce humaine, la diffusion des gènes obéit majoritairement à des contraintes culturelles¹.

2.2. Des langues sans origine ?

Pas plus que l'origine de l'univers n'est une question qui appartient à la chimie, l'origine du langage ne relève proprement de la linguistique. Comme le remarque justement Rossitza Kyheng (2006), « savoir que le langage est une généralisation des langues, et non pas une faculté, renvoie d'emblée la question des origines du langage vers d'autres disciplines : savoir quand, où et comment l'homme a acquis la faculté de parler est un problème anthropologique complètement externe à l'objet de la linguistique. Du point de vue linguistique, le langage est un *phénomène social* et Saussure précise, à juste titre, que ce phénomène existe « depuis le premier jour même où une société humaine a parlé » (ELG, p. 163) ».

À propos de l'origine du langage, le matérialisme antique (Lucrèce, Vitruve) insistait sur la nécessité externe et sur les fonctions internes. Avec les Lumières, la tradition du matérialisme antique l'a si bien emporté que les débats d'alors nous passionnent encore et les auteurs en vue reprennent les arguments traditionnels : par exemple, les sociétés humaines ont besoin de coalitions pour assurer leur unité, et c'est le langage qui le permet. Toutefois, les coalitions solides que l'on voit chez les félins et les canins, les politiques de *lobbying* chez les primates, montrent que les langues articulées ne sont aucunement indispensables pour assurer la concorde dans le groupe social ; par ailleurs, les mammifères pratiquent très rarement la mise à mort du congénère, et, nous le constatons chaque jour, le langage sert aussi à appeler au meurtre.

Si nous n'avons pas accès aux origines ni aux débuts des langues, nous pouvons tenter de caractériser des conditions, dont aucune n'est causale ni déterminante. La création des langues a contribué à faire passer l'humanité du règne de la nécessité naturelle à celui de la « liberté » sociale ; elles ne sauraient donc être jugées à l'aune de cette nécessité.

Bref, si le langage est un moment de l'évolution, les langues n'en sont pas moins des créations historiques. Il n'est pas exclu qu'elles soient de part en part artificielles, cumulant des innovations transmises et par là des expériences archaïques² et des « visions du monde ». L'émergence du langage résulte peut-être de la mystérieuse rencontre d'un organe du langage (jusqu'ici introuvable)³ et d'une pression évolutive, mais plus certainement de la création sociale et de la transmission de systèmes de signes vocaux. Cette création affranchit pour une part les hommes des pressions de l'environnement naturel ; elle institue ou renforce celles de l'entour culturel, le faisant passer du temps « darwinien » de l'évolution au temps « lamarckien » de l'histoire. Aussi estimons-nous que les langues échappent à une « explication » de type darwinien.

La question de l'origine du langage ne se pose pas si l'on admet que le langage est une création culturelle : son histoire n'est autre que celle des langues, et se confond avec celle des sociétés humaines. Au demeurant, l'opposition entre nature et culture semble aussi invalide que routinière : le feu est un phénomène naturel, et la maîtrise du feu une conquête culturelle qui s'est universalisée. Si la faculté de langage est naturelle au sens où elle a évidemment des substrats organiques, ces substrats ne sont pas des causes, et cette faculté ne s'exerce que dans la vie sociale des langues particulières : « l'individu réalisant sa faculté au moyen de la convention sociale qui est la langue » (Saussure, 1972, p. 419, note 63).

Le langage n'a pas de fonction, car il n'est pas un instrument. On pourrait arguer qu'il sert à s'adapter au milieu, mais chez l'homme le milieu est essentiellement sémiotique, puisque la culture est une formation sémiotique. Bref, le langage servirait à s'adapter à un milieu dont il constitue une part éminente : en d'autres termes, il servirait à s'adapter à lui-même...

Si les langues humaines sont des formations culturelles, bref des œuvres, transmises avec les autres formations culturelles, comme les techniques ou les règles d'alliance, la question de l'émergence du langage devient inséparable de celle de la phylogenèse de la culture, ou plutôt des cultures dans leur diversité : les langues sont toujours indéfiniment variables.

¹ Par exemple, la variabilité du chromosome Y est quatorze fois plus faible chez les bédouins polygames du Sinaï que chez les agriculteurs monogames de la vallée du Nil.

² En témoignent par exemple les phraséologies.

³ Les chimpanzés disposent des aires de Broca et de Wernicke, avec la même prédominance gauche que chez l'homme.

2.3. La fable des fourmis

D'une utilisation constatée du langage, le néo-darwinisme a fait une fonction à valeur adaptative, puis de cette fonction une cause. Rompant avec ce postulat qui pérennise l'image du langage-instrument, l'apologue qui suit met en scène non pas les abeilles comme la fable de Mandeville, mais les fourmis, si chères aux théoriciens de la sociobiologie (O. E. Wilson, Pierre Jaisson, notamment).

Par sa glande postpharingienne, chaque fourmi secrète des hydrocarbures et s'en oint à chaque toilette. Au cours de ses interactions quotidiennes avec les autres fourmis (par léchage ou trophallaxies, notamment), elle échange avec elles des hydrocarbures. Ainsi, son odeur est partagée par la colonie tout entière. En bref, l'odeur coloniale est un composé des odeurs de chacune des fourmis, et chacune en est enduite. Cette odeur évolue historiquement avec les décès et les naissances. Les groupes séparés divergent — de même que les groupes linguistiques. Les fourmis étrangères qui ne portent pas le visa colonial sont exterminées ; celles qui en sont artificiellement enduites sont tolérées. En somme, le visa colonial est une création socio-historique : mélange social d'hydrocarbures individuels, ce résultat chimique de l'interaction la conditionne aussi.

Le parallèle avec la langue est tentant. Comme le visa colonial, la langue est stabilisée par des échanges quotidiens de tous les membres de la société. La multiplication des échanges crée une unification linguistique et une intercompréhension. Un sociobiologiste ajouterait : qui parle un peu la langue est accepté, qui ne la parle pas est rejeté comme un barbare. En somme, la langue est immanente au social et transcendante à l'individu ; à son échelle historique, l'individu ne peut prétendre la modifier, sauf à prétendre à l'autorité d'un dictateur ou d'un onomatothète ; cependant, comme l'a bien vu Humboldt, chaque usage individuel la remanie quelque peu.

La fable des fourmis s'arrête cependant là, car elles n'ont pas de truchement, sans quoi peut-être nous n'aurions pu les contenir. Comme nos langues sont trop récentes et connaissent trop d'échanges pour être devenues intraduisibles, ce n'est pas tant leur traductibilité qui étonne, mais la volonté même de traduire. La curiosité envers d'autres groupes, la xénophilie, non moins attestée que la xénophobie, est peut-être liée à la prohibition de l'inceste ; l'exogamie en tout cas lui est certainement liée. De fait, à présent, les deux tiers des hommes emploient plus d'une langue chaque jour : le multilinguisme est la règle majoritaire et plusieurs langues cohabitent en nous.

2.4. Le langage comme milieu

Plutôt qu'un instrument, le langage est une part éminente du milieu où nous vivons : dirait-on que l'air est un instrument des oiseaux ? L'enfant naît environné de la langue qu'il a déjà entendue *in utero* et à laquelle il réagit déjà sélectivement — en tétant avec une énergie accrue que l'on mesure avec un biberon à capteurs. Il s'y adaptera progressivement par l'apprentissage, et en usera pour s'adapter au monde socialisé qui l'entoure.

Comme l'a justement affirmé Auroux, la langue n'est pas dans le cerveau. Ajoutons que les substrats anatomiques de l'activité linguistique ne la déterminent pas, et qu'en somme le véritable organe du langage est la société. Or la société n'a pas d'origine : elle a sans doute connu des évolutions depuis *Homo Erectus*, mais jamais des individus ne se sont ligués dans un premier contrat social : jamais une Ève africaine n'est partie en quête d'un chanceux Adam. La parenté et l'alliance sont en quelque sorte, plus que nos géniteurs, nos premiers parents.

Corrélativement, le langage n'a pas d'origine, car il est à l'origine, sinon de tout, du moins des mythes d'origine, néo-darwiniens par exemple. Le langage est un milieu et non une simple faculté : c'est pourquoi, dans la phylogenèse, aussi loin que l'on croie remonter, il n'apparaît pas après l'homme. Ils vont toujours ensemble, et se définissent l'un l'autre.

La conception du langage comme milieu s'éloigne également des théories *internalistes* et *externalistes*. On sait que le cognitivisme a maintenu voire renforcé l'opposition entre l'interne et l'externe, pour privilégier l'interne — en fait le niveau neuronal — en reconnaissant, sans plus, qu'il s'adapte à l'externe, qui serait le social (*cf.* Changeux, 2000, justifiant ainsi la plasticité organique du cerveau).

Sauf toutefois à considérer le langage comme une réalité purement individuelle et psychologique, on ne peut le considérer comme interne à l'instar des chomskiens et des autres tenants de l'individualisme méthodologique.

Au sein même du cognitivisme orthodoxe, des voies externalistes ont été explorées, tant en philosophie de l'esprit (Drestke) qu'en anthropologie cognitive. Les thèses du *stockage externe* formulée par Merlin Donald (1991) et celle de l'*externalisme* soutenue par Auroux (1998)¹ sont toutes deux intéressantes, mais elles auraient plus de sens si la localisation spatiale était déterminable et déterminante. La validité de l'opposition interne / externe reste en effet discutable : dans un couplage, l'interaction est déterminante, et par deux mouvements corrélatifs conduit tant à une subjectivisation qu'à une objectivation. Le langage est d'abord externe à l'individu et la compétence linguistique résulte en quelque sorte d'une intériorisation du social — si bien que l'apprentissage linguistique est déjà un implicite contrat social. Au demeurant, l'externe remanie l'interne de manière déterminante, tant anatomiquement dans l'épigenèse cérébrale que phénoménologiquement dans l'histoire personnelle.

Ni interne ni externe, la langue est ainsi un lieu du couplage entre l'individu et son environnement, parce que les signifiants sont externes (bien que reconstruits dans la perception) et les signifiés internes (bien que construits à partir d'une doxa externe). Comme le langage fait partie du milieu dans lequel nous agissons, c'est dans des pratiques diversifiées, dont témoignent les discours et des genres, que nous nous lions à notre environnement. Mais il est aussi peuplé de « choses » absentes, et dans l'expérience de l'altérité, du passé, de l'étranger, la culturalisation de l'enfant a lieu — non moins sinon plus que dans l'expression d'une expérience individuelle limitée au *hic et nunc*.

3. Conditions d'émergence du sémiotique

De la tradition philosophique, le cognitivisme néo-darwinien a conservé le privilège absolu donné au langage sur les langues : s'il existe une philosophie du langage, il n'existe pas de philosophie des langues. De la tradition religieuse, il a conservé le privilège donné au langage, moyen de la révélation, sur tous les autres systèmes de signes.

Plus que l'origine du langage, il nous paraît plus utile d'explorer les conditions de l'émergence du sémiotique et de la constitution propre de l'entour humain, où le langage occupe évidemment une place éminente mais non exclusive.

3. 1. Le couplage

Le couplage du vivant et de son environnement est la condition universelle de l'évolution biologique. Nous souhaitons tout à la fois relativiser et spécifier l'opposition entre *Umwelt* et *Welt*, tels que ces concepts sont définis par Uexküll, de manière à caractériser la spécificité sémiotique de l'environnement humain².

Les « états internes » des sujets humains sont des *présentations* — non des représentations, car ils apparaissent dans des couplages spécifiques entre l'individu et son entour mais ils ne *représentent* pas pour autant cet entour ou ce couplage³. Le substrat, notamment physique, de l'entour, le *Welt* venant en arrière-plan, nous l'appellerons *arrière-monde*.

L'entour est composé des niveaux présentationnel et sémiotique des pratiques. Le niveau physique n'y figure pas en tant que tel, mais en tant qu'il est perçu, c'est-à-dire dans la mesure où il a une incidence sur les présentations (« d'objets » ou de signifiants) ; aussi nous empruntons à Thom le terme de *phéno-physique*. Soit :

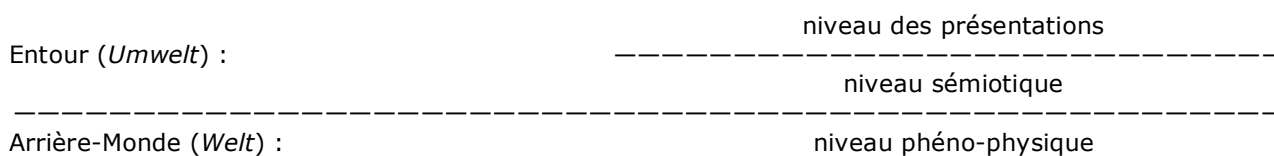


Figure 1 : La place de l'entour humain

¹ L'externalisme soutient le caractère artificiel et externe de l'intelligence humaine et Auroux en conclut de façon convaincante que l'esprit est « d'essence historique et empirique ». Soit, mais allons plus loin : l'esprit n'a pas d'essence et n'a d'autre existence que celle de la culture, tout à la fois interne et externe. On ne peut faire passer la différence entre nature et culture entre le cerveau et l'esprit. Pour naturaliser la culture, la philosophie de l'esprit (*philosophy of mind*) veut réduire l'esprit au cerveau et parvient donc à son objectif réductionniste en détruisant son objet.

² Dans ce paragraphe nous résumons des éléments (l'auteur, 2001a) nécessaires à la compréhension de ce qui suit.

³ J'adapte le concept de *présentation*, repris de Brentano, le maître de Husserl.

À ces trois niveaux, en jeu dans toute pratique, puisqu'on peut définir une pratique comme un mode réglé de leur interrelation, on peut faire correspondre, selon l'importance prépondérante qu'ils prennent, trois praxéologies : (i) la praxéologie représentationnelle comprend les arts de mémoire, le raisonnement, l'effort mémoriel, etc. ; (ii) la praxéologie sémiotique concerne la génération et l'interprétation des performances sémiotiques ; (iii) la praxéologie physique intéresse en premier l'activité technique et productive.

Si la division croissante et la spécialisation des pratiques restent caractéristiques de l'humanité, nous avons besoin en revanche d'une conception de la pratique qui rende compte du sémiotique. On ne peut plus opposer le langage et l'action — selon le préjugé partagé qui oppose les paroles aux actes, ce serait en rester à une vision purement représentationnelle du langage, déjà critiquée par la pragmatique. À chaque type de pratique sociale correspond un discours qui se divise en genres textuels oraux ou écrits ; tout texte relève d'un genre et appartient au niveau sémiotique d'un cours d'action, en quoi il est bien une performance sémiotique. En somme, par une spécification croissante, on obtient les correspondances suivantes :

<i>Instances sociales</i>	<i>Instances linguistiques</i>
Type de pratique	Discours
Pratique	Genre
Cours d'action	Texte

Tableau 1 : *Instances sociales et linguistiques*

3.2. Les ruptures catégorielles

Le niveau sémiotique de l'entour humain se caractérise par quatre décrochements ou ruptures d'une grande généralité et qui semblent diversement attestés dans toutes les langues décrites, si bien que l'on peut leur conférer par hypothèse une portée anthropologique.

- (i) La rupture *personnelle* oppose à la paire interlocutive JE/TU — nous employons des capitales pour résumer les diverses manières de désigner les protagonistes de l'interlocution représentée — une troisième personne, qui se définit par son absence de l'interlocution (fût-elle présente physiquement) : IL, ON, ÇA¹.
- (ii) La rupture *locale* oppose la paire ICI/LÀ à un troisième terme, LÀ-BAS, ou AILLEURS qui a également la propriété définitoire d'être absent du *hic et nunc*.
- (iii) La rupture *temporelle* oppose le MAINTENANT, le NAGUÈRE, et le FUTUR PROCHE au PASSÉ et au FUTUR. Il convient sans doute de distinguer la zone circonstancielle du présent de l'énonciation représentée, marquée par des futurs et passés proches, du passé éloigné, connu indirectement et souvent légendaire, et du futur éloigné de toute façon conjectural.
- (iv) Enfin, la rupture *modale* oppose le CERTAIN et le PROBABLE au POSSIBLE et au à l'IRRÉEL. On pourra bien entendu opposer à l'intérieur de ces catégories le conditionnel à l'irréel, etc. ; mais seul nous importe ici le point que les langues articulent ces catégories.

Ces ruptures catégorielles sont généralement grammaticalisées, et font donc l'objet de choix incessants et obligatoires des locuteurs, tout énoncé devant être situé dans au moins une des zones qu'elles délimitent.

3.3. Les trois zones anthropiques

On peut noter que les positions homologues sur les axes de la personne, du temps, du lieu et du mode sont fréquemment combinées ou confondues : en français, par exemple, les emplois modaux du futur et de l'imparfait sont légion, le futur antérieur a également une valeur modale, etc.

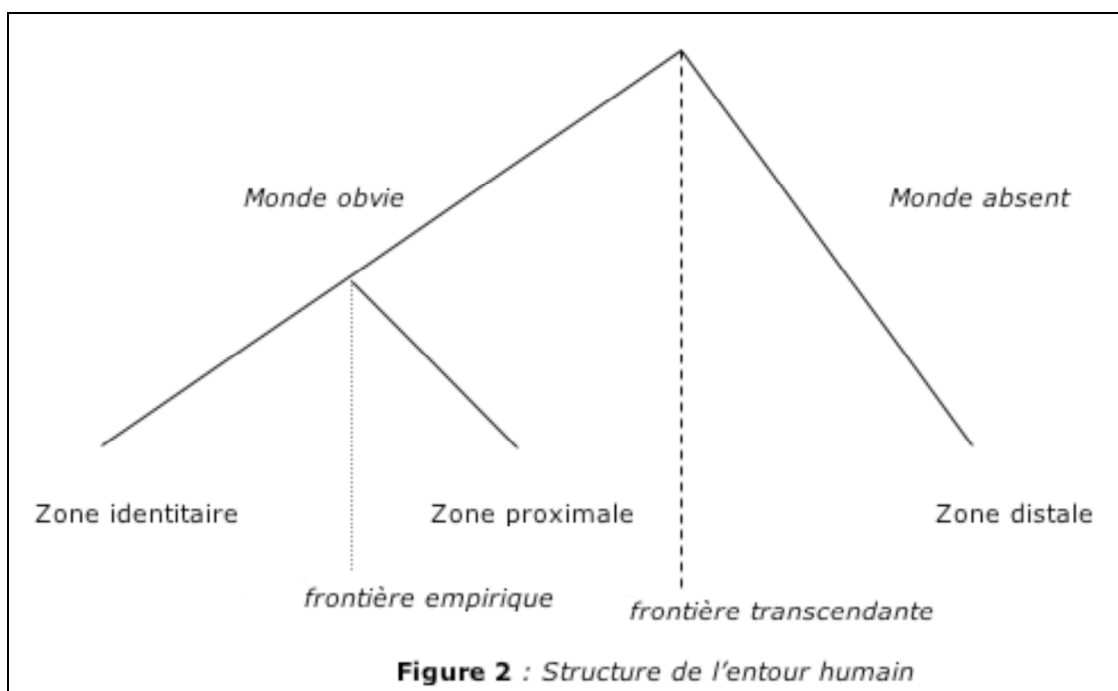
Les homologues entre ces ruptures permettent de distinguer trois zones : une de coïncidence, la zone *identitaire* ; une d'adjacence, la zone *proximale* ; une d'étrangeté, la zone *distale*.

¹ Les grammairiens arabes appellent la troisième personne l'*absent*. Nous suivons ici, sans entrer dans les détails techniques, la présentation de Benveniste (1966, pp. 225-236), qui distingue deux oppositions : celle de subjectivité, qui sépare JE et TU, et celle de personnalité, qui sépare ces personnes subjectives du IL.

	<i>Z. identitaire</i>	<i>Z. proximale</i>	<i>Z. distale</i>
<i>Personne</i>	JE, NOUS	TU, VOUS	IL, ON, ÇA
<i>Temps</i>	MAINTENANT	NAGUÈRE BIENTÔT	PASSÉ FUTUR
<i>Espace</i>	ICI	LÀ	LÀ-BAS AILLEURS
<i>Mode</i>	CERTAIN	PROBABLE	POSSIBLE IRRÉEL
Frontières		↓ <i>fr. empirique</i>	↓ <i>fr. transcendante</i>

Tableau 2 : Zones et frontières de l'entour

La principale rupture sépare les deux premières de la troisième. En d'autres termes, l'opposition entre zone identitaire et zone proximale est dominée par l'opposition qui sépare ces deux zones prises ensemble à la zone distale. Ainsi se distinguent un monde obvie (formé des zones identitaire et proximale) et un monde absent (établi par la zone distale)¹. On peut alors figurer ainsi la structure de l'entour humain :



Par rapport aux langages des animaux, la particularité des langues réside sans doute dans la possibilité de parler de ce qui n'est pas là : la zone distale. Sur l'axe de la personne, cela permet de parler des absents. L'homologation des décrochements les situe de préférence dans un autre temps (ancêtres, postérité, envoyés à venir), d'autres lieux et d'autres mondes (héros, dieux, esprits). Sur l'axe du temps, cela ouvre les aires de la tradition et de l'avenir ; sur ceux de l'espace et du mode, celle de l'utopie.

