



HAL
open science

Compréhension morphosyntaxique dans une situation de communication audiovisuelle : incidence du choix d'un avatar parlant auprès d'enfants présentant une surdité

Léa Dussere

► To cite this version:

Léa Dussere. Compréhension morphosyntaxique dans une situation de communication audiovisuelle : incidence du choix d'un avatar parlant auprès d'enfants présentant une surdité. Médecine humaine et pathologie. 2020. hal-03870258

HAL Id: hal-03870258

<https://hal.univ-lorraine.fr/hal-03870258>

Submitted on 24 Nov 2022

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-memoires-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



**UNIVERSITÉ
DE LORRAINE**



**DÉPARTEMENT
D'ORTHOPHONIE
NANCY**

Faculté de Médecine



FACULTÉ de MÉDECINE
NANCY

**UNIVERSITÉ DE LORRAINE
FACULTÉ DE MÉDECINE
DÉPARTEMENT D'ORTHOPHONIE**

MÉMOIRE présenté pour l'obtention du
CERTIFICAT DE CAPACITÉ D'ORTHOPHONISTE

par

Léa Dussere

soutenu le : **15 juin 2020**

**Compréhension morphosyntaxique dans une situation de
communication audiovisuelle : incidence du choix d'un avatar parlant
auprès d'enfants présentant une surdité**

Mémoire dirigé par : Mme PIQUARD-KIPFFER Agnès Maître de conférences en
Psychologie et ergonomie, UL

Mr OUNI Slim Maître de conférences en informatique, UL

Président de jury : Mr KABUTH Bernard Professeur en pédopsychiatrie, UL

Asseseurs : Mme GERARD-XEMAIRE Nathalie Orthophoniste
Mme AUDOIN-LATOURETE Nicole Formatrice en français et
Sciences de l'Education UL

REMERCIEMENTS

Je tiens tout d'abord à remercier sincèrement Madame Piquard-Kipffer Agnès, ma directrice de mémoire, qui m'a tout d'abord accueillie en stage recherche au Loria, m'a fait découvrir les projets de l'équipe Multispeech et en acceptant de m'encadrer, m'a permis de réaliser ce mémoire. Je la remercie de m'avoir tant aidée.

Je remercie également Monsieur Ouni Slim, mon deuxième directeur de mémoire qui m'a accordé beaucoup de son temps pour me faire découvrir le projet Dynalips et les avatars, qui m'a aidé à construire mon étude et qui a toujours pris le temps de répondre à mes messages.

Je remercie Madame Gerard-Xemaire Nathalie, Madame Doyen Karine et Madame Gigant Aurélie ainsi que les enfants de m'avoir accueillie dans leur cabinet ou dans leur classe et de m'avoir accordé du temps pour que je puisse réaliser cette étude.

Je remercie également Madame Bourgeois Annie, Madame Peltier Cloé et Madame Allain Nathalie avec qui j'ai échangé, parfois longuement, à propos de mon étude mais que je n'ai finalement pas pu rencontrer en raison de la crise sanitaire.

Je remercie le Professeur Kabuth Bernard d'avoir accepté de présider mon jury et Madame Audoin-Latourte Nicole et Madame Gerard-Xemaire Nathalie d'avoir accepté d'intégrer mon jury. Je vous remercie et je suis reconnaissante de l'intérêt que vous avez porté à ce projet.

Enfin je remercie mes proches et ma famille pour le soutien apporté durant cette longue mais très riche période.

DECLARATION SUR L'HONNEUR

Je soussigné (e), Mademoiselle DUSSERE Léa, inscrit(e) à l'Université de Lorraine, atteste que ce travail est le fruit d'une réflexion et d'un travail personnels et que toutes les sources utilisées ont été clairement indiquées. Je certifie que toutes les utilisations de textes préexistants, de formulations, d'idées, de raisonnements empruntés à un tiers sont mentionnées comme telles en indiquant clairement l'origine.

Conformément à la loi, le non-respect de ces dispositions me rend passible de poursuites devant la commission disciplinaire et les tribunaux de la République Française.

Fait à Vandoeuvre les Nancy, le 20 mai 2020

DUSSERE Léa

A handwritten signature in black ink, appearing to read 'dusserre', followed by a period.

SOMMAIRE

Remerciements	2
Introduction	8
Partie 1 : Ancrage théorique.....	11
Chapitre 1 Développement langagier d'un enfant présentant une surdité	12
1 Développement langagier chez l'enfant « tout-venant ».....	12
1.1 Développement langagier général	12
1.2 Développement morphosyntaxique	14
2 Conséquences de la surdité sur le développement du langage.....	15
2.1 Cadre général à propos de la surdité.....	15
2.2 Modifications dans la construction du langage oral chez l'enfant malentendant.....	17
Chapitre 2 Intervention globale dans une déficience auditive	18
1 Gain prothétique	18
1.1. Physiologie cochléaire	19
1.2. Implant cochléaire	20
1.3. Prothèses amplificatrices	21
1.3.1. Par voie aérienne	22
1.3.2. Par voie osseuse	22
1.4. Prothèses à transposition de fréquence	23
2. Communication alternative augmentée sans aide technique.....	23
2.1. Parole visuelle.....	23
2.2. Lecture labiale	25
2.3. Langage parlé complété (LPC).....	27
Chapitre 3 Evolutions technologiques : apports à la communication alternative augmentée ..	29
1 Moyens de communication alternative et augmentée	29
1.1. Naissance des outils de CAA avec aide technique	29

1.2.	Apport de la technologique dans les CAA avec aide technique	29
2	Avatars parlants	30
	Partie 2 : Partie méthodologique	34
	Chapitre 1 Problématique & Hypothèses	35
1.	Problématique.....	35
2.	Hypothèses	35
	Chapitre 2 Méthodologie.....	37
1	Caractéristiques de la population étudiée	37
2	Critères de sélection	37
	Critères d'inclusion :	37
	Critères de non-inclusion :	37
	Critères d'exclusion :	38
3	Présentation du matériel expérimental	38
3.1	Choix des modalités.....	38
3.2	Choix des indices morphosyntaxiques	39
3.3	Choix des items	40
3.4	Création du programme informatique de passation.....	41
4	Recrutement	43
5	Déroulement de la passation	44
6	Collecte et analyse statistique prévue des données	44
6.1	Nombre d'erreurs et du temps de latence par modalité : hypothèses 1, 1A et 1B	45
6.2	Indices morphosyntaxiques : hypothèse 2	47
6.3	Préférence de la modalité : hypothèse 3	47
	Partie 3 : Résultats.....	49
	Chapitre 1 Cohorte	50
1	Taille de l'échantillon.....	50
2	Caractéristiques de l'échantillon	50

2.1	Age.....	50
2.2	Degré de surdité sans prothèse auditive	50
2.3	Type d'appareillage auditif.....	51
Chapitre 2 Résultats		52
1	Correction des réponses et du temps de latence pour les trois modalités (« humain », « avatar » et « audio »).....	52
2	Correction des réponses et du temps de latence pour chaque modalité (Hypothèse 1)...	52
3	Correction des réponses : phrases reconnues en modalité « humain » (Hypothèse 1A).	54
4	Correction des réponses : phrases reconnues en modalité « humain » mais pas en modalité « audio » (Hypothèse 1B).....	55
5.	Correction des réponses : en fonction des indices morphosyntaxiques (Hypothèse 2) ..	57
6.	Correction des réponses : de la modalité préférée des sujets (Hypothèse 3)	58
Discussion		60
1	Analyse des résultats	60
1.1	Première hypothèse : réponses et du temps de latence pour chaque modalité	60
1.2	Hypothèse 1A : phrases reconnues en modalité « humain »	61
1.3	Hypothèse 1B : phrases reconnues en modalité « humain » mais pas en modalité « audio »	61
1.4	Deuxième hypothèse : indices morphosyntaxiques.....	62
1.5	Troisième hypothèse : préférence des sujets	63
2	Analyse de l'étude	63
3	Limites de l'étude.....	64
3.1	Population.....	64
3.2	Matériel.....	64
3.2.1	Programme	64
3.2.2	Vidéos.....	65
3.2.3	Corpus	66
3.2.4	Planches d'images	68

3.3	Passation de l'épreuve	68
4	Intérêts de l'étude	69
4.1	Recherches complémentaires	69
4.2	Applications.....	69
	Conclusion.....	70
	Bibliographie.....	72
	Annexes.....	77

INTRODUCTION

Dans le domaine de l'orthophonie, comme dans beaucoup d'autres domaines médicaux et paramédicaux, les **progrès récents de l'informatique** permettent aujourd'hui **d'améliorer les soins et l'accompagnement des personnes**. Les dispositifs numériques se développent, via des **outils de communication alternative et augmentée**. Certains d'entre-eux sont développés pour pallier des déficits. C'est le cas pour la dyslexie, où, dans un principe d'intervention visant à développer des techniques de compensation, se développe l'utilisation d'applications pour ordinateur comme des traitements de texte, des scripteurs ou encore des logiciels de lecture vocale.

Pour pallier des déficits de la communication orale dans le cadre de troubles acquis, tels que l'aphasie mais également de troubles développementaux (dysphasie, surdité congénitale) différents types d'aides existent : des pictogrammes, des signes ou encore des codes comme le langage parlé complété. Les aides numériques sont plus rares. Elles sont généralement axées sur le versant production de la parole et plus rarement sur le versant compréhension, qui met pourtant en difficulté les personnes atteintes de surdité ou de troubles du langage.

Au sein du projet de recherche Dynalips, élaboré dans le cadre de l'équipe Multispeech d'Inria-Loria Grand Est, (Ouni, 2015), des avatars parlants ont été développés et proposent ainsi de **nouvelles aides numériques** à la compréhension orale. A partir d'une bande son donnée, un avatar prend la forme d'un visage en mouvement qui reproduit les mouvements labiaux d'un locuteur en fonction du phonème prononcé et offre des indices visuels supplémentaires à la compréhension du message. Cette technologie, principalement utilisée dans le cadre du divertissement, notamment celui des jeux-vidéos ainsi que de l'animation 3D, pourrait permettre le développement d'outils de communication à destination d'enfants à besoins particuliers. En effet, cet avatar parlant apporte des **indices langagiers supplémentaires** à la bande sonore originale, utiles à la compréhension du langage.

Néanmoins, l'efficacité de cet outil auprès d'enfants n'a, à ma connaissance pas encore été testée. Mon mémoire s'inscrit dans cette démarche : étudier l'intérêt de certains de ces avatars parlant en situation de compréhension orale auprès d'enfants, âgés de 6 et 10 ans, atteints de surdité sévère à profonde et scolarisés en classe primaire. A cet âge, alors que les stratégies de compréhension langagières sont décrites comme organisées chez l'enfant entendant, en compréhension comme en expression (Kail & Fayol, 2015; Rondal, 2018), elles

sont fragilisées chez l'enfant sourd. En effet, les informations langagières reçues chez un enfant atteint de surdité sont limitées et partiellement traitées (Hage et al., 2006), créant des **retards dans l'acquisition du langage** : un stock lexical moins important (Le Normand, 2004) ainsi qu'un retard d'acquisition des données morphosyntaxiques (Goda, 1964).

L'étude expérimentale proposée, à travers l'observation d'une situation proche de celle d'une séance d'orthophonie, va examiner plus particulièrement chez les enfants le niveau de compréhension d'indices morphosyntaxiques.

L'objectif est de comparer **les résultats obtenus à une même épreuve de compréhension selon trois modalités présentées**. Dans tous les cas, le message oral proviendra d'un même locuteur humain, cependant le contenu visuel variera. Ainsi l'enfant sera confronté, soit à un locuteur humain original en audio-visuel, soit à un message audio et un message visuel porté par un avatar parlant, soit à un message uniquement auditif (enregistrement vocal). Les critères d'évaluation de la compréhension de l'enfant seront la précision (scores) et la rapidité (temps) des réponses apportées.

L'intérêt de cette étude est d'**examiner le niveau d'intelligibilité de l'avatar parlant** en observant si la compréhension d'un message est supérieure, lorsqu'elle est délivrée par l'avatar parlant que lorsqu'elle est délivrée par une bande sonore. Si tel est le cas, cette technologie pourrait être utilisée pour développer de nouveaux outils de communication alternative et augmentée.

M'appuyant sur la revue de littérature présentée dans la première partie de mon mémoire, en particulier les travaux qui observent que différents indices visuels, présents lors de situations de communication orale fournissent une aide notable (Dumont & Calbour, 2002; Mcgurk & Macdonald, 1976; Ouni & Gris, 2018; Piquard-Kipffer, 2014, 2016), dans ce mémoire je m'intéresse aux enfants atteints de surdité. Les hypothèses que je formule sont les suivantes :

Hypothèse 1 : L'ajout d'un canal visuel devrait permettre à l'enfant de mieux prendre en compte les indices morphosyntaxiques. Ainsi, ses stratégies de compréhension devraient être plus précises et l'accès à la signification des messages devrait être plus rapide. Nous nous attendons donc à observer un nombre de réponses incorrectes moins important ainsi que des temps de réponses plus courts dans les modalités « avatar » et « humain » que dans la modalité « audio ».

Hypothèse 1A : Les items bien reconnus par l'enfant (réponses correctes) en modalité « humain » seront mieux reconnus en modalité « avatar » qu'en modalité « audio ». Les scores des réponses correctes seront plus importants.

Hypothèse 1B : Pour les items bien reconnus par l'enfant (réponses correctes) en modalité « humain », mais incorrectement reconnus en modalité « audio » (réponses incorrectes), nous nous attendons à observer des scores de réponses correctes plus importants en modalité « avatar » qu'en modalité « audio ».

Hypothèse 2 : La modalité « avatar » devrait être plus aidante pour certains indices morphosyntaxiques que pour d'autres, en termes de scores de réponses correctes, comparée à la modalité « audio ». Nous nous attendons donc à observer une différence de scores en modalités « avatar » versus « audio » plus importante pour certains indices morphosyntaxiques que pour d'autres.

Hypothèse 3 : L'avatar serait bien accepté des enfants. Ainsi, nous nous attendons à ce que les enfants interrogés aient une préférence pour l'avatar, qui a l'aspect d'un personnage de dessin animé, par rapport aux deux autres modalités (locuteur humain et enregistrement vocal).

La seconde partie de mon mémoire présentera les aspects expérimentaux de mon étude. Sur le plan méthodologique, la même épreuve sera administrée à une population d'une quarantaine d'enfants : 20 enfants normo-entendants (groupe contrôle) du même niveau scolaire - du CE1 au CM1- et 20 enfants présentant une surdité sévère à profonde et bénéficiant d'un gain prothétique apporté par tout type d'appareillage (groupe cible). Après la description des items sélectionnés pour construire notre épreuve de compréhension morphosyntaxique ainsi que du protocole de passation, nous présenterons les résultats obtenus. Dans un premier temps, l'étude des performances globales des enfants à travers les trois modalités permettra de repérer les différences entre chacune d'entre-elles. Dans un second temps, l'analyse des résultats obtenus pour chaque item du corpus permettra d'approfondir les différences intra-modalités et d'affiner notre étude. Enfin, nous analyserons ces résultats en fonction de nos trois hypothèses formulées en amont de notre étude et réaliserons des tests pour évaluer les éventuels indices de significativité statistique de ces résultats. Suite à cela, nous pourrions discuter de leur intérêt au vu de l'usage de ces outils dans une situation orthophonique.

PARTIE 1 :
ANCRAGE THEORIQUE

CHAPITRE 1

DEVELOPPEMENT LANGAGIER D'UN ENFANT PRESENTANT UNE SURDITE

Dans le cadre de l'élaboration de mon étude, je me suis appuyée sur une revue de la littérature afin d'élaborer un protocole à destination d'enfants d'âges scolaires, plus particulièrement ceux présentant une surdité, le plus pertinent possible, prenant en compte leurs spécificités langagières. Pour ce faire, je me suis basée sur la description réalisée par les auteurs, concernant tout d'abord l'acquisition du langage chez un enfant normo-entendant ainsi que les différences observées chez un enfant présentant une surdité. Ce sont donc ces données que je vais présenter ici, permettant à la fois d'appréhender le développement « ordinaire » du langage, indispensable en orthophonie, et de pouvoir comprendre les points de décrochage ou les fragilités d'enfants à besoins particuliers comme les enfants sourds, ainsi que les facteurs explicatifs et les aides compensatrices.

1 Développement langagier chez l'enfant « tout-venant »

1.1 Développement langagier général

Le développement langagier de l'enfant se construit dès sa naissance, voire avant. En effet, lors du dernier trimestre de grossesse, le système auditif du bébé est fonctionnel et lui permet de percevoir les variations mélodiques, l'accentuation et le rythme des voix (de Boysson-Bardies, 1996).

Durant sa première année de vie, l'enfant est baigné dans **un environnement linguistique** riche qui le prépare au langage (Machart, 2013). Il développe, à partir de capacités innées et de son environnement, un système organisé et modulé par les propriétés typiques de sa langue maternelle (Bertoncini & Boysson-Bardies, 2000). En effet, le nouveau-né dispose de capacités de discrimination du matériau langagier. Selon Bertoncini et Mehler (1980, p.860), « l'apprentissage débiterait par la mise en œuvre de mécanismes prédéterminés permettant à l'enfant, dans des conditions favorables, de développer des capacités phonologiques qui vont se focaliser sur l'appréhension des règles et des contrastes pertinents dans la langue de l'entourage ». Ainsi plusieurs études ont été réalisées afin de définir ces mécanismes prédéterminés chez le nourrisson. Selon une étude d'Eimas (1971), les nourrissons, dès 1 mois,

perçoivent la distinction entre différentes occlusives voisées (b, d, g) et non voisées (p, t, k). Ils sont donc dotés d'une perception catégorielle, caractéristique d'un mode de traitement strictement langagier. Ensuite l'expérience de Moffit (1971) témoigne que le jeune enfant de 5 mois discrimine la place d'articulation puisqu'il distingue deux sons du langage qui ne diffèrent que sur la transition ascendante ou descendante du second formant du segment initial. Enfin l'expérience de Fodor, Garrett et Brill (1975) montre que sur une série de trois syllabes (pi, pa, ku), des nourrissons âgés de 3 mois sont sensibles à l'identité du phonème initial, peu importe l'environnement vocalique qui lui est adjoint. Ainsi le nourrisson possède différentes capacités perceptives lui permettant de discriminer et de catégoriser les différents phonèmes de la langue de son entourage à partir du voisement, de la place d'articulation ainsi que du mode d'articulation, et ce malgré les nombreuses variations acoustiques de chaque phonème. Vers 6 mois, l'enfant est donc capable de discriminer correctement les voyelles de sa langue et il est capable de faire de même pour les consonnes vers 12 mois.

A ces capacités phonologiques s'ajoutent également des capacités de communication telles que le regard avec le contact visuel puis l'attention conjointe, ou encore le babillage. Ces capacités sont appelées les « pré-requis » au langage oral et à la communication.

Si le bébé produit des sons dès les premières semaines de sa vie, il ne réalise un contrôle articulaire que vers l'âge de 7 mois (Bertoncini & Boysson-Bardies, 2000). Durant la première année, les mécanismes de traitement perceptif s'organisent en un système modulé par des propriétés prosodiques et phonologique de la langue (Bertoncini & Boysson-Bardies, 2000). L'enfant commence à babiller vers l'âge de 10 mois, on note alors l'émergence d'une organisation phonologique qui permet entre 18 et 24 mois un accroissement brutal du lexique.

La première étape dans la construction de son lexique est tout d'abord de pouvoir segmenter la parole continue en unités-mots afin de pouvoir ensuite reconnaître différents mots familiers (Bertoncini & Boysson-Bardies, 2000). Cette capacité arrive généralement avant 1 an.

Entre 8 et 12 mois apparaissent généralement **les premiers mots** d'un enfant. Ce sont bien souvent des mots simples bisyllabiques qui désignent des personnes ou des objets. Selon Brent & Siskind (2001), l'enfant identifie des lexèmes : unité minimale de signification appartenant au lexique (Larousse, 1867), proposés et répétés régulièrement par ses parents. Ce peuvent être des lexèmes se rapportant aux membres de la famille proche (/maman/, /papa/, /papi/, /mamie/, /tata/, etc.) ou des objets du quotidien de l'enfant (/doudou/, /bibi/ pour biberon, /gâteau/, /pipi/, etc.)

Jusqu'à 3 ans, Bassano (2000) parle de « développement lexical précoce » car c'est une période cruciale de la constitution du lexique où l'enfant « découvre que tout chose peut-être nommée » (Bassano, 2000, p. 241).

Très vite, selon Rondal (1998), lorsque le jeune enfant accumule des connaissances du monde qui l'entoure, il ne peut exprimer tout ce qu'il désire à l'aide d'un seul mot. Il produit alors des successions de lexèmes en rapport avec la scène qu'il est en train de vivre au moment où il parle, par exemple « encore gâteau » ou « papa prend balle ». C'est à cet instant qu'on note **le début du langage combinatoire**.

1.2 Développement morphosyntaxique

Si l'ordre des mots énoncés par l'enfant est d'abord aléatoire, l'enfant repère par la suite certains mots dans le langage de l'adulte, qu'il reprend en leur attribuant une place fixe, souvent en début d'énoncés. Toujours selon Rondal (2018), il construit alors ses énoncés à partir « d'un mot-pivot » toujours à la même place suivi d'un mot non-pivot qu'il fait varier.

A partir de 30 mois, l'enfant organise ses énoncés avec des structures simples mais correctes : sujet, verbe, compléments. Il est alors capable d'utiliser différentes relations sémantiques manifestant par exemple l'existence ou la disparition, la récurrence, l'attribution, la possession, la localisation ou encore le bénéfice. Ces énoncés lui permettent de se faire comprendre dans l'instant par une personne ayant connaissance du contexte. L'enfant va par la suite pouvoir quitter « l'ici et du maintenant » et développer son langage grâce à l'adjonction graduelle de mots grammaticaux. Concernant les adverbes et les prépositions, l'enfant commence par maîtriser les marqueurs de possessions (« à moi ») aux alentours de 2 ans. Viennent ensuite les adverbes de lieu (« dedans, dessus, devant, derrière » ...) vers 3 ans puis les prépositions de lieu (« à, dans, sous, sur » ...) vers 3 ans et demi. Enfin les adverbes et prépositions de temps ne sont maîtrisés que vers 5 ans et demi.

Concernant les articles, apparaissent en premier lieu les articles indéfinis « un / une » puis définis « le / la » marquant le genre vers 3 ans, puis les articles marquant le nombre « des », « les » vers 4 ans. L'emploi des différents articles est généralement maîtrisé aux alentours de 6 ans. L'emploi des pronoms débute généralement avec « moi » à l'âge de 2 ans, puis apparaissent « je » et « tu » vers 2 ans et demi, « il » vers 3 ans et enfin « lui » et « eux » vers 4 ans. Certains pronoms possessifs plus complexes n'apparaissent quant à eux que vers l'âge de 6 ans. Enfin concernant l'utilisation de flexions verbales, l'emploi de l'infinitif présent et de l'indicatif présent à partir de 3 ans est observé. Vers 4 ans, l'utilisation des auxiliaires être et avoir

permettent la production du passé composé ainsi que l'expression du futur proche (va + infinitif). Le futur simple ainsi que l'imparfait de l'indicatif sont produits vers 5 ans et le conditionnel à partir de 6 ans.

En parallèle, l'enfant va produire des énoncés de plus en plus complexes. Vers 3 ans et demi, il commence à produire des subordonnées relatives ou complétives sans pronom relatif ou conjonction de subordination. Ce n'est qu'aux alentours de 4 ans que ses productions seront correctes. Il commence à utiliser des propositions circonstancielles de cause et de conséquences vers 4 ans et demi et des propositions circonstancielles de temps vers 5 ans et demi. De 4 à 6 ans environ, l'enfant produit des phrases négatives avec une structure négative et intègre progressivement l'ordre correct des mots, c'est-à-dire le renversement de l'ordre habituel. Cette opération prend plusieurs années du fait de la complexité de la construction des énoncés négatifs. **A 6 ans, l'enfant maîtrise** donc une grande partie des **règles morphosyntaxiques**. Un développement linguistique se poursuit néanmoins, notamment jusqu'à 10 ans afin de maîtriser la production de certains pronoms possessifs, de certains adverbes de temps et d'espace, certaines flexions verbales telles que le conditionnel passé ou le plus-que-parfait, ou encore la concordance des temps et les phrases déclaratives à la voix passive. La voix passive est d'abord comprise avant d'être produite. La voix passive comprend les phrases passives renversables et celles non renversables : les énoncés passifs non renversables sont très rarement utilisés dans le sens inverse. C'est le cas par exemple avec la phrase « la glace est mangée par le garçon. » qui n'a pas de sens si elle est renversée puisque la glace ne peut pas manger le garçon. A l'inverse l'énoncé « le camion est doublé par la voiture » est passif renversable puisque le sujet peut être inversé. Ainsi vers 4 ou 5 ans, l'enfant peut comprendre les phrases passives non renversables, mais ne les produit néanmoins que vers l'âge de 7 ou 8 ans. Il faut également attendre 7 ans avant qu'il ne comprenne les énoncés passifs renversables.

2 Conséquences de la surdité sur le développement du langage

2.1 Cadre général à propos de la surdité

La surdité est définie dans le champ médical comme étant un handicap sensoriel qui se définit par une déficience de l'audition plus ou moins sévère. Il existe différents **degrés de surdité** (Cf. Tableau 1) en fonction de la perte tonale de chaque oreille et les conséquences relatives à la surdité diffèrent en fonction de l'atteinte :

Niveau de surdité	Perte tonale	Conséquences
Surdité légère	Entre 21 et 40 dB	Difficultés à percevoir les voix faibles ou lointaines et les conversations, surtout dans un environnement bruyants
Surdité moyenne	Entre 41 et 70 dB	Perception difficile de la parole notamment en conversations de groupes, qui induit de la fatigue.
Surdité sévère	Entre 71 et 90 dB	Perception de la parole et des sons limitée aux sons forts et très proches. Une aide auditive est indispensable.
Surdité profonde	Entre 91 et 120 dB	La majorité des sons ne sont pas perçus. La communication orale sans appareil auditif est impossible. Certains sons, extrêmement forts, sont audibles.
Surdité totale ou cophose	Supérieure à 120 dB	Aucun son n'est perçu.

Figure 1. Tableau : Niveaux de surdité et incidences sur la langage et la communication (D'après (www.laboratoires-unisson.com, 2020))

Il existe également **différents types de surdités**, selon le lieu de la lésion. Les surdités de transmission concernent une atteinte au niveau de l'oreille externe ou moyenne. La perte tonale moyenne qu'elle engendre n'excède pas 60 dB. On parle alors de surdité moyenne, voire de surdité légère si la perte tonale est inférieure à 40 dB. Elle ne nécessite pas toujours l'utilisation d'une remédiation prothétique.

Les surdités de perception sont, quant à elles, dues à une atteinte de l'oreille interne ou de ses afférences. Elles sont le plus souvent incurables et nécessitent une **remédiation prothétique : un appareillage auditif ou un implant cochléaire**. Une surdité de perception peut aller jusqu'à une perte tonale de 120 dB, c'est-à-dire une surdité totale, ou cophose. L'individu ne dispose plus d'aucun reste auditif.

Enfin il est question de surdité mixte lorsqu'un individu est atteint simultanément de surdités de perception et de transmission.

Selon le degré d'atteinte, la surdité a **de multiples conséquences**. Le dictionnaire d'orthophonie (Brin-Henry et al., 2011, p. 263) évoque un « trouble de la communication préverbale chez le nourrisson avec incidences développementales, absence ou retard de langage, troubles de la parole et de la voix, difficultés d'intégration scolaire et/ou sociale ». En effet, la déficience sensorielle engendre de multiples difficultés. En premier lieu, l'ouïe assure une fonction vitale : alerter. Une déficience auditive non prise en charge peut donc se révéler dangereuse puisque l'individu peut manquer des indices sonores tels que des alarmes ou des cris qui le préviendrait d'un danger potentiel.

2.2 Modifications dans la construction du langage oral chez l'enfant malentendant

Une déficience auditive engendre également **des troubles de la communication, de la parole et de la voix**. En effet, tout-petit, l'enfant interagit avec son environnement, tout d'abord à travers des gestes (le pointage) et des regards (l'attention conjointe) ainsi que par des cris. En effet dans les deux premiers mois, avant que le larynx ne descende progressivement à partir de 3 mois, le nourrisson communique par des cris différenciés lui permettant de manifester des états émotionnels variés (Prat, 2007). C'est grâce à la distinction faite par la mère des différents cris que le bébé apprend à les moduler et à s'ajuster pour obtenir une réponse satisfaisante. Ainsi, petit à petit, en comprenant que ses actions et sa voix provoquent des effets sur son entourage, il construit son langage par imitation. Il ajuste et développe alors ses productions grâce au feed-back que lui renvoie son entourage.

Chez l'enfant malentendant, le défaut de perception de ce feed-back ne lui permet pas de progresser aussi vite qu'un enfant normo-entendant. Un nombre plus important de répétitions est donc nécessaire pour qu'il puisse assurer et différencier ses productions orales. Les informations reçues telles que les mots, les phrases mais aussi la prosodie et l'intonation, sont donc **limitées et partiellement traitées** (Hage et al., 2006). C'est donc cette difficulté à percevoir et traiter les informations langagières qui vont créer des retards dans l'acquisition du langage chez les personnes atteintes de surdit .

Les enfants sourds possèdent tout d'abord **un stock lexical** moins important. Selon une étude évaluant le lexique de production d'enfants sourds profonds munis d'un implant cochléaire (Le Normand, 2004), concernant la morphologie lexicale, la moyenne du nombre de mots différents pour un enfant implanté depuis 36 mois est d'environ 30 mots, ce score équivaut à la moyenne d'un enfant âgé entre 27 et 30 mois. Les sujets de cette étude ont été implanté entre 21 et 78 mois. 70% d'entre eux ont été implantés avant 4 ans. On note donc un décalage pour chacun d'eux d'au minimum 2 ans. Ce décalage perdure puisque la moyenne du stock lexical est d'environ 12 000 mots chez un individu de 18 ans atteint de surdit  contre 30 000 mots chez un normo-entendant du m me  ge (Machart, 2013).

Concernant l'acquisition **de la morphosyntaxe**, des études montrent que les enfants atteints de surdit  utilisent principalement des noms et des verbes et peu d'adjectifs, d'adverbes, de mots-outils (Goda, 1964), d'auxiliaires, de pronoms, de pr positions, de quantificateur et de pronoms ind finis (Brannon Jr, 1968). Geffner et Freeman (1980) ajoutent que les structures syntaxiques sont simplifi es chez les enfants sourds : Sujet – Verbe – Objet et que cette syntaxe varie peu dans leur discours.

CHAPITRE 2

INTERVENTION GLOBALE DANS UNE DEFICIENCE AUDITIVE

Mon étude concerne des enfants présentant une surdité. A partir d'une revue de littérature, je vais donc à présent détailler les parcours de soin ainsi que les différents moyens mis en place pour pallier la déficience auditive dans le but d'observer les forces et les faiblesses de ces moyens.

1 Gain prothétique

Selon l'OMS (2017) 360 millions de personnes souffrent d'une **déficience auditive incapacitante** c'est-à-dire supérieure à 40 dB, ce qui représente plus de 5% de la population mondiale. Les aides auditives sont donc un **enjeu de société majeur**. Il existe différents moyens techniques pour permettre à une personne atteinte de déficience auditive de recouvrer une partie de son audition. Selon Lina-Granade, Gallego, Thai-Van, & Truy, (2010, p.6), « L'objectif de l'appareillage est de compenser autant que faire ce peut les déficits du système auditif du patient en améliorant les seuils auditifs ainsi que l'intelligibilité des sons de parole »-. Le choix de l'appareil se fait en fonction du degré de surdité de la personne, de la localisation de son atteinte ainsi que de l'environnement de la personne et de ses choix personnels. La grande majorité des remédiations prothétiques sont réalisées dans le cadre d'une surdité de perception, c'est-à-dire lors d'une atteinte de l'oreille interne : bien souvent de la cochlée : organe interne de l'audition.

La surdité se traduit par **une élévation des seuils d'audition** : l'individu ne perçoit plus les sons en dessous d'une certaine intensité. Dans le cas d'une atteinte cochléaire, elle se traduit également par **un phénomène de distorsion** : une distorsion du codage de l'intensité, c'est-à-dire que le seuil de douleur est plus proche du seuil d'intelligibilité que la normale ; une distorsion du codage de la fréquence, c'est-à-dire que les bruits de l'environnement, principalement en basses fréquences, interfèrent avec les sons de la parole, généralement de haute et moyenne et fréquences et parasitent la compréhension ; enfin la distorsion de la dimension temporelle : une difficulté pour les individus à percevoir des sons brefs, c'est-à-dire inférieurs à 200 msec (Krahe & Alegria, 2007).

1.1. Physiologie cochléaire

L'oreille interne assure une phase de perception et de transduction du son ainsi que la transmission nerveuse du son aux voies auditives centrales et au cortex primaire et associatif. L'oreille interne est composée des labyrinthes osseux externe et interne. Le labyrinthe osseux comprend le vestibule, les canaux semi-circulaires et la cochlée spiralée. Les canaux semi-circulaires interviennent dans le maintien de l'équilibre et dans l'orientation du corps (McFarland, 2009).

La cochlée est la **partie responsable de l'audition**. Elle est localisée dans l'os temporal et communique avec l'oreille moyenne par deux orifices fermés par des membranes appelés la fenêtre ovale, qui limite la rampe vestibulaire et la fenêtre ronde qui clôt la base de la rampe tympanique et permet l'équipression. La cochlée est constituée de trois canaux enroulés en spirale autour d'un axe osseux : la rampe tympanique, la rampe vestibulaire et le canal cochléaire.

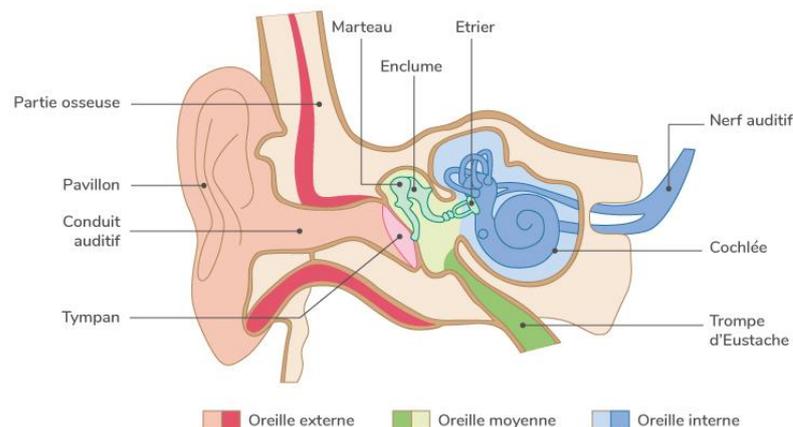


Figure 2 : schéma : description complète de l'oreille (Site internet www.bivea.fr)

La cochlée abrite l'organe de Corti, situé entre le canal cochléaire et la rampe tympanique. Cet organe contient les cellules sensorielles ciliées externes et internes. Les cellules ciliées externes permettent la sensibilité et la sélectivité du son en fréquence. Les cellules ciliées internes, quant à elles, permettent de transformer l'énergie acoustique en énergie électrique qui est ensuite transmise au nerf auditif (Krahe & Alegria, 2007). La cochlée réalise l'analyse des sons incidents en codant les différentes fréquences du son. Chaque fréquence est codée à un endroit particulier de la cochlée. Les hautes fréquences sont extraites à la base de la cochlée tandis que les basses fréquences sont extraites à l'apex de la cochlée. L'onde sonore aérienne est donc codée en fonction de sa fréquence et transformée en stimulus mécanique grâce à la variation de la pression de l'air sous forme de mouvement vibratoire (Krahe & Alegria, 2007).

Ce mouvement fait vibrer la membrane basilaire. L'amplitude de cette vibration permet de coder l'intensité du signal sonore tandis que le lieu de la vibration sur la membrane permet de coder la fréquence du signal. Le signal est ensuite transformé en signal électrique par les cellules ciliées internes et transmis au nerf auditif.

Une altération d'une partie de la cochlée entraîne une **rupture** dans le processus de **perception du son** provoquant une surdité de perception. Aujourd'hui, il existe des moyens de remédiation spécifiques aux atteintes cochléaires.

1.2. Implant cochléaire

L'implant cochléaire est un dispositif électro-acoustique permettant de recouvrer une fonction auditive chez des personnes atteintes d'une surdité neurosensorielle c'est-à-dire un dysfonctionnement de l'oreille interne.

Dans le cadre d'une surdité de perception, l'oreille interne est endommagée, en particulier les cellules ciliées externes et internes, entraînant une perte auditive sévère voire totale en cas de disparition totale des cellules ciliées.

Dans les années 1980, des appareils permettant de remplacer une cochlée atrophiée sont développés. On les appelle les implants cochléaires. Ils analysent et recodent les sons à la place de la cochlée et envoient l'information aux fibres nerveuses cochléaires (Wathour et al., 2019). Ils se composent d'une partie interne et d'une partie externe. La partie interne comprend une antenne réceptrice ainsi que des électrodes insérées dans la rampe tympanique de la cochlée. La partie externe se compose d'un boîtier contenant un processeur vocal, un contour d'oreille muni d'un microphone et une antenne émettrice (Krahe & Alegria, 2007). Le son est enregistré par le microphone. Le processeur vocal récupère les informations sonores et transforme l'énergie acoustique en énergie électrique. L'énergie est alors transmise par les antennes jusqu'aux électrodes via des impulsions analogiques ou digitales. Le signal est ensuite transmis au nerf auditif ainsi qu'au cerveau qui le traite.

Depuis, les implants cochléaires n'ont cessé de se perfectionner. Les implants multicanaux actuels peuvent comporter jusqu'à une vingtaine d'électrodes. Ils permettent également aux enfants sourds profonds de recouvrer une partie de leur audition et de développer un langage oral. On estimait en 2007 le nombre d'implants dans le monde à 40 000 (Krahe & Alegria, 2007) et on estime aujourd'hui que 45 000 nouveaux implants cochléaires sont implantés chaque année dans le monde dont 1 400 en France (Rosier, 2017).

L'implant cochléaire est préconisé dans le cadre de surdités profondes, lorsqu'une prothèse auditive classique ne permet pas un gain auditif suffisant. Dans une surdité congénitale, l'indication pour une implantation chez un enfant est recommandée durant la période pré-linguistique, dans l'idéal dès 12 mois, afin de permettre à l'enfant de bénéficier d'un bain linguistique conséquent qui favorisera le développement de son langage oral.

Dans le cadre d'une surdité acquise, on constate de meilleurs résultats avec une implantation rapide après la survenue de la surdité.

L'implant cochléaire ne suffit pas lui-même à recouvrer l'audition, il nécessite une **prise en charge pluridisciplinaire** longue afin de régler correctement l'appareil et d'apprendre à percevoir et à reconnaître les sons.

Dans certains cas, la remédiation auditive ne nécessite pas de cophochirurgie, c'est à dire d'implantation cochléaire. C'est le cas pour les surdités de transmission où la cochlée n'est pas endommagée ainsi que pour certaines surdités de perception où un appareillage classique suffit à recouvrer un gain auditif suffisant. Également en cas de surdité sévère voire profonde, soit car le patient ou sa famille la refuse, soit car elle est impossible à cause d'une malformation ; ou bien encore à la suite d'un échec de chirurgie. Dans ce cas, les prothèses amplificatrices sont proposées.

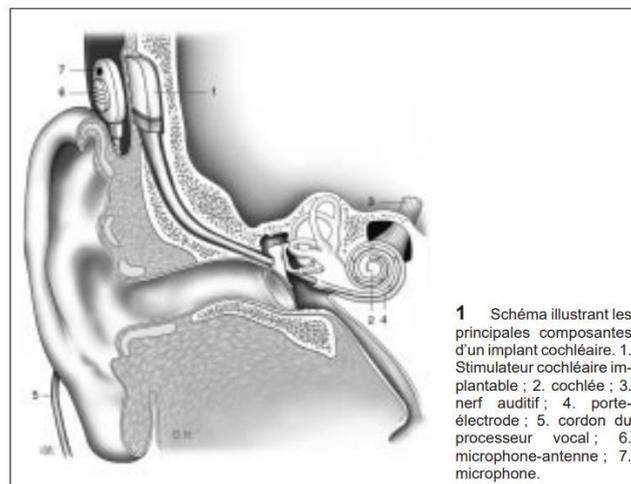


Figure 3 : schéma : principales composantes d'un implant cochléaire (Dauman et al., 1998,p.2)

1.3. Prothèses amplificatrices

Pour une surdité de transmission ou dans certains cas de surdité de perception, le patient se voit proposer une **remédiation amplificatrice**. Cette restauration fonctionnelle se fait à partir de prothèses amplificatrices, le plus souvent par voie aérienne, mais aussi par voie non

aérienne. Les prothèses amplificatrices, via une amplification du son perçu, agissent sur la perte d'audition, le phénomène de recrutement, l'élargissement des filtres cochléaires et l'accroissement des effets de masques.

1.3.1. Par voie aérienne

Les prothèses amplificatrices par conduction aérienne les plus répandues sont les « contours d'oreille » et les appareils intra-auriculaires. Les sons de l'environnement sont captés par un microphone et transformés et amplifiés par un microprocesseur. Le son est ensuite transmis par l'écouteur (Wathour et al., 2019).

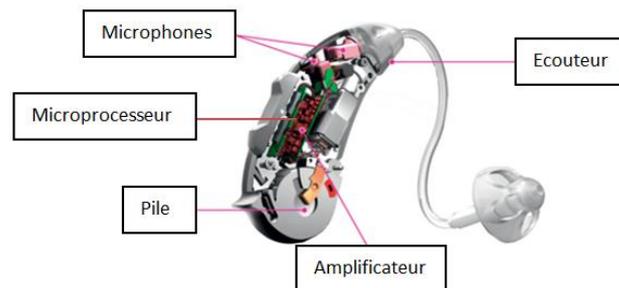


Figure 4 : schéma : principales composantes d'une prothèse amplificatrice contours d'oreille

(Site internet www.auditionimplantscontour.e-monsite.com)

1.3.2. Par voie osseuse

Lorsque les prothèses amplificatrices n'utilisent pas la voie aérienne, elles peuvent utiliser la voie osseuse. Cette technique repose sur un principe de transmission d'information auditive aux liquides cochléaires par ébranlement de la boîte crânienne. Selon Truy, Gallego, Bouccara, & Tringali (2012, p.1), « l'oreille interne est en partie constituée d'un os dense entouré d'une structure osseuse plus souple, la mastoïde. Lorsque l'on vient faire vibrer le crâne, l'élasticité de la mastoïde transmet la vibration à l'oreille interne. Cette énergie vient mettre en mouvement les liquides cochléaire et ainsi transmettre et transformer l'onde vibratoire en influx nerveux ».

La conduction osseuse est indiquée dans certaines surdités de transmissions ou mixtes ainsi que dans des surdités unilatérales. Il s'agit soit de vibrateurs externes ou d'une stimulation osseuse directe grâce à des prothèses implantées. Il est aussi possible de venir apporter des stimulations directement au niveau des osselets ou de la fenêtre ronde grâce à des implants d'oreille moyenne : cette remédiation est proposée pour certaines surdités neurosensorielle, transmissionnelles ou mixtes.

1.4. Prothèses à transposition de fréquence

Hormis les implants cochléaires et les prothèses amplificatrices il existe un troisième type de remédiation auditive : les prothèses à transposition de fréquence. Elles ne sont utilisées que dans des cas particuliers. Lorsque la perception dans les hautes fréquences est trop détériorée, l'amplification du son ne permet pas l'intelligibilité. Ces prothèses permettent alors de convoier les informations contenues dans les portions supérieures du spectre, plus aigües, vers les fréquences plus graves, mieux préservées.

Si les appareils auditifs sont une grande partie dans la prise en charge de la surdité ils sont rarement isolés. A l'aide auditive s'ajoutent des supports visuels à la communication tels que la lecture labiale ou le langage parlé complété pour accompagner la personne atteinte de surdité dans sa compréhension.

2. Communication alternative augmentée sans aide technique

La communication alternative et augmentée fait référence au domaine de la pratique clinique qui tend à apporter des compensations aux déficits et incapacités des individus souffrant de troubles de la communication (Alm & Parnes, 1995). Ces compensations peuvent comporter - ou non - des aides techniques. Les techniques de communication alternative augmentée sans aide technique utilisent les sens intacts de l'individu, le plus souvent la vue avec la parole visuelle ou la lecture labiale ou en se servant des mains de l'individu comme c'est le cas avec la langue des signes ou la langue parlée complétée.

2.1. Parole visuelle

D'après Dumont & Calbour (2002), la perception de la parole se fait à travers plusieurs étapes. Nous devons d'abord trier les sons de la parole parmi les autres sons parasites environnants. Puis, il est nécessaire d'extraire des invariants phonologiques en prenant en compte les variations induites par le locuteur : son articulation, son débit, son accent, etc...Enfin, le signal filtré doit être segmenté pour en extraire le message linguistique envoyé. A toutes ces étapes s'ajoutent plusieurs complexités parmi lesquels le chevauchement des indices acoustiques des phonèmes ou la variabilité des indices acoustiques d'un phonème en fonction du contexte, vocalique par exemple.