Le rapport entre l'individu et la société est l'une des formes que prend pour l'humanité le couplage biologique de l'organisme avec l'environnement. Mais la zone proximale, où par exemple les congénères sont reconnus pour tels, appartient vraisemblablement aussi à l'entour

¹ Les termes *obvie* et *absent* témoignent d'une concession purement temporaire au paradigme positiviste : la question de l'absence est naturellement irrecevable pour le positivisme — et, plus précisément, dans le domaine des sciences du langage, pour le représentationnalisme.

des autres mammifères, sans avoir le même statut, faute de zone distale. En effet, la zone distale reste spécifique de l'entour humain, sans doute parce qu'elle est établie par les langues¹.

Que l'entour humain contienne des espaces distincts du *hic et nunc*, cela peut être mis en rapport avec la théogonie et la cosmogonie, deux activités propres à notre espèce, et auxquelles nous devons aussi bien les sciences que les religions. Le cosmos et les univers divins sont des présentations de la zone distale, sans substrat perceptif immédiat. Ces deux types de créations sans cesse continuées s'appuient notamment sur les décrochements de personne, de temps, de lieu et de mode.

Le contenu des zones varie évidemment avec les cultures, et *a fortiori* les pratiques sociales. La zone identitaire n'est pas nécessairement celle d'un Ego, et peut être instanciée par un groupe, un ancêtre totémique, une nation, etc. ; corrélativement, l'Ego peut parfois occuper la zone proximale (« Je est un autre », écrit Rimbaud) voire la zone distale (chez certains mystiques).

La frontière empirique est marquée dans la grammaire des langues par ce que les grammairiens appellent la *zone inaliénable* : elle est peuplée d'« objets » qui exigent ou permettent des constructions réfléchies ou des datifs éthiques. C'est le cas évidemment des parties du corps, mais aussi des vêtements, des parures, voire des animaux familiers ou des véhicules habituels².

3.4. Les deux médiations

Revenons aux trois niveaux de la pratique, pour préciser leurs rapports avec les trois zones de l'entour. Chacune des zones de l'entour est susceptible de trois sortes de descriptions : au niveau de ses substrats physiques, de ses corrélats sémiotiques, et de ses corrélats présentationnels.

On distingue souvent en sémiotique l'axe de la représentation (sujet/objet) et l'axe de la communication (sujet₁/sujet₂). Cette distinction repose sommairement sur deux traditions sémiotiques majeures : la tradition aristotélicienne et la tradition augustinienne que Bühler a tenté d'unir dans son modèle du signe.

Dans un modèle de la pratique qui tienne compte des performances sémiotiques, elle est dépassée dans la distinction entre la *médiation sémiotique*, qui spécifie le rôle des signes dans la cognition humaine et la *médiation symbolique*³ qui rend compte des relations entre les trois zones anthropiques.



Figure 3 : Les médiations sémiotique et symbolique

¹ En elle-même, l'évidence que les langues permettent de parler de ce qui n'est pas là a souvent été rappelée. Hockett (1960) parle à ce propos de *displacement*. Plaçant la discussion au niveau des représentations, Gopnik (1982) parle de représentations à connexions causales indirectes, Fodor (1980) d'information inférée, Gärdenfors (1994b) de représentations détachées (*detached*). La zone distale est en somme la source de présentations sans substrat perceptif immédiat. Dans les termes familiers de la philosophie, la zone proximale est celle de l'empirique, et la zone distale celle du transcendant.

² Ainsi, en français, on pourra dire : *j'ai les mains rougies, j'ai les chaussettes tirées, j'ai une soupe grillée*.

³ L'expression *médiation symbolique* est reprise de Geertz, 1973. L'opposition entre signe et symbole varie grandement selon les auteurs. Ici, nous considérons les symboles comme un sous-ensemble des signes, ceux qui peuvent être indépendants de la situation, les signes linguistiques en premier lieu.

Les problèmes constitutifs du cognitivisme peuvent alors être reformulés en fonction des rapports entre les trois niveaux de la pratique, phéno-physique, sémiotique et présentationnel¹ : ils définissent l'axe de la cognition, qui transforme la perception du monde physique en (re)présentations mentales et réciproquement. La cognition humaine se caractérise ainsi par la médiation sémiotique entre le niveau phéno-physique et le niveau (re)présentationnel. En maintenant l'autonomie relative du niveau sémiotique, la médiation sémiotique permet la médiation symbolique qui articule les zones anthropiques. Mais elle ne la détermine pas, et chaque culture représente la médiation sémiotique en fonction de ses croyances, d'où par exemple l'efficacité de la magie, croyance qui dépend de la zone distale, mais détermine la relation entre niveau présentationnel et niveau phéno-physique, par le biais de pratiques sémiotiques.

Alors que la philosophie du langage se préoccupe des relations entre le monde phéno-physique et les représentations, la sémiotique et la linguistique ont à traiter du rapport dynamique entre les trois zones de l'entour, c'est-à-dire de la médiation symbolique. Les parcours d'énonciation et de compréhension consistent en passages constants d'une zone à l'autre. Ces passages sont orientés par des valeurs (esthétiques, éthiques, thymiques — euphoriques ou dysphoriques). L'activité d'évaluation dépend notamment de la zone de l'entour qui est valorisée au moment de la production ou de l'interprétation.

La cognition est ordinairement définie comme un processus de médiation sémiotique : les opérations le plus souvent évoquées sont l'abstraction de classe ou de types et la catégorisation d'objets ou d'occurrences. Mais cette conception aboutit ordinairement à un face-à-face entre le concept et l'objet et ne tient guère compte du contexte et des performances sémiotiques. Prendre en considération la médiation symbolique permet en revanche de restituer le caractère actif et critique de toute création de connaissance.

4. Les conditions du couplage sémiotique

Nous pouvons distinguer trois types de conditions de couplage : les conditions d'*identification* qui objectivent les stimuli endogènes ou exogènes en objets valués ; les conditions de *socialité*, associées à la zone proximale ; les conditions de *dé-limitation*, qui ouvrent l'entour au-delà du *hic et nunc* et sont donc associées à la zone distale.

L'interprétation comme la compréhension qui en dépend sont des couplages du sujet avec l'entour. Comme un couplage suppose des modifications et des adaptations réciproques, l'interprétation comprend une activité perceptive (sur la perception sémantique, cf. l'auteur, 1991, ch. VIII). Le concept de couplage peut s'entendre ici de trois façons, selon qu'il concerne des performances sémiotiques, des congénères ou des objets non sémiotiques. Dans tous les cas, le couplage a lieu au sein d'une pratique socialisée.

4.1. Conditions d'identification et conquête de l'identité

La permanence de l'objet caché, jadis étudiée par Piaget, coïncide avec l'accès du jeune enfant à la fonction symbolique. On peut distinguer trois phases de cet accès : le sens se définit par un rapport des signes aux objets présents ; puis aux objets absents (qui sont *a fortiori* des artefacts sémiotiques), enfin aux objets inexistantes ou « abstraits ».

À la permanence de l'objet caché, on peut lier la permanence ou du moins l'objectivation récurrente de l'Ego ; ainsi, la reconnaissance de l'image spéculaire s'acquiert chez l'enfant à 19-20 mois, avec l'accès à la fonction symbolique. Ultérieurement, l'acquisition des catégories actantielles exprimées par les cas et constructions linguistiques permet de multiplier les guises du sujet et de décliner les variétés d'objets. Hormis quelques remarques incidentes en éthologie, on a peu étudié les catégories actantielles chez les animaux supérieurs. Selon Lorenz, le jeune corbeau classe les objets de son environnement en prédateurs, proies, et circonstants : du moins, devant un objet inconnu, il s'assure que ce n'est pas un prédateur, puis une proie, et enfin s'en désintéresse ou se pose dessus.

Chez l'homme, on peut faire l'hypothèse d'une catégorisation actantielle relative à trois sortes d'actions : polémique vs irénique, transformative vs communicative, sémiotique vs physique. Les types d'objets se définissent par la position du sujet dans une catégorie actantielle : ainsi un prédateur-ennemi sera l'agent d'une fonction polémique dont Ego est l'objet ; une proie-ennemi sera l'objet d'une fonction polémique dont Ego est agent. Un adjuvant est défini par sa relation irénique à un Ego engagé dans une fonction ; un opposant

¹ Bien que le cognitivisme classique ne reconnaisse que deux sphères, représentationnelle et physique, pour s'efforcer d'ailleurs de réduire la première à la seconde.

par sa relation polémique au même sujet. Un circonstant par sa neutralité à l'égard des fonctions en cours. La catégorisation des fonctions et des actants conditionne la production et l'interprétation de récits, qu'ils soient événementiels ou mythiques. Dans cette mesure, elle est une condition d'accès à la textualité.

4.2. Conditions de socialisation

- L'*exposition aux signes* a d'abord lieu *in utero* : des travaux de Jacques Mehler et de ses collaborateurs ont récemment montré que le nouveau-né, déjà habitué à la langue maternelle, réagit par des succions accentuées aux propos tenus dans cet idiome.
- Jadis décrite par Vygotsky, la *triangulation désignative* lie l'enfant, l'objet qu'il désigne par ostension et l'adulte qu'il consulte du regard. Cyrulnik (1995) y voit à sa suite la naissance du sens : en bref, quelque chose n'est pour moi que dans la mesure où il est pour quelqu'un. Montrer à quelqu'un quelque chose, c'est déployer simultanément les deux dimensions de l'ostension et de l'adresse ; cela suggère que les représentations, loin de se cantonner à une sphère privée, restent indissociables de l'échange.
- La *triangulation contractuelle* apparaît dans la règle du jeu : le jeu est une condition de socialisation, et pour l'enfant les jeux de langage codifiés permettent de sortir des genres idiosyncrasiques et d'apprendre ainsi la langue par ses usages communs. En tant que règle des jeux de langage, les genres témoignent, dans l'usage linguistique même, de la normativité fondamentale du social (car tout texte, toute parole relève d'un genre). Par les répétitions inlassables de l'usage, cette normativité « s'objective » en Langue ou en Loi.

Si la règle est contractuelle, tout contrat, matrimonial ou économique, et plus généralement toute promesse réciproque reposant sur une dimension fiduciaire, suppose un tiers-garant, représenté ou non par une personne et agissant sur le mode de la prescription et / ou de l'interdiction. Le tiers-garant est une puissance distale : par exemple, dans l'échange monétaire, l'effigie garantit le bon aloi ; dans l'échange matrimonial, un officier ou officiant quelconque, présent ou invoqué, assure le respect des prescriptions et prohibitions toujours à l'œuvre, même si elles varient selon les lieux et les époques. Dans le contrat, la dimension de l'adresse est un couplage entre identitaire et proximal, mais l'ostension passe sous le régime du distal : l'objet du contrat est ordinairement disjoint dans le temps et dans le mode, puisque le contrat prévoit un échange futur.

4.3. Conditions de dé-limitation

Trois facteurs favorisent la sortie du *hic et nunc* et l'institution d'une zone distale : l'élaboration d'images mentales et l'évocation de stimuli absents ; le rêve et les états de conscience altérée (hallucinations) ; enfin la fiction, qui décrit des domaines inaccessibles à l'expérience immédiate et les institue par là.

L'existence d'une zone distale est sans doute une condition de socialité qui dépasse le simple regroupement proximal de la horde : l'efficace qu'on lui prête, par la crainte et l'invocation, transforme la horde animale en groupe humain. La Loi semble généralement conçue comme une puissance distale, ce qui témoigne de la transcendance du social.

L'humanité a redoublé dans ses hiérarchies sociales celles des primates, de manière que la puissance physique s'allégorise en pouvoir symbolique. En effet, le contrat social exige une triangulation : il s'appuie et se légitime toujours par une force distale, et les puissants excipent de leur totem, de leur lignée divine ou héroïque, et ils emploient pour chanter leur gloire bardes, griots et sherpas présidentiels.

Abstraction et développement de la triangulation contractuelle, la *triangulation rituelle* met en relation le sujet ou le groupe avec des objets distaux, sous la réaction d'une puissance distale. Tant l'ostension que l'adresse sont ainsi symbolisées, comme on le voit dans les rituels. On doit alors compléter la théorie de la constitution dialectique du sujet dans l'intersubjectivité¹. À la dialogique (au sens philosophique du terme), qui se fonde sur une intersubjectivité passablement spéculaire et s'y épuise, il convient de substituer alors une constitution « trilogique », qui détermine la règle du jeu de dialogue et l'instaure².

¹ Cette thèse de Schleiermacher, d'ascendance platonicienne, a connu grâce à Bakhtine un nouvel essor, sous le nom de *dialogique*.

² Cette « trilogique » semble dérivable à partir de Humboldt : « L'imagination créatrice n'est pas seulement quelque chose qui s'accomplit totalement dans le rapport sujet/objet, mais elle se voit d'emblée

En bref, on aurait la variation suivante :

	<i>Ostension</i>	<i>Adresse</i>
<i>Désignation</i>	objet présent	tiers présent
<i>Contrat</i>	objet absent	tiers présent
<i>Rite</i>	objet absent	tiers absent

Tableau 3 : De la désignation au rite

NB : L'objet rituel est présent, mais il tient lieu d'un autre qu'il symbolise et qui est absent ; les participants au rite sont présents, mais s'adressent à un tiers absent. C'est là une règle apparemment générale (cf. Descola, 1992, ch. XX, sur les chamanes achuar).

La première triangulation peut être décrite comme la mise en place et confirmation de la frontière empirique. La seconde structure l'espace proximal de la société en consignnant à l'individu droits et devoirs. La troisième enfin trace la frontière transcendante.

Nous ne prétendons pas que la première triangulation soit dans la phylogenèse à l'origine des deux autres, quoique l'usage des termes de filiation pour désigner les dieux soit fort général (de la Mère des Animaux chamanique à la Déesse-Mère anatolienne, voire au Nom-du-Père lacanien). En revanche, leur succession dans l'ontogenèse semble vraisemblable, et le passage de la première triangulation à la dernière peut se décrire comme une « conquête » progressive de l'absence.

Toutes les conditions de couplage (identification, socialisation et dé-limitation) reposent sur l'usage des signes linguistiques dans trois de leurs propriétés majeures. (i) Comme l'usage des signes saussuriens est indifférent à la présence de ce que nous appelons leurs référents, ils autorisent la monstration non ostensive d'objets cachés ou absents rendue possible parce que les signifiés suscitent des images mentales, d'où la thèse de leur valeur représentationnelle. (ii) Ils semblent combinables en prédications modalisées, et donc en formules de prescription et d'interdiction — décontextualisées, car valables en tout contexte. (iii) Ils semblent combinables en textes, qui transforment les visions du monde en conceptions du monde, et permettent ainsi, là encore, de passer du proximal au distal. En fait, l'image de signes qui se combinent en texte demeure une simplification grammaticale : le signe linguistique n'a pas d'existence empirique, il n'est qu'un passage d'un texte oral ou écrit.

4.4. Des sortes de couplage aux types de parcours sémiotiques

La conception du couplage qui vient d'être exposée conduit à reconsidérer le principe même des typologies des signes habituelles en sémiotique. Des spécialistes européens de la communication animale ont noté qu'il « manque » aux singes — et peut-être à certains primatologues cognitivistes — la notion de signe saussurien (cf. Vauclair et Fagot, 1993). On sait que la tradition philosophique anglo-saxonne privilégie deux types de signes, *l'index* (cf. les indexicaux du positivisme logique) et *l'indice*, signe indiciaire de tradition augustinienne. Bien que l'un soit canonique pour sa référence et que l'autre repose sur l'inférence, aucun de ces deux signes n'a de lien nécessaire avec des langues.

Après un inlassable et maternel conditionnement, les chimpanzés les plus talentueux paraissent se servir comme nous de signes linguistiques : certains ont acquis des vocabulaires de plusieurs centaines de mots — plus que certains collégiens — et des capacités syntaxiques équivalentes à celles d'un enfant de deux ans. Mais rien ne permet cependant d'affirmer qu'ils se servent des mots comme nous le faisons : tout laisse à penser qu'ils se servent des signes saussuriens comme des indices ou des index¹. De fait, Savage-Rumbaugh reconnaît que 97 % des performances sémiotiques de Kanzi sont « purement pragmatiques ». Même des singes moins talentueux ou moins conditionnés vérifient les principes de la pragmatique, qui, issue du positivisme logique, se cantonne délibérément à la situation².

située dans le rapport à *l'Autre*. La création du non-réel est une création *commune* des hommes créatifs. L'enjeu est 'd'enflammer l'imagination par l'imagination' » (GS, II, p. 127, tr. in : Trabandt, 1999, p. 155).

¹ Certains auteurs ont d'ailleurs observé chez les chimpanzés la triangulation ostensive, nous dépossédant ainsi d'un de ces multiples propres de l'homme qui disparaissent les uns après les autres avec les progrès de la primatologie.

² Les chimpanzés et les bonobos ont conscience de la pertinence (au sens de Sperber et Wilson). Brillants interactionnistes, ils mettent en œuvre le principe de charité de Davidson, attribuent correctement les intentions et respectent mieux que nous les maximes de Grice. Excellents psychologues, rusés politiciens, surtout en captivité où ils se désennuient par des querelles de préséances, ils restent de pauvres écolo-

L'index et l'indice se satisfont de la médiation sémiotique : l'index apparie une présentation d'objet et un signe ; l'indice, deux représentations d'objets, dont l'une, antécédente, est promue au rang de signe « naturel ». En revanche le symbole — j'entends par là le signe saussurien — suppose une médiation symbolique, par le principe de la langue qui prescrit des relations contextuelles et exclut les autres termes du même paradigme.

Le signe saussurien ne se réduit pas à l'association du signifiant et du signifié — qui relèverait simplement de la mémoire épisodique, comme le suggère Merlin Donald. Il se caractérise par sa dépendance contextuelle qui lui permet de varier en fonction des contextes et donc des textes, comme par son emploi *différentiel* avec les autres signes qui résulte de son engagement dans un système linguistique. Ce double engagement, syntagmatique et paradigmatique, que ne connaissent ni l'index ni l'indice, place le signe saussurien sous le double régime de l'absence¹ et de la co-présence : le signe saussurien n'a pas de sens déterminable hors contexte. Au palier supérieur, il permet le déploiement de textes, dans une « grande syntagmatique » qui met en œuvre des paradigmes de formes textuelles, comme les fonctions narratives. De fait, comme le soulignait Gärdenfors, nous attendons toujours que Kanzi nous raconte une histoire autour d'un feu de camp : on ne peut construire de récit qu'avec des symboles et non avec des indices ou des index.

Les conventions systématiques que concrétisent les symboles permettent l'autonomie relative des performances linguistiques à l'égard des situations et donc l'institution de situations nouvelles. Ces conventions sont nécessaires à l'institution de la Loi, qui non seulement est indissoluble de son énoncé, mais indépendante de toutes les situations, tout en gouvernant leur souvenir et leur advenir.

Nous ne postulons aucune « nature » du signe : tout signe peut être l'objet d'un usage indexical, indiciaire ou symbolique, et ce sont les parcours interprétatifs qui le qualifient. Le même son, le même geste, le même tracé peuvent être interprétés sur le mode indexical, indiciaire ou symbolique. Par exemple, les icônes peuvent être interprétées comme des index, mais aussi comme des symboles².

L'indice ne nécessite pas de système linguistique et semble à la portée de tous les animaux supérieurs. Prototype du signe référentiel, l'index est généralement considéré comme primaire. Le symbole reste un apanage humain.

Peu importe ici que l'ontogenèse récapitule cette gradation ; retenons que le signe saussurien permet un passage à l'absence et résulte d'une *distalisation*. Pour préciser comment, voici une comparaison raisonnée qui tient compte de quatre pôles dédoublés : celui du *signifiant* (divisé en *Signifiant 1*, celui du signe examiné, et *Signifiant 2*, celui du signe voisin dans le contexte), celui du *signifié* (divisé en *Signifié 1*, celui du signe examiné, et en *Signifié 2*, celui du signe voisin dans le contexte), celui de l'*objet* (divisé en *Objet 1*, source de l'inférence, et *Objet 2* but de l'inférence) ; enfin, celui de l'*adresse*, articulé par la distinction entre *Destinateur* et *Destinataire*.

	<i>Index</i>	<i>Indice</i>	<i>Symbole</i>
<i>Signifiant 1</i>	Présent	Présent	Présent
<i>Signifiant 2</i>	Néant	Néant	Présent/Absent
<i>Signifié 1</i>	Présent	Présent	Absent
<i>Signifié 2</i>	Néant	Néant	Présent/Absent
<i>Objet 1</i>	Présent	Présent	Néant
<i>Objet 2</i>	Néant	Présent/Absent	Néant
<i>Destinateur</i>	Présent	Présent/Absent	Présent/Absent
<i>Destinataire</i>	Présent	Présent	Présent/Absent

Tableau 4 : Typologie des signes

gues, comme le note Robin Dunbar (1993, p. 759), faute, sans doute, d'abstraire la notion de milieu des alentours perceptibles.

¹ Un terme évoque préférentiellement son antonyme, comme l'ont montré depuis la fin du XIX^e siècle les expériences d'amorçage conduites en psychologie.

² Une belle expérience de Martine Cornuéjols a montré que le temps de reconnaissance d'une icône était proportionnel au nombre de syllabes du terme correspondant (2000, p. 313).

Par tous ses traits définitoires, l'index demeure sur le mode de la présence et n'a de contexte que situationnel¹. L'indice admet en revanche l'absence de l'objet inféré, mais il faut cependant distinguer les relations d'inférence entre percepts d'objets et les relations inférentielles entre signes, car ces dernières sont contraintes par les discours et genres textuels et peuvent à la différence des premières, être rapportées à un destinataire². Enfin, le symbole linguistique, pour sa part, fonctionne en contexte, sans relation nécessaire de co-présence avec des « objets », ni avec des co-énonciateurs appartenant à une situation originaire (comme dans le cas exemplaire de l'écrit).

4.5. Le substrat neuronal du symbolique et de la zone distale

La compétence symbolique n'est nullement câblée dans des zones dédiées : elle trouve ses substrats anatomiques au cours de la socialisation humaine ; en cas de lésion, on note des récupérations par utilisation d'autres substrats : les compétences symboliques des personnes devenues aveugles, sourdes ou muettes en témoignent.

Le substrat physiologique du distal semble toutefois lié au développement exceptionnel et récent chez l'homme du cortex préfrontal, où précisément se traite la perception des objets absents³.

Le thalamus assure la connexion entre le cerveau limbique, siège de la mémoire et des émotions, et le lobe préfrontal où se produisent les anticipations. Cette connexion apparaît chez les mammifères, se précise chez les primates non humains, et prend une place importante chez l'homme. Elle permet de lier la mémoire et l'anticipation, et participe donc de la constitution du temps humain : « le cerveau humain est, entre tous, le plus capable de traiter et d'articuler des informations relatives à des choses absentes, à des phénomènes disparus et à des événements passés ou à venir » (Cyrulnik, 2000). L'empan temporel, la rémanence du passé et l'anticipation du futur sont évidemment nécessaires à l'intelligence narrative : ainsi le temps est-il, sinon le premier des objets absents, du moins leur site. Corrélativement, des expériences récentes conduites par Mazoyer & al. ont mis en évidence que la compréhension de récits mobilise, bien plus que des listes de mots et des phrases litigieuses, un grand nombre d'aires cérébrales, en particulier la région préfrontale gauche (cf. Changeux, 2002, p. 190). On a remarqué par ailleurs que les patients souffrant de lésions du cortex préfrontal ont de grandes difficultés à planifier des actions à long terme. En somme, qu'il s'agisse de planifier des actions ou de comprendre des histoires, d'agir ou d'interpréter, les mêmes zones cérébrales sont mobilisées : elles concourent à la perception des objets absents.

On connaît l'opposition formulée par Saussure entre les relations linguistiques *in praesentia* fondées sur la compatibilité, et les relations *in absentia* fondées sur l'incompatibilité. Par exemple, à une place donnée de la chaîne syntagmatique, on ne peut avoir que *refaire*, ou *défaire*⁴. Ainsi l'absence, entendue comme présence niée (en termes logiques), ou inhibée (en termes neuropsychologiques), reste au fondement de l'activité de langage, car toute énonciation suppose, à chaque choix d'un signe, l'exclusion des signes du même paradigme qui pourraient occuper la même place. En cela, la négation précède l'affirmation, ou en termes plus précis l'inhibition globale accompagne l'activation locale⁵.

Dans la théorie des zones anthropiques, la zone distale, sans substrat perceptif immédiat, est établie et configurée par l'activité sémiotique. L'énonciation consiste alors à passer du distal absent au signe proximal présent, par une inhibition qu'on nomme ordinairement *actualisation*. En d'autres termes, le choix d'un signe, décrit comme une activation, s'accompagne de l'inhibition de son antonyme et des autres signes appartenant à la même classe.

¹ Ainsi les signes linguistiques dits indexicaux, comme les pronoms ou les démonstratifs, ne sont-ils pas des index.

² Les premières font l'objet des théories des « signes naturels ». Saint Augustin a tenté une synthèse en affirmant que dans l'Écriture les choses signifiaient, de manière à légitimer l'interprétation allégorique.

³ Cf. Houdé, 1997, ch. 2, sur la persistance des objets disparus, le cortex préfrontal et l'inhibition.

⁴ Les expériences des associationnistes au siècle dernier ont souligné le caractère fondamental de la relation d'antonymie et les expériences contemporaines sur le *priming* ou amorçage n'ont fait que les confirmer.

⁵ Si l'action énonciative, comme les autres formes d'action, peut se définir par un défaut temporaire d'inhibition, en logique apophantique l'affirmation pourra être conçue comme le résultat d'une somme de négations. En d'autres termes, la présence pourrait être définie comme une sommation d'absences refusées.

Ce processus fondamental de sélection paradigmatique est caractéristique des langues humaines, par opposition aux langages animaux. Il est lié à la conquête de l'absence par notre espèce, bref à ce que l'on pourrait appeler la phylogenèse de la zone distale.

Au palier du signe, la sélection paradigmatique éclaire un phénomène plus général, à rapporter à la perception sémantique (cf. l'auteur, 1991, ch. 8). Cette perception hiérarchise trois types de données : des formes, des fonds, et l'arrière-plan des formes et des fonds, c'est-à-dire les paradigmes des autres formes et fonds concurrents qui relie la perception présente au corpus des expériences linguistiques passées. Ainsi, à la différence de la perception animale, la perception humaine, culturalisée, laisse comme telle une grande place aux processus descendants. Bref, elle agit sur le mode de la hantise, qui est sans doute un des propres de l'homme.

En somme, on peut supposer ces trois stades d'émergence des types de signes¹ :

Indices	Animaux supérieurs
Index	Anthropoïdes
Symboles	Humains

Tableau 5 : Émergence des types de signes

Les indices, indépendamment de l'établissement de relations causales, sont établis par la mémoire épisodique. Le conditionnement suffit à les instituer. Les index supposent l'ostension et la triangulation désignative (cf. *supra*). Les symboles supposent l'établissement de paradigmes, condition et effet tout à la fois de la textualité.

4. 6. Le déplacement

Revenons à présent au fonctionnement sémiotique définitoire des symboles.

On a souvent rapproché les symboles et les monnaies, fût-ce pour les opposer. Les titres à une participation au sacrifice et donc à la répartition se nommaient *symbola*.² Si la monnaie est partie prenante du sacrifice, le symbole monétaire, comme la victime dont il tient lieu, devient l'objet d'une *substitution avec déplacement* qu'il convient de détailler. Sans référence prédéterminée, cette relation entre la victime et son prix n'a d'autre opérativité que celle du sacrifice : la valeur semble fixée, mais seulement dans un ordre sacré, institué par une Loi hiératique. En outre, le référent n'est pas l'objet présent, comme le voudraient les théories indexicales de la signification, mais bien l'objet absent.

Enfin, la relation de « mise en valeur » n'est pas unique : la monnaie en offrande vaut pour une « victime » au second degré, car le sacrifice animal résulte lui-même d'un évitement du sacrifice humain. Si l'on tenait à la référence, il faudrait parler de référence immédiate (l'animal domestique) et de référence différée (la victime humaine) : ce qui distingue le sacrifice du meurtre pur et simple est en effet un évitement, car la vengeance, dès lors qu'elle est ritualisée, voire symbolisée dans le sacrifice, frappe à côté³.

La référence du symbole se réduit de fait à un déplacement qu'il serait inutile de déclarer « métonymique » sans préciser sa fonction : il reconnaît et instaure une différence qualifiée, qui comme telle donne sens⁴. Soit :

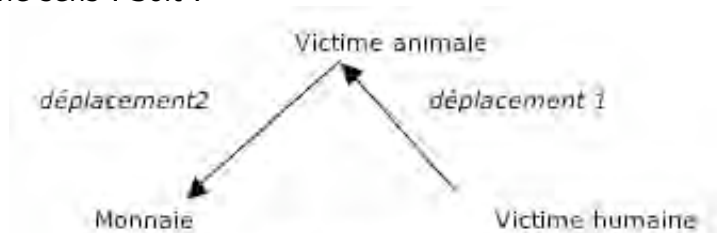


Figure 4 : Un triangle sacrificiel

¹ Nous ne traitons pas ici des icônes.

² L'autre sens, plus notoire, désigne les parties d'un même objet, par exemple les perles d'un bracelet. Dans les romans antiques, elles permettent aux proches longtemps séparés de se reconnaître enfin.

³ Les exemples ethnologiques sont multiples : ainsi, chez les Tchoukches, on se vengeait jadis en tuant le frère de l'offenseur.

⁴ Elle repose sur l'opposition entre la zone identitaire et la zone proximale à laquelle appartient la victime animale, et sur la contradiction entre le monde humain et le monde divin qui entraîne, pour la résoudre, la mise à mort de la victime humaine.

Les déplacements diffèrent par les extensions qu'ils opèrent : alors que toute monnaie peut être offerte en sacrifice, les victimes animales sont toujours choisies en fonction de certaines restrictions (couleur, sexe, etc.), et la victime humaine est un individu choisi (émissaire, roi, etc.). Par une double extension, on passe ainsi du symbole hiératique, purement hétéronome, au symbole monétaire, purement isonome, et de l'objet singulier à l'objet quelconque fonctionnant comme équivalent général. On aurait alors :

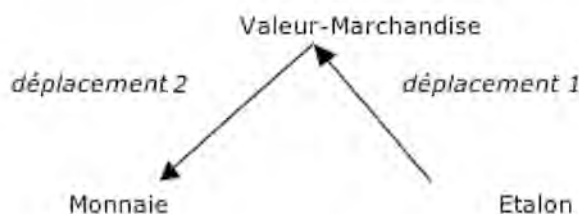


Figure 5 : Le triangle de la valuation

Ici, c'est la valeur étalonnée de la marchandise (déplacement 1) qui permet son intégration à l'échange monétaire (déplacement 2).

Or la conception traditionnelle du signe, qui perdure d'Aristote à Ogden et Richards, et qui fut heureusement ruinée par Saussure, a pu être représentée par un triangle homologue (cf. l'auteur, 1990) :

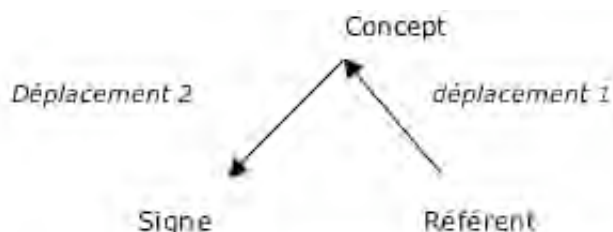


Figure 6 : Le triangle sémiotique traditionnel

En ce cas, le déplacement est conçu comme une représentation : le signe représente le concept, et le concept représente le référent. L'orientation de la relation de représentation *inverse* ainsi celle du déplacement. Toutefois, le référent n'est qu'un préjugé, c'est-à-dire une réification de la doxa et autres normes sociales. La concrétisation de la doxa instituée, par figement, un signifié, et le passage du signifié au signifiant, définitoire de la sémosis, qualifie le signifiant comme tel.

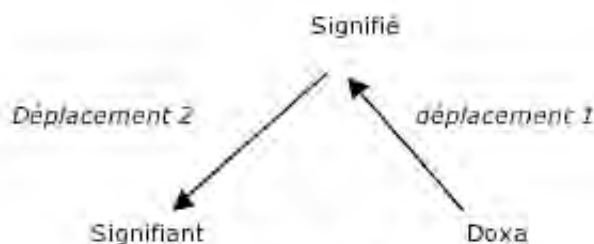


Figure 7 : Le triangle du symbole linguistique

Dans les termes de la théorie des zones anthropiques, le premier déplacement correspond au passage du distal (la doxa comme Loi) aux zones identitaire et proximale¹. La sémosis (relation signifiant/signifié) est alors décrite comme un couplage entre ces deux zones, ce qui confirme le statut du signe linguistique comme objet transitionnel. Si l'on voulait filer la comparaison entre signe et monnaie, c'est la doxa comme Loi qui jouerait pour le signe linguistique le rôle de l'étalon pour la monnaie.

Qualifiée dans les deux cas par le déplacement, la différence entre le symbolisant et le symbolisé permet d'éviter la perte du sens : leur identité équivaldrait à une tautologie, qui, comme l'affirme justement Wittgenstein, est dépourvue de sens. En outre, on peut estimer

¹ Le premier déplacement correspondrait en termes hjelmsléviens à la relation de la substance du contenu à la forme du contenu, le second à celle qui unit le plan du contenu au plan de l'expression.

que la perte des différences qualifiées détruit la société (cf. Girard, 1998, p. 77-82). Appliqué au signe linguistique, l'idéal sémiotique de l'univocité équivaldrait non pas à la fixation de la signification, mais à la perte du sens.

Nous ne posons évidemment pas que les signes et les monnaies fonctionnent de même, ni que la théorie du sacrifice en soit la clé : nous rappelons simplement le principe général que les relations entre les zones du monde empirique sont déterminées par leur relation au monde transcendant du distal.

Si l'on prend le signe pour un objet physique comme un autre, comme font les positivistes, de Russell à Morris, sa référence devient évocatoire, donc inférentielle ; mais l'on n'a pas pris garde que le signe positiviste fonctionne comme le signe magique : si je prononce le nom, j'évoque le dieu au sens où je suscite sa présence, et où l'invocation entraîne l'évocation¹. Ce *si-alors* au fondement de toute magie partage la même forme que le signe indiciaire : *si* fumée, *alors* feu ; à cela près que l'inférence est prospective plutôt que rétrospective. Dans ce modèle, le signifiant et le signifié sont considérés comme deux objets en relation temporelle nécessaire : le mot, objet signifiant, suscite la chose, objet signifié ; et par transfert magique, la chose suscitée devient la source de l'évidence référentielle².

Paradoxalement, ce modèle du signe convient à l'animisme qui unit dans un même champ d'interactions les lois naturelles et les lois divines, tout comme au réductionnisme cognitif qui réduit les fonctionnements sémiotiques à des interactions physiques, conformément à son programme de naturalisation du sens.

Même l'indexicalité supposée de la signification, qui fonde la théorie de la dénotation directe chez Carnap et dans le positivisme logique, reste fondée sur un déplacement symbolique, puisque la « chose » indexée et le signe qui l'indexe, qu'il soit gestuel ou verbal, gardent des statuts différents³.

4.7. Des rythmes à la textualité

Le principe sémiotique de déplacement ne doit pas nous surprendre : le *si-alors* qui définit l'indice n'est que la réification d'une transformation, et toute unité sémiotique est un moment de stabilisation relative dans une série de transformations. On pourrait bien entendu le relier à un axe métonymique ; si cette caractérisation paresseuse n'explique rien, elle a le mérite de rappeler que les tropes sont des moments de transformations textuelles, premières par rapport aux relations sémiotiques qui les synthétisent et les décontextualisent. Tout modèle du signe est en effet une simplification et une réduction d'un modèle du texte. Le mot isolé n'existe pas⁴ : d'une part il appartient toujours à un paradigme qui a été institué en discours, et d'autre part un mot n'est qu'un passage d'un texte oral ou écrit.

La capacité de métamorphoses et de transposition s'illustre dans les actes de langage dits performatifs. Par exemple, si l'on dit d'une coupe de vin « Ceci est mon sang », ce

¹ D'où le tabou lexical et l'euphémisation nécessaires à la tranquillité humaine.

² « Si nous analysons, résume Frazer, les principes de la pensée sur lesquels est basée la Magie, nous trouverons qu'ils se résolvent à deux : le premier c'est que tout semblable appelle son semblable, ou qu'un effet est similaire à sa cause ; le second, c'est que deux choses qui ont été en contact à un certain moment continuent d'agir l'une sur l'autre, alors même que ce contact a cessé. » (Frazer, 1981, p. 41). Frazer nomme le premier principe « Loi de similitude » et le second « Loi de contact ou de contagion » ; ils seront repris par Freud et Mauss. La référence somme ces deux principes : par une relation de similitude, qui est celle de la représentation voire de l'iconicité ; et par une relation d'imposition ostensive qui continue à agir alors même que le référent a disparu (la rigidité du nom propre selon Kripke en est un exemple éclatant).

³ Par exemple, Pierre Jacob (1997) écrit la phrase : « Le sens du mot *banane* est une banane ». Si les italiques désignent le signifiant *banane*, pourquoi l'expression en romain « une banane », qui prétend en donner le sens, le contiendrait-il hors de tout contexte, sauf à le montrer ostensivement ? Ce mot pourrait parfaitement avoir pour acception un petit sac porté à la ceinture, ou une mèche frontale enroulée et outrageusement gominée, etc. Le sens de « banane » en romain n'est pas moins problématique que celui de *banane* en italiques. Si la tautologie rassure, certes, c'est qu'elle est vide de sens ; mais elle ne résoud aucun problème. Or quand Jacob emploie le mot *banane*, sans guillemets, il présuppose que ce mot n'est pas un mot, mais bel et bien un objet, c'est-à-dire qu'il est substitué à son référent, comme si sa thèse était démontrée du seul fait de sa formulation, ce qui reste caractéristiquement magique.

⁴ On a longtemps cherché le monosyllabe initial et les mots primitifs – les primitives sémantiques à la Wierzbicka héritent de ces spéculations prélinguistiques – sans s'aviser que le mot isolé perd le statut de symbole et ne peut fonctionner que comme un signal.

déplacement condense en signe des discours antérieurs¹. Il marque en outre l'incidence de la zone distale sur les autres². Si la textualité est une dimension propre aux performances symboliques, il convient donc de formuler des hypothèses sur les conditions de son émergence. Le rythme a tout d'abord une valeur hédonique (attestée aussi chez les chimpanzés) qui vaut pour elle-même et semble donc sans valeur adaptative. Produit par la voix ou le geste (claquement de mains, frappements de pieds), il organise des boucles sensori-motrices et semble ainsi au fondement de la sémiotisation du temps vécu : il permet de l'aménager à court terme par la prévisibilité au sein de la cellule rythmique et à moyen terme par la réitération de cette cellule.

Parmi les universaux empiriques, les seuls qui comptent dès lors que l'on quitte l'anthropologie philosophique *a priori*, les musiques ne sont pas moins universelles que les langues, les danses et les chants, très généralement associées. Le mythe est d'ailleurs souvent chanté³ voire dansé. En linguistique, la tradition grammaticale a certes gravement sous-estimé le rôle de la prosodie, qui intègre à la parole des facteurs mélodiques.

Toutefois, dans leurs études sur la métrique française, Roubaud et Lusson ont pu montrer que ses règles étaient une extension des propriétés de la langue. Mais qu'elles soient prosodiques ou métriques, ces affinités peuvent être interprétées en un autre sens : la langue serait la spécialisation de formes musicales.

Dans l'ontogenèse, non seulement les vocalisations rythmiques précèdent l'apprentissage de la parole, mais elles le conditionnent et l'enfant apprend à parler par répétition de formes mélodiques⁴. La prédication apparaît à partir de l'acquisition de contours prosodiques entendus et répétés (cf. Bertoncini et Boysson-Bardies, 2000, pp. 104-106).

Outre la conventionnalité paradigmatique et la contextualité syntagmatique, la troisième caractéristique du symbole saussurien est d'être doublement articulé : or le principe de la double articulation se résume à la sémantisation de groupes phoniques. Le nombre des groupes étant évidemment plus élevé que celui des phonèmes, qui ne dépassent jamais quelques dizaines, les capacités sémantiques sont élevées, voire indéfinies, si l'on tient compte des variations contextuelles.

Au palier du texte, comme les cellules rythmiques, les formes sémantiques sont aussi l'objet de réitérations avec variations⁵. On peut caractériser le récit, forme ordinaire d'organisation des mythes, par ses structures de variation. Ces structures sont mémorisables, dans la mesure où la mémoire est une réeffectuation, qui se guide par les normes de récurrence phonique et sémantique — dont les isotopies sémantiques et phoniques sont la forme la plus simple. Au palier textuel, l'obsession sémantique prend ainsi le relais du rythme prosodique au palier de la période.

4.8. L'abstraction

L'abstraction est une condition générale de passage au distal. Si les oiseaux et les primates obtiennent de bons scores, les humains se flattent des meilleures capacités d'abstraction, définie comme capacité de sélectionner des formes et de les reconnaître dans un autre contexte, voire dans une autre modalité sensorielle. Comme l'ont montré les travaux de la *Gestalt* sur les images virtuelles, poursuivis naguère par ceux de Kanisza, la reconnaissance de formes transposées est en outre créatrice, et il semble que nos illusions, dues au caractère « descendant » ou imaginaire de nos perceptions, l'emportent nettement sur celles de nos cousins simiens.

Si ces travaux portent pour l'essentiel sur l'abstraction visuelle, l'abstraction auditive garde un rôle fort important pour reconnaître les transpositions dans les variations. Ainsi, des enfants de deux mois peuvent reconnaître le même mot prononcé de douze manières différentes par la même personne ou par douze personnes différentes. Ils l'emportent ainsi sur des chimpanzés du même âge. En outre, quand les singes se plaisent à produire des rythmes, ils n'utilisent pas la syncope, ni le contretemps, qui sont des jeux sur le rapport entre type abstrait et occurrence.