Ainsi un normo-entendant, en plus des indices acoustiques, s'appuie, lorsqu'il le peut, sur des **indices visuels** transmis par son locuteur : l'articulation, les mimiques, les gestes, le contexte. Ces indices visuels sont encore plus précieux chez une personne présentant une surdité. Ils contribuent à son acquisition du langage oral (Dodd et al., 1998). La lecture labiale lui permet de pallier le manque d'informations sonores intégrées, par certaines informations visuelles pour comprendre le sens du message transmis. Tous ces indices visuels forment la parole visuelle, et cette dernière s'est développée au moyen de nouvelles techniques pour aider au mieux la personne sourde (Pandzic et al., 1999).

La **parole visuelle** entre en jeu dans la **perception audiovisuelle de la parole**. La perception audiovisuelle de la parole est « un processus complexe dans lequel les informations initialement séparées entre les modalités sensorielles auditives et visuelles sont combinées avec la connaissance linguistique antérieure, stockée en mémoire à long terme » (Wathour et al., 2019, p. 7). Elle est utilisée par tous les locuteurs en situation d'échanges en face à face. En effet lors d'un échange téléphonique, la parole visuelle n'entre pas en compte dans la compréhension, ce qui explique en partie le fait qu'il est parfois plus difficile de comprendre notre locuteur dans ce contexte.

La vision est utilisée dans la communication à différents niveaux. Tout d'abord, elle donne des informations relatives aux émotions du locuteur grâce à ses mimiques faciales : ses yeux, sa bouche, ses sourcils. Le regard donne du sens à nos messages. Il a également une fonction phatique, c'est à dire qu'il permet de réguler la conversation sans communiquer un message. L'attention visuelle permet de traiter les informations non verbales telles que les expressions du regard, du visage et de la posture renforcer, exprimées lors d'interactions afin de détecter les intentions du locuteur, son acceptation ou son rejet par exemple, dans le but de se rassurer ou d'adapter sa réponse (Cozolino, 2014).

Elle complète également la compréhension du message oral distribué. En effet dans un environnement bruyant ou en cas de déficience auditive grâce à la lecture labiale. Selon une étude de Sumbly & Pollack (1954), la lecture labiale augmente la performance de la perception de la parole par les entendants, surtout quand le signal acoustique est masquée par le bruit. D'autres études ont été réalisées dans ce sens : notamment par Massaro & Cohen (1995) et Reisberg, Mclean, & Goldfield (1987). Ce phénomène a également été étudié et observé en langue française par Benoit et son équipe (1994). De plus elle améliore également la compréhension d'un signal clair mais comprenant un contenu sémantiquement complexe (Reisberg et al., 1987).

Au-delà d'un indice supplémentaire pour la compréhension, la parole visuelle peut également constituer une gêne à la compréhension lorsqu'elle ne coïncide pas avec le message audio. C'est un fait que l'on constate en cas de décalage son et image dans un film par exemple. Ce phénomène a notamment été étudié par McGurk & Macdonald (1976). Ils ont démontré que la lecture labiale peut influencer le traitement de la parole en présentant un stimulus auditif en contradiction avec le signal apporté par la lecture labiale. Les résultats montrent que dans ce cas, le système perceptif fusionne les informations des deux modalités pour en créer un percept unifié original, ne correspondant ni au stimulus auditif, ni au stimulus visuel : c'est ce qu'on appelle l'Effet McGurk.

Ainsi de nombreuses études prouvent que la lecture labiale fait partie intégrante des processus de perception et de traitement de la parole également chez les normo-entendants.

2.2. Lecture labiale

La lecture labiale est définie par la « perception visuelle du langage à l'aide de la reconnaissance des mouvements articulatoires du locuteur » (Brin-Henry et al., 2011, p.141). Cette perception nécessite un **apprentissage** et différentes compétences car les personnes qui la pratiquent se heurtent à plusieurs difficultés.

Tout d'abord la parole est un flux sonore continu, le labiolecteur afin de pouvoir décoder une parole doit segmenter ce flux en unités discrètes. Il doit ensuite utiliser des compétences dites analytiques pour identifier ces unités discrètes en se construisant des invariants lui permettant d'extraire des informations langagières. Pour cela, il identifie des gestes articulatoires appelés visèmes. Mais il n'existe pas un geste articulatoire par phonème. Concernant les 15 consonnes de la langue française, elles sont toutes regroupées en seulement **6 visèmes** (Wathour et al., 2019), voir le tableau ci-dessous.

Mouvements de lèvres CONSONNES

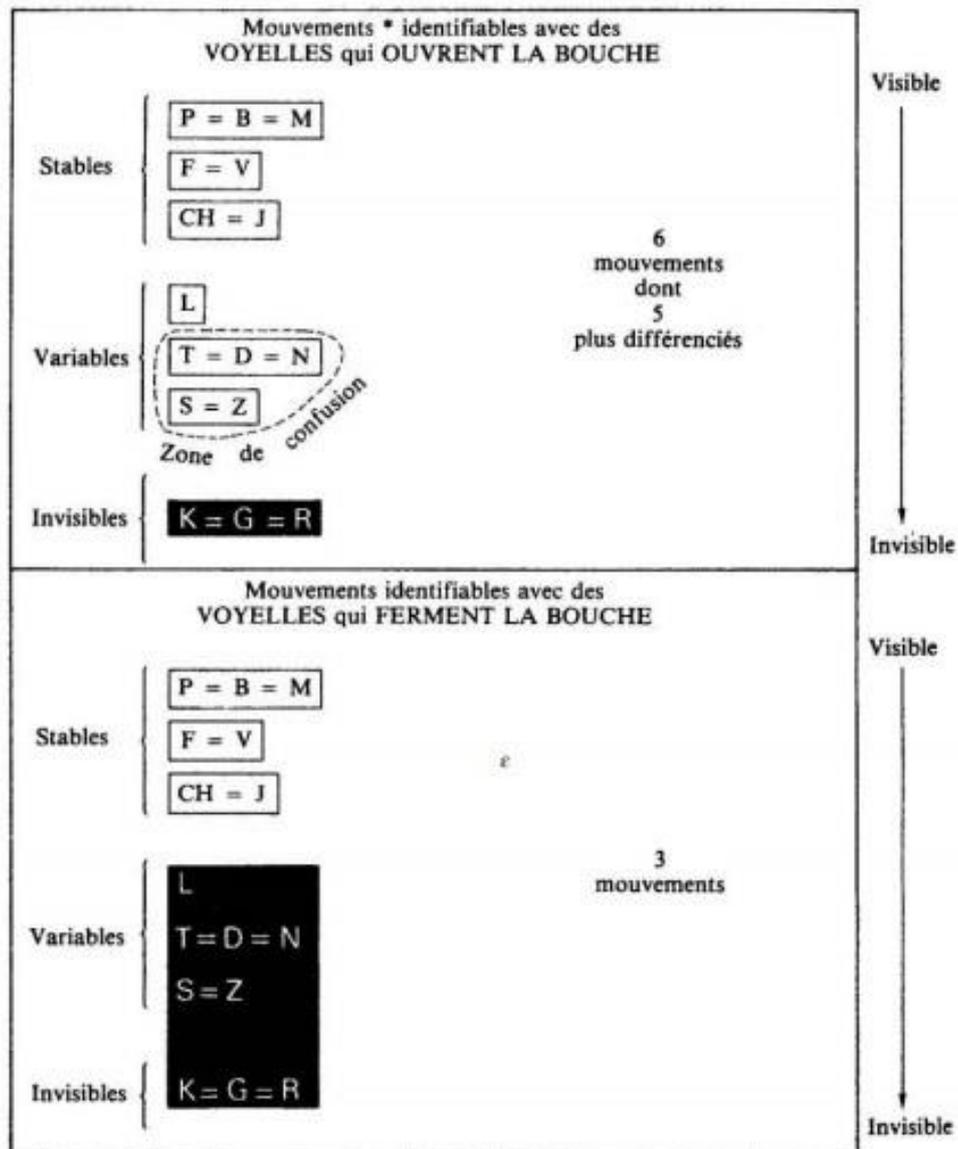


Figure 5 : Tableau des mouvements de lèvres, consonnes (Istria et al., 1982, p.174)

Les voyelles comptent 7 images articulatoires différentes pour 13 phonèmes. Nous avons donc à faire à plusieurs sosies labiaux : des sons ou des mots ayant la même image labiale. C'est le cas par exemple pour les sons [y] et [u] qui ont la même image articulatoire ou encore [a] et [ã]. A cette première difficultés s'ajoutent des phénomènes de coarticulation modifiant l'image articulatoire, l'articulation du locuteur ou encore des facteurs externes de visibilité tels que la position, la luminosité ou la distance. C'est pourquoi le labiolecteur doit faire preuve d'une compétence de synthétisation : c'est-à-dire une capacité à compléter et interpréter les signes et

les indices à l'aide de la suppléance grâce aux connaissances linguistiques, situationnelles et contextuelles syntaxiques dont il dispose. Cette suppléance mentale varie d'un individu à l'autre en fonction de sa capacité à générer des inférences et de ses compétences comportementales à utiliser des stratégies de négociation et d'anticipation pour mieux comprendre.

Cependant, **la lecture labiale** et les compétences qu'elle requière ne sont pas innées. Elle nécessite **un apprentissage auprès d'un orthophoniste**. Cet apprentissage est long puisqu'il dure plusieurs mois à raison d'une à deux séances par semaine. Dans la langue française on identifie trois méthodes d'apprentissage : analytique, globale-synthétique et mixte. La méthode analytique apprend à reconnaître chaque visème afin d'automatiser leur reconnaissance pour aller à la compréhension tandis que la méthode globale-synthétique privilégie la compréhension du message global par tous les indices. La suppléance mentale est constamment travaillée. Enfin la méthode mixte se repose sur la perception, l'interprétation et la complétion : on débute par un travail analytique pour rapidement aller à un travail plus global.

2.3. Langage parlé complété (LPC)

Il existe d'autres supports visuels pour aider la communication orale. Parmi elle se trouve la langue parlée complétée. Ce système tire son origine du docteur R. Orin Cornett qui invente à la fin des années 60 le « Cued Speech » qui est la version anglophone de la langue française parlée complétée (LPC). Cette méthode utilise la main en complément de la lecture labiale à travers des clés manuelles où la forme de la main code les phonèmes consonantiques et la position de la main code les phonèmes vocaliques. Il existe donc 8 configurations pour les consonnes et 5 positions pour les voyelles où configuration et position sont effectuées simultanément permettant ainsi un codage syllabique. Elle a été inventée dans le but de permettre aux personnes atteintes de surdité de « percevoir sans ambiguïté le langage parlé » (Leybaert, 2011). Il crée ainsi une « troisième voie, entre une communication basée sur l'oralisme strict et une communication basée sur la langue des signes » (Leybaert, 2011). En effet selon lui l'oralisme strict se basant sur les restes auditifs de l'individu et la lecture labiale, cela ne permettait pas une compréhension totale en raison des ambiguïtés provoquées par la lecture labiale. De même le problème majeur de la langue des signes réside dans le fait qu'elle est une langue à part entière avec une phonologie, un lexique et une morphosyntaxe propre. Elle est donc le fruit d'un apprentissage long et coûteux et ne permet pas l'ouverture sociale avec un plus grand nombre. Ainsi le LPC est une forme visuelle du français parlé qui s'apprend rapidement. Le deuxième avantage étant qu'il facilite l'apprentissage de la lecture puisqu'il

aide à « construire des épisodes d'attention conjointe, au cours desquels se développent les représentations lexicales, et le début de la grammaire » (Leybaert & Colin, 2007, p. 250). Également, le fait que les parents codent en LPC permet à l'enfant de différencier des formes phonologiques proches et ainsi de construire des représentations phonologiques précises et ainsi de développer sa morphosyntaxe (Leybaert & Colin, 2007).

CHAPITRE 3

EVOLUTIONS TECHNOLOGIQUES : APPORTS A LA COMMUNICATION ALTERNATIVE AUGMENTEE

Notre étude se basant sur l'évaluation d'un dispositif numérique utilisé dans le cadre de la communication, nous sommes amenés à faire l'état des lieux des outils techniques existants à ce sujet afin de comprendre tout l'enjeu du développement de ce type de procédés. Avant d'étudier la littérature existante concernant l'avatar parlant, nous ferons donc un rapide état des autres dispositifs existants.

1 Moyens de communication alternative et augmentée

Les personnes atteintes de surdit   disposent de deux types de moyens de communication alternative et augment  e (CAA) : avec ou sans aide technique. Le LPC, la langue des signes ou encore la lecture labiale font partie des moyens de CAA sans aide technique et n  cessitent un apprentissage comme nous l'avons d  crit pr  c  demment.

1.1. Naissance des outils de CAA avec aide technique

La communication alternative et augment  e avec aide technique comprend tous les outils mis    disposition pour faciliter la communication et le langage. Certains outils existent depuis de nombreuses d  cennies, ils ne n  cessitent pas d'apprentissage particulier et sont compris par la majorit   des locuteurs. Ce sont par exemple l'utilisation de pictogrammes, de l'  criture, d'images, de tableau de communication ou photos.

1.2. Apport de la technologique dans les CAA avec aide technique

Ces derni  res d  cennies, les communications alternatives et augment  es avec aides techniques ont   volu   gr  ce    la technologie offrant aux personnes sourdes des nouveaux moyens pour communiquer    travers des   crans (Antoine & Maurel, 2007). C'est le cas de la synth  se vocale par exemple, permettant de saisir un message    l'aide d'un clavier phon  tique, en langue des signes ou par ic  nes avec l'alphabet dactylologique. Le message est ensuite vocalis   par l'appareil (Moudenc & Emerard, 2003). On note aussi le d  veloppement des aides    la prise de notes sous forme de logiciels rempla  ant un clavier traditionnel par un clavier st  notype qui utilise une   criture abr  g  e form  e de signes conventionnels.

2 Avatars parlants

Les aides techniques de communication alternative et augmentée telles que l'utilisation de signes, d'images ou de synthèses vocales permettent au locuteur en difficulté de communication et à son interlocuteur de faciliter les interactions en multipliant les canaux de réception et de production.

En situation d'interaction directe, les locuteurs peuvent **s'ajuster en fonction du feedback** fourni par leur interlocuteur si bien qu'en cas d'incompréhension, le message est redonné de différentes manières jusqu'à la compréhension des deux interlocuteurs.

Néanmoins, **en situation de message différé**, le locuteur en difficulté, par exemple atteint d'une déficience auditive, ne peut pas bénéficier de cette reformulation et de cette attente de compréhension. C'est le cas lorsqu'il écoute un message vocal, audio ou une vidéo. Dans le cas d'un message audio, **l'individu ne dispose que du message auditif pur** et ne peut s'aider de la parole visuelle. Dans le cas de vidéos, il existe plusieurs possibilités. Soit une personne est filmée, mais dans ce cas les conditions pour que l'individu obtienne un nombre suffisant d'indices visuels dépendent des critères tels que la lumière, la qualité de l'image ou encore l'orientation du visage. Soit le message est écrit, mais cette condition nécessite la maîtrise du langage écrit et ôte l'aspect confortable et rassurant de l'humain et du message verbal (Walker et al., 1994). En effet dès la naissance, le nourrisson montre une **préférence pour les schémas ressemblant à un visage** (Bond, 1972). Dès deux mois ils commencent à traiter les expressions faciales et à différencier les caractéristiques visuelles spécifiques du visage (Carey, 1992).

Aussi, les interfaces numériques ont évolué en ce sens et ont été **humanisée**. En 1974, Parke a conçu le modèle du visage entier. Ce modèle a ensuite été amélioré, notamment concernant la production vocale par Cohen et Massaro (1995).

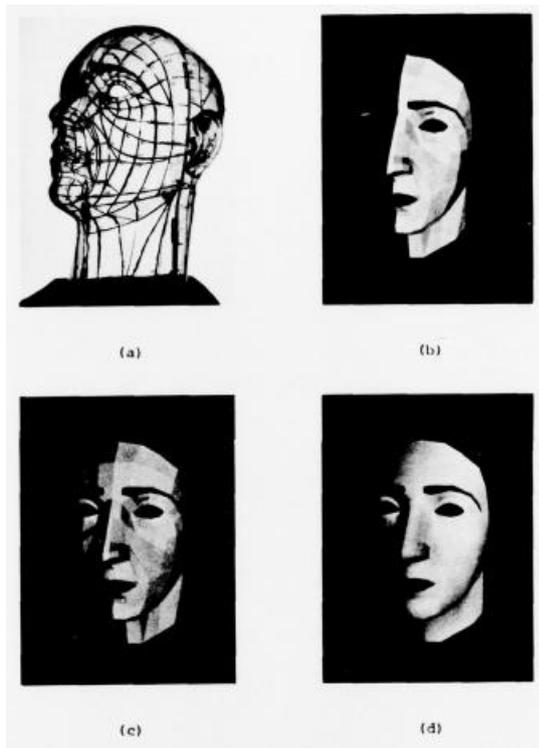


Figure 6 : Modèle du visage entier (Parke, 1974, p.7)



Figure 7 : Modèle de Cohen et Massaro (1995, p.3)

Ces visages numériques, aussi appelés avatars, ont été créés pour délivrer des messages informatiques, dans le but de rendre les interfaces plus faciles et plus confortables à utiliser (Walker et al., 1994). Différentes études ont été réalisées avec des avatars. En juin 1998, Pandzic les a étudiés à travers 190 sujets : des adultes anglophones étudiants ou employés de l'Université Princeton. Aucun sujet de cette étude ne présentait de déficience auditive. Ces études témoignent que la présence de visages animés apporte des avantages à différents niveaux. Tout d'abord concernant la compréhension, elle aide les utilisateurs à mieux comprendre le message dans des conditions bruyantes (Panzic et al., 1999). De plus, les visages animés sont plus attrayants et permettent de rendre les temps d'attente plus acceptables (Panzic et al., 1999). Une étude de Walker et al a également été menée en 1994 sur le personnel d'un laboratoire de recherche informatique d'une entreprise industrielle. **Les sujets étaient donc des adultes, anglophones, diplômés et ne présentant pas de troubles de la communication.** Le but était de comparer une interface informatique sous la forme d'un texte, avec deux interfaces informatiques présentées sous la forme de visages, l'un d'expression neutre et l'autre d'expression sévère. Les résultats montrent que les sujets adultes de l'étude utilisant l'interface avec le visage parlant ont fait moins d'erreurs, ont passé plus de temps sur le questionnaire et écrit plus de commentaires que les sujets ayant reçu les messages écrits (Walker et al., 1994).

D'autres études ont étudié l'aspect de l'avatar, montrant alors que **plus l'avatar ressemble à un humain, plus il est perçu comme fiable par les sujets de l'étude** (Van Mulken et al., 1999). N'ayant aucune information sur l'échantillon de sujets sur lequel l'étude de Van Mulken se repose, nous ne pouvons pas relier les conclusions de l'étude à une population particulière.

Ainsi en 2008, un avatar numérique a permis la création d'un outil à destination des enfants porteurs de handicaps tels qu'une surdité, une dysphasie ou une dyslexie (Piquard-Kipffer et al., 2010). Des livres numériques ont été imaginés et développés à l'aide d'une tête codeuse Datha-Inria dans le cadre de l'action de développement technologique Handicom de l'INRIA. Cette collection de trois livres permet à l'enfant de suivre une histoire à travers différents supports : des images illustrant tous les éléments du texte, le texte apparaissant à l'écrit et oralisé au moyen d'une tête parlante et/ou codeuse qui raconte l'histoire. L'enfant peut s'aider, s'il choisit cette option de lecture, des signes LPC ainsi que des mouvements labiaux reproduit par la tête. Ce projet, présenté dans un documentaire scientifique (<https://interstices.info/je-peux-voir-les-mots-que-tu-dis/>) a donné lieu à plusieurs études, notamment dans le cadre de mémoires de recherche en orthophonie. Parmi ceux-ci, Gabet et Thiollier, (2012) avaient pour objectif, à partir de phrases de différents niveaux de complexité morphosyntaxique issues du prototype « Nina fête son anniversaire », de comparer le niveau de compréhension verbale d'enfants âgés entre 4 et 11 ans et présentant une surdité profonde, provenant soit d'une locutrice humaine utilisant le LPC, soit d'un avatar codant. Les habiletés de compréhension étaient testées à travers la capacité de répétition des enfants des phrases vues et entendues. Les résultats ont montré que les enfants percevaient mieux les marques morphosyntaxiques avec l'aide du L.P.C. Malgré des différences suivant les enfants, elles ont tout de même observé de meilleurs résultats avec le visage humain plutôt que la tête codeuse numérique.



Figure 8 : Visuel du livre numérique « Nina fête son anniversaire » (A. Piquard-Kipffer, 2011, p.7, prototypes Inria)

Les **avatars parlant** sont donc un **enjeu majeur dans le développement de la communication audio-visuelle numérique**. Il semble essentiel de développer, en parallèle de l'apparence de ces visages animés, leur articulation et la synchronisation de la parole avec les lèvres afin de limiter au maximum les phénomènes de perception de la parole audiovisuelle incongrue ou effet McGurk. Pour ce faire, les chercheurs de l'équipe Multispeech du laboratoire INRIA Nancy-Grand Est, dans le cadre du projet Dynalips porté par Slim Ouni (<https://dynamips.loria.fr>), ont conçu un **outil informatique permettant la synchronisation des mouvements labiaux d'un personnage avec sa parole**. L'innovation technique apportée par ce projet est portée en particulier sur les phonèmes ainsi que sur les transitions (Shadows, 2018). Un algorithme d'animation, appelé articulographe a été réalisé qui permet de modifier la partie inférieure du visage en exploitant l'influence de points de contrôle (Ouni & Gris, 2018). Ainsi l'outil synchronise une parole avec un visage virtuel permettant au locuteur de compléter sa compréhension à l'aide de la lecture labiale. Il existe un visage féminin ainsi qu'un visage masculin, visibles ci-dessous.



Figure 9 : Exemples d'avatars parlant (Site internet dynamips.com)

Le projet Dynalips propose sa technologie dans le but d'améliorer la qualité de l'animation dans les domaines des jeux, du divertissement, du cinéma ainsi que dans la réalisation d'aide à la communication pour malentendants ou personne en cours d'acquisition d'une langue par exemple. Il nous semble donc intéressant de mesurer le niveau d'intelligibilité de cet avatar parlant sur une population cible : des enfants sourds en cours d'acquisition du langage, c'est-à-dire entre 6 et 10 ans.

PARTIE 2 :
PARTIE METHODOLOGIQUE

CHAPITRE 1

PROBLEMATIQUE & HYPOTHESES

1. Problématique

Concernant la nouvelle version de l'avatar parlant issu du projet Dynalips, plus précise quant à l'articulation reproduite, à ma connaissance aucune étude n'a encore testé son utilisation auprès de personnes sourdes, adultes ou enfants. Je propose donc d'étudier l'intérêt de l'usage de cet avatar parlant en situation de compréhension orale auprès d'enfants sourds âgés de 6 et 9 ans, présentant une surdité sévère à profonde et scolarisés en classe primaire.