¹ Les sacrements sont techniquement des signes, tant en théologie qu'en sémiotique.

² Les exemples pourraient être multipliés ; cf. les intersignes de la mythologie celtique.

³ En Occident, la poésie fut chantée jusqu'au XVI^e.

⁴ Il n'est pas exclu que les phonèmes jouent un rôle analogue aux notes — ni qu'ils soient choisis comme elles en fonction de leurs capacités discriminatives pour l'ouïe humaine.

⁵ Nous avons évoqué naguère (2001b) une conception prosodique de la sémantique textuelle, à propos des rythmes sémantiques et de la théorie des transformations (métamorphismes et transpositions).

L'abstraction semble nécessaire non seulement à l'organisation de la temporalité, mais à la projection de formes dans les matériaux ouvrés : déjà les bifaces d'*Erectus*, voici plus d'un million d'années, témoignent d'un souci de la symétrie et d'un goût de la forme pure qu'il reste difficile de rapporter à une considération utilitaire (cf. Wynn, 1993, p. 774). L'évolution de l'art, ou du moins des figurations du Paléolithique, va sans doute, comme l'a suggéré Leroi-Gourhan, de signes abstraits à des représentations qui nous paraissent réalistes — ce qui est contraire à toutes les théories empiristes de l'abstraction. Des signes abstraits demeurent d'ailleurs très généralement associés aux représentations que nous jugeons réalistes : moins spectaculaires, ces groupes de points et lignes stéréotypées sont peut-être cependant à l'origine des idéogrammes et phonogrammes¹.

Enfin, même si l'obstination répétitive n'est aucunement absente du monde animal, les comportements obsessionnels semblent un des propres de l'homme, et l'on pourrait opposer le mode compulsif de l'action animale au mode obsessionnel de l'action humaine. Si l'on revient à la typologie générale des comportements humains, on peut distinguer les *activités* spontanées, incoercibles, liées à la zone identitaire (se gratter), les *actions* permettant le couplage avec la zone proximale (chasser, se déplacer), et les *actes* permettant le couplage avec la zone distale (cf. l'auteur, 1999). Les actes typiques que sont les rituels, par la planification, et leur retour répété et invariable, témoignent tout à la fois d'une structure obsessionnelle et de la relation avec des objets absents. De fait, les objets distaux sont eux-mêmes le produit sémiotique et représentationnel des actes évocatoires. L'abstraction est une distalisation, et les sciences mêmes relèvent de la zone distale, où elles rivalisent avec les religions.

5. Evolution et anthropologie linguistique

Même quand elle débouche sur une sémiotique pour laquelle l'homme et le langage trouvent leur « origine » l'un dans l'autre, une anthropologie linguistique ne peut être fondée seulement sur les catégories générales du langage.

On étudie la phylogenèse et l'ontogenèse, mais il n'existe pas de terme synthétique pour désigner l'émergence des cultures ; cette lacune révèle un problème crucial, car l'humanisation prolonge l'hominisation. L'émergence des cultures est une question qui exige une coopération interdisciplinaire : le développement du monde sémiotique. Nous utiliserons le terme de *périgenèse* pour désigner la genèse de l'entour humain.

5.1. Genèse du sémiotique

On ne peut comprendre l'émergence du langage qu'au sein de l'*évolution générale du sémiotique*. Au cours de la périgenèse, il faut noter le développement conjoint des langues, des sépultures et des « arts » ou modes de décor mobilier ou pariétal.

En étudiant les contraintes anatomiques sur la phonation, Liberman avait émis l'hypothèse que la capacité de parole de l'*homo sapiens sapiens* — à qui l'on attribue environ 100 000 ans d'ancienneté, pouvait ne remonter qu'à 50 000 ans ; il est revenu récemment sur cette datation tardive. Il reste que c'est au paléolithique moyen que l'on trouve les premières sépultures indubitables, notamment dans la culture moustérienne. Elles annoncent la révolution symbolique du paléolithique supérieur (40 000-10 000) qui voit l'extension du peuplement à l'ensemble de la planète et sans doute la différenciation des langues et des cultures. Apparaissent les hiérarchies sociales, les sépultures collectives et des formes que l'on considère aujourd'hui comme « artistiques » : art mobilier, parures, sanctuaires en grotte. On assiste à une sémiotisation de l'environnement.

Les économies de chasse spécialisées qui se développent alors supposent une programmation des stockages sur l'année et en quelque sorte l'organisation du futur : sa découverte est une conquête sémiotique.

5.2. La conquête du distal : de l'hominisation à l'humanisation

La formation des langues a généralement été conçue comme l'émergence progressive d'une classe de fonctions, diversement nommées : *symbolique* (dans la tradition sociologique de Durkheim et de Mauss), *mythique*, *narrative*. Chris Knight, Mark Turner, Merlin Donald leur attribuent une *fonction* sociale de cohésion du groupe, par des manifestations (danses, chants, etc.) fondées sur la capacité de représentation. Toujours avancée sous la pression du néodarwinisme, l'explication fonctionnelle nous semble un leurre car l'avantage adaptatif lui-même reste souvent

¹ Si l'on suit Leroi-Gourhan, cela compose une sorte d'idéographie sexuelle, qui est peut-être à l'origine de la pensée différentielle, c'est-à-dire de la pensée tout court.

invoqué sans être problématisé. En quoi l'évocation et la création d'objets absents sert-elle à s'adapter à l'environnement ? Par exemple, le mythe n'est pas moins meurtrier que salvateur, et Cassirer a pu ainsi décrire le nazisme comme l'irruption d'un mythe dans l'histoire.

Dans l'hypothèse que l'émergence du langage s'accompagne d'une conquête de l'absence, cette conquête reste sans fonction. L'évocation ou constitution de l'absence peut évidemment trouver toutes sortes d'utilités. Par exemple, dans des économies de horde où s'établit une division du travail entre les âges et les sexes, femmes et enfants pratiquant la cueillette et les hommes la chasse, on peut évoquer la spécialisation de pratiques, la formation de lexiques spécialisés, la description d'itinéraires pour s'éloigner du camp et en revenir. Ou encore, l'élaboration des techniques et leur transmission supposent l'usage du langage, comme l'a souvent rappelé Leroi-Gourhan. Mais ces utilisations bienvenues ne sont pas nécessairement des causes, et d'ailleurs l'objectif des sciences de la culture n'est pas d'expliquer d'un point de vue causal, mais de reconstituer des conditions.

Par rapport aux systèmes de communication des animaux, le langage permet une maîtrise de l'absence, par son autonomie à l'égard de la situation. Mais cela n'entraîne pas que nos ancêtres se soient simplement émancipés du *hic et nunc*. La création de la zone distale n'est pas une simple extension : elle remanie structurellement le couplage avec la situation dans la mesure où le rapport entre l'identitaire et le proximal est sous la rectio du distal. En d'autres termes, les rapports au sein du monde obvie sont déterminés par les rapports entre le monde obvie et le monde absent.

Il reste que cette absence est en fait peuplée de signes bien présents. Le niveau sémiotique est constitué de performances complexes : danses, parures, récits, chants, etc. À l'œuvre dans toutes les pratiques socialisées, il est déterminant dans certaines : jeux, fêtes, rites, etc., dont les manifestations créent de nouvelles coordonnées spatio-temporelles, celles des terrains de jeux, des espaces sacrés, etc.

Les performances complexes supposent une planification de l'action et donc un essor de l'imagination, responsable des intentions et des désirs à moyen et long terme — car Éros doit beaucoup à Logos. Or le support des images mentales, scripts et scénarios engagés par l'imagination est précisément le cortex préfrontal qui « traite » les objets absents.

Les performances sémiotiques ont pour caractéristique leur caractère téléologique, permis par la démarcation claire de leur début et de leur fin et sans doute lié à leur stylisation. Les capacités de planifier de l'action et celles de produire et d'interpréter des récits reposent sur des médiations sémiotiques communes. La singularité d'un texte ou d'une autre performance sémiotique réside dans le fait qu'il est tout à la fois action énonciative et action énoncée, narration et récit, *historia* et *res gestae*.

Dépassant la simple tension narrative vers un but lointain, qui permet le déploiement de chaînes d'action sémiotiques, le récit peut se découpler d'avec la situation et passer de l'événement au mythe. En effet les relations entre contexte et situation vont en proportion inverse (l'auteur, 1998), et, dans la mesure où elle est une extension de la contextualité, la textualité prend toute sa dimension quand elle s'autonomise à l'égard de la situation, et/ou suscite de nouvelles situations (la pratique de la lecture, par exemple).

La relation entre le développement du distal et celui de la textualité engage à revenir à la thèse du fondement mythique des cultures. Cette thèse de Vico, spéculative à son époque, pourrait trouver des prolongements aujourd'hui. Alors que le marxisme vulgaire avait accrédité la thèse que l'idéologie reflétait les rapports de production, on en vient à penser que l'idéologie — j'entends ici le patrimoine sémiotique — les conditionne. Par exemple, Jacques Cauvin (1994) établit que la sédentarisation vient après et non avant la « révolution symbolique » qui au Moyen-Orient se traduisit entre 10 000 et 9 500 ans avant notre ère, par l'apparition des premières représentations féminines. Elle aboutit assez vite à celles de la Déesse-Mère et du Dieu-Taureau, cela *avant* l'apparition de l'agriculture. En bref, la création de la Déesse-Mère n'est pas un reflet de l'agriculture mais une condition de son développement.

Les refus de certaines pratiques productives et modes de développement sont souvent l'effet d'interdits divers, et le respect de la nature se justifie par exemple par la crainte de troubler divers dieux ou esprits¹.

¹ Par exemple, certains responsables brésiliens avaient rêvé de mettre les Indiens au travail pour leur faire « exploiter la forêt », ce qui suppose une destruction de leur panthéon.

Depuis Max Weber, les historiens du capitalisme, en revanche, ont souvent souligné que la conception instrumentale de la nature ne peut se comprendre que dans un monde laïcisé ou séparé de Dieu (c'est ainsi qu'ils ont souligné le rôle de la Réforme).

En somme, la pratique reflète l'idéologie, dont elle est une part. Comme toute pratique comporte un niveau sémiotique, elle peut revêtir un aspect mythique. La thèse du primat de la base économique sur la superstructure idéologique suppose que l'économie s'édifie en silence et sans signes, alors même qu'elle est inspirée par des mythes, comme celui de la croissance, et repose sur une base fiduciaire purement irrationnelle.

La phylogénèse des cultures conduit à une sémiotisation générale de l'environnement. Dans le cas de l'espace, on relève une interprétation mythique des singularités naturelles : sources, rochers, montagnes (cf. le rôle central d'Ayers Rock dans la pensée aborigène), etc., qui cartographient des mythes — il suffit de lire une carte du Péloponnèse pour rappeler cette évidence. Les territoires, les aires de chasse ou de cueillette, les frontières déterminent autant de zones valuées subtilement hiérarchisées.

Dans le domaine temporel, on passe du temps darwinien de l'évolution, lent et à discontinuités externes, au temps lamarckien de la culture, rapide, valué, à discontinuités endogènes, revêtant diverses formes : généalogique, calendaire, historique. Comme celle de l'espace, la culturalisation du temps se traduit par la création de zones valuées, d'où la théorie des âges de l'humanité, celle des stades historiques, etc.

La conquête du distal s'affirme complètement par le statut médiateur des signes. Par exemple, l'hypothèse chamanique reprise aujourd'hui par des préhistoriens comme Jean Clottes, fait des parois des grottes des séparations d'avec l'au-delà, et des mains dessinées par crachis des traces de « passage ». Mais surtout, l'art pariétal, à la différence de l'art mobilier et de la parure, témoigne d'une autonomisation des signes à l'égard du *hic et nunc*. Les grottes ornées ne sont pas ordinairement des lieux d'habitation, mais vraisemblablement des sanctuaires. L'œuvre peinte se détache dans l'espace comme dans le temps. Par opposition aux fétiches, comme les parures, qui restent associés à ceux qui les créent et qui les portent, elle s'affirme comme une idole¹.

L'autonomisation des signes et l'extension du distal ont une grande importance dans la constitution d'un monde sémiotique en expansion. C'est dans l'art pariétal que l'on trouve les premières abstractions graphiques (lignes, groupes de points, etc. dont procèdent peut-être les figurations ultérieures). Si ces signes sont codifiés, ils forment peut-être une idéographie, et, pour autant que l'on fasse commencer l'histoire avec l'écriture, on pourrait la faire reculer aux premiers tracés pariétaux.²

Avec l'écriture, on entre dans l'histoire, celle des historiens. D'une façon plus profonde qu'il n'y paraît, l'écriture introduit un nouveau type de temporalité. Elle permet de définir des intervalles, des mesures, des inventaires. L'apparition de documents (et non plus seulement de monuments) ouvre une dimension critique : le lecteur peut faire sécession dans le temps et dans l'espace, et l'écrit devient lui-même un témoin du distal (comme en témoignent les religions du Livre).

L'objectivation des documents permet en outre les débats et conjectures, le développement de la réflexivité qui érige en sciences des savoirs et des techniques. Enfin, la grammatisation permet une maîtrise sociale et politique des langues. Nous avons retracé ailleurs les étapes de la « révolution symbolique » continuée qui conduisent de l'écriture à l'imprimerie, puis à la numérisation (cf. 2001b, ch. 2). Retenons que le monde « virtuel » est sans doute un développement ultime et gigantesque de l'inscription pariétale, mais devenu familier et peuplé de fétiches : les démiurges y pullulent, dans le crépuscule des idoles.

Bref, les cultures, et au premier chef les langues, permettent à l'humanité de passer d'une évolution continue à une évolution discontinue et cumulative. À supposer que les conditions environnementales aient jamais eu le rang de causes déterminantes, elles le perdent, et l'humanité a pu s'adapter à peu près à tous les milieux — quitte à les adapter dangereusement à elle.

¹ Sur la distinction entre les fétiches, qui peuplent la frontière empirique, et les idoles, qui occupent la frontière transcendante, cf. l'auteur, 2001b.

² L'histoire des cultures commence quand les sociétés deviennent « chaudes », au Paléolithique supérieur, et l'on pourrait peut-être faire reculer l'histoire aux premiers graphismes, bien avant l'écriture.

Plus exactement, elle a modifié à son usage la notion même de milieu, car l'environnement humain, tout à la fois naturel et culturel, se compose d'un milieu physique et d'un entour sémiotique et représentationnel. Cette conquête de la liberté, évidemment liée à l'autonomie du sémiotique, fait de l'histoire le facteur déterminant de l'évolution. De façon concordante, la transmission du patrimoine sémiotique, par les règles d'alliance notamment, détermine ou du moins contraint fortement celle du patrimoine génétique.

Références bibliographiques

- Andler (D.) & al. 1992. *Philosophie et cognition — Colloque de Cerisy*. Bruxelles : Mardaga.
- Auroux (S.). 1998. *La raison, le langage et les normes*. Paris : P.U.F.
- Benveniste (E.). 1966. *Problèmes de linguistique générale*. Paris : Gallimard.
- Bergougnoux (G.). 2005. « L'origine du langage, mythes et théories ». in : Hombert (J.-M.), (ed.). pp. 14-41.
- Bertoncini (J.) & Boysson-Bardies (B. de). 2000. « La perception et la production de la parole avant deux ans ». in : Kail (M.) & Fayol (M.), (eds.). *L'acquisition du langage*. Paris : P.U.F., t. I, Ch. IV.
- Bickerton (D.). 1990. *Language and Species*. Chicago : University of Chicago Press.
- Bickerton (D.). 1996. *Language and Human Behavior*. Londres : University College London Press.
- Boehm (C.). 1988. « Vocal communication of Pan troglodytes : « triangulating » to the origin of spoken language ». in : Wind (J.) & al., (eds.). *Language Origin — A Multidisciplinary Approach*. Dordrecht : Kluwer.
- Brossut (R.). 1996. *Phéromones et communication chimique chez les animaux*. Paris : Belin.
- Bruner (J.). 1983. *Le développement de l'enfant, savoir-dire, savoir-faire*. Paris : P.U.F.
- Bruner (J.). [1983] 1987. *Comment les enfants apprennent à parler*. Paris : Retz.
- Cauvin (J.). 1994. *Naissance des divinités, naissance de l'agriculture*. Paris : Éditions du CNRS.
- Cavalli-Sforza (L.) & al. 1994. *The History and Geography of Human Genes*. Princeton : Princeton University Press.
- Changeux (J.-P.). [1970] 2000. « L'inné et l'acquis dans la structure du cerveau ». in : *La Recherche*, 331, pp. 19-31.
- Changeux (J.-P.). 2002. *L'homme de vérité*. Paris : Odile Jacob.
- Chomsky (N.). 1984. « La connaissance du langage ». in : *Communications*, 40, pp. 7-34.
- Clairis (C.). 2005. *Vers une linguistique inachevée*. Louvain : Peeters.
- Clottes (J.). 1998. « La détermination des figures animales et humaines dans l'art paléolithique européen ». in : *Voyage en préhistoire*. Paris : La Maison des roches, pp. 153-188.
- Cornuéjols (M.). 2000. *Sens du mot, sens de l'image*. Paris : L'Harmattan.
- Cuxac (C.). 2005. « Des signes et du sens ». in : Hombert (J.-M.), (ed.). pp. 196-214.
- Cyrułnik (B.). 1995. *La naissance du sens*. Paris : Hachette.
- Cyrułnik (B.). 2000. *Les nourritures affectives*. Paris : Odile Jacob.
- Dahbi (A.), Jaisson (P.), Lenoir (A.) & Hefetz (A.). 1998. « Comment les fourmis partagent leur odeur ». in : *La Recherche*, 314, pp. 32-35.
- Descola (P.). 1992. *Les lances du crépuscule*. Paris : Plon.
- Dessales (J.-L.). 2000. *L'origine du langage*. Paris : Hermès.
- Donald (M.). 1991. *Origins of the Modern Mind*. Cambridge (MA) : Harvard University Press.
- Donald (M.). 1993. « Précis of *Origins of the Modern Mind* : Three stages in the evolution of culture and cognition ». in : *Behavioural and Brain Sciences*, 16, pp. 737-791.
- Duclos (J.) & Joulian (F.). 1998. *La culture est-elle naturelle ?* Paris : Éditions Errance.
- Dumont (L.). 1991. *L'idéologie allemande*. Paris : Gallimard.
- Edelman (G.). 1992. *Biologie de la conscience*. Paris : Seuil.
- Fodor (J.). 1980. *Representations*. Cambridge (MA) : M.I.T. Press.
- Fodor (J.). 1990. *A theory of content and others essays*. Cambridge (MA) : M.I.T. Press.
- François (F.). 1993. *Pratiques de l'oral*. Paris : Nathan.
- Frazer (J.-G.). [1890-1915] 1981. *Le Rameau d'or, Le roi magicien dans la société primitive, Tabou et les périls de l'âme*. Paris : Robert Laffont, Coll. « Bouquins ».
- Gamkrelitze (T.). 2005. « La théorie glottogonique de Marr et l'isomorphisme structural entre les codes génétique et linguistique ». in : Sériot (P.), (ed.). pp. 139-142.
- Gärdenfors (P.). 1994a. « Speaking about inner environment ». in : *Of thoughts and Words - Proceedings of the Nobel symposium on the relation between language and mind* [à paraître].
- Gärdenfors (P.). 1994b. *Cued and detached representations in animal cognition* [ms, 15 pages].
- Geertz (C.). 1973. *The Interpretation of Cultures*. New York : Basic Books.
- Girard (R.). [1972] 1998. *La violence et le sacré*. Paris : Hachette.
- Glaserfeld (E. von). 1977. « Linguistic communication : theory and definition ». in : Rumbaugh (D.-M.), (ed.). *Language Learning by a Chimpanzee*. New York : Academic Press, pp. 55-71.
- Gopnik (M.). 1982. « Some distinctions among representations ». in : *Behavioural and Brain Sciences*, 5, pp. 378-379.
- Greenberg (J.-H.). 1987. *Language in the Americas*. Stanford : Stanford University Press.

- Gusdorf (G.). 1966. *Les sciences humaines et la pensée occidentale, t. I : De l'histoire des sciences à l'histoire de la pensée*. Paris : Payot.
- Hagège (C.). 1982. *La structure des langues*. Paris : P.U.F.
- Hermann (B.) & Hummel (S.), (eds.). 1994. *Ancient DNA : recovery and analysis of genetic material from paleontological, archaeological, museum, medical and forensic specimens*. Berlin : Springer.
- Hocart (A.-M.). [1927] 2005. *Au commencement était le rite*. Paris : La Découverte.
- Hockett (C.-F.). 1960. « The origin of speech ». in : *Scientific American*, 203, 3, pp. 88-96.
- Hombert (J.-M.). 2005. *Aux origines des langues et du langage*. Paris : Fayard.
- Houdé (O.). 1997. *Rationalité, développement et inhibition*. Paris : P.U.F.
- Hovelacque (A.). 1876. *La linguistique*. Paris : Reinwald.
- Huizinga (J.). 1954. *Homo ludens*. Paris : Gallimard.
- Humboldt (W. von). 1974. *Introduction à l'œuvre sur le kavi et autres essais*. Paris : Seuil [tr. Pierre Caussat].
- Humboldt (W. von). 1995. *Le dix-huitième siècle. Plan d'une anthropologie comparée*. Lille : P.U.L., Éd. Jean Quillien [cf. 1903. *Gesammelte Schriften*, 1, Éd. A. Leitzmann. Berlin : Behr, pp. 377-410].
- Humboldt (W. von). 2000. *Sur le caractère national des langues et autres écrits sur le langage*. Paris : Seuil [introduction, traduction et édition par Denis Thouard].
- Hurford (J.-R.), Studdert-Kennedy (M.) & Knight (C.). 1998. *Approaches to the Evolution of Language*. Cambridge : Cambridge University Press.
- Jacob (P.). 1997. *Pourquoi les choses ont un sens*. Paris : Odile Jacob.
- Jisa (H.). 2005. « La langue façonne-t-elle le monde ? ». in : Hombert (J.-M.), (ed.). pp. 254-272.
- Johnson (M.). 1992. « Philosophical implications of cognitive semantics ». in : *Cognitive Linguistics*, 3-4, pp. 345-366.
- Johnson-Laird (P.-N.). 1983. *Mental Models*. Cambridge : Cambridge University Press.
- Jucquois. 1986. « Aspects anthropologiques de quelques notions philologiques ». in : *Cahiers de l'institut de linguistique de Louvain*, 12, 1-2, pp. 183-248.
- Kail (M.) & Fayol (M.), (eds.). 2000. *L'acquisition du langage*. Paris : P.U.F. [2 tomes].
- Kyheng (R.). 2006. *Le langage, faculté, ou généralisation des langues ? Une enquête saussurienne*. Site : <http://www.revue-texto.net>
- Langaney (A.). 1988. *Les Hommes*. Paris : Armand Colin.
- López-García (Á.). 2005. *The grammar of genes — How the Genetic Code Resembles the Linguistic Code*. Berne, Berlin, Bruxelles, Francfort Main, New York, Oxford, Wien : P. Lang.
- Lorblanchet (M.). 1988. « De l'art pariétal des chasseurs de rennes à l'art rupestre des chasseurs de kangourous ». in : *L'anthropologie*, 92, 1.
- Lotman (I.). 1998. *La sémiotique*. Limoges : Pulim.
- Marr (I.-N.). [2005] 1925. « Sur l'origine du langage ». in : Sériot (P.), (ed.). 2005. pp. 383-387.
- Mascie-Taylor (C.), (ed.). 1988. *Biological Aspects of Human Migration*. Cambridge : Cambridge University Press.
- Pollock (J.-Y.). 1997. *Langage et cognition*. Paris : P.U.F.
- Popper (K.). 1972. *Objective Knowledge : an Evolutionary Approach*. Oxford : Oxford University Press.
- Pottier (B.). 1962. *Systématique des éléments de relation*. Paris : Klincksieck.
- Rastier (F.). 1987. *Sémantique interprétative*. Paris : P.U.F.
- Rastier (F.). 1989. *Sens et textualité*. Paris : Hachette.
- Rastier (F.). 1990. « La triade sémiotique, le trivium et la sémantique linguistique ». in : *Nouveaux actes sémiotiques*, 9.
- Rastier (F.). 1991. *Sémantique et recherches cognitives*. Paris : P.U.F.
- Rastier (F.). 1992. « Réalisme sémantique et réalisme esthétique ». in : *TLE*, 10, pp. 81-119.
- Rastier (F.). 1993. « La sémantique cognitive — Éléments d'histoire et d'épistémologie ». in : *Histoire, Épistémologie, Langage*, XV, 1, pp. 153-187.
- Rastier (F.). 1997. « Herméneutique matérielle et sémantique des textes ». in : Salanskis (J.-M.), Rastier (F.) & Scheps (R.), (eds.). *Herméneutique : textes, sciences*. Paris : P.U.F.
- Rastier (F.). 1998. « Le problème épistémologique du contexte et l'interprétation dans les sciences du langage ». in : *Langages*, 129, pp. 97-111.
- Rastier (F.). 1999. « Action et récit ». in : *Raisons Pratiques*, 10, pp. 173-198.
- Rastier (F.). 2001a. « L'action et le sens. — Pour une sémiotique des cultures ». in : *Journal des Anthropologues*, 85-86, pp. 183-219.

- Rastier (F.). 2001b. *Arts et sciences du texte*. Paris : P.U.F.
- Rastier (F.). 2004. « Sciences de la culture et post-humanité ». in : *Actes 2003-2004 du groupe de travail STP (du Sujet ; Théorie et Praxis) de la Maison des Sciences de l'Homme de Paris. Volume VII*. Paris : Éditions de la MSH, pp. 7-27 [Rééd. <http://www.revue-texto.net>]
- Rastier (F.) & Bouquet (S.), (ed.). 2002. *Une introduction aux sciences de la culture*. Paris : P.U.F.
- Rastier (F.), Cavazza (M.) & Abeillé (A.). 1994. *Sémantique pour l'analyse*. Paris : Masson.
- Ruhlen (M.). 1997. *L'origine des langues*. Paris : Belin.
- Salanskis (J.-M.). 1997. « Herméneutique et philosophie du sens ». in : Salanskis (J.-M.), Rastier (F.) & Scheps (R.), (eds.). *Herméneutique : textes, sciences*. Paris : P.U.F.
- Saussure (F. de). [1916] 1972. *Cours de linguistique générale*. Paris : Payot.
- Saussure (F. de). 2002. *Écrits de linguistique générale*. Paris : Gallimard [ed. Rudolf Engler et Simon Bouquet].
- Saussure (F. de). 1997. *Deuxième cours de linguistique générale (1908-1909)*. Oxford : Pergamon.
- Savage-Rumbaugh (E.-S.) & Rumbaugh (D.-M.). 1993. « The emergence of language ». in : Gibson (K.-R.) & Ingold (T.), (eds.). *Tools, Language and Cognition in Human Evolution*. Cambridge : Cambridge University Press, pp. 86-108.
- Savage-Rumbaugh (E.-S.), Shanker (S.-G.) & Taylor (T.-J.). 1998. *Apes, Language, and the Human Mind*. Oxford : Oxford University Press.
- Schleiermacher (F.). 1987. *Herméneutique*. Genève : Labor et Fides.
- Sériot (P.), (ed.). 2005. « Un paradigme perdu : la linguistique marriste ». in : *Cahiers de l'ILSL*, 20.
- Sjölander (S.). 1993. « Some cognitive breakthroughs in the evolution of cognition and consciousness, and their impacts on the biology of language ». in : *Evolution and Cognition*, 3, pp. 3-10.
- Sperber (D.). 1996. *Explaining Culture — A Naturalistic Approach*. Oxford : Blackwell.
- Sperber (D.). 2001. *Vers une lecture sans écriture ?* Bibliothèque publique.
- Sperber (D.) & Orrigi (G.). 2005. « Pourquoi parler, comment comprendre ? ». in : Hombert (J.-M.), (ed.). pp. 236-254.
- Steinthal (H.). 1851. *Der Ursprung der Sprache*. Berlin : Cohn-Weigert.
- Stewart (J.), Scheps (R.) & Clément (P.). 1997. « Phylogénèse de l'interprétation ». in : Salanskis (J.-M.) & al., (eds.). *Herméneutique : textes, sciences*. Paris : P.U.F.
- Tattersall (I.). 1999. *L'émergence de l'homme*. Paris : Gallimard.
- Thomason (S.) & Kaufman (T.). 1988. *Language Contact, Creolization and Genetic Linguistics*. University of California : University of California Press.
- Thouard (D.), (ed.). 2000. *Humboldt - Sur le caractère national des langues et autres écrits sur le langage*. Paris : Seuil.
- Trabant (J.). 1999. *Traditions de Humboldt*. Paris : Éditions de la MSH.
- Trombetti (A.). [1905] 1962. *L'unità d'origine del linguaggio*. Bologna : Scuola Grafica « Civitas Dei ».
- Turner (M.). 1996. *The Literary Mind*. Oxford : Oxford University Press.
- Uexküll (J. von). [1934] 1956. *Mondes animaux et mondes humains*. Paris : Denoël.
- Vauclair (J.). 1990. « Primate cognition : from representation to language ». in : Parker (S.-T.) & Gibson (K.-R.), (eds.). « *Language* » and intelligence in monkeys and apes. Cambridge : Cambridge University Press, pp. 312-329.
- Vauclair (J.) & Fagot (J.). 1993. « Can a saussurian ape be endowed with episodic memory ? ». in : *Behavioural and Brain Sciences*, 16, pp. 772-773.
- Victorri (B.). 1999. « La place de la fonction narrative dans l'émergence du langage et la structure des langues ». in : *Théorie Littérature Enseignement*, 17, pp. 23-38.
- Victorri (B.). 2005. « Les « mystères » de l'émergence du langage ». in : Hombert (J.-M.), (ed.). pp. 212-234.
- Winnicott (D.-W.). [1971] 1975. *Jeu et réalité. L'espace potentiel*. Paris : Gallimard.
- Wolpoff (M.) & Caspari (R.). 1997. *Race and Human Evolution. — A fatal attraction*. New York : Simon & Schuster.
- Wynn (T.). 1993. « Archaeological Evidence for mimetic mind and culture ». in : *Behavioural and Brain Sciences*, 16, p. 774.
- Yau (S.-C.). 1992. *Création gestuelle et début du langage*. Hong Kong : Langages croisés.



Mai 2006

1. Introduction

Continuel sujet de perplexité pour les linguistes : comment, d'une même faculté de langage propre aux humains, peut jaillir l'immense diversité des langues ? 6000 sont en effet parlées sur la planète. Cette question resterait à jamais insoluble sans un recours aux autres sciences. Si un coin du voile commence à être levé, c'est que l'archéologie, la paléoanthropologie, la génétique des populations, la psychoneurologie, l'éthologie et d'autres disciplines viennent, avec la linguistique, instruire le dossier. Pas à pas, elles analysent le présent et remontent le temps.

2. Où commence la diversité ? Le langage et les langues

La faculté de langage est inscrite au génome humain, sans qu'il y ait à proprement parler de « gènes du langage », mais simplement une implication de certains gènes dans des processus qui culminent dans le langage. Elle est considérée aujourd'hui par les neuropsychologues comme reposant sur la spécialisation d'aires cérébrales qui contribuent par ailleurs à d'autres aspects de la cognition. Grâce notamment aux techniques de l'imagerie cérébrale, on peut entrevoir qu'il existe une grande hétérogénéité dans l'exploitation de la faculté de langage : complexité et flexibilité en sont les caractéristiques principales. Et ce qui est observable sur les populations actuelles peut, dans une certaine mesure, être extrapolé sur des *Homo sapiens* un peu plus anciens, puisque leurs capacités langagières semblent avoir été similaires (Mandel, 2003, p. 48 ; Hublin, 2003, p. 26, 2005, p. 116 ; Nespoulous, 2005, pp. 440, 447-448). Selon l'hypothèse la plus vraisemblable des paléoanthropologues, cette aptitude de langage moderne se serait mise en place par paliers successifs, soit dès l'émergence de notre espèce, il y a environ 150 000 ans, soit lors de l'explosion culturelle il y a 50 000 ans (Coupé, 2005, p. 169). Il se trouve que le canal visuo-gestuel et le canal vocal-auditif sont également susceptibles de construire du sens. Mais, par avantage adaptatif, un langage articulé à caractère vocal pourrait avoir émergé à travers la longue durée (voir Annexes 1, Encadré 2 : « Quatre domaines de divergence »), à partir d'un protolangage rudimentaire de *Homo ergaster* et de *Homo erectus*. Ce protolangage se présenterait comme un système de communication peu performant, en étroite dépendance avec le « ici et maintenant » (Hublin, 2003, p. 30 ; Hombert, 2005, pp 462-463). Dès lors, la nette distinction opérée par un noyau de linguistes tels que André Martinet (1968, p. XI) et Claude Hagège (2003, p. 154) entre aptitude au langage et l'une de ses cristallisations : les langues, apparaît comme pleinement justifiée. Les généticiens des populations et les paléodémographes, pour leur part, nous apprennent que les premiers peuplements d'*Homo sapiens* étaient très réduits, les espaces géographiques probablement restreints, et les interactions humaines faibles. Ce qui présuppose une diversité linguistique elle aussi très ténue (Coupé, 2005, p. 183). Ainsi, notre dispositif cérébral tout comme les conditions d'évolution des *Sapiens* semblent favoriser l'éclosion de la diversité. Avec cette toile de fond éclairante, le linguiste peut, à son tour, apporter sa pièce à l'édifice.

3. Langues, dialectes et idiolectes

C'est un fait notoire que les 6000 langues planétaires se subdivisent en unités plus petites, les dialectes et sous-dialectes, dont les locuteurs ne se comprennent pas toujours entre eux. La distinction entre langue et dialecte n'étant pas toujours claire, ne serait-ce qu'en raison de l'intervention des États. Le mandarin est, en réalité, un dialecte officialisé. Il en va de même du français, issu du « francien » naguère parlé dans le Bassin parisien. Dans bien des cas, on baptise « langue », à l'inverse, une abstraction représentant un continuum de parlers et de dialectes sur un espace donné. Il en va ainsi du tcherkesse, au Caucase, ou du quetchua,

l'ancienne langue des Incas. Les dialectes se fragmentent à leur tour en unités encore plus petites qui aboutissent aux idiolectes, systèmes linguistiques propres aux individus. Les abstractions « langue » et « dialecte » ne sauraient conduire à une réification de ces concepts pour tous ceux qui acceptent de les voir de manière dynamique comme un perpétuel ajustement d'idiolectes (premiers degrés des contacts linguistiques), indéfiniment combiné avec une constante application du code, ou une dialectique typique des comportements humains. Comme l'énonce Salikoko S. Mufwene (2001, p. 2) :

To be sure, a communal language is an abstraction inferred by the observer. It is an extrapolation from I [internalized] –language whose speakers communicate successfully with each other most of the time.

Ce n'est que par abus de langage métaphorique (« guerre des langues », « langues dominantes », « langues dominées » etc.) qu'on peut être amené à les considérer comme des entités évoluant indépendamment des sujets parlants.

4. Vie, mort et renaissance des langues

Chacun sait ou croit savoir, par ailleurs, que les langues « naissent », « meurent », et occasionnellement « renaissent ». Mais c'est encore une métaphore. L'une des multiples que les linguistes utilisent par emprunt à la biologie, en l'occurrence au vitalisme du XIXe siècle, par analogie avec les espèces vivantes, ou en puisant dans le fonds religieux, tout en gardant à l'esprit les limites de cette transposition. Elle ne doit pas faire perdre de vue que le second cas (la mort) est plus facile à détecter que le premier. La métaphore biologique a d'ailleurs ses failles : une même langue peut mourir dans un lieu et survivre dans un autre. Le français est mort dans le Maine mais pas au Québec ; le yidiche est mort en Alsace, mais vit en Israël, en Argentine ou à New York. Une nouvelle langue peut apparaître par scission au sein d'une langue donnée, et dans ce cas échapper à la conscience de plusieurs communautés de locuteurs, dont les usages divergent peu à peu jusqu'à franchir une limite, au-delà de laquelle ils cessent de se comprendre. Au début de ce siècle, Antoine Meillet (1965, réédition), aimait à rappeler qu'à la fin de l'Empire romain, une longue chaîne d'hommes, dont chacun avait eu le sentiment de parler latin, était en fait en train de donner naissance à toute la gamme des langues romanes.

Dans un passé récent et aujourd'hui même, nous assistons dans plusieurs régions du monde à d'autres formes de « naissance » : les créoles (au sens large, typologique du terme) et les pidgins. Les créoles (au sens étroit, socio-historique du terme, cf. Hazaël-Massieux, 2005, pp. 3-17), nés au cours des colonisations européennes des XVIIe-XVIIIe siècles, sont le parler de descendants d'esclaves noirs mis en contact avec la langue de leurs maîtres, anglais ou français par exemple. Par extension, on peut dénommer créoles et créolisation des langues ainsi qu'un processus qui impliquent une dynamique de contacts très intenses, mais ces concepts ne font pas l'unanimité. Dans cette perspective, le français d'Abidjan, le joual, ou le cap-verdien sont des créoles, voire pour certains qui poussent le trait, l'anglais ou le chinois. Une telle position est notamment combattue par Robert Chaudenson (2000) et Mufwene (1986, 2001). Les pidgins, attestés historiquement avant les créoles, sont le parler plus rudimentaire de communautés de langue différente mais forcées de commercer entre elles. Leurs systèmes sont plus simples, leur fonction de communication restreinte, et en outre, ces parlers n'ont pas le statut de langues maternelles. Toutefois, dans certains cas, ils évoluent vers des pidgins étendus (« expanded pidgins »), véritables *lingua franca*, plus complexes, et parfois même natives.

Ils sont dès lors analogues aux créoles (cf. Mufwene, 2001, p. 7). Quoi qu'il en soit, il s'agit, dans les deux cas, de la manifestation extrême d'une situation, où le contact prolongé entre deux ou plusieurs langues conduit à une réorganisation complète de divers systèmes linguistiques. Comme l'énonce encore Mufwene (2001, p. 192) :

Creole vernaculars are not outcomes of abnormal, unusual, or unnatural developments in language evolution. Rather, they make more evident restructuring processes that must have taken place in the evolution of other languages.

En outre, vifs sont les débats autour de l'ontogénèse et de la phylogénèse d'un côté (cf. Bickerton, 1981), et des substrats de l'autre, suscités par cette situation d'observatoire en quelque sorte. Il se produit, comme le décrit Hagège (1985, pp. 29-41), sur fond, encore mal cerné, des facultés cognitives propres à l'espèce humaine. A un autre extrême, plus rare, il

arrive qu'une langue « morte » renaisse du fait d'une volonté politique et idéologique. Le cas le plus spectaculaire est l'hébreu moderne. La situation quasi-expérimentale de cet idiome ayant cessé d'être vernaculaire depuis au moins 1 500 ans et réanimé par un puissant volontarisme oblige l'observateur à un recentrage vers l'homme architecte de sa langue. Elle s'inscrit dans un contexte historique qui l'autorise (nationalisme et création d'un État) et un contexte culturel (courant « millénariste » et courant « esthétique » hérité de l'époque des Lumières) (cf. Szulmajster-Celnikier, 1997, pp. 141-142, 145).

4.1. Le nombre de langues : équilibre et rupture

Qu'en était-il dans le lointain passé ? Robert M.-W. Dixon (1997) affirme, en s'inspirant de la biologie évolutionniste, que l'histoire de la diversité des langues, à un niveau macroscopique, a été ponctuée de longues périodes d'équilibre, au cours desquelles le nombre de langues était relativement constant ou en augmentation notable, entrecoupées de phases de rupture qui voyaient mourir de nombreuses langues et en naître d'autres. La première phase d'équilibre aurait duré plus de 40 000 ans, de -50 000 à -10 000 avant le présent, à la fin du Paléolithique supérieur (cité par Peyraube, 2005, pp. 452-53). Daniel Nettle (2002) évalue globalement, sur la base de considérations démographiques, le nombre de langues d'alors entre 1 667 et 9 000. La première période de rupture remonte au début du Néolithique, lorsqu'apparaissent, en divers points du globe de manière parallèle, les premières sociétés d'agriculture et d'élevage, soit entre -8 500 et -3 000 ans. Des langues en remplacent d'autres, mais leur nombre demeure à peu près stable. Une deuxième période d'équilibre s'est alors instaurée, qui a perduré jusqu'à un passé récent, celui des grandes conquêtes maritimes européennes et chinoises du XVe siècle. Une poussée démographique a eu, à cette époque, pour corollaire un accroissement considérable de la diversité linguistique. Il est important de souligner que pour certains paléoanthropologues, éthologues et sémiologues, les premières traversées maritimes des *Sapiens* pourraient constituer un indice comportemental de l'émergence du langage. Au-delà de la préhistoire, les conquêtes maritimes pourraient être interprétées comme des événements susceptibles de changer la direction d'une trajectoire évolutive (en termes mathématiques, comme les « attracteurs » de la théorie du chaos). Toujours est-il que cette phase d'équilibre a probablement cessé aux XVe et XVIe siècles, période de la seconde rupture, toujours en cours. Par conséquent, le nombre de langues n'aurait cessé de croître jusqu'au XVIe siècle, pour commencer à décliner à partir de cette date (Peyraube, 2005, p. 455).