2. Hypothèses

M'appuyant sur la revue de littérature présentée dans la première partie de mon mémoire, notamment les travaux de Dumont & Calbour, 2002; Massaro & Cohen, 1995; Pandzic et al., 1999; Reisberg et al., 1987 concernant l'aide apportée par les aides visuelles à la compréhension orale et ceux de Piquard-Kipffer (2014) concernant l'acceptation d'une tête parlante par des enfants sourds, les hypothèses que je formule sont les suivantes :

Hypothèse 1 : L'ajout d'un canal visuel devrait permettre à l'enfant de mieux prendre en compte les indices morphosyntaxiques. Ainsi, ses stratégies de compréhension devraient être plus précises et l'accès à la signification des messages devrait être plus rapide. Nous nous attendons donc à observer un nombre de réponses incorrectes moins important ainsi que des temps de réponses plus courts dans les modalités « avatar » et « humain » que dans la modalité « audio ».

Hypothèse 1A : Les items bien reconnus par l'enfant (réponses correctes) en modalité « humain » seront mieux reconnus en modalité « avatar » qu'en modalité « audio ». Les scores des réponses correctes seront plus importants.

Hypothèse 1B : Pour les items bien reconnus par l'enfant (réponses correctes) en modalité « humain », mais incorrectement reconnus en modalité « audio » (réponses incorrectes), nous nous attendons à observer des scores de réponses correctes plus importants en modalité « avatar » qu'en modalité « audio ».

Hypothèse 2 : La modalité « avatar » devrait être plus aidante pour certains indices morphosyntaxiques que pour d'autres, en termes de scores de réponses correctes, comparée à la modalité « audio ». Nous nous attendons donc à observer une différence de scores en modalités « avatar » versus « audio » plus importante pour certains indices morphosyntaxiques que pour d'autres.

Hypothèse 3 : L'avatar serait bien accepté des enfants. Ainsi, nous nous attendons à ce que les enfants interrogés aient une préférence pour l'avatar, qui a l'aspect d'un personnage de dessin animé, par rapport aux deux autres modalités (locuteur humain et enregistrement vocal).

CHAPITRE 2

METHODOLOGIE

1 Caractéristiques de la population étudiée

La cohorte de mon étude est composée d'enfants âgés de 6 à 10 ans présentant une surdité sévère à profonde, c'est-à-dire une perte tonale moyenne comprise entre 71 et 119 décibels.

Dans le cadre d'une surdité, la pose d'implant cochléaire est préconisée avant l'âge de 6 ans. En effet, les recherches indiquent que ce sont durant les six premières années de vie que les conditions biologiques pour le développement du langage sont optimales (Eric, 1967). De plus c'est à partir de 7 ans que la compréhension métadiscursive se développe chez l'enfant (Coquet, 2006). Nous considérons donc qu'à partir de 6 ans, l'enfant, présentant une surdité, est capable de discriminer et comprendre des indices morphosyntaxiques, indices choisis pour l'évaluation de la compréhension des enfants. De plus, il y est confronté à l'école ou dans son environnement personnel.

2 Critères de sélection

Au sein de cette cohorte d'enfants, des critères de sélection ont été définis.

Critères d'inclusion :

Les enfants que nous rencontrerons pratiqueront un mode de communication oral ou bilingue comprenant le mode de communication oral. Leur maîtrise du langage oral français doit leur permettre de comprendre des phrases simples ou complexes contenant différents indices morphosyntaxiques.

Ces enfants peuvent bénéficier d'un gain prothétique apporté par un appareillage auditif de type contour d'oreille ou bandeau ou encore d'un implant cochléaire unilatéral ou bilatéral.

Critères de non-inclusion :

Ne pourront participer à cette étude les enfants présentant des troubles de la vision non corrigés. En revanche s'ils portent des lunettes, des lentilles, ils peuvent être inclus dans cette étude.

Critères d'exclusion :

Ne pourront participer à cette étude les enfants présentant une cophose bilatérale, c'est-à-dire ne disposant d'aucun reste auditif.

Les réponses données par les enfants ne seront prises en compte que si la passation a lieu dans son intégralité, correspondant aux 60 items présentés.

3 Présentation du matériel expérimental

L'étude nécessite l'utilisation d'un ordinateur permettant de lancer le programme contenant les messages vidéo et audio, d'une connexion à internet et d'une enceinte.

Un matériel « papier » est à disposition devant l'enfant, se composant d'un porte-documents contenant 60 planches, chacune d'elles étant composée de quatre images dont une correspond à une des 60 phrases entendues.

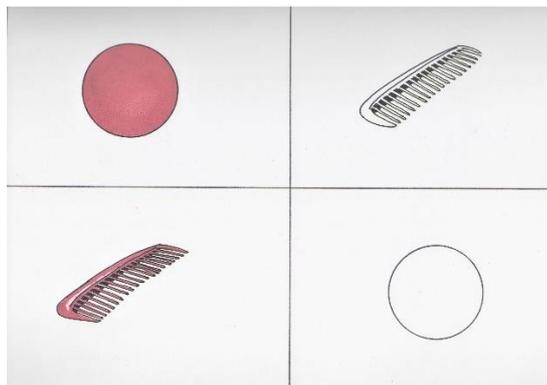


Figure 10 : Exemple de planche pour la phrase « Le peigne n'est pas rouge » de l'E.CO.S.SE (Lecocq, 1998)

3.1 Choix des modalités

J'ai enregistré et filmé dans le laboratoire Loria - Inria Grand Est un adolescent prononçant les 20 phrases du corpus choisi (voir Annexe). J'ai isolé la bande sonore pour pouvoir l'associer à différentes modalités visuelles. Ainsi cette étude compare trois modalités. Dans chaque modalité la bande sonore est identique, ce sont les modalités visuelles ajoutés à la bande sonore qui varient.

- **Modalité « humain »** : la bande sonore est associée à un locuteur humain original filmé
- **Modalité « avatar »** : la bande sonore est associée à un avatar parlant
- **Modalité « audio »** : la bande sonore est accompagnée d'un écran noir.



Figure 11 : Visuels des trois modalités (« audio », « humain », « avatar »)

L'objectif de cette étude étant est de tester l'intelligibilité visuelle d'un avatar parlant, nous avons décidé de comparer cet avatar parlant à deux autres modalités visuelles : un locuteur humain et un fond noir. La modalité « humain » représente la situation la plus écologique possible pour une compréhension langagière où les indices visuels sont les plus présents et naturels. A l'inverse la modalité « sans visuel » représente la situation la plus difficile pour une compréhension langagière : c'est-à-dire lorsque le locuteur ne peut s'appuyer sur aucun indice visuel pour comprendre le message. Ainsi il s'agit d'observer où la modalité « avatar parlant » se situe par rapport à ces deux situations opposées pour évaluer l'intelligibilité de l'avatar parlant.

3.2 Choix des indices morphosyntaxiques

Nous avons choisi des éléments morphosyntaxiques comme critères d'évaluation puisque ces aspects sont liés à des compétences phonologiques, lexicales et interactionnelles que l'enfant s'approprie progressivement. La morphosyntaxe est un domaine qui émerge au niveau de l'expression orale dès l'âge de 2 ans (Kail & Fayol, 2015). Il s'est agi de sélectionner des indices morphosyntaxiques normalement maîtrisés par les enfants sourds de cet âge, qui sont nécessaires à la bonne compréhension du message oral. A partir des travaux décrits par Rondal (2018), notre choix s'est porté sur trois indices morphologiques et trois indices syntaxiques.

Les indices morphologiques sont les suivants :

- Des indices marquant le **genre** : **il/elle**. En effet l'emploi du pronom « il » apparaît vers 3 ans et la distinction de genre se fait également à environ 3 ans.
- Des indices marquant le **nombre** : **le/les**, qui sont acquis vers 4 ans. Les flexions de l'indicatif présent sont maîtrisées à partir de 3 ans.
- **Clitiques accusatifs** : **le/la**. A 6 ans, l'emploi des différents articles tels que les clitiques accusatifs est généralement maîtrisé. Un clitique est un pronom qui, dans une phrase, désigne le « sujet » du verbe à l'infinitif. Les articles marquant le genre « le / la » sont généralement acquis vers 3 ans. Les pronoms

Les indices syntaxiques sont les suivants :

- **Forme négative** : L'enfant commence à produire des phrases négatives à 4 ans. On considère son acquisition complète vers l'âge de 6 ans.
- **Propositions relatives** en qui et avec effacement de relative : L'enfant produit des subordinées relatives ou complétives à partir de 4 ans.
- **Adverbe de lieu** : « devant » et « derrière » qui émergent vers 3 ans.

3.3 Choix des items

Un corpus a été constitué, comprenant 20 items. Chaque item est présenté trois fois dans trois modalités. Les items ont été choisis à partir de planches d'images existantes de l'ECOSSE ou du test ELO.

Le corpus comprend :

- 10 phrases contenant un indice morphologique : 2 clitiques accusatifs, 4 traitant du genre et 4 traitant du nombre
- 10 phrases contenant un indice syntaxique : 4 formes négatives, 4 propositions relatives et 2 adverbes de lieu (voir figure 11).

<i>Indices morphosyntaxiques</i>	<i>Nombre</i>
Morphologie : 10	
Clitiques accusatifs	2
Genre	4
Nombre	4
Syntaxe : 10	
Forme négative	4
Propositions relatives	4
Relatives en qui	2
Effacement de relatives	2
Prépositions	2

Figure 12 : Tableau récapitulatif des différents indices morphosyntaxiques du corpus et de leur nombre

Les clitiques « le » et « la » sont évalués à partir des deux phrases suivantes « La fille le regarde » et « la fille la regarde ».

La distinction **du nombre** dans une phrase est testée à partir des quatre phrases suivantes : « l'ours dort », « les ours dorment », « le garçon joue avec le train » et « les garçons jouent avec le train ». Pour la deuxième phrase « les ours dorment », l'enfant peut s'appuyer sur deux

indices, l'article « les » ainsi que la flexion verbales [dɔʁm] différente du singulier [dɔʁ]. Pour la seconde phrase « les enfants jouent avec le train » l'enfant ne peut s'appuyer que sur l'article « les ».

La distinction **du genre** est testée à travers quatre items : « il est tombé », « elle est tombée », « il va partir travailler », « elle va partir travailler ». Dans ces quatre propositions, le seul indicateur du genre est le pronom personnel « il » ou « elle ».

La forme négative est testée à partir de deux phrases « le peigne n'est pas rouge » et « le chien ne boit pas ». Le corpus comprend également les formes positives de ces items « le peigne est rouge » et « le chien boit » pour s'assurer de la compréhension de la phrase en elle-même indépendamment des indices de négation.

Les adverbes de lieu « devant » et « derrière » sont évalués à partir des deux items suivants : « le crayon est derrière la boîte » et « le crayon est devant la boîte ».

Enfin les phrases contenant **des propositions relatives** présentes dans le corpus sont « le cercle qui est dans l'étoile est jaune », « l'étoile qui est dans le cercle est jaune », « le cercle dans l'étoile est jaune » et « l'étoile dans le cercle est jaune ». **Les propositions sans élément relatif** « qui » sont d'ordinaire compris et employés avant les propositions relatives en qui.

3.4 Création du programme informatique de passation

Avec l'aide de Monsieur OUNI, un programme a été créé permettant de recueillir les réponses de chaque sujet. Ainsi ce programme se compose de plusieurs pages :

- 1^{ère} page : Présentation de l'étude

Expérience de perception

Merci pour votre participation à cette étude...

Une personne va prononcer des phrases de longueurs variables. Nous allons vous présenter des vidéos et vous faire écouter ces phrases. **Il ne s'agit pas de vous évaluer, c'est le système qu'on évalue.**

Le déroulement de l'expérience se fait comme suit :

1. Réglage du niveau sonore de votre ordinateur
2. Saisie d'informations (nom, prénom, âge, etc...)

(ces informations resteront confidentielles et ne seront jamais divulguées. Ces informations seront anonymisées lors de l'analyse des données. Vous avez le droit d'accès, de rectification et de suppression de ces informations à tout moment.)

3. Lancement de l'évaluation.

Veuillez rester attentif pour l'ensemble des présentations durant cette expérience.

Etape 1: les réglages

Figure 13 : 1^{ère} page de l'expérimentation

- 2^{ème} page : Réglage du niveau sonore



Etape de réglage du niveau sonore

Voici la procédure à suivre:

- Veillez maximiser la taille du navigateur durant l'évaluation.
- Réglez le volume de votre ordinateur de façon à ce que l'écoute soit la plus confortable possible.




Vous ne pouvez pas modifier le volume sur le lecteur vidéo, vous devez régler le volume de votre PC

⚠ Attention, il vous est demandé de ne pas modifier le volume une fois le test démarré

Etape 2: Saisie d'information

Figure 14 : 2^{ème} page de l'expérimentation

- 3^{ème} page : Prise de renseignements personnels. Les résultats sont ensuite anonymisés. Le nom permet à ce moment-là d'associer le sujet à un ordre de passage en raison des difficultés techniques et des fichiers parfois incomplets.

Avant de commencer l'évaluation, veuillez renseigner les champs suivants

Nom* :	<input style="width: 60%;" type="text"/>
Prénom* :	<input style="width: 60%;" type="text"/>
Genre :	féminin <input checked="" type="radio"/> masculin <input type="radio"/>
Âge :	<input style="width: 20px;" type="text" value="5"/>
Classe :	CP <input checked="" type="radio"/> CE1 <input type="radio"/> CE2 <input type="radio"/>
Ecole* :	<input style="width: 60%;" type="text"/>
Surdité :	<input style="width: 60%;" type="text" value="légère"/>
Audition :	normo-entendant <input checked="" type="radio"/> malentendant <input type="radio"/>
Si surdité, type d'appareillage : (appareil auditif / implant cochléaire)* :	
	<input style="width: 60%;" type="text"/>

Etape 3: Commencer l'évaluation

Figure 15 : 3^{ème} page de l'expérimentation

- 4^{ème} page : une page avant le départ permettant de s'assurer que le sujet est prêt et à bien compris les consignes.



Figure 16 : 4^{ème} page de l'expérimentation

- 5^{ème} page : Première question sur les 60 du corpus.



Figure 17 : 5^{ème} page de l'expérimentation

4 Recrutement

Nous avons recruté notre population en prenant contact avec des écoles spécialisées ainsi que des orthophonistes de Lorraine. Ce sont des instituteurs spécialisés ou des orthophonistes qui se sont adressés à des parents d'enfants dont le profil correspondait aux critères définis pour cette étude (cf. chapitre 2, page 39) parmi leur classe ou patientèle. L'accord des parents (ou tuteurs légaux) a été demandé à partir d'un formulaire de consentement parental que nous leur avons fourni.

5 Déroulement de la passation

J'ai rencontré chaque enfant individuellement soit dans son lieu de scolarisation soit dans le cabinet de son orthophoniste, dans une pièce calme. L'entretien a duré une demi-heure. La phase d'investigation comprenait la réalisation de l'épreuve par les enfants sous ma supervision ainsi que la collecte des données.

J'ai accueilli chaque enfant face à l'écran d'ordinateur et me suis assis à côté de lui. J'ai réglé le son grâce à une enceinte que j'ai placée devant l'enfant et me suis assurée qu'il entendait correctement. J'ai ensuite expliqué à chaque enfant la consigne suivante « *Tu vas entendre des petites phrases assez courtes, ensuite je vais te montrer 4 images et tu devras me montrer l'image qui correspond à ce que tu as entendu. Sur l'écran tu vas voir qui parle mais on ne voit pas toujours la même personne, parfois c'est un garçon, parfois c'est un personnage de dessin animé, un avatar, et parfois on entend quelqu'un parler mais on ne le voit pas. Tu peux t'aider de l'image pour bien comprendre la phrase. Tu ne peux entendre la phrase qu'une seule fois, prends ton temps, si tu ne sais pas ce n'est pas grave, dis-le-moi et on passe à la suivante.* ». Je me suis assurée que l'enfant avait bien compris la consigne avant de lancer le programme.

A la première page du programme j'ai demandé à l'enfant s'il avait des questions et s'il était prêt à commencer. A la deuxième page, j'ai testé à nouveau le son et me suis assurée que l'enfant entendait la voix, voyait l'image et comprenait la consigne. J'ai demandé quelques informations à l'enfant pour remplir certains champs de la page numéro 3, à savoir son prénom, son âge et sa classe. J'ai noté également le degré de sévérité de sa surdité transmis par son orthophoniste ou son instituteur ainsi que son type d'appareillage auditif.

Durant les 60 items, l'enfant a entendu et vu chaque énoncé, puis je lui ai montré une planche de quatre images, je l'ai laissé réfléchir puis lorsqu'il m'a pointé l'image qu'il a choisi j'ai cliqué sur le numéro de l'image correspondant sur le programme avant de valider pour passer à l'item suivant. A la fin de l'épreuve j'ai demandé à chaque enfant « *As-tu préféré répondre aux questions lorsque tu voyais le vrai garçon, lorsque tu voyais l'avatar, tu sais le personnage de dessin animé ou juste lorsque tu entendais la voix ?* ». J'ai noté sa réponse dans un cadre avant de valider et terminer l'épreuve.

6 Collecte et analyse statistique prévue des données

Les réponses aux 60 items (20 phrases oralisées dans 3 modalités) de l'épreuve d'appariement image / énoncé (voir annexe 1) ainsi que le temps de latence pour chaque item

ont été collectés via un programme créé spécialement dans le cadre de mon mémoire. Pour chaque sujet, j'ai reçu un fichier .pad totalisant anonymement les réponses fournies par le sujet ainsi que le temps mis par l'individu pour répondre à chaque item. La collecte des données a été organisée de façon à pouvoir répondre à mes hypothèses. Ainsi j'ai réalisé des tableaux pouvant répertorier au mieux les réponses des enfants afin de pouvoir réaliser une analyse statistique descriptive ainsi qu'une comparaison des données à l'aide des tests de Student, de comparaison des proportions et de Chi².

6.1 Nombre d'erreurs et du temps de latence par modalité : **hypothèses 1, 1A et 1B**

Hypothèse 1 :

A partir des données récoltées, j'ai organisé les variables pour définir le nombre d'erreurs par modalité (H, A, O). Une fois toutes les données recueillies, j'ai calculé l'effectif total des erreurs, la moyenne par modalité, l'écart-type et la médiane pour pouvoir comparer chaque modalité en termes d'erreurs.

Concernant le temps, j'ai calculé également le temps total par modalité, ainsi que la moyenne de temps de latence par modalité et j'ai répertorié toutes ces données dans un tableau de recueil des données anonymisées (voir tableau ci-dessous).

Sujet (numéro attribué par ordre de passation)	Erreurs par modalité			Nombre d'erreurs total	Temps par modalité min : sec			Temps total
	Humain (H)	Avatar (A)	Audio (O)		Humain (H)	Avatar (A)	Audio (O)	
1								
EFFECTIF TOTAL								
MOYENNE								
ECART-TYPE								
MEDIANE								

Figure 18 : Tableau de recueil des scores et temps obtenus par modalité

Afin de m'assurer de la significativité des données recueillies, j'ai comparé les moyennes de nombre d'erreurs ainsi que les moyennes de temps de réponse obtenues dans chaque modalité à l'aide du test de Student.

Hypothèses 1A :

J'ai isolé tous les items où la réponse en modalité « humain » est correcte afin d'observer et de comparer les résultats obtenus dans les modalités « avatar » et « audio ». J'ai ensuite calculé la moyenne du nombre de réponses correctes pour les modalités « avatar » et « audio » par rapport au nombre total d'items.

Pour finir, j'ai comparé ces deux moyennes à l'aide du test de Student afin d'évaluer si la différence était statistiquement significative.

Hypothèse 1B :

J'ai établi cinq catégories pour faciliter la lecture et l'analyse des résultats ultérieurs, en répertoriant :

- le nombre de réponses incorrectes dans les trois modalités (catégorie A),
- le nombre de réponses correctes dans les trois modalités (catégorie B)

Pour répondre à mon **hypothèse 1B**, j'ai également répertorié :

- le nombre de réponses correctes en modalité « humain » et incorrectes en modalité « audio » (catégorie C). Parmi cette dernière catégorie j'ai calculé le nombre de réponses où la modalité « avatar » est correcte (catégorie E) et le nombre de celles où la modalité « avatar » est incorrecte (catégorie F). Enfin les réponses ne correspondant à aucune des catégories précédentes ont été répertoriées dans la catégorie D « Autres ».

Pour chacune de ces six catégories, j'ai calculé l'effectif total, la moyenne et le pourcentage parmi la totalité des 60 phrases entendues par les enfants (voir tableau ci-dessous). Ces calculs me permettront d'identifier les items pour lesquels la modalité « avatar » aura – selon mes hypothèses - un impact positif sur la compréhension (catégorie E).

Sujet (numéro attribué par ordre de passation)	Nombre de phrases incorrectes dans 3 modalités A	% p(A)	Nombre de phrases correctes dans 3 modalités B	% p(B)	Nombre de phrases correctes en modalité humain et incorrectes en modalité audio C	% p(C)	Autres D	% p(D)	Parmi C, nombre de phrases où modalité Avatar correcte E	% p(E)	% p(E)/C	Parmi C, nombre de phrases où modalité Avatar incorrecte F	% p(F)	% p(F)/C
EFFECTIF TOTAL														
MOYENNE														
POURCENTAGE														

Figure 19 : Tableau de recueil des scores obtenus par réponses

*p(A) : pourcentage de la catégorie A

Afin de m'assurer de la significativité des données recueillies, j'ai comparé les moyennes de chaque catégorie pour chaque modalité à l'aide du test de Student et les proportions de chaque catégorie pour chaque modalité à l'aide du test de comparaison de proportions.

6.2 Indices morphosyntaxiques : hypothèse 2

Pour répondre à ma deuxième hypothèse, j'ai relevé le nombre d'items corrects en modalités « avatar » et le nombre d'items corrects en modalité « audio » pour chaque indice morphosyntaxique. J'ai calculé la différence pour chaque indice entre ces deux modalités (voir tableau ci-dessous).

Sujet (numéro attribué par ordre de passation)	Nombre de phrases correctes en fonction de l'indice morphosyntaxique													
	indices morphologiques						indices syntaxiques						TOTAL	
	Clitique accusatif		Nombre		Genre		Forme négative		Proposition relative		Préposition			
	A	O	A	O	A	O	A	O	A	O	A	O	A	O
EFFECTIF TOTAL														
DIFFERENCE (A - O)														

Figure 20 : Tableau de recueil des scores obtenus par indice morphosyntaxique

*A : modalité « avatar »

*O : modalité « audio »

Une fois la différence de chaque indice calculée, j'ai utilisé le test Chi² afin de m'assurer de la significativité des données recueillies.