4.2. Les causes d'extinction

Les causes de la mort des langues sont aussi diverses que celles qui président à leur naissance. Certaines disparues sont demeurées célèbres en raison de l'impact de leur civilisation (cas du sumérien), de l'intérêt de leur écriture (l'étrusque) ou simplement de leur découverte récente (le hittite, voir Annexes 2, « Photo de divers types d'écriture »). Bien plus nombreuses, sont celles qui se sont évanouies sans laisser de trace. La cause d'extinction la plus courante est le contact prolongé entre deux langues. On a vu qu'il peut aussi être à la source de nombreux idiomes, et qu'il constitue l'un des moteurs du changement linguistique. Mais il peut apparaître encore comme un facteur d'acculturation, insidieuse ou agressive. Les langues disparaissent ainsi à la suite de processus de transformation, de substitution et d'extinction (Hagège, 2000, pp. 93-94). Les raisons profondes de cette disparition sont le plus souvent d'origine extralinguistique : politiques, économiques, culturelles, sociales ou liées à des catastrophes naturelles, ou provoquées par l'homme (Hagège, 2000, chap. VII).

4.3. Un crescendo dans l'extinction ?

Le phénomène semble s'être accéléré depuis le XVIe siècle, d'abord en raison de l'expansion et de la domination européennes, puis, plus récemment, sous la pression de la mondialisation des échanges et de l'urbanisation de populations rurales et dispersées. Selon les estimations, quelques centaines, voire plus d'un millier de langues risquent de mourir à court ou moyen terme. Pour Hagège (2000), une langue disparaît tous les quinze jours. À ce rythme, dans un siècle, la moitié de notre patrimoine pourrait s'évanouir. Exit le cupeño, le matipu, le sikiana ou autre yugh, qui viendront rejoindre les lointains gaulois, thrace, illyrien et celtibère.

Des familles entières, voire des macrofamilles, seront bientôt rayées de la carte. C'est notamment le cas de la famille des langues iénisséennes, dont le khet, parlé en Sibérie centrale,

est le dernier survivant. C'est encore le cas, bien que dans une moindre mesure, de la macro-famille khoïsan, parlée en Afrique du Sud et en Tanzanie. Certaines aires géographiques sont plus touchées que d'autres. Ainsi des langues paléo-sibériennes (dites encore paléo-asiatiques, ou hyperboréennes, dispersées sur l'immense territoire sibérien, surtout au Kamtchatka, comprenant le chukchi et le giliak), des langues indiennes d'Amérique du Nord, des langues aborigènes d'Australie. Processus naturel inévitable, ou produit de civilisations irresponsables ? Les opinions divergent. « Language vitality and endangerment as aspect of language evolution », énonce Mufwene, en tête de chapitre (2001, chap. VIII), dépassionnant le débat. « Halte à la mort des langues » dit, à rebours, le titre d'un récent livre de Hagège, déjà cité. Le premier a une approche écologique des phénomènes linguistiques, qui s'exprime en termes adaptatifs, de compétition et sélection langagière, de coûts et profits pour le locuteur. La vie et la mort des langues sont le résultat d'une dynamique d'intégration des individus et des groupes dans le corps socio-économique environnant. Dans le même esprit, Georges Mounin (1992, pp. 156-158) relativise l'attachement des locuteurs à la survie de leur langue ainsi que la corrélation entre langue, vision du monde et identité ethnique et culturelle. Pour le second, nourri par l'hypothèse Sapir-Whorf, reprise par les ethnologues et un certain nombre de linguistes comme James Matisoff (1980) et Dixon (1980), la valeur historique et anthropologique des visions du monde portées par ces langues disparaît avec elles, en même temps que des informations précieuses sur notre patrimoine linguistique. Leur perte est donc irréparable. Cependant, les décisions de la communauté parlante sont seules pertinentes.

5. Le foisonnement linguistique : trois modes de classification

Si l'on cherche à remonter dans le temps – à l'échelle des linguistes, la profondeur maximale se situe au Néolithique, de -8 000 à -10 000 ans avant le présent - comment appréhender le foisonnement des langues ? Quelles règles dynamiques sous-tendent cette turbulence ? (Voir Annexes 1, Encadré 1 : « La diversité des langues »).

5.1. La classification généalogique

Les linguistes ont longtemps privilégié l'approche généalogique. C'est aussi la mieux documentée. La métaphore de la « naissance » a pour corollaire celles de « langues mères » et de « langues filles » - bien que la seule réalité observable soit un processus insensible de diversification et que seul le recul de l'histoire permette de poser, avec certaines marges d'incertitude, des « familles » de langues (comme la famille tibéto-birmane), des « sous-familles » ou « branches » (comme la branche des langues slaves). La classification généalogique est fondée sur la statistique lexicale portant sur un vocabulaire de base, et sur l'identification de correspondances phonétiques régulières. Calquée sur l'évolutionnisme, cette méthodologie a donné lieu à d'importantes avancées au XIXe siècle et surtout au XXe siècle, à commencer par les domaines indo-européen et sémitique. Cependant, il montre aussi ses limites (Voir Annexes 2, Cartes couleur : « Les langues austronésiennes... » ; « Le Caucase : mosaïque linguistique... »). Certaines langues demeurent toujours inclassées, comme le japonais – dont on ignore s'il faut le rattacher à l'ensemble altaïque – ou sont des isolats, seuls représentants d'une famille et sans relation démontrée avec aucune autre, comme le sumérien ou aujourd'hui le basque ou le bouroushaski. On reviendra plus loin (sections 5.2 et 6) sur les limites de cette classification, et on s'arrêtera (section 7) sur l'exemple du proto-indo-européen.

5.2. La classification aréale

À ce modèle vertical et arborescent, se superpose et parfois s'oppose la classification horizontale dite aréale (par aires géographiques). C'est un concept venu de l'ethnologie, tandis que le premier est issu de la biologie. Il fait apparaître un mode d'appareillage de type sériel. Au sein de la diversité, les faits homologues observés ne s'expliquent pas nécessairement par une origine commune. Ils peuvent être simplement dus au voisinage et aux relations continues qui s'ensuivent. Les considérations géographiques ont l'avantage d'être moins abstraites que celles de type généalogique. En l'absence de données historiques, comme c'est souvent le cas, origine commune et contagion sont deux hypothèses également envisageables. D'où une révision de certaines reconstructions généalogiques. Ainsi, le modèle généalogique semble insuffisant et impropre à rendre compte de la véritable évolution historique de nombreux ensembles linguistiques, comme celui des langues bantoues, au sein du domaine négro-africain, ou du quetchua, au sein du domaine amérindien, ou du turc et du mongol, dans le domaine altaïque. Ainsi encore, l'unité du groupe ouralo-altaïque n'est plus reconnue, celle du chamito-sémitique fait l'objet de dissensions : famille ou union ? Évolution divergente ou convergente ? Voici plus d'un siècle, Johannes

Schmidt critiquait déjà la méthode comparative du modèle arborescent, pour proposer que les innovations dialectales se répandent comme des ondes (Wellen-Theorie). La théorie fut reprise dans les années 1930 par Nicolas S. Troubetzkoy (1939), qui voyait les langues se grouper en « maillons de chaîne », ou encore en portions d' « arc-en-ciel », idée adoptée plus tard par Roman Jakobson dans le domaine de la phonologie.

Cette approche commence aujourd'hui à porter ses fruits. Elle ne fait pas obstacle aux concepts de structure et de système (au sens saussurien où une langue est un système où tout se tient).

Bertil Malmberg (1990, p. 16) a analysé dans ce cadre le phénomène d'affaiblissement d'un système linguistique, que ce soit à la périphérie (géographique, sociale) ou par rencontre avec un autre système (bilinguisme). On a pu aussi évoquer l'idée de cellules aux membranes perméables réagissant les unes avec les autres, ou encore d'osmose (Mufwene 2001, p. 15).

Comme le formule Mounin (1992, p. 152), la « vieille tradition de l'impénétrabilité réciproque des grammaires » est à dépasser. C'est ainsi que l'on peut rendre compte, par exemple, de la curieuse unité du domaine balkanique (le bulgare et le grec notamment présentent des homologies). Cette méthode a aussi permis de mettre en évidence (Haudricourt, 1954), en Extrême-Orient, l'existence d'une vaste chaîne, le tibéto-birman et le chinois étant apparentés au miao (langue sino-tibétaine parlée en Chine du Sud, au Vietnam, au Laos et en Thaïlande) et à l'austro-asiatique (famille de quelque 150 langues, dont le khmer et le vietnamien), ceux-ci au thaï, ce dernier au malayo-polynésien (dit encore austronésien, famille comprenant notamment l'indonésien et les langues polynésiennes. Voir Annexes 2, carte « Les langues austronésiennes »). De même, de grandes homologies de traits grammaticaux et phonologiques entre les langues paléo-sibériennes et certaines langues avoisinantes (comme entre le chukchi et l'eskimo, ou entre le giliak et le coréen ou le japonais). Les affinités entre des groupes tels que l'ouralo-altaïque et le dravidien (sud de l'Inde) peuvent également s'expliquer en termes de communication prolongée et intime autant qu'en terme de familles (Voir Annexes 2, Schémas couleur : « Les langues ouraliennes »).

La théorie aréale révèle ses limites lorsque les ressemblances entre langues observées s'étendent sur d'immenses aires, parfois discontinues, ce qui laisse le choix entre de nombreuses hypothèses. Les contacts entre langues peuvent être à la source de la réduction de la diversité (par un processus de convergence), ou au contraire engendrer la diversité. On a tenté d'expliquer la multiplicité des langues de Mélanésie en les faisant dériver de mélanges assez récents entre les langues indonésiennes et les langues papoues, qui ont donné naissance à des variétés disparates d'indonésien pidginisé avec des substrats papous dans chaque groupe insulaire. En tout état de cause, la parenté doit aujourd'hui équilibrer les considérations sur divergence, convergence et mélange, sans privilégier aucun de ces processus.

5.3. La classification typologique

Aux côtés de la méthode généalogique et de l'approche aréale, un troisième mode de classification s'est constitué. C'est la méthode dite typologique. Elle range les langues en types en se fondant sur les similarités de structures linguistiques. On peut ainsi distinguer avec Hagège (1982, pp. 9-10) les langues à distinction verbo-nominale (comme les langues indo-européennes et sémitiques), à relative indistinction verbo-nominale (comme le nootka de l'île de Vancouver ou le santali d'Inde), ou mixtes de ce point de vue (comme les langues sino-tibétaines). Puis les langues à construction ergative, marquant formellement l'agent (comme le basque ou langues du Sud du Caucase), celles à construction accusative, marquant formellement le patient (comme le français ou le malais), ou encore des langues mixtes de ce point de vue (comme les langues océaniques ou australiennes). On distingue aussi, depuis le début du XIXe siècle, les langues essentiellement agglutinantes (suffixes mis bout à bout, comme dans les langues ouraliennes), flexionnelles (verbes conjugués, noms déclinés, comme en russe), isolantes (mots le plus souvent d'une seule syllabe, invariables et inanalysables, comme le chinois) ou encore les langues à très riche système consonantique (comme les langues caucasiennes) ou vocalique (comme les dialectes alémaniques et les langues polynésiennes).

Élaborée par les frères Schlegel, puis par Wilhelm von Humboldt, cette classification par types a considérablement évolué. Elle a montré qu'en certaines aires, une grande diversité généalogique s'accompagne d'un ensemble de traits typologiques communs : il en est ainsi dans la sous-région continentale de l'Asie du Sud-Est (Chine du Sud, Est de l'Inde), en Afrique

du Sud et de l'Est, sur la Côte Nord-Ouest et le Plateau limitrophe d'Amérique du Nord (qui comprend les plateaux des régions de Colombie et du Colorado et les hauts plateaux mexicains à l'est de la Sierra Madre). Mais force est de constater que familles de langues et types linguistiques coïncident rarement. Dans l'aire tibéto-birmane, la proximité généalogique ne va pas de pair avec la proximité typologique. Et la classification par types n'a pu éliminer l'arbitraire dans le nombre de critères typologiques retenus.

La recherche typologique a cependant apporté un enseignement fondamental : en dépit de la vaste pluralité des langues, les différences restent contenues dans certaines limites. Elles renvoient à des ressemblances de fond, qui désignent les facultés langagières d'*Homo sapiens sapiens*. Ces ressemblances ne sont pas plus interprétables, à l'heure actuelle, que l'expansion de l'art paléolithique dans les mêmes formes et quasi simultanément en Tanzanie, Afrique de l'Est, Afrique du Sud, Australie, Inde et Amérique ; pas plus interprétables que l'apparition de systèmes de numération à base 3, 4, 5, 10 ou 20 selon différents points du Globe - ou encore que l'invention des différents types d'écriture (Voir Annexes 2, « Photo de divers types d'écriture »).

6. La reconstruction : jusqu'où ?

Avec ce fond commun en tête, les généalogistes sont tentés de reconstruire des « langues-ancêtres », ou proto-langues. Des macrofamilles, ou superfamilles, voire comme Meritt Ruhlen (1997), une proto-langue mondiale (« proto-world »). Certains se risquent plus loin encore, en tentant de construire des modèles du « proto-sapiens » (de Grolier, 1988, p. 203).

Ainsi, dès 1903, Holger Pedersen, tenta de reconstituer une langue-mère commune dite nostratique, aujourd'hui toujours d'actualité mais très discutée. Vladislav M. Illitch-Svitytch (1976) et Aaron B. Dolgopolski (1986) reprirent le flambeau vers 1960, la faisant remonter au début du Néolithique. Six familles en dériveraient, dont l'indo-européen et le sémitique (Jucquois, 1996, pp. 152-154). Un autre unificateur, célèbre et moins contesté, est Joseph H. Greenberg, (2003) qui a notamment reconstruit la superfamille eurasiatique, au sein de laquelle se rangerait l'indo-européen. Mais l'unificateur Ruhlen, en collaboration avec John Bengston, a proposé, quant à lui, pour asseoir le proto-world, un ensemble de 27 racines mondiales, mots qui se retrouvent dans les 12 macrofamilles et qui regroupent, selon lui, toutes les langues du monde pour désigner à peu près le même référent.

Cette hypothèse est loin d'être admise par la communauté des linguistes. Trois arguments notamment paraissent la réfuter. Le premier est d'ordre extralinguistique. Les modèles mathématiques de type probabiliste émanant de certains chercheurs indiquent en effet qu'en suivant la méthode de Ruhlen, les chances de trouver des cognats dans toutes les langues sont de 100 % (Boë, cité par Coupé, 2005, p. 173). Le second est d'ordre linguistique. Au-delà de ce qu'on appelle le premier degré, c'est-à-dire dès qu'on aborde le stade de la superfamille, les données linguistiques, mais aussi culturelles et historiques se raréfient. Selon la formule de Guy Jucquois (1996, pp. 152-154), « on glisse progressivement des « correspondances » aux « ressemblances » ». La méthode comparative poussée à ses extrémités et vers ces glissements conduit à travailler « sous le niveau du bruit », c'est-à-dire vers une indétermination foncière entre ressemblances accidentelles et ressemblances dues à une histoire commune entre les langues (Boë cité par Coupé, 2005, p. 173). Le troisième argument est encore d'ordre linguistique. La plupart des chercheurs estiment aujourd'hui que l'applicabilité de la méthode comparative devient nulle après environ 7 000 à 10 000 ans de séparation. S'inspirant de la datation de restes organiques par le carbone 14, Morris Swadesh (1952, pp. 452-453) avait inauguré la méthode lexicostatistique de datation des langues dite glottochronologique. Elle est utilisée pour calculer la durée de séparation entre des langues apparentées et se fonde sur le taux de remplacement du vocabulaire de base par périodes de 1000 ans. À sa suite, Hagège, ainsi que Georges André Haudricourt, et parallèlement Lionel Bender ont proposé une théorie cyclique de l'évolution des langues. Celles-ci évolueraient, en transformant leurs lexiques et leurs structures à un rythme relativement constant, tant et si bien qu'au bout de la longue période indiquée, elles se seraient entièrement renouvelées. Que l'on songe simplement à l'exemple phonétique du mot signifiant « langue » dans le domaine indo-européen. « Le prototype que l'on reconstruit, écrit Martinet (1986, pp. 128-130) a la forme « d₁ng₁hweH₂ ». Par une série de transformations régulières à travers les siècles, l'auteur montre qu'il a pu aboutir à des formes aussi contrastées que le latin *lingua*, l'anglais *tongue*, le

russe *jazyk* et le tokharien *käntwa*. Par ces arguments, l'horizon des linguistes se trouve en tout cas sérieusement raccourci. À supposer même que l'on puisse démontrer un jour l'existence d'une langue mère de toutes les langues, cette langue reconstruite ne saurait se confondre avec le concept de « langue originelle ». Antonio Tovar (1954) décrivait voilà plus d'un demi-siècle, l'illusion de ce qu'il appelait « le vieux fantôme de l'arbre généalogique ».

Dans la reconstruction à l'infini, dans certaines tentatives de macro-comparaisons se profile en réalité la recherche de la langue adamique, d'avant Babel. Il en est de même de la « quête d'une langue parfaite » évoquée par Umberto Eco (1992). Vaine quête du Graal... Des disciplines comme les mathématiques (théorie des probabilités), l'archéologie, la paléontologie et l'éthologie viennent ébranler l'ancienne utopie. Il n'est pas à exclure en effet que, dans un lointain passé, plusieurs langues mères se soient succédées ou encore aient coexisté, et qu'une seule ait survécu par la suite (Coupé, 2005, p. 186). De nombreux scénarios sont envisageables.

7. Le proto-indo-européen : une marge d'indétermination

Pour redescendre plus modestement de l'arbre généalogique au niveau des familles du premier degré, la prudence s'impose encore. Il est vraisemblable que les similitudes identifiées au cours des tentatives de reconstruction soient davantage le fait de la procédure utilisée, de la succession des inférences, que de la nature réelle de la proto-langue qu'on pense avoir caractérisée. Comme l'énonce Jucquois (1996, p. 156) :

La comparaison génétique aboutit ainsi à une illusion partielle puisqu'elle focalise en un temps originel les résultats de la comparaison supposés être synchrones par la simplicité qui résulte pourtant du simple aboutissement de la procédure de comparaison.

Ne serait-ce que pour la seule famille indo-européenne, la plus anciennement documentée historiquement, l'incertitude demeure entière. Les conditions initiales étaient-elles homogènes ou hétérogènes ? Déterminer une famille de langues est une chose (cela relève de la taxinomie), chercher une langue ancêtre en est une autre (on passe au domaine de la reconstitution). Faut-il donc poser une langue mère unique, comme le croient certains ? Faut-il plutôt voir, à l'origine, un amas linguistique qui a donné naissance au groupe de langues ayant par la suite subi un phénomène de convergence, sous le double effet du contact et du temps ? C'est la thèse de Troubetzkoy (1939) reprise par Tovar et d'autres. Ou bien peut-on imaginer une forme originellement dialectale, donc multiple ? À ces incertitudes ontologiques s'ajoutent les difficultés de localisation spatiale du foyer originel supposé : foyer anatolien selon l'archéologue Colin Renfrew (1989). C'est la thèse dominante actuellement, en raison de données archéologiques et génétiques convergentes, mais l'auteur reconnaît lui-même que son hypothèse ne peut tenir que si elle n'entre pas en conflit avec les données de la comparaison des langues), russe du sud-est selon l'archéologue Maria Gimbutas ainsi que Martinet (1986), boréal en Eurasie septentrionale selon Nicolas D. Andreïev, et foyers multiples entre la Mer du Nord et la Mer Caspienne selon Troubetzkoy. Sans parler des incertitudes sur la date d'émergence du proto-indo-européen : elle varie, selon les auteurs de -8 500 à -6000 avant le présent. « Ceci demeure, quoi qu'il en soit, une pure supposition. Il aurait tout aussi bien pu être parlé il y a 15 000 ans », renchérit, pour sa part, Alain Peyraube (2005, p. 344).

8. Monogenèse et polygenèse : le débat n'est pas tranché

Dans l'état actuel des connaissances, il est impossible de se prononcer sur la monogenèse ou la polygenèse. Les avancées de la génétique et de la paléoanthropologie tendent à privilégier aujourd'hui l'hypothèse de l'évolution unirégionale, dite « Out of Africa », ou « Arche de Noé ». Celle-ci propose que l'*Homo sapiens sapiens* actuel serait originaire d'Afrique de l'Est et issu d'une seule population d'*Erectus* (en ayant partout supplanté ce dernier). L'autre hypothèse, multirégionale, dite « Candélabre » préconise qu'une population d'*Erectus* a évolué de façon parallèle et indépendante pour donner naissance à des groupes de « *Sapiens régionaux* » sur toute la planète. La majorité des données génétiques va ainsi dans le sens d'une origine récente et africaine de notre espèce. Toutefois, l'étude de la diversité génétique de certaines régions du génome contredit cette proposition (Quintana-Murci & Hombert, 2005, pp. 312-313, 315, 317). En réalité, les résultats obtenus dépendent des marqueurs génétiques choisis, et par conséquent, la reconstruction génétique est liée à la procédure utilisée. Quoi qu'il en soit, si la thèse unirégionale l'emportait, elle ne serait pas pour autant incompatible avec la polygenèse linguistique, comme on l'a vu plus haut (section 6).

Ce n'est, encore une fois, que de l'étroite collaboration entre les préhistoriens, les archéologues, les généticiens et les linguistes que pourront surgir des éléments de réponse.