6.3 Préférence de la modalité : hypothèse 3

Pour répondre à ma troisième hypothèse, j'ai organisé dans un tableau (voir tableau ci-dessous) les réponses obtenues à la fin de chaque passation à la question « *As-tu préféré répondre aux questions lorsque tu voyais le vrai garçon, lorsque tu voyais l'avatar, tu sais le personnage de dessin animé ou juste lorsque tu entendais la voix ?* ». La notion « avatar » est préalablement développé avec l'enfant et je me suis assurée de sa bonne compréhension. Lorsque l'enfant notifiait avoir préféré plusieurs modalités, j'ai attribué 1 à chaque modalité concerné. Si l'enfant n'a préféré aucune modalité j'ai attribué 1 à la case « Aucune préférence ».

Sujet (numéro attribué par ordre de passation)	Situation préférée				TOTAL
	"humain"	"avatar"	"audio"	"pas de préférence"	
Effectif total					
Pourcentage					

Figure 21 : Tableau de recueil des réponses obtenues pour la préférence des enfants

Suite à cela, j'ai également utilisé le test de comparaison des proportions qui me permet d'accepter ou de rejeter ma troisième hypothèse, en définissant si la différence entre les pourcentages calculés est statistiquement significative ou non.

PARTIE 3 :

RESULTATS

CHAPITRE 1

COHORTE

La cohorte initialement prévue était composée de deux groupes : un groupe d'enfants présentant une surdité ainsi qu'un groupe contrôle d'enfants normo-entendants. En raison de la crise sanitaire du covid-19 ayant entraîné la fermeture des écoles le 13 mars 2020 et le confinement le 17 mars 2020, les expérimentations ont été suspendues et n'ont pu reprendre à temps. La cohorte finale n'est donc composée que des enfants rencontrés avant cette date, à savoir 10 enfants présentant une surdité.

1 Taille de l'échantillon

L'échantillon est composé de 10 sujets, tous résidant dans la région Grand-Est. Parmi eux, 2 ont abandonné en cours de passation, seulement 8 ont donc terminé la passation. Les données recueillies se basent sur ces 8 enfants.

2 Caractéristiques de l'échantillon

2.1 Age

<i>Age</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>Total</i>
Effectif	1	3	2	1	3	10

Figure 22 : Tableau de recueil des âges de l'échantillon

Parmi les 10 enfants rencontrés, le plus jeune a 6 ans, trois enfants sont âgés de 7 ans, deux enfants sont âgés de 8 ans, un enfant est âgé de 9 ans et les trois plus âgés ont 10 ans. La moyenne d'âge du groupe est de 8 ans.

2.2 Degré de surdité sans prothèse auditive

<i>Degré de surdité</i>	<i>Sévère</i>	<i>Profonde</i>	<i>Total</i>
Effectif	1	9	10

Figure 23 : Tableau de recueil du degré de surdité de l'échantillon

Parmi les enfants rencontrés, la majorité est atteinte d'une surdité profonde, ce qui correspond à une perte tonale moyenne comprise entre 71 et 90 décibels. Seulement un enfant

de l'étude est atteint de surdité sévère, c'est-à-dire une perte tonale moyenne moins importante, comprise entre 51 et 70 décibels.

2.3 Type d'appareillage auditif

<i>Type d'appareillage</i>	<i>Prothèses amplificatrices</i>		<i>Implants cochléaires</i>	<i>Non-appareillé</i>	<i>Total</i>
	<i>Par voie aérienne</i>	<i>Par voie osseuse</i>			
Effectif	7	1	2	0	10

Figure 24 : Tableau de recueil des types d'appareillage auditif de l'échantillon

Les enfants rencontrés sont tous appareillés et ce depuis plusieurs années. Ils ont bénéficié d'un suivi auditif et/ou orthophonique ainsi que d'une éducative auditive leur permettant d'intégrer et comprendre les sons entendus ainsi que les indices visuels de la parole. La majorité possède des prothèses amplificatrices : sept enfants portent des contours d'oreilles, l'amplification se fait alors par voie aérienne. Un enfant porte un bandeau, dans ce cas, l'amplification se fait par voie osseuse. Les deux autres enfants portent des implants cochléaires.

CHAPITRE 2

RESULTATS

1 Correction des réponses et du temps de latence pour les trois modalités (« humain », « avatar » et « audio »)

La moyenne des **réponses incorrectes** pour la totalité des phrases (60 phrases oralisées) est de 26 et le **temps moyen de passation** de 14 minutes et 4 secondes.

	Nombre de réponses incorrectes (/60)	Temps de passation
Moyenne	26	14 min et 04 sec

Figure 25 : Tableau de recueil des moyennes du nombre de réponses incorrectes et du temps de passation pour les 60 items

2 Correction des réponses et du temps de latence pour chaque modalité (**Hypothèse 1**)

Concernant le **nombre de réponses incorrectes** (ou erreurs) pour les 20 phrases oralisées, la moyenne des erreurs en modalité « audio » (9,6/20) est supérieure à celle de la modalité « avatar » (8,9/20), elle-même supérieure à la modalité « humain » (7,9/20).

Modalité	H	A	O	Total
Nombre d'erreurs /160 items*	63	71	77	211
Effectif	8	8	8	24
Moyenne	7,9	8,9	9,6	8,8
Ecart-type	3,5	4,1	3,4	

Figure 26 : Nombre d'erreurs : Récapitulatif des effectifs, moyenne et écart-type pour chaque modalité

*160 : correspond au nombre de 20 phrases du corpus multiplié par les 8 sujets.

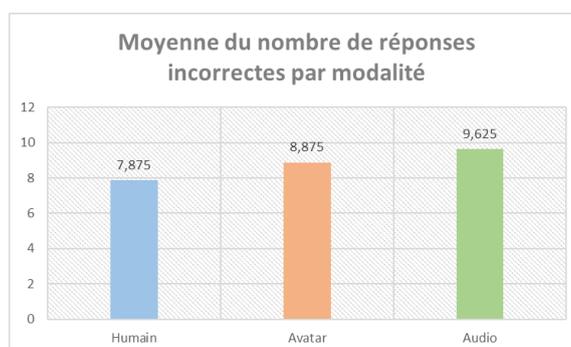


Figure 27 : Graphique : moyenne du nombre de réponses incorrectes par modalité

Concernant **le temps de latence**, nous relevons, dans l'ordre décroissant, une moyenne du temps de réponse plus importante pour la modalité « humain » : 292 secondes (presque 5 minutes), suivie de près par celle de la modalité « audio » : 290 secondes. Et enfin, la moyenne la moins élevée est celle de la modalité « avatar » : 267 secondes (4 minutes et 27 secondes).

Modalité	Humain	Avatar	Audio	Total
Temps (en sec)	2336	2139	2316	6791
Effectif	8	8	8	24
Moyenne	292	267,4	289,5	283
Ecart-type	44,4	35,2	45,2	

Figure 28 : Temps de latence : Récapitulatif des effectifs, moyenne et écart-type pour chaque modalité

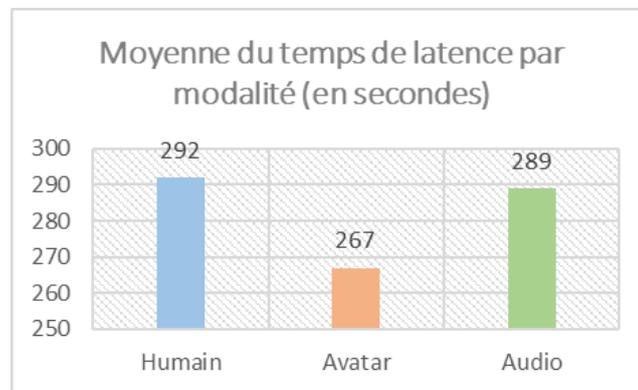


Figure 29 : Graphique : moyenne du temps de latence par modalité

Pour examiner **la significativité statistique** des résultats, nous avons réalisé le test de Student (calcul de la valeur p-value) à partir du site internet www.biostatgv.sentiweb.fr.

La comparaison des **moyennes du nombre d'erreurs** entre les modalités « avatar » et « audio » montre que la valeur p-value est d'environ 0,71. ($p > 0,05$).

La comparaison des moyennes du nombre d'erreurs entre les modalités « avatar » et « humain », montre que la valeur p-value est d'environ 0,63 ($p > 0,05$).

La valeur p-value pour la comparaison des moyennes du nombre d'erreurs entre la modalité « avatar » et la moyenne du nombre d'erreurs total est d'environ 0,92 ($p > 0,05$).

La comparaison des **moyennes du temps de latence** entre les modalités « avatar » et « audio » montre que la valeur p-value est d'environ 0,33 ($p > 0,05$).

A la comparaison des moyennes du temps de latence entre les modalités « avatar » et « humain », la valeur p-value est d'environ 0,27 ($p > 0,05$).

Enfin la comparaison des moyennes du temps de latence entre la modalité « avatar » et la moyenne du temps de latence total montre que la valeur p-value est d'environ 0,42. ($p > 0,05$).

En résumé aucune des différences observées pour toutes ces moyennes n'est statistiquement significative.

3 Correction des réponses : phrases reconnues en modalité « humain » (Hypothèse 1A)

Pour répondre à mon hypothèse 1A, je suis partie du préambule que la situation de compréhension est optimale lorsque nous avons accès au visage du locuteur, ce qui correspond à la modalité « humain » de notre étude (catégorie H). Nous cherchons à démontrer l'intelligibilité de l'avatar, c'est-à-dire dans ce cas une meilleure compréhension avec la modalité « avatar » qu'avec la modalité « audio ». Parmi les données recueillies, après avoir isolé les phrases où le sujet a répondu correctement en modalité « humain » (catégorie H), j'ai observé les résultats en modalités « avatar » et « audio ».

Parmi les phrases où la réponse était correcte en modalité « humain », la moyenne du nombre de réponses correctes en modalité « avatar » est supérieure à celle en modalité « audio » : $m_{\text{avatar}} = 8,6 > m_{\text{audio}} = 7,9$.

Nombre de phrases /160*	Réponse en modalité « humain » correcte : H	Parmi H, réponse en modalité « avatar » correcte	Parmi H, réponse en modalité « audio » correcte
Phrases concernées	96	69	63
Moyenne**	12	8,6	7,9

Figure 30 : Tableau récapitulatif des effectifs et des moyennes de chaque type de réponses

*160 : correspond au nombre de 20 phrases du corpus multiplié par les 8 sujets.

** : la moyenne est calculée par rapport à l'effectif de 8 sujets

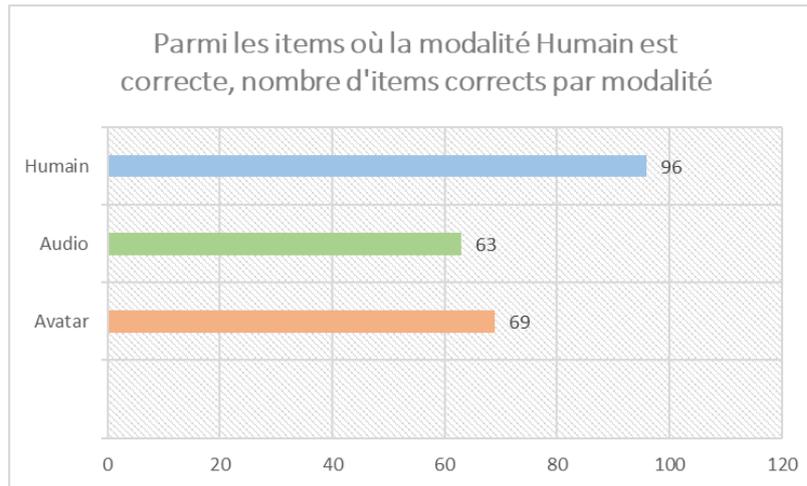


Figure 31 : Graphique récapitulatif du nombre de réponses correctes par modalité

La différence obtenue entre m« avatar » et m« audio » est évaluée à l'aide du test de Student. La valeur p-value obtenue est environ égale à 0,62 ($p > 0,05$). Les résultats obtenus ne sont pas statistiquement significatifs.

4 Correction des réponses : phrases reconnues en modalité « humain » mais pas en modalité « audio » (**Hypothèse 1B**)

La base du préambule utilisé pour répondre à mon hypothèse 1B est la même que pour mon hypothèse 1A, à savoir, la situation de compréhension est optimale lorsque nous avons accès au visage du locuteur, ce qui correspond à la modalité « humain » de cette étude. De plus j'ai estimé que la situation inverse, c'est-à-dire la situation de compréhension la plus difficile correspondait à la modalité « audio » où le locuteur ne bénéficie d'aucune aide visuelle. Ainsi parmi les données recueillies, les phrases où le sujet avait répondu correctement en modalité « humain » et incorrectement en modalité « audio » sont regroupées dans la catégorie C. Cette situation représente 21% des données recueillies. La catégorie E représente le nombre de réponses correctes en modalité « avatar » parmi les phrases de la catégorie C. La catégorie F représente le nombre de réponses incorrectes en modalité « avatar » parmi les phrases de la catégorie C.

La catégorie A représente les phrases dont les réponses étaient correctes dans les trois modalités. La catégorie B représente les phrases dont les réponses étaient incorrectes dans les trois modalités. Enfin la catégorie D représente les phrases dont les réponses dans les trois modalités étaient représentées par une autre configuration que pour les autres catégories et qui n'étaient pas interprétables.

Nombre de phrases /160*	Modalités H, A et O correctes (A)	Modalités H, A et O incorrectes (B)	Modalité H correcte & Modalité O incorrecte (C)		Autres (D)	Total
			Modalité A correcte (E)	Modalité A incorrecte (F)		
Phrases concernées	52	32	34		42	160
			17	17		
Pourcentage	33%	20%	21%		26%	100%
			10,5%	10,5%		

Figure 32 : Tableau récapitulatif des effectifs et des proportions de chaque type de réponses

*160 : correspond au nombre de 20 phrases du corpus multiplié par les 8 sujets.

A partir des 160 phrases observées durant les expérimentation (20 phrases du corpus pour 8 sujets), j'ai relevé 52 phrases pour la catégorie A soit 33% de l'ensemble des phrases observées. La catégorie B comprend 32 phrases soit 20% de l'ensemble des phrases observées.

Afin de répondre à mon hypothèse 1B, j'utilise les résultats obtenus en catégorie C, à savoir 34 phrases concernées, soit 21% de l'ensemble des phrases observées. Parmi cette catégorie C, 17 phrases concernent la catégorie E et 17 phrases concernent la catégorie F soit 10,5% de l'ensemble des phrases observées pour chacune.

Enfin 42 phrases, soit 26% des phrases observées, concernent la catégorie D.

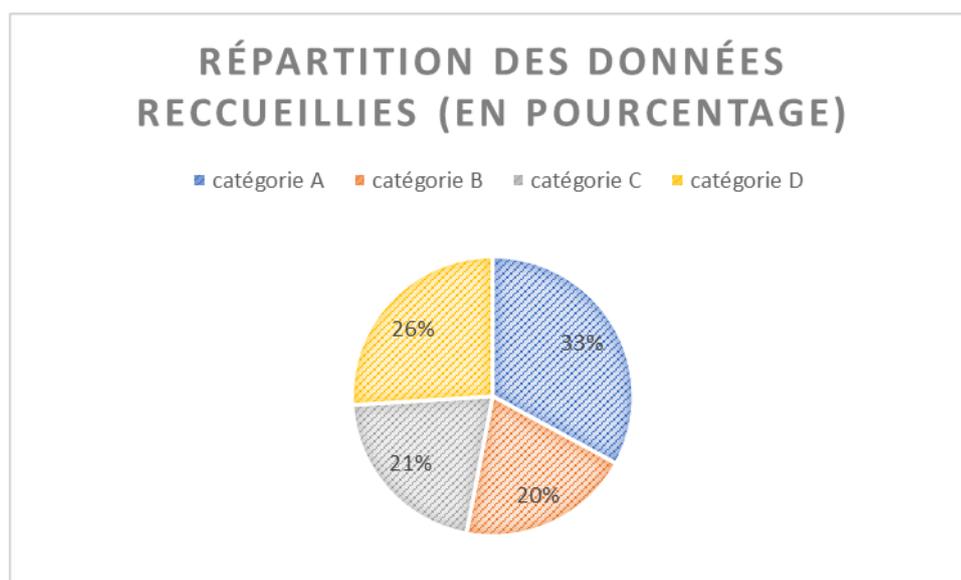


Figure 33 : Graphique en secteur de répartition des données recueillies

J'ai utilisé le test de comparaison des proportions afin d'évaluer la significativité statistiques des résultats obtenus. La valeur de l'écart-réduit entre les catégories C et E est égale à 2,60 au seuil de 5%, soit une valeur supérieure à la valeur 1,96. On peut donc rejeter l'hypothèse que les pourcentages sont égaux. On peut donc affirmer que la **différence entre la catégorie C et la catégorie E est statistiquement significative.**

5. Correction des réponses : en fonction des indices morphosyntaxiques (Hypothèse 2)

Sujet (numéro attribué par ordre de passation)	Nombre de phrases correctes en fonction de l'indice morphosyntaxique													
	indices morphologiques						indices syntaxiques						TOTAL	
	Clitique accusatif		Nombre		Genre		Forme négative		Proposition relative		Préposition			
	A	O	A	O	A	O	A	O	A	O	A	O	A	O
Effectif total /160*	4	8	20	20	20	19	27	22	8	5	10	9	89	83
Différence (A - O)	-4		0		1		5		3		1		6	

Figure 34 : Tableau récapitulatif des effectifs et des différences de réponses correctes chaque indice

*160 : correspond au nombre de 20 phrases du corpus multiplié par les 8 sujets.

En fonction de chaque indice morphosyntaxique, j'ai relevé le nombre d'items corrects en modalités « avatar » et « audio ». J'ai calculé la différence pour chaque indice entre ces deux modalités. Les différences de scores entre modalité « avatar » et modalité « audio » pour chaque indice morphosyntaxique sont relevées par ordre décroissant :

Indice morphosyntaxique	Différence (modalité « avatar » - modalité « audio »)
Forme négative	+5
Propositions relatives	+3
Genres	+1
Prépositions	+1
Nombres	0
Clitiques accusatifs	-4

Figure 35 : Tableau récapitulatif des différences entre les modalités « avatar » et « audio » par indice morphosyntaxique

La différence de scores (modalité « avatar » - modalité « audio ») est plus importante pour la forme négative, puis pour les propositions relatives et enfin pour les indices de genre et les prépositions. Aucune différence n'est notée pour les indices de nombre. Enfin pour les clitiqes accusatifs, les résultats corrects sont plus élevés en modalité « audio ».

L'utilisation du test Chi² permet de comparer la différence de scores pour un nombre de groupes supérieur à deux, afin de savoir si mes données recueillies sont statistiquement significatives. La valeur p (p-value) obtenue est 0.79. Cette valeur est supérieure à 0,05, elle ne permet donc pas d'assurer que la différence de résultats entre les catégories morphosyntaxiques est statistiquement significative.

6. Correction des réponses : de la modalité préférée des sujets (**Hypothèse 3**)

Modalité préférée du sujet	Avatar	Humain	Audio	Aucune préférence	Total
Effectif	5	4	3	1	13
Proportion	38%	31%	23%	8%	100%

Figure 36 : Tableau récapitulatif des effectifs et des proportions des modalités préférées par les sujets

A la fin de chaque passation, chaque enfant a pu choisir sa ou ses modalité(s) préférée(s). Ainsi, sur un effectif total de 13 réponses, 5 enfants ont préféré la modalité « avatar », 4 enfants ont préféré la modalité « humain », 3 enfants ont préféré la modalité « audio » et un enfant n'a eu aucune préférence pour une ou plusieurs modalités.

La proportion de préférence pour la modalité « avatar » est donc supérieure avec 38%, suivie par la préférence de la modalité « humain » 31%, la préférence pour la voix seule 23% et enfin aucune préférence pour 8% des sujets.

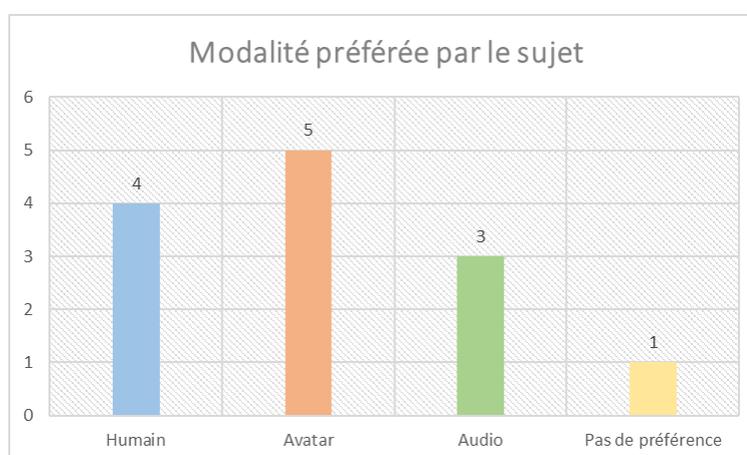


Figure 37 : Graphique : Modalité préférée par les sujets

La fiabilité des données recueillies a été évaluée à l'aide du test de comparaison de proportion. La valeur de l'écart-réduit obtenue entre les proportions $p(A)=0,38$ et $p(H)=0,31$ est égale à 0,41 au seuil de 5%.

La valeur de l'écart-réduit entre les proportions $p(A)=0,38$ et $p(O)=0,23$ est égale à 0,85 au seuil de 5%.

La valeur de l'écart-réduit entre la proportion $p(A)=0,38$ et $p(\ll \text{aucune préférence} \gg)=0,08$ est égale à 1,86 au seuil de 5%.

Toutes les valeurs de l'écart-réduit obtenues sont inférieures à 1,96. On ne peut donc pas rejeter les hypothèses que les pourcentages sont égaux, ce qui signifie que ces différences ne sont pas statistiquement significatives.

DISCUSSION

1 Analyse des résultats

1.1 Première hypothèse : réponses et du temps de latence pour chaque modalité

Concernant le nombre de réponses incorrectes relevées en fonction des modalités, nous observons que les moyennes **du nombre de réponses incorrectes en modalités « avatar » et « humain » sont inférieures à celle en modalité « audio »** :

$$m(H) = 7,9 < m(A) = 8,9 < m(O) = 9,6.$$

Les résultats obtenus en calculant la moyenne du temps de latence par modalité indique que la moyenne du temps de latence en modalité « avatar » est inférieure à celles des modalités « audio » et « avatar » ainsi qu'à la moyenne de temps de latence total :

$$m(A) = 267 < m(\text{Total}) = 283 < m(O) = 289 < m(H) = 292.$$

L'hypothèse 1 est donc validée. Néanmoins nous ne pouvons généraliser ces résultats à l'échelle de la population puisque les Tests de Student réalisés, n'ont pas montré de significativité statistique, ni entre les différences de nombre d'erreurs, ni de temps de latence entre la modalité « avatar » et les autres modalités.

Si l'hypothèse 1 est validée, certains résultats sont cependant à nuancer. En effet si la **moyenne du nombre de réponses incorrectes en modalité « avatar » est inférieure à celle de la modalité « audio », elle reste supérieure à celle de la modalité « humain »** : $m(A) = 8,9 > m(H) = 7,9$. La moyenne du **nombre de réponses incorrectes en modalité « avatar » est également légèrement supérieure à la moyenne du nombre de réponses incorrectes des 3 modalités** : $m(A) = 8,9 > m(\text{Total}) = 8,8$.

La compréhension avec la modalité « avatar » n'est donc pas aussi bonne qu'avec le locuteur original, ce qui indique que si l'avatar apporte des indices visuels permettant une meilleure compréhension, son intelligibilité n'est pas aussi performante qu'avec le locuteur humain.

De plus les résultats révèlent que si le temps de réponse est inférieur en modalité audio-visuelle « avatar » à celui obtenu en modalité « audio », ce n'est pas le cas pour la modalité audio-visuelle « humain » où le temps relevé est plus important que pour les deux autres modalités. Cette tendance nous interroge sur les raisons de la rapidité en modalité « avatar ».

L'intelligibilité et le caractère ludique et agréable de l'avatar peuvent être des raisons à cette rapidité. Ces questionnements pourront être précisés en poursuivant l'étude.

1.2 Hypothèse 1A : phrases reconnues en modalité « humain »

Concernant l'analyse des items dans les trois modalités, j'ai estimé que la situation de compréhension était optimale lorsque le locuteur a accès au visage du locuteur, ce qui correspond à la modalité « humain » de cette étude. Les réponses correctes en modalité « humain » ont été regroupées sous le nom de sous-échantillon H. A partir de ces données, j'ai comparé la moyenne du nombre de réponses correctes en modalités « avatar » et « audio ».

Un plus grand nombre de réponses correctes en modalité « avatar » est observé : 69 contre 63 pour la modalité « audio ». En termes de moyenne, la moyenne de réponses correctes en modalité « avatar » (8,6) est supérieure à la moyenne de réponses correctes en modalité Audio (7,9).

L'hypothèse 1A est validée. Lorsque la réponse est correcte en modalité « humain », cela signifie que la phrase est comprise et que l'indice morphosyntaxique est maîtrisé par l'enfant. Ainsi, en comparant les modalités « audio » et « avatar » pour ce sous-échantillon H, il est assuré qu'il n'y a pas de biais de non-maîtrise de l'indice morphosyntaxique. Ces résultats indiquent que les indices visuels apportés par l'avatar permettent à l'enfant de saisir le sens de l'item.

A l'aide du test de Student, j'ai évalué la significativité statistique de la différence de moyennes d'items corrects parmi le sous-échantillon H en modalités « avatar » et « audio » mais la valeur p-value obtenue ne permet pas d'affirmer que les résultats représentent la population. Les résultats observés sont encourageants mais il faudrait obtenir un nombre de réponses plus important pour avoir un échantillon suffisant pour représenter la population.

1.3 Hypothèse 1B : phrases reconnues en modalité « humain » mais pas en modalité « audio »

Pour aller plus loin dans l'analyse des items dans les trois modalités, j'ai estimé que la modalité « humain » de cette étude correspondrait à la situation de compréhension optimale, c'est-à-dire lorsque nous avons accès au visage du locuteur, et que la modalité « audio » correspondrait à la situation de compréhension la plus difficile puisque le locuteur ne bénéficie d'aucune aide visuelle. Ainsi parmi les données recueillies, j'ai isolé les cas où le sujet a

répondu correctement à une phrase avec la modalité « humain » et incorrectement à la même phrase avec la modalité « audio ». Cette situation représente la catégorie C et représente 21% des données recueillies. Afin de répondre à mon hypothèse 1B, j'ai utilisé les catégories E et F. 17 phrases concernent la catégorie E et 17 phrases concernent la catégorie F soit 10,5% de l'ensemble des phrases observées pour chacune. Ces deux catégories permettent de répondre aux questions suivantes : la présence de l'avatar a-t-elle joué un rôle dans la compréhension de l'item et donc les indices visuels apportés par l'avatar ont-ils été suffisants pour permettre la compréhension de l'item ? Ce fut le cas dans 50% des cas, c'est-à-dire pour 17 items.

L'hypothèse 1B est validée puisque nous observons un nombre positif de phrases dont la modalité « avatar » est correcte dans la catégorie E, qui représente le nombre de réponses correctes en modalité « avatar » et « humain » et incorrectes en modalité « audio ». En effet les scores obtenus dans cette catégorie, c'est-à-dire en modalité « avatar » témoigne que les indices visuels apportés par l'avatar parlant ont permis aux sujets de répondre correctement à la phrase alors que le message auditif seul ne suffisait pas. Ceci signifie que dans 50% des cas, lorsque l'enfant maîtrise l'item mais que le message sonore n'est pas suffisant, l'apport de l'avatar lui permet de comprendre le message.

1.4 Deuxième hypothèse : indices morphosyntaxiques

Les nombres de réponses correctes en modalité « avatar » et « audio » ont été calculés en fonction de l'indice morphosyntaxique concerné. Nous avons pu calculer la différence de scores entre les modalités « audio » et « avatar » pour chaque indice morphosyntaxique.

Des écarts différents en fonction des indices morphosyntaxiques ont été relevés. Concernant les indices syntaxiques, ces résultats indiquent que l'avatar a été plus pertinent pour les indices de négation puisqu'il a permis dans cinq cas de répondre correctement tandis que la réponse en modalité « audio » était incorrecte. Pour les propositions relatives, l'avatar a permis la compréhension de la phrase dans trois cas. Pour les prépositions, la modalité « avatar » compte une réponse correcte de plus qu'en modalité « audio ». Concernant les indices morphologiques, la pertinence de l'avatar semble moins présente puisque hormis pour les indices de genre où l'on observe que la modalité « avatar » compte une réponse correcte de plus qu'en modalité « audio », aucune différence n'est relevée pour les phrases contenant un indice de nombres. On relève même pour les clitiques accusatifs une différence négative de 4, c'est-à-dire que les réponses ont plus souvent été incorrectes en modalité « avatar » qu'en modalité « audio ».

L'hypothèse 2 est donc validée puisque l'ajout de la modalité visuelle « avatar » est **plus aidant** pour certains indices, notamment les **indices syntaxiques de négation et de propositions**.

Cependant, l'indice de confiance p-value obtenu à partir du test Chi² indique que les différences obtenues entre chaque indice morphosyntaxique ne sont pas statistiquement significatives et donc représentatives de la population.

1.5 Troisième hypothèse : préférence des sujets

J'ai donc relevé leurs réponses et calculé la proportion pour chaque préférence. Une légère préférence est constatée pour la modalité « avatar » pour 38% des sujets, suivi par 31% des sujets pour la modalité « humain » et enfin 23% pour la modalité « audio ». 8% des sujets interrogés n'ont aucune préférence.

L'hypothèse 3 est validée puisque la modalité préférée **par les sujets est la modalité « avatar »**.

La valeur de l'écart réduit entre les proportions calculée à l'aide du test Chi², indique que la différence obtenue entre chaque proportion n'est pas statistiquement significative.

2 Analyse de l'étude

En analysant chaque phrase dans les trois modalités, nous constatons que 20% des phrases ont été incomprises dans les trois modalités. Ainsi, même en modalité « humain », c'est-à-dire en situation « optimale », l'item est échoué. Plusieurs explications peuvent être apportées : soit le format vidéo dépourvu de contexte et différent d'un échange en face à face classique ne permet pas l'accès au sens de l'item, soit l'indice morphosyntaxique présent ou d'autres indices présents dans la phrase n'ont pas permis la compréhension de l'item, soit l'enfant atteint de surdit  est d mun  par ces deux param tres. Ces 20% « échoués » ne permettent donc pas d'évaluer l'intelligibilit  de l'avatar, qui est l'objectif principal de notre  tude.

Nous constatons  galement **que 32,5% des items ont  t  compris dans les trois modalit s**. Ainsi, pour cet  chantillon, le canal auditif suffit   l'enti re compr hension de l'item. Le canal visuel n'apporte donc pas d'indices suppl mentaires n cessaires   la compr hension de l'item. Cet  chantillon non plus ne permet donc  galement pas d' valuer l'intelligibilit  de l'Avatar.

Ces données peuvent permettre d'affiner le corpus pour présenter aux sujets des items adapté à la population interrogée. J'ai étudié en détails les phrases concernées afin de connaître précisément les indices et phrases les moins adaptées du corpus. Ce point sera abordé dans la partie concernant les limites de cette étude.

3 Limites de l'étude

3.1 Population

Les prises de contact pour la passation de cette étude, à savoir la mise en contact avec des instituts spécialisés et des orthophonistes, ont débuté en septembre 2019. Les premières expérimentations ont eu lieu en octobre 2019 et se sont poursuivis jusqu'en février 2020. Ainsi à cette période 10 sujets ont été rencontrés, 8 rencontres ont abouti à une passation complète. Environ 15 enfants présentant une surdité devaient être rencontrés entre mars et mai 2020. Or mi-mars, la crise sanitaire du Covid-19 a engendré des mesures de confinements stricts ne permettant pas le déroulement d'entretiens pour mon étude. Ainsi les prévisions de données n'ont pas été atteintes et l'extension de notre étude à une population contrôle d'enfants normo-entendants, prévu en avril 2020 n'a pas pu être réalisée.

Les données recueillies des enfants rencontrés sont encourageantes mais le petit nombre ne permet pas d'obtenir des résultats statistiquement significatifs. Ils sont pourtant essentiels pour pouvoir construire une étude fiable sur laquelle s'appuyer ensuite pour la création d'outils. Les données actuelles incitent donc à poursuivre cette étude sur de nouveaux sujets sourds ainsi qu'avec une population contrôle normo-entendante afin de comparer les tendances recueillies sur notre population d'enfants présentant une surdité avec une population contrôle normo-entendante.

3.2 Matériel

3.2.1 Programme

Le programme crée pour l'étude comporte quelques biais et limites dans l'exploitation de nos résultats. Notamment, le temps de latence calculé pour chaque item n'est pas assez précis, l'idéal serait d'ajouter un chronomètre qui débute à la fin du visionnage de la vidéo et se termine à l'instant où le sujet donne sa réponse, c'est-à-dire lorsqu'une réponse est cochée. Or actuellement le programme relève l'heure exacte à la seconde près à laquelle un nouvel item démarre, c'est-à-dire lorsqu'une nouvelle page s'affiche à l'écran. Ceci permet de calculer le

temps passé sur chaque item mais inclus le temps de visionnage de vidéos et ne prend pas en compte le potentiel temps mis par l'ordinateur pour déclencher la vidéo, en raison du débit internet.

En effet un deuxième biais est le fait que l'expérimentation est soumise à une connexion internet puisque le programme s'ouvre via un URL internet qui permet à la fin de l'entretien d'envoyer automatiquement les résultats à Mr Ouni. Or dans certains cas, la connexion est soumise à certains ralentissements et certaines vidéos prennent quelques secondes avant de démarrer. Cette attente est désagréable pour le sujet, et lui demande une attention très soutenue puisque la vidéo peut démarrer à tout moment, il est alors possible que le sujet ne soit pas attentif au visuel lorsque la vidéo de quelques secondes démarre.

Un autre détail à améliorer serait de pouvoir intégrer au programme la planche d'images correspondant à l'item. En effet actuellement le programme diffuse une vidéo contenant l'item puis le sujet désigne une image parmi quatre sur une planche contenue dans un porte-document. Le sujet doit donc alterner son regard entre l'écran et le porte-document. L'idéal serait donc que l'intégralité des informations soit sur l'écran pour faciliter la tâche du sujet.

3.2.2 Vidéos

L'expérimentation comprend 60 vidéos. Les 60 extraits contiennent le même enregistrement sonore. La qualité de cet enregistrement pourrait être amélioré notamment sur le nombre de décibels. En effet augmenter l'intensité du signal permettrait une meilleure intelligibilité.

Concernant les images, 20 extraits montrent le locuteur original, 20 extraits montrent l'avatar et les 20 derniers extraits contiennent un écran noir. Il existe une différence de proximité entre le locuteur humain et l'avatar. Le locuteur original est filmé d'assez loin, l'image montre son buste, contrairement à l'avatar qui est plus proche, on ne voit que son visage et son cou. On observe donc un biais sur la lisibilité des indices visuels présents chez le locuteur humain par rapport à l'avatar. Il serait pertinent de filmer à nouveau le locuteur humain de façon à observer exactement la même distance entre le locuteur humain et l'avatar.

3.2.3 Corpus

En comparant les réponses obtenues aux différentes phrases, j'ai calculé la moyenne et l'écart-type du nombre de réponses correctes. La moyenne du nombre de réponses correctes sur les 20 phrases du corpus, toutes modalités confondues, est égale à 13,5 et l'écart-type est égal à 5,8. (voir tableau ci-dessous)

Indices	Phrases	Nombre de réponses correctes	Nombre de réponses incorrectes	Ecart des réponses correctes à la moyenne
CA	La fille le regarde.	6	18	-1,5ET
CA	La fille la regarde.	16	8	M+
N	L'ours dort.	17	7	M+
N	Les ours dorment.	18	6	+1ET
N	Les garçons jouent avec le train.	14	10	M
N	Le garçon joue avec le train.	13	11	M
G	Il est tombé.	19	5	+1,5ET
G	Elle est tombée.	16	8	M+
G	Il va partir travailler.	13	11	M
G	Elle va partir travailler.	13	11	M
FN	Le peigne n'est pas rouge.	16	8	M+
FN	Le peigne est rouge.	19	5	+1ET
FN	Le chien ne boit pas.	22	2	+1,5ET
FN	Le chien boit.	22	2	+1,5ET
Re	Le cercle qui est dans l'étoile est jaune.	5	19	-1,5ET
Re	L'étoile qui est dans le cercle est jaune.	4	20	-1,5ET
Re	Le cercle dans l'étoile est jaune.	4	20	-1,5ET
Re	L'étoile dans le cercle est jaune.	3	21	-2ET
Pr	Le crayon est derrière la boîte.	14	10	M
Pr	Le crayon est devant la boîte.	16	8	M+
Effectif total		270	210	
Moyenne		13,5	10,5	
Ecart-type		5,8	5,8	

Figure 38 : Tableau récapitulatif des écarts des réponses correctes à la moyenne pour chaque phrase du corpus

Nous constatons alors que certaines phrases sont échouées par rapport à la moyenne.

En effet les phrases contenant une proposition relative se situent les quatre entre -1,5 et -2 écart-types de la moyenne du nombre de réponses correctes avec un nombre de réponses correctes situé entre 3 et 5. Si les propositions relatives sont difficiles à maîtriser dans l'apprentissage de la morphosyntaxe, les phrases choisies sont particulièrement difficiles car elles représentent une situation figées de formes géométriques et de couleurs. Il aurait été plus judicieux d'utiliser des phrases exprimant une action, par exemple « La voiture qui double le camion est rouge. » ou « Le chat courant après le chien est noir. ». De plus les quatre phrases présentées pour cet indice se ressemblent et sont évaluées à travers la même planche d'images. Ainsi cette planche a été présentée aux sujets 12 fois durant l'expérience sur 60 items. Il serait préférable de distinguer les deux couples de phrases à travers deux planches d'images distinctes afin de ne pas lasser et perdre le sujet.

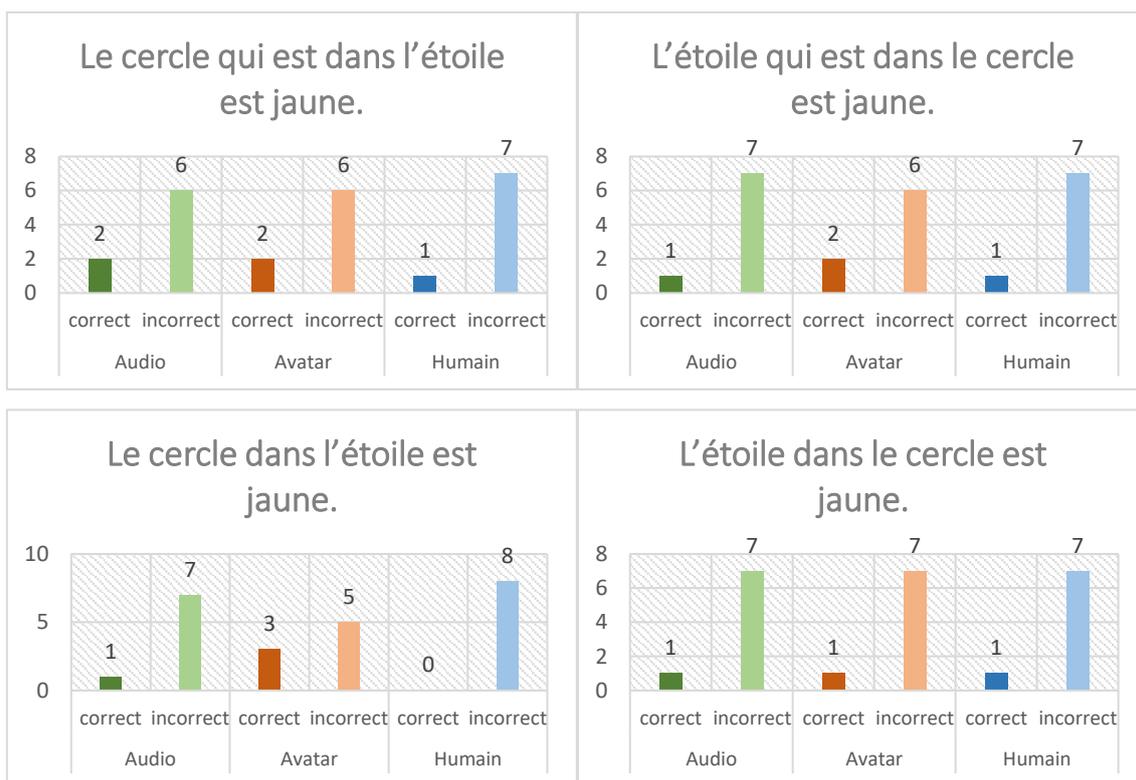


Figure 39 : Diagrammes récapitulatifs du nombre de réponses correctes et incorrectes par modalité pour chaque phrase contenant une proposition relative.

La phrase « La fille le regarde. » contenant le clitique accusatif « le » se situe à 1 écart-type sous la moyenne avec seulement 6 réponses correctes. Pour autant la phrase « la fille la regarde. » indique 16 réponses correctes. Nous pouvons en déduire que l'erreur ne se trouve pas au niveau de la compréhension du sens de l'indice morphosyntaxique mais plutôt à son identification. Il est probable qu'une partie des erreurs soit causée par une confusion de « le » identifié « se ». En effet les visèmes des phonèmes [s] et [l] sont proches et peu différenciables, notamment lorsqu'ils sont articulés avec la voyelle [ə] (Istria et al., 1982). En privilégiant une voyelle plus ouverte telle que [a] ou une consonne plus visible et stable telle que [m], par exemple « la » ou « me » on permet une meilleure lecture du visème et ainsi une meilleure compréhension de l'indice morphosyntaxique.

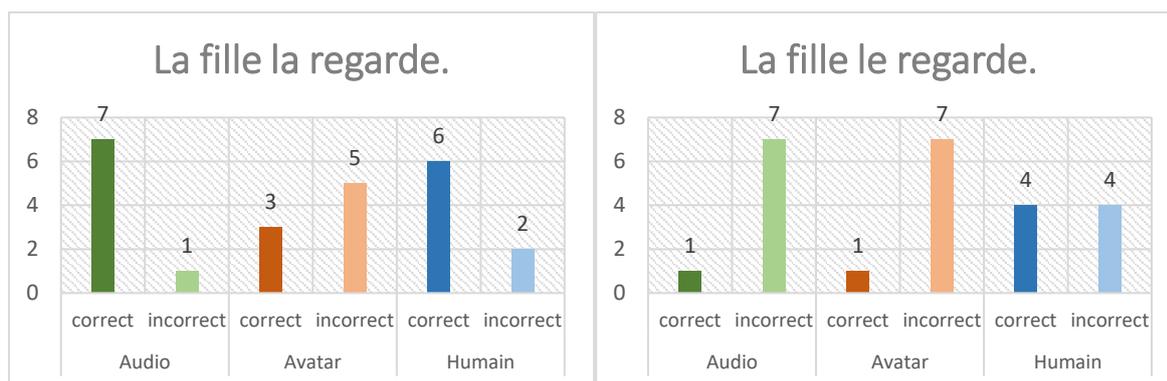


Figure 40 : Diagrammes récapitulatifs du nombre de réponses correctes et incorrectes par modalité pour chaque phrase contenant un clitique accusatif.

3.2.4 Planches d'images

Certains détails concernant les planches d'images présentées pourraient être améliorés afin de faciliter la passation. Les planches sont issues de deux tests, ainsi certaines planches ont été réalisées en format paysage tandis que les autres sont en format portrait. Cette différence nécessite de devoir tourner le porte document régulièrement pour proposer au sujet la planche d'images correspondante dans le sens de lecture. Il serait donc intéressant de réaliser de nouvelles planches en respectant un unique format, paysage ou portrait.

Enfin la majorité des images est en noir et blanc, la présence de couleurs pourrait permettre d'apporter plus d'intérêts à l'épreuve ainsi que plus de facilité.

3.3 Passation de l'épreuve

L'épreuve comprend 60 items et dure entre 20 et 30 minutes suivant l'enfant. La tâche à réaliser est toujours la même, à savoir, écouter et regarder un court extrait vidéo et choisir une image parmi quatre correspondante à la phrase entendu précédemment. L'entretien peut donc paraître long et répétitif pour l'enfant. La quantité d'items est peut-être trop important puisque 2 enfants sur les 10 rencontrés ont préféré arrêter la passation avant la fin.

De plus, le temps de réponse étant estimé à partir de l'heure de départ de chaque page, il est difficile de faire une pause entre deux items puisque cela fausserait le temps de latence. Si le calcul du temps de réponse est modifié, l'enfant pourrait prendre son temps entre deux items ce qui serait plus confortable pour lui.

4 Intérêts de l'étude

4.1 Recherches complémentaires

Au vu des données recueillies lors de cette étude, il semble important de tout d'abord poursuivre les entretiens actuels afin d'obtenir plus de données qui pourront amener à des résultats statistiquement significatifs.

Il serait également pertinent de réaliser cette étude sur différentes populations. Une étude réalisée auprès d'une population d'enfants entendants sans troubles du langage afin d'obtenir des données contrôles était envisagée et est reportée. Cette étude pourrait également s'effectuer à l'aide d'une bande sonore détériorée afin de recréer des conditions sonores altérées où la modalité visuelle prend en charge une partie importante de la compréhension langagière (Sumbly & Pollack, 1954).