9. L'hypothèse de l'aire ancestrale

Il existe, en revanche, un certain consensus sur l'hypothèse plus modeste, de l' « aire ancestrale », énoncée dans l'entre-deux guerres par Edward Sapir. Si l'on veut identifier le foyer originel d'une famille linguistique, disait-il, la zone de plus grande ancienneté historique correspond à celle de la plus grande divergence. Ainsi, c'est en Mélanésie, entre l'Archipel Bismarck et les Nouvelles Hébrides que se trouve la plus grande diversité des langues austronésiennes. Cela semblerait indiquer que le proto-indonésien était parlé là voici des millénaires. Or, l'archéologie vient à l'appui de cette thèse, car c'est aussi en Mélanésie qu'ont été trouvés les assemblages d'objets façonnés les plus anciens de cette partie du monde. En Afrique, c'est la région située au Nord de la zone bantoue et au Sud du Sahara (nord de la frontière Cameroun-Nigéria) qui est la plus diversifiée linguistiquement. On suppose donc, en suivant ce raisonnement, que c'était là le foyer originel des Proto-Bantous. Or, cela semble être conforté par les données génétiques fondées sur l'analyse de certaines protéines. Dans le domaine amérindien du Nord, la plus grande diversité des langues subarctiques conduit à l'hypothèse d'une migration athabasque à partir de cette aire.

10. Gènes et langues : corrélation ?

On a vu que la génétique, n'échappe pas, quant aux hypothèses de l'origine uni- ou multi-régionale des *Sapiens*, à des problèmes liés à la procédure utilisée. De la même façon, les corrélations entre langues, populations et cultures à travers le monde sont difficiles à tracer. Les contrées notoires pour leur diversité linguistique ne recouvrent pas d'hétérogénéité ethnologique qu'on serait en droit d'attendre. Des exemples contraires comme les Bochimans d'Afrique du Sud ou les Berbères d'Afrique du Nord ne sont que des exceptions. Les récentes découvertes de l'école de Luigi Cavalli-Sforza (1996) concernant la correspondance entre certains groupes linguistiques et humains témoignent de la complexité des phénomènes. Non seulement gènes et langues ne coïncident pas toujours, mais les résultats, là encore, divergent en fonction des marqueurs génétiques choisis. Le chromosome Y (révélateurs de l'histoire des hommes) montre une corrélation plus forte avec la linguistique que l'ADN mitochondrial (révélateur de l'histoire des femmes). Or, selon le premier marqueur, ainsi que les marqueurs classiques, les Basques, dont la langue est un isolat, sont aussi isolés génétiquement. En revanche, selon le second marqueur, les populations basque et indo-européenne sont très homogènes génétiquement. Histoire migratrice ou histoire culturelle différentes ? On l'ignore. Il en est de même pour les populations et les langues sames. Selon des analyses raffinées d'ADN mitochondrial, pas de coïncidence entre gènes et langues : les gènes sont proches de ceux portés par les populations de langues indo-européennes, tandis que les langues appartiennent à la famille ouralienne. Dans le sens inverse, les Tunisiens, Egyptiens et Ethiopiens, pourtant géographiquement séparés par plusieurs milliers de kilomètres, appartiennent à la même macro-famille linguistique afro-asiatique (celle-ci n'est pas toutefois reconnue par tous) et se révèlent très proches génétiquement (Quintana-Murci & Hombert, 2005, pp. 324-325). La prudence s'impose donc en ces matières. D'autant plus que l'équation langue=race - et toute l'idéologie mortifère relative aux supériorités et infériorités supposées - rôde désormais comme un spectre. Quoi qu'il en soit, les avancées de l'école Cavalli-Sforza, aussi intéressantes que prometteuses, plaident autant pour un apparentement aréal ancien que pour un ancêtre commun.

11. Homogénéité et hétérogénéité linguistique

Chaque idiome recèle ses clefs de la diversité linguistique. La parfaite homogénéité d'une langue est une vue de l'esprit. En réalité, l'hétérogénéité et la variation caractérisent l'être même de la langue - comme l'ont montré les travaux importants de Martinet, d'Uriel Weinreich (1968), de Hagège (1993), William Labov (1994) et Mufwene. Pour le premier, structuraliste et fonctionnaliste, et inventeur du concept de « synchronie dynamique », variation et changement linguistique se font *dans* et *par* la structure (1964, pp. 521-532) ; « une langue change parce qu'elle fonctionne » résume-t-il (1990, p. 155). Ce en quoi il est très proche de la biologie évolutionniste non fixiste, qui lit les structures et les fonctions comme résultat d'une histoire, et encore en devenir. Le second consacre un volume entier (1993) à la dynamique des langues, mettant à l'avant-scène l'activité consciente et inconsciente des locuteurs-constructeurs.

Pour le dernier, plutôt ancré dans la sociolinguistique (2001, p. 24) :

The agency of change lies definitely within the behavior of individual speakers, and causation partly in the mutual accommodations they make to each other while they are more intent on communicating effectively than on preserving idiolectal, dialectal or language boundaries.

Les causes de l'hétérogénéité sont des plus diverses : paramètres sociaux, de sexe, d'âge, d'habitat, d'ethnie, de discours, de style... Aujourd'hui, grâce à la paléogéographie, on peut ajouter un nouveau facteur de variation : la taille du groupe de locuteurs. Celle-ci conditionnerait en effet les structures linguistiques mêmes d'une langue donnée (Bocquet-Appel, cité par Coupé, 2005, p. 190). L'hétérogénéité est à l'oeuvre jusque chez le locuteur, chacun de nous n'en finissant pas de réorganiser notre langue. Dans tout processus de communication, opère un double mouvement d'homogénéisation et d'hétérogénéisation. Des zones de stabilité relative ainsi que d'instabilité sont détectables en synchronie.

12. Conclusion

La question des origines n'est désormais plus taboue pour les linguistes. On peut s'en réjouir car cette problématique est abordée par bien d'autres sciences, naturelles et humaines. Loin de constituer une vaine quête du Graal, comme le serait celle d'une « langue parfaite », elle s'avère féconde, dans notre champ, lorsqu'elle s'inscrit dans une démarche pluridisciplinaire. Néanmoins, vu la complexité des données, la prudence méthodologique doit constamment guider nos pas. Ceci d'autant plus que toute remontée aux sources premières n'est pas exempte de passion, traversée par les mythologies, et les idéologies qui nous environnent. Qu'il s'agisse de l'origine de l'univers, de la vie, de l'homme, ou du langage et des langues, elle est sporadiquement sous-jacente, et il ne manque pas de chercheurs (à l'instar de l'astrophysicien Hubert Reeves, du cosmologue Jayant Vishnu Narlikhar, du chimiste Jacques Reisse et du paléoanthropologue Yves Coppens) pour stigmatiser ses intrusions dans la science. Dans le domaine linguistique, on observe que les théories sur le monogénisme et le polygénisme ont toutes deux été des vecteurs d'idéologies douteuses, conduisant quelquefois au racisme.

Tour à tour désacralisée et libérée des tabous, cette délicate question peut prendre toute sa place comme l'une des clés explicatives de la diversité linguistique. Toutefois, notre dispositif cérébral, les conditions d'évolution des *Sapiens* ainsi que leur histoire migratrice ne sauraient rendre compte à eux seuls, bien entendu, de la diversité observée. L'usage que font de la langue les locuteurs en perpétuelle dynamique adaptative, et en incessante situation de contact, ainsi que la remise en question progressive des valeurs conventionnelles constitutives des langues, en tant qu'institutions sociales, demeurent un très puissant moteur du changement linguistique et de la diversité.

L'attention portée à la pluralité linguistique, a en outre l'avantage de mettre en relief, sur fond d'universalisme (l'unité écologique et psycho-physiologique du genre humain, et le fait que l'on puisse traduire d'une langue à l'autre) un relativisme linguistique fondamental. Faire voir ce que, selon la formule de Jakobson reprise par Hagège, les langues « obligent » à dire ou « empêchent de dire », c'est aussi prendre conscience que les structurations spécifiques des idiomes imposent aux sujets parlants des représentations diversifiées du monde qu'il importe d'analyser dans leur richesse, non de réduire à l'unique.

Références bibliographiques

- Bickerton (D.). 1981 : *Roots of Language*. Ann Arbor : Karoma.
- Boë (L.-J.), Bessière (P.) & Vallee (N.). 2003. « When Ruhlen's « mother tongue » theory meets the null hypothesis ». in : *Proceedings of the XVth International Congress of Phonetic Sciences*. Barcelone.
- Chaudenson (R.). 2000. « Peut-on décrire un créole sans référence à sa genèse ? Pronoms et adjectifs dans les créoles français ». in : *Langages*. Paris : Larousse.
- Cavalli-Sforza (L.). 1996. *Gènes, peuples et langues*. Paris : Odile Jacob.
- Coupé (C.). 2005. « L'impossible quête de la langue mère ». in : *Aux origines des langues et du langage*. Paris : Fayard, pp. 161-195.
- Dixon (R.-M.-W.). 1980. « The Endangered Languages of Australia, Indonesia and Oceania ». in : Robin & Uhlenbeck (eds.). *Endangered Languages*. Oxford, New York : Berg, pp. 229-255.
- Dixon (R.-M.-W.). 1997. *The Rise and Fall of Languages*. Cambridge : Cambridge University-Press.
- Dolgopolski (A.-M.). 1986. « A Probabilistic Hypothesis Concerning the Oldest Relationships Among the Languages Families in the Northern Eurasia ». in : Shevoroskin & Markey (eds.). *Typology, Relationship and Time*. Ann Arbor : Karoma.
- Eco (U.). 1992 : « La quête d'une langue parfaite dans l'histoire de la culture européenne ». [Leçon inaugurale au Collège de France, le 2 octobre 1992].
- Greenberg (J.-H.). 2003. *Les langues indo-européennes et la famille eurasiatique*. Paris : Belin [trad. fr.].
- Grolhier (E. de). 1988. « Aux origines du langage ». in : Brenot (P.), (dir.). *Les origines*. Paris : L'Harmattan, pp. 187-222.
- Hagège (C.). 1982. *La structure des langues*. Paris : P.U.F., Coll. « Que sais-je ? ».
- Hagège (C.). 1985. *L'homme de paroles*. Paris : Fayard.
- Hagège (C.). 1993. *The Language Builder : an Essay on the Human Signature in Linguistic Morphogenesis*. Amsterdam : Benjamins.
- Hagège (C.). 2000. *Halte à la mort des langues*. Paris : Odile Jacob.
- Hagège (C.). 2003. « Le langage et les langues entre le biologique et le social ». in : Changeux (J.-P.), (dir.). *Gènes et culture*. Paris : Odile Jacob, pp. 153-172.
- Haudricourt (G.-A.). 1954. « Comment reconstruire le chinois archaïque ». in : *Word*, 10.
- Hazaël-Massieux (M.-C.). 2005. « Au sujet de la définition des créoles ». in : *La Linguistique*, 41, 1 : *Les créoles*, pp. 3-17.
- Hombert (J.-M.). 2005. « Conclusion ». in : Hombert (J.-M.), (dir.). *Aux origines du langage et des langues*. Paris : Odile Jacob, pp. 462-465.
- Hublin (J.-J.). 2003. « Évolution des hominidés et origine du langage ». in : Changeux (J.-P.), (dir.). *Gènes et culture*. Paris : Odile Jacob, pp.17-33.
- Hublin (J.-J.). 2005. « La langue des premiers hommes ». in : Hombert (J.-M.), (dir.). *Aux origines des langues et du langage*. Paris : Fayard, pp. 102-117.
- Illitch-Svitytch (V.-M). 1976 : *Opyt sravnenija nostratičeskix jazykov* [Étude comparative des langues nostratiques]. Moscou : Nauka.
- Jucquois (G.). 1996. « Le nostratique ». in : *La Linguistique*, 32, 2. Paris : P.U.F., pp.152-154.
- Labov (W.). 1994. *Principles of Linguistic Change : Internal Factors*. Oxford : Blackwell.
- Mandel (J.-L.). 2003. « Maladies monogénétiques et troubles de la cognition ». in : Changeux (J.-P.), (dir.). *Gènes et culture*. Paris : Odile Jacob, pp. 35-56.
- Malmberg (B.). 1990. « Linguistique aréale et recherches comparatives ». in : *Mémoire de la SLP*. Nouvelle Série, tome I. Paris : Klincksieck, p. 16.
- Martinet (A.). 1964. « Structural variation in language ». in : *Proceedings of the 9th International Congress of Linguists*, pp.521-532.
- Martinet (A.). 1968. « Préface ». in : Martinet (A.), (dir.). *Le Langage. Encyclopédie de la Pléiade*. Paris : Gallimard, pp. 6-12.
- Martinet (A.). 1986. *Des steppes aux océans. L'indo-européen et les « Indo-Européens »*. Paris : Payot.
- Martinet (A.). 1990. « Changements linguistiques ou fonctionnement du langage ? ». in : *La Linguistique*, 26, 2, pp. 153-158.
- Matisoff (J.A.). 1980. « Endangered Languages of Mainland Southeast Asia ». in : Robin & Uhlenbeck (eds.). *Endangered Languages*. Oxford, New York : Berg, pp. 189-228.

- Meillet (A.). 1965. « Le problème de la parenté des langues ». in : Meillet (A.). *Linguistique historique et linguistique générale*. Paris : Honoré Champion, pp. 76-109 [rééd.].
- Mounin (G.). 1992. « Sur la mort des langues ». in : *La Linguistique*, 28, 2, pp. 149-158.
- Mufwene (S.-S.). 1986. « Les langues créoles peuvent-elles être définies sans allusion à leur histoire ? ». in : *Études créoles*, 9, 1, pp. 135-150.
- Mufwene (S.-S.). 2001. *The Ecology of Language Evolution*. Cambridge : Cambridge University Press.
- Nespoulous (J.-L.). 2005. « La désintégration du langage dans le cerveau lésé de l'aphasique » in : Hombert (J.-M.), (dir.). *Aux origines du langage et des langues*. Paris : Odile Jacob, pp. 428-449.
- Nettle (D.). 2002. *Linguistic Diversity*. Oxford : Oxford University Press.
- Peyraube (A.). 2005. « Comment classer les langues ». in : Hombert (J.-M.), (dir.). *Aux origines du langage et des langues*. Paris : Odile Jacob, pp. 328-345.
- Peyraube (A.) 2005. « L'homme aux 6000 langues ». in : Hombert (J.-M.), (dir.). *Aux origines du langage et des langues*. Paris : Odile Jacob, pp. 450-461.
- Quintana-Murci (L.) & Hombert (J.-M.). 2005. « Gènes et langues : une évolution parallèle ? ». in : Hombert (J.-M.), (dir.). *Aux origines du langage et des langues*. Paris : Odile Jacob, pp. 308-327.
- Renfrew (C.). 1989. *Archaeology and Language. The Puzzle of Indo-European Origin*. London : Penguin Press.
- Ruhlen (M.). 1997. *L'origine des langues. Sur les traces de la langue-mère*. Paris : Belin [trad. fr.].
- Swadesh (M.). 1952. « Lexicostatistic Dating of Prehistoric Ethnic Contacts ». in : *Proceedings of the American Philosophical Society*, 96, pp 452-463.
- Szulmajster-Celnikier (A.). 1997. « À propos de *Langue et idéologie : les mots étrangers en hébreu moderne* - de Michel Masson ». in : *La Linguistique*, 33, 1, pp. 141-145.
- Tovar (A.). 1954. « Linguistics and Prehistory ». in : *Word*, 10, p. 348.
- Trubetzkoy (N.-S.). 1939. « Gedanken über das Indogermanenproblem ». in : *Acta Linguistica*, I, 2, pp 81-89. Copenhague.
- Weinreich (U.), Labov (W.) & Herzog (M.). 1968. « Empirical foundations for a theory of language change ». in : Lehman & Malkiel (eds.). *Directions for Historical Linguistics*. Austin : University of Texas Press.

Annexes 1

Encadré 1 : La diversité des langues

La diversité des langues est inégalement répartie sur la surface du Globe. Certaines régions sont connues pour leur foisonnement : Amérique du Sud, Nouvelle Guinée, province du Plateau au Nigéria, province d'Oaxaca au Mexique, certains territoires du Nord-Est de l'Inde. En Afrique, ce sont les langues bantoues, appartenant à la famille Niger-Congo, qui dominent par leur nombre (900) et celui de leurs locuteurs (160 millions). En revanche, les langues d'Europe, et singulièrement d'Europe occidentale, présentent un faible degré de diversité. C'est le résultat de conditions historiques particulières. D'une part, les contacts de populations et de leurs langues ont fait de l'Europe une aire caractérisée par des emprunts multiples et multidirectionnels. D'autre part, l'édification de normes linguistiques communes y a été très précoce et très générale. C'est en Europe que la réduction, voire l'élimination des variations dialectales a été le plus efficace. Mais on trouve aussi des exemples de larges groupes à l'homogénéité frappante sans intervention du politique. C'est le cas de la famille australienne aborigène, qui connaît un continuum dialectal étendu sur des distances considérables. À mi-chemin, le groupe slave se propage en maillons de chaîne. Divergence et convergence semblent y avoir joué un rôle équilibré.

C'est la famille indo-européenne qui compte le plus grand nombre de locuteurs. Elle est suivie par la famille sino-tibétaine, dans laquelle le chinois occupe une place centrale par sa masse numérique. Sa forme la plus répandue, le mandarin, détient de ce point de vue un record mondial (610 millions de locuteurs natifs). À l'autre extrémité du spectre, certaines langues en péril comme les langues paléo-sibériennes, ne comptent plus que quelques centaines de locuteurs (400 pour le youkaghir, langue de Sibérie arctique, à l'Est de la Léna ; une centaine seulement pour le khet, dans le bassin du Ienisséi).

En termes d'espace, certaines familles et groupes de langues s'étendent sur des aires impressionnantes : il en est ainsi de la famille austronésienne (Voir carte « Les langues austronésiennes »), de la famille indo-européenne, ou du groupe samoyède (langues ouraliennes). On voit aussi des langues peu parlées mais dispersées sur de vastes territoires : telles le yidiche, le romani (ou tzigane), ou encore l'arménien. Parfois, à l'inverse, un pullulement de langues se concentre dans un périmètre très réduit. Ainsi l'ensemble caucasien est confiné et morcelé dans les monts du Caucase, région nommée « mont des langues » par les géographes arabes du Xe siècle (Voir carte « Le Caucase : mosaïque linguistique »). C'est parfois pour faire face à une diversité ressentie comme trop forte que s'élaborent des *lingua franca* (au sens où l'anglais est devenu la langue des scientifiques), des *koine*, dialectes véhiculaires, des langues supralocales et fédératrices, comme le swahili en Afrique de l'Est, ou encore des pidgins comme le bichelamar (beach-la-mer) aux Nouvelles Hébrides et en Australie.

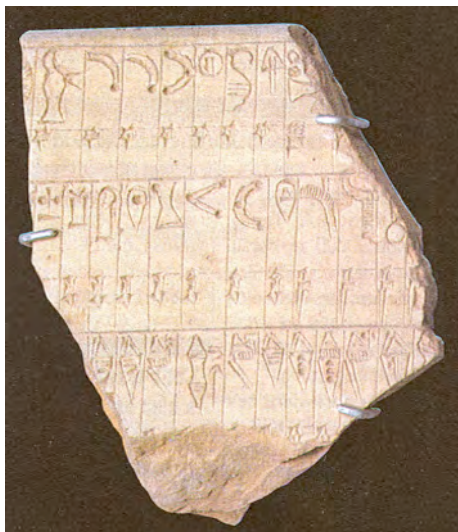
Encadré 2 : Quatre domaines de divergence

Les principaux domaines de divergence entre les langues sont par ordre croissant de source de divergence la syntaxe, la phonologie, le lexique et la morphologie. Sur une base commune d'organisation et d'expression de contenus plus ou moins universels propre à *Homo sapiens sapiens*, les langues divergent d'abord par leur structuration syntaxique. Les divergences sont plus fortes en phonologie, dans les limites imposées par la constitution des organes de la parole et de l'audition. Sur le plan lexical, les différences se manifestent non seulement entre les langues, mais aussi d'un individu à l'autre. C'est la morphologie qui constitue le champ de plus forte différenciation et partant, le domaine où les coïncidences interlinguistiques se font rares.

Les neuropsychologues, grâce à la technique d'imagerie cérébrale, viennent aujourd'hui conforter la réalité de l'architecture cérébrale postulée par les linguistes : la phonétique, la phonologie, le lexique et la syntaxe renvoient bel et bien à des représentations et des processus distincts dans le cerveau humain.

Annexes 2

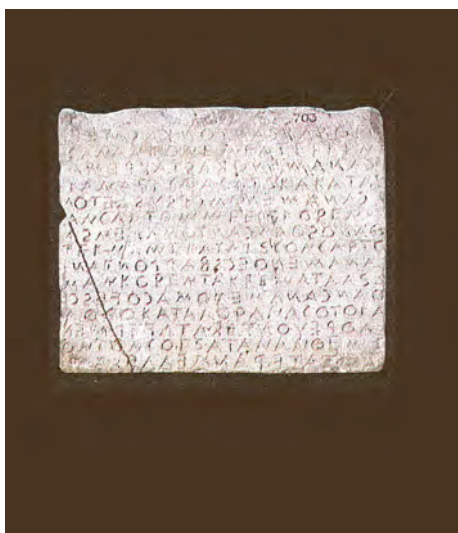
Image 1 : 7 photos de divers types d'écriture [1a à 1g]



(1a) Inscription assyrienne, 680 av. J.-C.



(1b) Inscription étrusque, Ve siècle avant J.-C.



(1c) Inscription dorienne de Crète, époque archaïque.



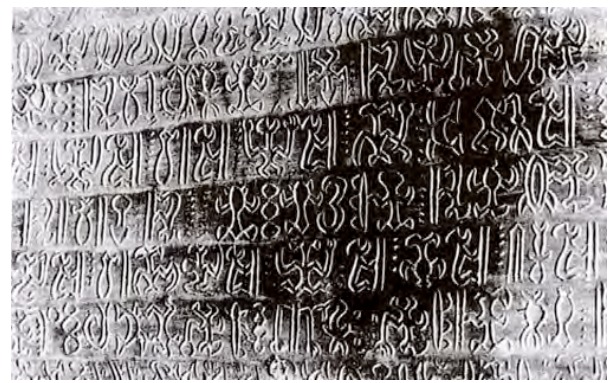
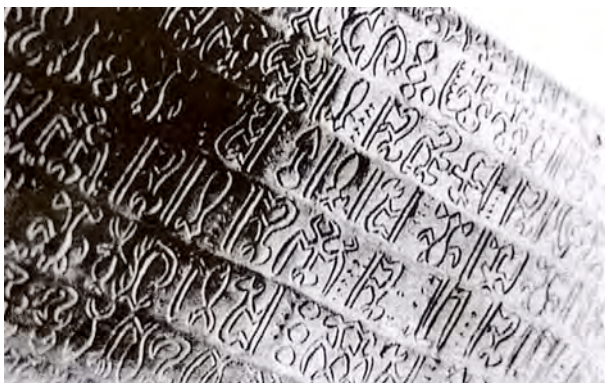
(1d) Hiéroglyphes hittites de la langue Iuwi, Xe-VIe siècle av. J.-C.
Écriture rarissime de type « boustrophédon »
(lecture horizontale en zigzag).



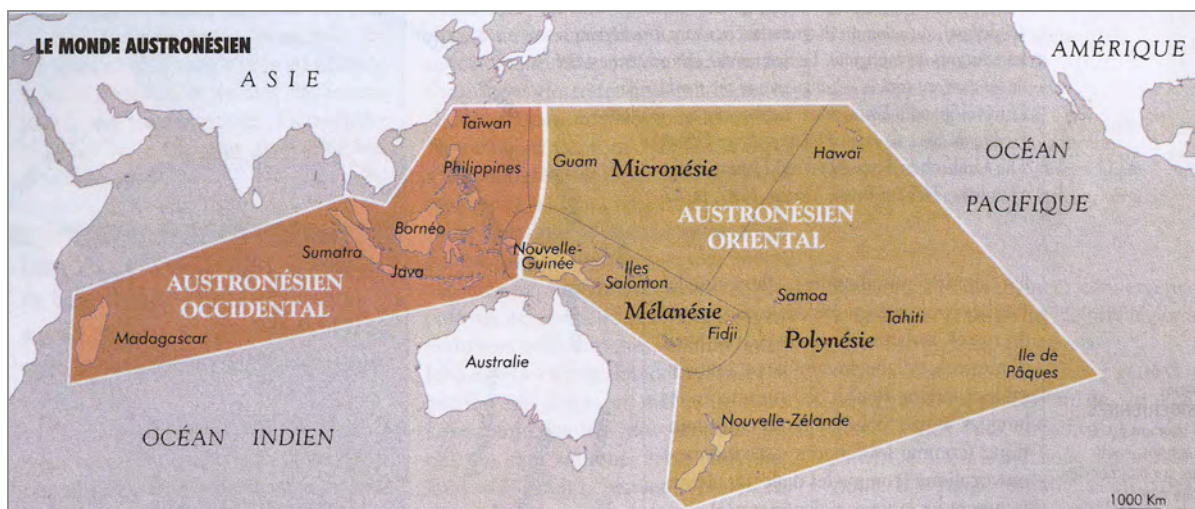
(1e) Inscription sumérienne, 2000 ans av. J.-C.



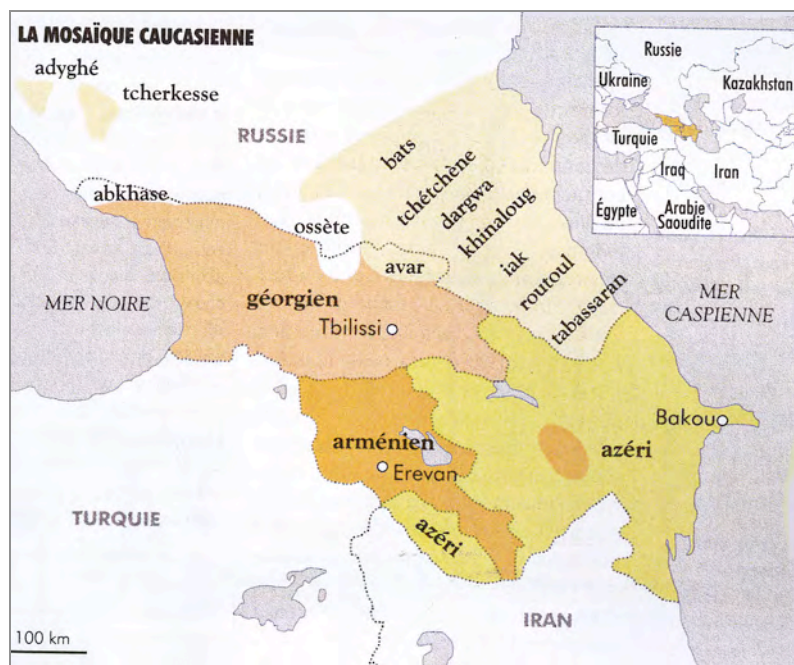
(1f) Inscription sumérienne, 2550 ans av. J.-C.



(1g) (1h) Hiéroglyphes de l'Île de Pâques (ici « Le bâton de Santiago »), en exploitation jusqu'au XIXe siècle, dont le déchiffrement est très récent. Écriture unique au monde de type « boustrophédon à inversion alternée » (lecture horizontale en zigzag, tête bêche d'une ligne à l'autre).



Carte 1 : Les langues austronésiennes : vaste extension d'environ 23 600 km. Pour expliquer la parenté entre ces langues, on en est réduit à échafauder des hypothèses.
 (Source : article « Languages of the world », Encyclopaedia Britannica).



Carte 2 : Le Caucase : mosaïque linguistique sur un espace réduit d'environ 1 200 X 600 km, et sur un relief très accidenté. On y distingue des langues des familles caucasienne, indo-européenne et altaïque.
 (Source : Hagège (C.). 1992. Le souffle de la langue. Paris : Odile Jacob, carte 3).

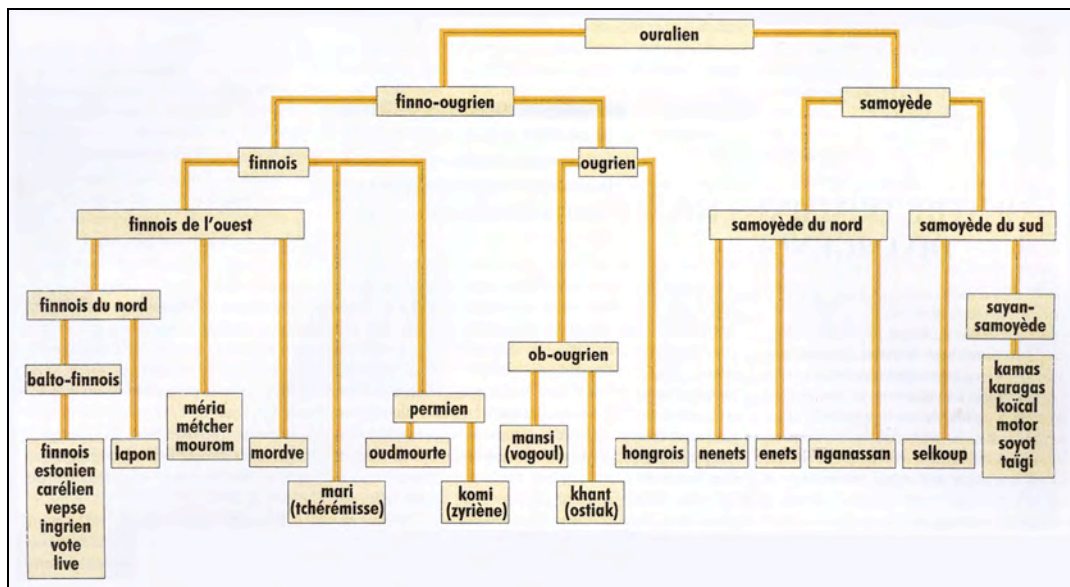


Schéma 1 : Les langues ouraliennes selon un diagramme arborescent, de type généalogique.
 (Source : article « Languages of the world », Encyclopaedia Britannica).

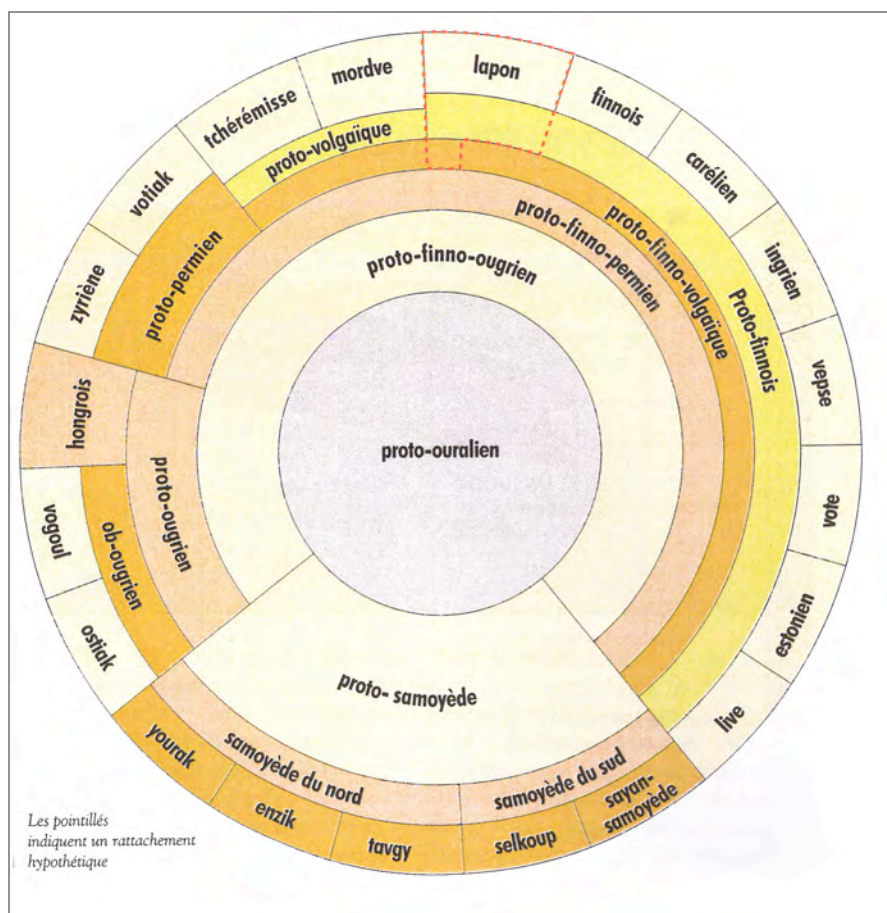


Schéma 2 : Les langues ouraliennes selon un diagramme circulaire, privilégiant les relations de contiguïté.
 (Source : Hajdù (P.) & Dometlos (P.). 1987. Die Uralischen Sprachen und Literaturen. Hamburg: Helmut Buske Verlag, p. 312).

Cette page sera l'occasion de remercier chaleureusement toutes celles et tous ceux qui ont contribué à l'élaboration du numéro 11 depuis plus d'un an.

Et tout d'abord **les auteurs**

- ▶ Sylvain Auroux
Laboratoire d'histoire des théories linguistiques
– UMR 7597 – CNRS, Paris 7, ENS-LSH, France
Courriel 1 : sylvain.auroux@ens-lsh.fr
Courriel 2 : auroux.sylvain@wanadoo.fr
- ▶ Luigi Luca Cavalli Sforza
Stanford University, USA
Courriel : cavalli@stanford.edu
- ▶ Georges Chapouthier
Université de Paris 6, France
Courriel : chapout@ext.jussieu.fr
- ▶ Jean-Paul Colin
Université de Franche-Comté, France
Courriel : janpol.colin@wanadoo.fr
- ▶ Michael C. Corballis
Department of Psychology,
University of Auckland, New Zealand
Courriel : m.corballis@auckland.ac.nz
- ▶ Marcio Alexandre Cruz
Université Paris 3 : Sorbonne Nouvelle
École doctorale 268 UPRES SYLED,
EA 2290, Paris, France
Courriel : marcioalexandreacruz@hotmail.com
- ▶ Jean-Louis Dessalles
École Nationale Supérieure des Télécom-
munications, Paris, France
Courriel : dessalles@enst.fr
- ▶ Donald Favareau
National University of Singapore, Singapore
Courriel : favareau@ucla.edu
- ▶ Béatrice Fracchiolla
Université de Toulon,
UFR de Lettres et Sciences Humaines, France
Courriel 1 : fracchiollaml11@yahoo.fr
Courriel 2 : bearfrac@yahoo.com
- ▶ Charles Goodwin
University of California
Los Angeles, USA
Courriel : cgoodwin@humnet.ucla.edu
- ▶ Dominique Klingler
Université de Paris 3,
Sorbonne Nouvelle, France
Courriel : dominique.klingler@tiscali.fr
- ▶ Laurent Métoz
Laboratoire HTL,
UMR CNRS 7597, ENS LSH, Lyon, France
Courriel : laurentmetoz@yahoo.fr
- ▶ Lisa Mikesell
University of California
Los Angeles, USA
Courriel : lmikesell@esrv.com
- ▶ Salikoko S. Mufwene
University of Chicago, Department of
Linguistics, USA
Courriel : s-mufwene@uchicago.edu
- ▶ Robert Nicolai
Université de Nice et Institut universitaire de
France, France
Courriel 1 : nicolai@unice.fr
Courriel 2 : nicolai@eva.mpg.de
Courriel 3 : nicolai@taloa.unice.fr
- ▶ François Rastier
Directeur de recherche CNRS, Paris, France
Courriel 1 : Lpe2@ext.jussieu.fr
Courriel 2 : Francois.Rastier@ifs.cur-archamps.fr
- ▶ Stéphane Robert
CNRS-LLACAN & INALCO, France
Courriel 2 : robert@vjf.cnrs.fr
- ▶ Eric A. Salzen
Psychology Department
Aberdeen University, Scotland
Courriel : e.salzen@abdn.ac.uk
- ▶ John H. Schumann
University of California
Los Angeles, USA
Courriel : schumann@humnet.ucla.edu
- ▶ Anne Szulmajster-Celnikier
Collège de France,
Chaire de Théorie Linguistique, Paris, France
Courriel : anne.szulmajster@college-de-france.fr
- ▶ Hongyin Tao
University of California
Los Angeles, USA
Courriel : tao@humnet.ucla.edu
- ▶ Daniel Véronique
Université de Paris 3, France
Courriel : Daniel.Veronique@paris3.sorbonne.fr

- ▶ Namhee Lee
University of California
Los Angeles, USA
Courriel : nhlee@ucla.edu
- ▶ Alison Wray
Cardiff University, Wales
Courriel : wraya@cardiff.ac.edu

Les membres du **comité scientifique** de Marges Linguistiques et tout spécialement :

- ▶ Michel Arrivé
Université de Paris X : Nanterre, France
Courriel : michel.arrive@wanadoo.fr
- ▶ Michèle Monte,
Université de Toulon, France
Courriel : monte@univ-tln.fr
- ▶ Claudine Moïse
Université d'Avignon, France
Courriel : claudine.moise@univ-avignon.fr
- ▶ François Rastier
Directeur de recherche CNRS, Paris, France
Courriel 1 : Lpe2@ext.jussieu.fr
Courriel 2 : Francois.Rastier@ifs.cur-archamps.fr
- ▶ Daniel Véronique
Université de Paris 3, France
Courriel : Daniel.Veronique@paris3.sorbonne.fr

Les **lecteurs extérieurs** qui ont accepté de lire et relire les textes que nous leur soumettons :

- ▶ Christian Abry
Institut de la Communication Parlée
UMR CNRS (INPG-Stendhal)
Université Grenoble 3, France
Courriel : abry@icp.inpg.fr
- ▶ Nathalie Auger
Université Paul Valéry, Montpellier 3, France
Courriel : Nathalie.Auger@univ-montp3.fr
- ▶ Jacqueline Authier-Revuz
Université Paris 3, Sorbonne Nouvelle, France
Courriel : jacqueline.authier@wanadoo.fr
- ▶ Claire Blanche Benveniste
Université de Provence, France
Courriel : claire.benveniste@freesbee.fr
- ▶ Olivier Bertrand
École Polytechnique
Courriel : olivier.bertrand46@wanadoo.fr
- ▶ Jean-Paul Bronckart
Université de Genève, Suisse
Courriel : jean-paul.bronckart@pse.unige.ch
- ▶ Marina Cavassilas
Université de Limoges
Chercheuse indépendante
Courriel : marina_cavassilas@yahoo.fr
- ▶ Georges Chapouthier
Université de Jussieu, Paris, France
Courriel : chapout@ext.jussieu.fr
- ▶ Jean-Louis Chiss
Université de Paris 3, France
Courriel : jlchiss@wanadoo.fr
- ▶ Michael C. Corballis
Department of Psychology,
University of Auckland, New Zealand
Courriel : m.corballis@auckland.ac.nz
- ▶ Boris Cyrulnik
Département d'Ethologie,
Université de Toulon, France
Courriel : boris.cyrulnik@wanadoo.fr
- ▶ Jean-Louis Dessalles
École Nationale Supérieure des Télécommunications, Paris, France
Courriel : dessalles@enst.fr
- ▶ Pierre Legoffic
Université Paris 3, Sorbonne Nouvelle, France
Courriel : Pierre.Legoffic@univ-paris3.fr
- ▶ Claudine Normand
Université de Bourgogne, France
Courriel 1 : normand.claudine@u-bourgogne.fr
Courriel 2 : normand.claudine@wanadoo.fr
- ▶ Salikoko S. Mufwene
University of Chicago, Department of
Linguistics, USA
Courriel : s-mufwene@uchicago.edu
- ▶ Berthille Pallaud
Université de Provence,
Cnrs, UMR 6057, Laboratoire LPL, France
Courriel : berthille.pallaud@lpl.univ-aix.fr
- ▶ Christian Puech
Université de Paris 3
Sorbonne Nouvelle, Paris, France
Courriel 1 : christian.puech@univ-paris3.fr
Courriel 2 : cpuech50@yahoo.fr
- ▶ Robert Nicolai
Université de Nice et Institut universitaire de
France, France
Courriel 1 : nicolai@unice.fr
Courriel 2 : nicolai@eva.mpg.de
Courriel 3 : nicolai@taloa.unice.fr
- ▶ Izabel Vilela
Université Fédérale de Goias, Brésil
Courriel : Izabel.Vilela@wanadoo.fr

Les membres du **comité de rédaction** de **Marges Linguistiques** et tout spécialement :

- ▶ Mireille Bastien
Université de Provence, France
Courriel : bastien@aixup.univ-aix.fr
- ▶ Véronique Magaud
Auteure indépendante, France
Courriel : magaudv@wanadoo.fr
- ▶ Béatrice Fracchiolla
Université de Toulon,
UFR de Lettres et Sciences Humaines, France
Courriel : bearfrac@yahoo.com
- ▶ Michel Santacroce
Cnrs, Université de Provence, France
Courriel : Michel.Santacroce@wanadoo.fr

L'ensemble des **correspondants** et **responsables techniques** de **Marges Linguistiques** et tout spécialement :

- ▶ Rachel Atkinson
Penguin Press, United Kingdom
Courriel : Rachel.Atkinson@uk.penguingroup.com
- ▶ Raymond Merzoug
Courriel : raymond.merzoug@free.fr
- ▶ Bémat Grossenbacher
La-Chaux-de-Fonds, Suisse
Courriel : bege@vtx.ch
- ▶ Marc Seassau
Auteur indépendant, France
Courriel : marc.seassau@free.fr
- ▶ Guenther Hauska
Université de Regensburg, Allemagne
Courriel : guenther.hauska@biologie.uni-regensburg.de
- ▶ Gianna Zei
Cnr, Italie
Courriel : zei@igm.cnr.it

Enfin, et pour conclure, nous remercions tout particulièrement notre collègue :

- ▶ Béatrice Fracchiolla
Université de Toulon,
UFR de Lettres et Sciences Humaines, France
Courriel : bearfrac@yahoo.com

qui a su diriger le numéro 11 de la Revue **Marges Linguistiques** avec talent, dynamisme et une rare énergie.