Une étude similaire pourrait également être envisagée pour une population d'individus atteints de troubles du langage tels que la dysphasie où les troubles de la compréhension peuvent être importants. La présence d'indices visuels supplémentaires pourrait alors être une aide pour cette population.

4.2 Applications

A la suite de ces études, si les résultats obtenus indiquent une amélioration significative de la compréhension grâce à l'avatar, comme l'étude que nous venons de réaliser le dévoile timidement, la technologie Dynalips pourrait permettre la création de nouveaux outils de communication.

En effet un avatar numérique dispose de nombreux avantages par rapport à un individu. Tout d'abord cette technologie propose un visage neutre, qui ne vieillit pas, adaptable et modulable en fonction des besoins esthétiques. De plus l'avatar ne commet pas d'erreur d'articulation et propose une articulation constante, le locuteur n'est alors pas soumis aux variations d'articulations interpersonnelles et peut s'entraîner à la lecture labiale avec un locuteur constant. Il évite également les difficultés liées à l'individu et au droit à l'image, ainsi que les questions sous-jacentes sur l'identité de l'individu.

CONCLUSION

Les personnes atteintes de troubles du langage comme c'est le cas pour les individus présentant une surdité ou un trouble dysphasique sont généralement handicapées dans leur maîtrise de la langue orale. Ce handicap crée de nombreuses difficultés supplémentaires dans leur milieu scolaire, professionnel et personnel. Les techniques de remédiation médicales et paramédicales ne leur permettent pas de recouvrer un mode de vie totalement identique aux autres. Les appareillages auditifs ne permettent que très rarement de recouvrer la totalité du gain auditif perdu par exemple. Des méthodes alternatives et outils de communication sont alors disponibles pour aider ces personnes à communiquer avec leurs entourages. Ces moyens nécessitent néanmoins soit un apprentissage, comme c'est le cas pour l'utilisation de la Langue des Signes Française ou du Langage Parlé Complété, soit du matériel, par exemple les tableaux de communication ou les tablettes interactives. Le plus simple pour ces personnes, qui ne nécessite ni matériel, ni apprentissage ou adaptation de la part du locuteur est la lecture labiale qui permet de compléter la compréhension altérée du langage oral. Néanmoins la lecture labiale n'est pas toujours possible. C'est le cas lorsque le message de l'interlocuteur est délivré à distance ou uniquement sous forme audio ou lorsque les conditions visuelles requises ne sont pas présentes, lorsque l'interlocuteur n'est pas face à son locuteur, lorsque l'éclairage n'est pas suffisant ou que les lèvres de l'interlocuteur sont masquées par une moustache par exemple. La lecture labiale doit aussi prendre en compte les variations d'articulation interindividuelles. Toutes ces conditions rendent parfois difficile voire impossible la lecture labiale.

De plus la situation sanitaire actuelle liée à l'apparition du covid-19 nous montre qu'il est nécessaire de trouver des solutions alternatives à nos méthodes de communication habituelles et de développer les outils de communication à distance. L'exercice orthophonique en est un exemple puisque la majorité des cabinets d'orthophonie libéraux ayant dû fermer durant 8 semaines entre mars et mai 2020, il a été nécessaire pour de nombreux orthophonistes de développer le télé-soin et les techniques de communication à distance.

L'utilisation de la technologie Dynalips, proposant un support visuel supplémentaire à un contenu auditif à travers un avatar parlant, pourrait apporter une solution à ces différents problèmes. Néanmoins, il est nécessaire d'évaluer la performance de cet avatar parlant à travers son intelligibilité pour savoir si cet outil peut aider les individus présentant une déficience auditive ou un trouble du langage. C'est dans ce but qu'a été élaborée cette étude qui avait pour objectif de rencontrer entre 20 et 30 enfants présentant une surdité sévère à profonde ainsi

qu'autant ne présentant aucune déficience auditive. Malheureusement la crise sanitaire du Covid-19 nous a contraint à interrompre nos entretiens après avoir rencontré 10 sujets.

Les données que nous avons tout de même pu recueillir sont encourageantes et répondent en partie aux hypothèses formulées avant le début de cette étude. Malgré cela, les sujets rencontrés ne sont pas assez nombreux pour que nos résultats soient statistiquement significatifs. Cette étude doit donc être poursuivie dans le but d'obtenir des données stables sur lesquelles s'appuyer pour assurer la fiabilité de l'avatar parlant.

Dynalips pourrait ainsi permettre la création et le développement de nouveaux outils numériques, à l'image de livres interactifs déjà créés en 2016 dans le cadre du projet de recherche Handicom (Piquard-Kipffer, 2016).

BIBLIOGRAPHIE

- Alm, N., & Parnes, P. (1995). Augmentative and alternative communication : Past, present and future. *Folia phoniatrica et logopaedica*, 47(3), 165-192.
- Antoine, J.-Y., & Maurel, D. (2007). Aide à la communication pour personnes handicapées et prédiction de texte. *TAL*, 48(2), 9-46.
- Bassano, D. (2000). La constitution du lexique : Le développement lexical précoce. *L'acquisition du langage*, 1, 137-168.
- Benoit, C., Mohamadi, T., & Kandel, S. (1994). Effects of phonetic context on audio-visual intelligibility of French. *Journal of Speech, Language, and Hearing Research*, 37(5), 1195-1203.
- Bertoncini, J., & Boysson-Bardies, B. de. (2000). La perception et la production de la parole avant deux ans. *L'acquisition du langage. Le langage en émergence, de la naissance à trois ans*, 1, 95-136.
- Bertoncini, J., & Mehler, J. (1980). La perception du langage chez le nourrisson : Quelques observations. *Reproduction Nutrition Développement*, 20(3B), 859-869.
- Bond, E. K. (1972). Perception of form by the human infant. *Psychological bulletin*, 77(4), 225.
- Brannon Jr, J. B. (1968). Linguistic word classes in the spoken language of normal, hard-of-hearing, and deaf children. *Journal of Speech and Hearing Research*, 11(2), 279-287.
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), B33-B44.
- Brin-Henry, F., Courier, C., Lederlé, E., & Masy, V. (2011). Dictionnaire d'orthophonie. *Isbergues, Ortho-Edition*.
- Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 335(1273), 95-103.

- Coquet, F. (2006). Prise en compte du versant compréhension dans l'évaluation et la prise en charge du langage oral chez l'enfant. *Rééducation orthophonique*, 44(227), 7-20.
- Cozolino, L. (2014). *The neuroscience of human relationships : Attachment and the developing social brain*. WW Norton & Company.
- Dauman, R., Carbonnière, B., Soriano, V., Berger-Lautissier, S., Bouyé, J., Debruge, E., Coriat, G., & Bébéar, J. (1998). Implants cochléaires chez l'adulte et l'enfant. *Encycl Med Chir, Elsevier, Paris, Oto-Rhino-Laryngologie*, 20.
- de Boysson-Bardies, B. (1996). *Comment la parole vient aux enfants*. Odile Jacob.
- Dodd, B., McIntosh, B., & Woodhouse, L. (1998). Early lipreading ability and speech and language development of hearing-impaired pre-schoolers. *Hearing by Eye (II): The Psychology of Speechreading and Auditory-visual Speech*, 229-242.
- Dumont, A., & Calbour, C. (2002). *Voir la parole : Lecture labiale, perception audiovisuelle de la parole* (Vol. 18). Elsevier Masson.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(3968), 303-306.
- Eric, L. (1967). *Biological foundations of language*.
- Fodor, J. A., Garrett, M. F., & Brill, S. (1975). Pi ka pu : The perception of speech sounds by prelinguistic infants. *Perception & Psychophysics*, 18(2), 74-78.
- Gabet, M., & Thiollier, H. (2012). *LPC et perception de la morphosyntaxe : Tête codeuse numérique versus locuteur humain*. ir00581a. <http://search.ebscohost.com/bases-doc.univ-lorraine.fr/login.aspx?direct=true&AuthType=ip,url,uid&db=ir00581a&AN=petale.univ.lorraine.ori.18293&lang=fr&site=eds-live&scope=site>
- Geffner, D. S., & Freeman, L. R. (1980). Assessment of language comprehension of 6-year-old deaf children. *Journal of communication disorders*, 13(6), 455-470.

- Goda, S. (1964). Spoken syntax of normal, deaf, and retarded adolescents. *Journal of Verbal Learning and Verbal Behavior*, 3(5), 401-405.
- Hage, C., Charlier, B., & Leybaert, J. (2006). *Compétences cognitives, linguistiques et sociales de l'enfant sourd : Pistes d'évaluation*. Editions Mardaga.
- Istria, M., Nicolas-Jeantoux, C., & Tamboise, J. (1982). *Manuel de lecture labiale:(exercices d'entraînement)*. Masson.
- Kail, M., & Fayol, M. (2015). *L'acquisition du langage. Vol. I, Vol. I.* P.U.F.
- Krahe, J. L., & Alegria, J. (2007). *Surdit  et langage : Proth ses, LPC et implants cochl aires*. Presses universitaires de Vincennes.
- Larousse, P. (1867). *Grand dictionnaire universel du XIXe si cle*. Larousse.
- Le Normand, M.-T. (2004). Evaluation du lexique de production chez des enfants sourds profonds munis d'un implant cochl aire sur un suivi de trois ans. *R ducation orthophonique*, 217, 125-140.
- Lecocq, P. (1998). *L' . co. S. Se une  preuve de compr hension syntaxico-s mantique (manuel et  preuve) : Deux volumes*. Presses Univ. Septentrion.
- Leybaert, J. (2011). *La langue fran aise parl e Compl t e : Fondements et Perspectives*.
- Leybaert, J., & Colin, C. (2007). Le r le des informations visuelles dans le d veloppement du langage de l'enfant sourd muni d'un implant cochl aire. *Enfance*, 59(3), 245-253.
- Lina-Granade, G., Gallego, S., Thai-Van, H., & Truy,  . (2010). *Appareillage auditif conventionnel par voie a rienne*.
- Machart. (2013).
- Massaro, D. W., & Cohen, M. M. (1995). Perceiving talking faces. *Current Directions in Psychological Science*, 4(4), 104-109.
- McFarland, D. H. (2009). *L'anatomie en orthophonie : Parole, d glutition et audition*. Elsevier Masson.

- Mcgurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748. <https://doi.org/10.1038/264746a0>
- Moffitt, A. R. (1971). Consonant cue perception by twenty-to twenty-four-week-old infants. *Child Development*, 717-731.
- Moudenc, T., & Emerard, F. (2003). *Synthèse vocale et handicap*. 58, 928-934.
- Ouni, S., & Gris, G. (2018). Dynamic lip animation from a limited number of control points : Towards an effective audiovisual spoken communication. *Speech Communication*, 96, 49-57. <https://doi.org/10.1016/j.specom.2017.11.006>
- Pandzic, I. S., Ostermann, J., & Millen, D. (1999). User evaluation : Synthetic talking faces for interactive services. *The visual computer*, 15(7-8), 330-340.
- Parke, F. I. (1974). *A parametric model for human faces*. UTAH UNIV SALT LAKE CITY DEPT OF COMPUTER SCIENCE.
- Piquard-Kipffer, A. (2014). *Critères d'évaluation d'un album numérique pour des enfants en difficulté de langage*.
- Piquard-Kipffer, A. (2016). *Un album numérique pour raconter une histoire avec un avatar narrateur*. XVIèmes rencontres internationales en orthophonie-Orthophonie et technologies innovantes.
- Piquard-Kipffer, A., Lelarge, D., Pierron, L., & Monnay, F. (2010). *Création de livres numériques pour enfants présentant des troubles du langage*.
- Prat, R. (2007). L'épaisseur du langage : Comment vient-elle au bébé? *Enfances Psy*, 3, 10-19.
- Reisberg, D., Mclean, J., & Goldfield, A. (1987). *Easy to hear but hard to understand : A lip-reading advantage with intact auditory stimuli*.
- Rondal, J.-A. (1998). *Votre enfant apprend à parler* (Vol. 78). Editions Mardaga.
- Rondal, J.-A. (2018). *De la sémantique à la morphosyntaxe. L'apprentissage implicite de la grammaire* (Ortho édition).

- Rosier, F. (2017, avril 4). Surdit   : La cl   du succ  s des implants. *Le Monde*.
https://www.lemonde.fr/medecine/article/2017/04/04/surdite-la-cle-du-succes-des-implants_5105745_1650718.html
- Shadows. (2018, juin 18). Annecy 2018 : Dynalips, solution de lip-sync automatis  e.
3DVF.com.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of america*, 26(2), 212-215.
- Truy, E., Gallego, S., Bouccara, D., & Tringali, S. (2012). *Proth  ses auditives amplificatrices par voie non a  rienne*.
- Van Mulken, S., Andr  , E., & M  ller, J. (1999). *An empirical study on the trustworthiness of life-like interface agents*. 152-156.
- Walker, J. H., Sproull, L., & Subramani, R. (1994). *Using a human face in an interface*. 85-91.
- Wathour, J., Nyssen, E., & Deggouj, N. (2019). *La lecture labiale en groupe : Pratique aupr  s des adultes implant  s cochl  aires et/ou appareill  s-30 s  ances d'exercices*. Elsevier Health Sciences.
- www.laboratoires-unisson.com. (2020). *La perte auditive*.

ANNEXES

Annexe 1 : Corpus

N°		Phrase	1	2	3	4
1	A	La fille le regarde.				
2	H	Il est tombé.				
3	O	Les ours dorment.				
4	A	Elle est tombée.				
5	H	Le garçon joue avec le train.				
6	A	Le cercle qui est dans l'étoile est jaune.				
7	H	Les garçons jouent avec le train.				
8	O	Le crayon est devant la boîte				
9	H	Le crayon est derrière la boîte.				
10	A	L'ours dort.				
11	H	Le chien boit.				
12	A	Le crayon est derrière la boîte.				
13	O	L'étoile dans le cercle est jaune.				
14	H	Le chien ne boit pas.				
15	A	Le crayon est devant la boîte.				
16	O	Le peigne n'est pas rouge.				
17	H	Il va partir travailler.				
18	A	Le cercle dans l'étoile est jaune.				
19	O	Elle va partir travailler.				
20	H	L'étoile qui est dans le cercle est jaune.				
21	A	Il est tombé.				
22	O	Le chien boit.				
23	H	Elle va partir travailler.				
24	O	Le cercle qui est dans l'étoile est jaune.				
25	H	Le peigne est rouge.				
26	A	Il va partir travailler.				
27	O	La fille le regarde.				
28	A	Les garçons jouent avec le train.				
29	O	La fille la regarde.				
30	O	Le crayon est derrière la boîte.				
31	A	La fille la regarde.				
32	O	Il est tombé.				
33	A	Le chien ne boit pas.				
34	H	L'étoile dans le cercle est jaune.				
35	O	Le garçon joue avec le train.				
36	H	Le peigne n'est pas rouge.				
37	O	Le chien ne boit pas.				
38	A	Le peigne est rouge.				
39	O	Il va partir travailler.				
40	H	L'ours dort.				
41	A	L'étoile qui est dans le cercle est jaune.				
42	H	Le cercle dans l'étoile est jaune.				
43	A	Le peigne n'est pas rouge.				
44	O	Les garçons jouent avec le train.				
45	H	Le cercle qui est dans l'étoile est jaune.				
46	A	Elle va partir travailler.				
47	O	Elle est tombée.				
48	A	L'étoile dans le cercle est jaune.				
49	H	Les ours dorment.				
50	O	L'étoile qui est dans le cercle est jaune.				
51	A	Le garçon joue avec le train.				
52	H	La fille le regarde.				
53	O	Le peigne est rouge.				
54	H	Le crayon est devant la boîte.				
55	O	Le cercle dans l'étoile est jaune.				
56	H	La fille la regarde.				
57	A	Le chien boit.				
58	O	L'ours dort.				
59	H	Elle est tombée.				
60	A	Les ours dorment.				

Figure 41 : Tableau récapitulatif des 20 phrases du corpus

Annexe 2 : Fichier .pad pour chaque passation

```
#data
date xx/xx/xx      heure : xx:xx:xx
resolution
dimension
IP :
navigateur :
version :
langue du formulaire : francais
nom & prenom :
Genre :
Age :
Classe :
Ecole :
Surdité :
Audition :
Appareillage :
fileXP : /ORTHOLEA/XP/exp_percep_LEA.xml
-----
#evaluation
#evaluation 1 type filename cat rep mode bruit nbplay nbpause nbend |q1| heure
1 video Stimuli/video/avatar/video_br_A_cat_CA_ph_1.mp4 CA Feuille1 image1 avatar 1 1 1 |Image 1| 12:52:21
1 video Stimuli/video/humain/video_br_H_cat_G_ph_7.mp4 G Feuille7 image2 humain 2 1 1 |Image 2| 12:52:40
1 video Stimuli/video/audio/video_br_O_cat_N_ph_4.mp4 N Feuille4 image1 audio 1 1 1 |Image 1| 12:52:53
1 video Stimuli/video/avatar/video_br_A_cat_G_ph_8.mp4 G Feuille8 image3 avatar 1 1 1 |Image 3| 12:53:06
1 video Stimuli/video/humain/video_br_H_cat_N_ph_6.mp4 N Feuille6 image1 humain 2 1 1 |Image 2| 12:53:30
1 video Stimuli/video/avatar/video_br_A_cat_Re_ph_15.mp4 Re Feuille15 image3 avatar 1 1 1 |Image 4| 12:53:54
1 video Stimuli/video/humain/video_br_H_cat_N_ph_5.mp4 CA Feuille5 image4 humain 2 1 1 |Image 1| 12:54:14
1 video Stimuli/video/audio/video_br_O_cat_Pr_ph_20.mp4 Pr Feuille20 image4 audio 1 1 1 |Image 4| 12:54:30
1 video Stimuli/video/humain/video_br_H_cat_Pr_ph_19.mp4 Pr Feuille19 image1 humain 2 1 1 |Image 1| 12:54:51
1 video Stimuli/video/avatar/video_br_A_cat_N_ph_3.mp4 N Feuille3 image4 avatar 1 1 1 |Image 4| 12:55:10
1 video Stimuli/video/humain/video_br_H_cat_FN_ph_14.mp4 FN Feuille14 image3 humain 2 1 1 |Image 3| 12:55:33
1 video Stimuli/video/avatar/video_br_A_cat_Pr_ph_19.mp4 Pr Feuille19 image1 avatar 1 1 1 |Image 1| 12:55:46
1 video Stimuli/video/audio/video_br_O_cat_Re_ph_18.mp4 Re Feuille18 image2 audio 1 1 1 |Image 4| 12:56:04
1 video Stimuli/video/humain/video_br_H_cat_FN_ph_13.mp4 FN Feuille13 image4 humain 1 1 1 |Image 4| 12:56:29
1 video Stimuli/video/avatar/video_br_A_cat_Pr_ph_20.mp4 Pr Feuille20 image4 avatar 1 1 1 |Image 4| 12:56:40
1 video Stimuli/video/audio/video_br_O_cat_FN_ph_11.mp4 FN Feuille11 image2 audio 1 1 1 |Image 3| 12:56:53
1 video Stimuli/video/humain/video_br_H_cat_G_ph_9.mp4 G Feuille9 image3 humain 2 1 1 |Image 3| 12:57:26
1 video Stimuli/video/avatar/video_br_A_cat_Re_ph_17.mp4 Re Feuille17 image3 avatar 1 1 1 |Image 2| 12:57:43
1 video Stimuli/video/audio/video_br_O_cat_G_ph_10.mp4 G Feuille10 image4 audio 1 1 1 |Image 4| 12:57:55
1 video Stimuli/video/humain/video_br_H_cat_Re_ph_16.mp4 Re Feuille16 image2 humain 1 1 1 |Image 4| 12:58:40

1 video Stimuli/video/avatar/video_br_A_cat_G_ph_7.mp4 G Feuille7 image2 avatar 1 1 1 |Image 2| 12:58:48
1 video Stimuli/video/audio/video_br_O_cat_FN_ph_14.mp4 FN Feuille14 image3 audio 1 1 1 |Image 3| 12:58:58
1 video Stimuli/video/humain/video_br_H_cat_G_ph_10.mp4 G Feuille10 image4 humain 1 1 1 |Image 4| 12:59:21
1 video Stimuli/video/audio/video_br_O_cat_Re_ph_15.mp4 Re Feuille15 image3 audio 1 1 1 |Image 2| 12:59:36
1 video Stimuli/video/humain/video_br_H_cat_FN_ph_12.mp4 FN Feuille12 image3 humain 2 1 1 |Image 3| 13:00:28
1 video Stimuli/video/avatar/video_br_A_cat_G_ph_9.mp4 G Feuille9 image3 avatar 1 1 1 |Image 3| 13:00:39
1 video Stimuli/video/avatar/video_br_A_cat_N_ph_5.mp4 N Feuille5 image4 avatar 1 1 1 |Image 1| 13:00:58
1 video Stimuli/video/audio/video_br_O_cat_CA_ph_1.mp4 CA Feuille1 image1 audio 1 1 1 |Image 1| 13:01:06
1 video Stimuli/video/audio/video_br_O_cat_CA_ph_2.mp4 CA Feuille2 image2 audio 1 1 1 |Image 2| 13:01:10
1 video Stimuli/video/audio/video_br_O_cat_Pr_ph_19.mp4 Pr Feuille19 image1 audio 1 1 1 |Image 1| 13:01:21
1 video Stimuli/video/avatar/video_br_A_cat_CA_ph_2.mp4 CA Feuille2 image2 avatar 1 1 1 |Image 2| 13:01:31
1 video Stimuli/video/audio/video_br_O_cat_G_ph_7.mp4 G Feuille7 image2 audio 1 1 1 |Image 2| 13:01:39
1 video Stimuli/video/avatar/video_br_A_cat_FN_ph_13.mp4 FN Feuille13 image4 avatar 1 1 1 |Image 4| 13:01:51
1 video Stimuli/video/humain/video_br_H_cat_Re_ph_18.mp4 Re Feuille18 image2 humain 2 1 1 |Image 4| 13:02:18
1 video Stimuli/video/audio/video_br_O_cat_N_ph_6.mp4 N Feuille6 image1 audio 1 1 1 |Image 1| 13:02:27
1 video Stimuli/video/humain/video_br_H_cat_FN_ph_11.mp4 FN Feuille11 image2 humain 2 1 1 |Image 2| 13:02:44
1 video Stimuli/video/audio/video_br_O_cat_FN_ph_13.mp4 FN Feuille13 image4 audio 1 1 1 |Image 4| 13:02:55
1 video Stimuli/video/avatar/video_br_A_cat_FN_ph_12.mp4 FN Feuille12 image3 avatar 1 1 1 |Image 3| 13:03:04
1 video Stimuli/video/audio/video_br_O_cat_G_ph_9.mp4 G Feuille9 image3 audio 1 1 1 |Image 3| 13:03:16
1 video Stimuli/video/humain/video_br_H_cat_N_ph_3.mp4 N Feuille3 image4 humain 2 1 1 |Image 4| 13:03:33
1 video Stimuli/video/avatar/video_br_A_cat_Re_ph_16.mp4 Re Feuille16 image2 avatar 1 1 1 |Image 4| 13:03:43
1 video Stimuli/video/humain/video_br_H_cat_Re_ph_17.mp4 Re Feuille17 image3 humain 1 1 1 |Image 2| 13:04:09
1 video Stimuli/video/avatar/video_br_A_cat_FN_ph_11.mp4 FN Feuille11 image2 avatar 1 1 1 |Image 2| 13:04:20
1 video Stimuli/video/audio/video_br_O_cat_N_ph_5.mp4 N Feuille5 image4 audio 1 1 1 |Image 1| 13:04:33
1 video Stimuli/video/humain/video_br_H_cat_Re_ph_15.mp4 Re Feuille15 image3 humain 1 1 1 |Image 2| 13:05:04
1 video Stimuli/video/avatar/video_br_A_cat_G_ph_10.mp4 G Feuille10 image4 avatar 1 1 1 |Image 4| 13:05:13
1 video Stimuli/video/audio/video_br_O_cat_G_ph_8.mp4 G Feuille8 image3 audio 1 1 1 |Image 3| 13:05:21
1 video Stimuli/video/avatar/video_br_A_cat_Re_ph_18.mp4 Re Feuille18 image2 avatar 1 1 1 |Image 4| 13:05:29
1 video Stimuli/video/humain/video_br_H_cat_N_ph_4.mp4 N Feuille4 image1 humain 2 1 1 |Image 1| 13:05:39
1 video Stimuli/video/audio/video_br_O_cat_Re_ph_17.mp4 Re Feuille17 image2 audio 1 1 1 |Image 4| 13:05:49
1 video Stimuli/video/avatar/video_br_A_cat_N_ph_6.mp4 N Feuille6 image1 avatar 1 1 1 |Image 1| 13:05:58
1 video Stimuli/video/humain/video_br_H_cat_CA_ph_1.mp4 CA Feuille1 image1 humain 2 1 1 |Image 1| 13:06:11
1 video Stimuli/video/audio/video_br_O_cat_FN_ph_12.mp4 FN Feuille12 image3 audio 1 1 1 |Image 3| 13:06:20
1 video Stimuli/video/humain/video_br_H_cat_Pr_ph_20.mp4 Pr Feuille20 image4 humain 2 1 1 |Image 4| 13:06:38
1 video Stimuli/video/audio/video_br_O_cat_Re_ph_17.mp4 Re Feuille17 image3 audio 1 1 1 |Image 2| 13:06:48
1 video Stimuli/video/humain/video_br_H_cat_CA_ph_2.mp4 CA Feuille2 image2 humain 2 1 1 |Image 2| 13:07:14
1 video Stimuli/video/avatar/video_br_A_cat_FN_ph_14.mp4 FN Feuille14 image3 avatar 1 1 1 |Image 3| 13:07:23
1 video Stimuli/video/audio/video_br_O_cat_N_ph_3.mp4 N Feuille3 image4 audio 1 1 1 |Image 4| 13:07:32
1 video Stimuli/video/humain/video_br_H_cat_G_ph_8.mp4 G Feuille8 image3 humain 2 1 1 |Image 3| 13:07:45
1 video Stimuli/video/avatar/video_br_A_cat_N_ph_4.mp4 N Feuille4 image1 avatar 1 1 1 |Image 1| 13:07:53

#end
com_general
```

Figure 42 : Visuel du fichier reçu par expérimentation

Annexe 3 : tableaux et graphiques

Sujet (numéro attribué par ordre de passation)	Erreurs par modalité			Nombre d'erreurs total	Temps par modalité (en minutes / en secondes)						Temps total	
	Humain (H)	Avatar (A)	Audio (O)		Humain (H)		Avatar (A)		Audio (O)			
1	5	6	8	19	05:59	359	05:37	337	05:07	307	16:08	968
2	4	6	4	14	04:14	254	03:37	217	04:10	250	12:01	721
3	8	6	8	22	03:42	222	03:52	232	04:20	260	11:54	714
4												
5	15	18	14	47	05:27	327	04:27	267	04:26	266	14:20	860
6	12	10	10	32	05:23	323	04:48	288	04:47	287	14:58	898
7												
8	7	9	13	29	04:58	298	04:07	247	04:35	275	13:40	820
9	6	12	13	31	05:10	310	04:23	263	04:30	270	14:03	843
10	6	4	7	17	04:03	243	04:48	288	06:41	401	15:32	932
EFFECTIF TOTAL	63	71	77	211 (3x70,33)	00:38:56	2336	00:35:39	2139	00:38:37	2316	01:52:36 (3x00:37:32)	
MOYENNE	7,875	8,875	9,625	26,375 (3x8,8)	00:04:52	292	00:04:27	267,375	00:04:50	289,5	00:14:04 (3x00:04:41)	
ECART-TYPE	3,51559028	4,22603538	3,27633561	9,999218719	0,03084225	44,412836	0,02445622	35,2169615	0,03136361	45,1635915	0,05958479	
MEDIANE	6,5	7,5	9	25,5	05:04	304	04:25	265	04:32	272,5	14:11	

Figure 43 : Tableau récapitulatif des réponses incorrectes et des temps de réponses par modalité

Sujet (numéro attribué par ordre de passation)	Nombre de phrases incorrectes dans 3 modalités A	% p(A)	Nombre de phrases correctes dans 3 modalités B	% p(B)	Nombre de phrases correctes en modalité humain et incorrectes en modalité audio C	% p(C)	Autres p(D)	% p(D)	Parmi C, nombre de phrases où modalité Avatar correcte D	% p(E)	% p(E)/C	Parmi C, nombre de phrases où modalité Avatar incorrecte E	% p(F)	% p(F)/C
1	1	0,05	9	0,45	5	0,25	5	0,25	2	0,1	0,4	3	0,15	0,6
2	2	0,1	12	0,6	2	0,1	4	0,2	1	0,05	0,5	1	0,05	0,5
3	6	0,3	11	0,55	1	0,05	2	0,1	1	0,05	1	0	0	0
4														
5	10	0,5	0	0	4	0,2	6	0,3	1	0,05	0,25	3	0,15	0,75
6	2	0,1	2	0,1	4	0,2	12	0,6	2	0,1	0,5	2	0,1	0,5
7														
8	3	0,15	3	0,15	8	0,4	6	0,3	5	0,25	0,625	3	0,15	0,375
9	4	0,2	3	0,15	8	0,4	5	0,25	3	0,15	0,375	5	0,25	0,625
10	4	0,2	12	0,6	2	0,1	2	0,1	2	0,1	1	0	0	0
EFFECTIF TOTAL	32		52		34		42		17			17		
MOYENNE	4	0,2	6,5	0,33	4,25	0,21	5,25	0,26	2,125	0,11	0,5	2,125	0,11	0,5
POURCENTAGE														

Figure 44 : Tableau récapitulatif des effectifs, moyennes et pourcentages par catégorie

Sujet (numéro attribué par ordre de passation)	Nombre d'items corrects en fonction de l'indice morphosyntaxique													
	Clitique accusatif		Nombre		Genre		Forme négative		Proposition relative		Préposition		TOTAL	
	AVATAR	AUDIO	AVATAR	AUDIO	AVATAR	AUDIO	AVATAR	AUDIO	AVATAR	AUDIO	AVATAR	AUDIO	AVATAR	AUDIO
EFFECTIF TOTAL	4	8	20	20	20	19	27	22	8	5	10	9	89	83
DIFFERENCE	-4		0		1		5		3		1		6	

Figure 45 : Tableau récapitulatif du nombre de réponses correctes par modalité pour chaque indice morphosyntaxique

Sujet (numéro attribué par ordre de passation)	Situation préférée				TOTAL
	avatar	audio	humain	pas de préférence	
1			1	1	2
2	1				1
3		1			1
4					
5	1	1	1		3
6			1		1
7					
8	1				1
9	1				1
10	1	1	1		3
Effectif total	5	3	4	1	13
Pourcentage	0,38	0,23	0,31	0,08	1

Figure 46 : Tableau récapitulatif des modalités préférées par les enfants

Annexe 4 : Graphiques représentant le nombre de réponses correctes et incorrectes par modalité pour chaque phrase du corpus.

Indice morphologique : clitique accusatif

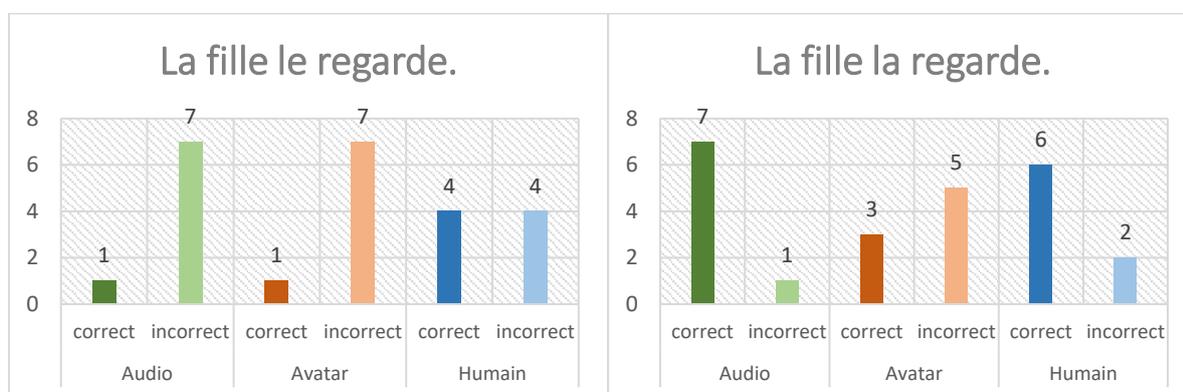


Figure 47 : Diagrammes récapitulatifs pour les clitiques accusatifs

Indices morphologiques de nombre :

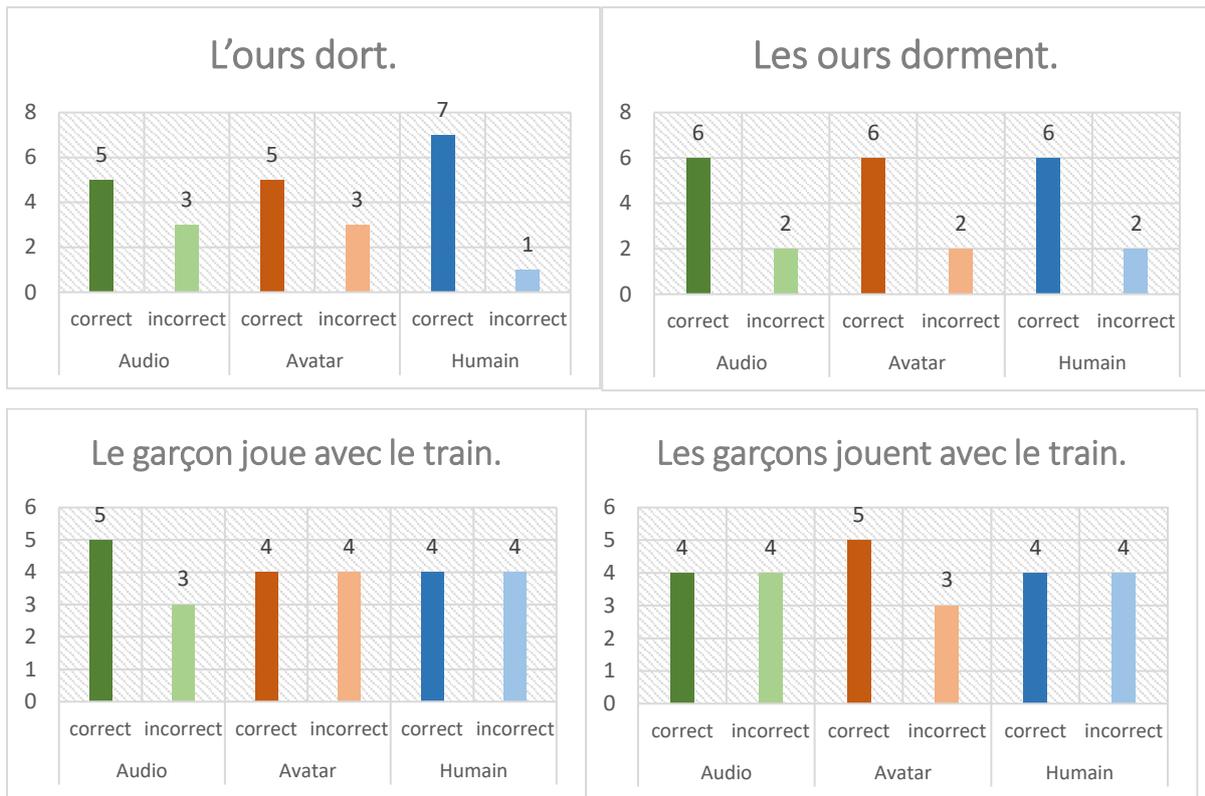


Figure 48 : Diagrammes récapitulatifs pour les indices de nombre

Indices morphologiques de genre :

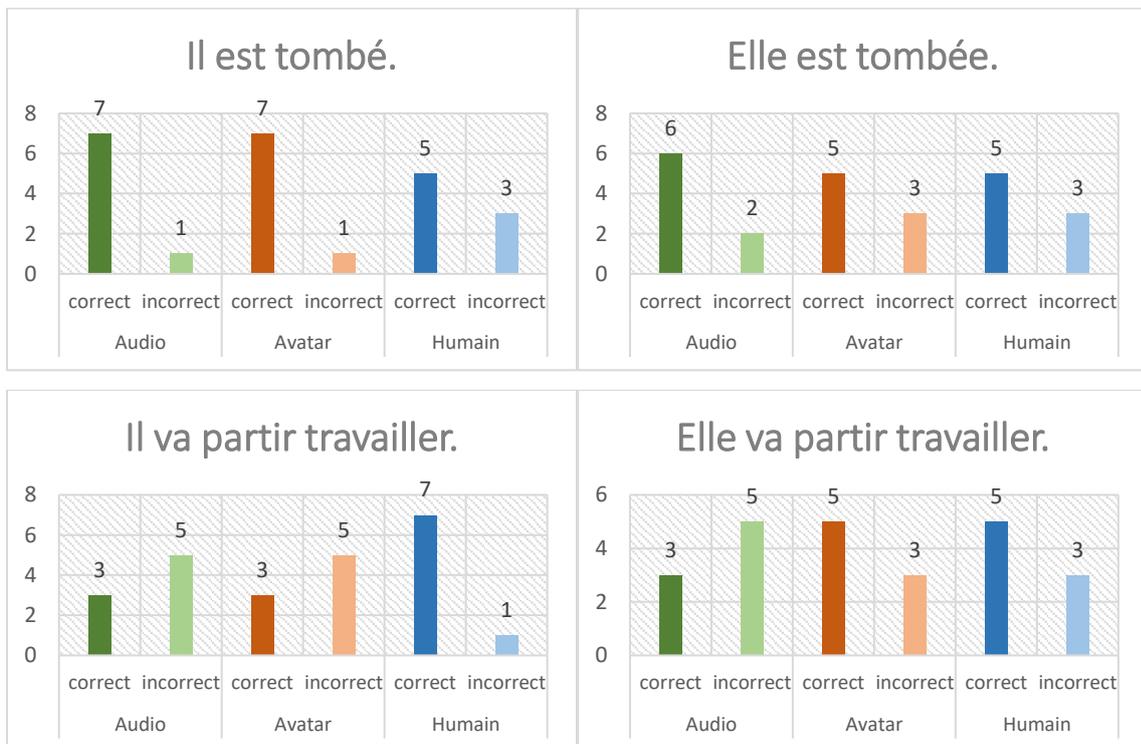


Figure 49 : Diagrammes récapitulatifs pour les indices de genre

Indices syntaxiques : forme négative

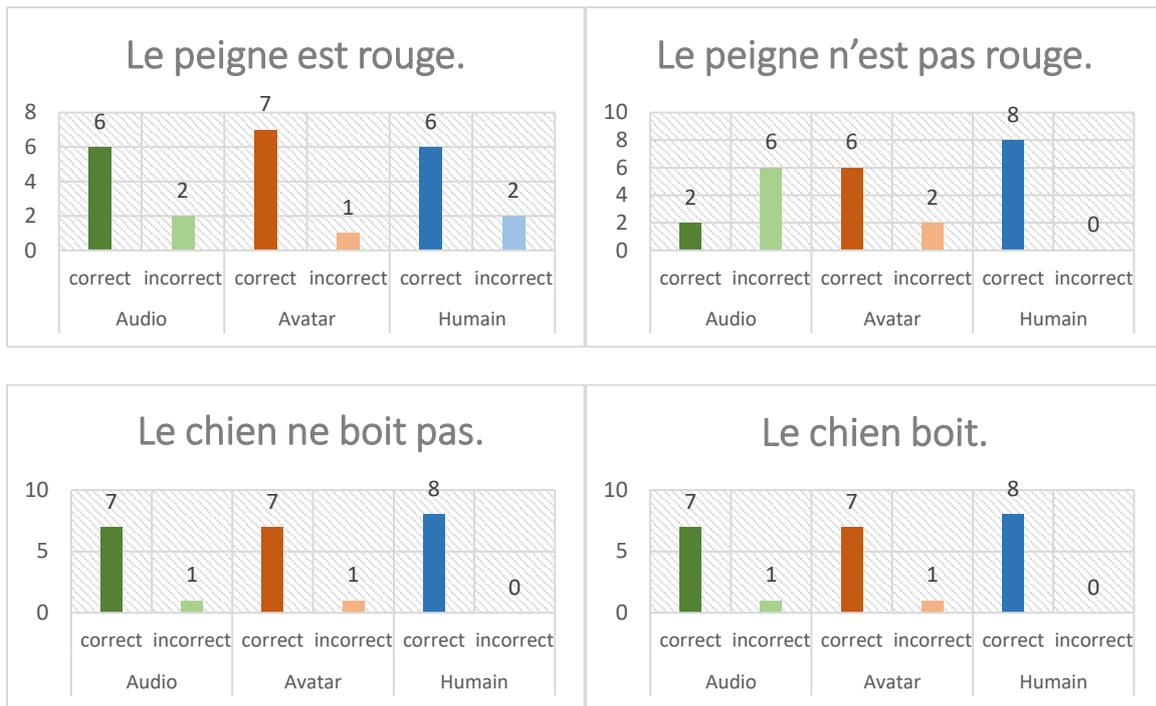


Figure 50 : Diagrammes récapitulatifs pour les formes négatives

Indices syntaxiques : propositions relatives

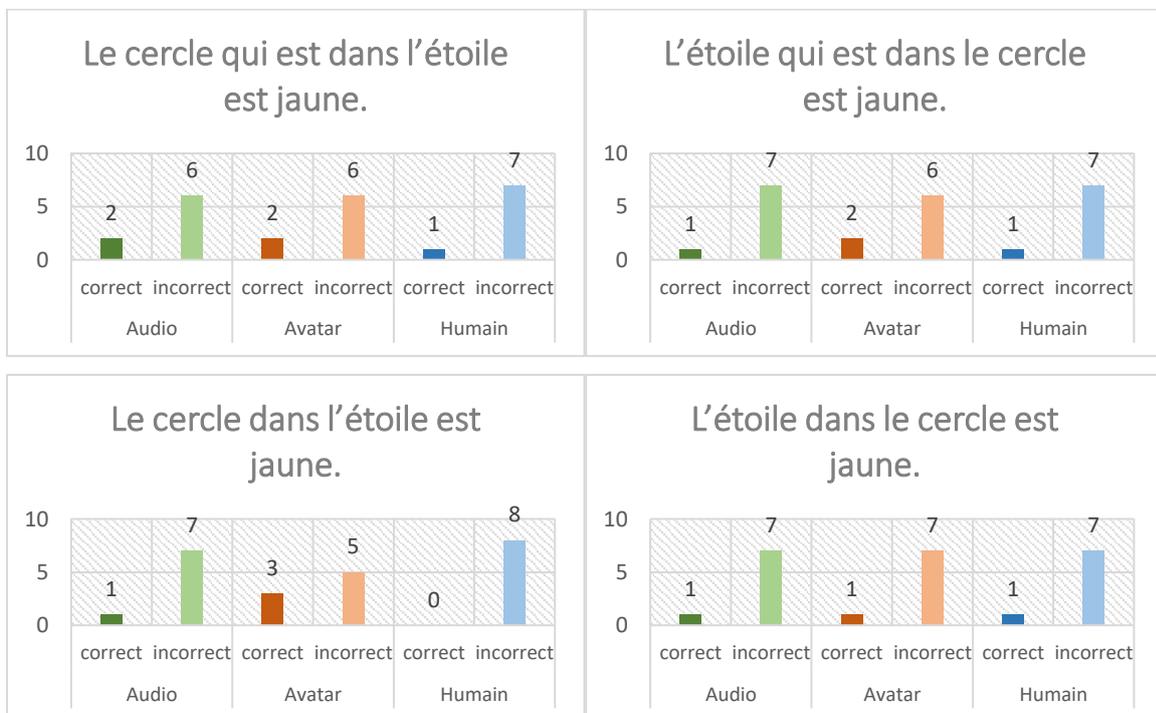


Figure 51 : Diagrammes récapitulatifs pour les propositions relatives

Indices syntaxiques : prépositions

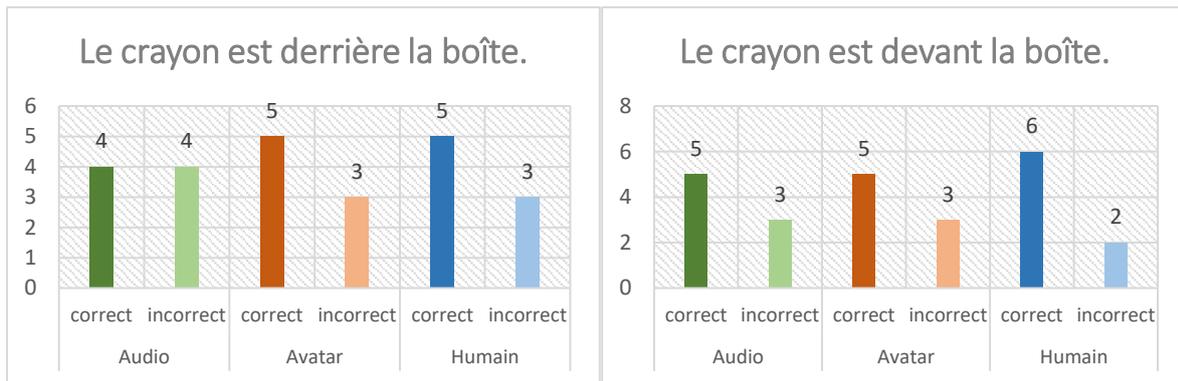


Figure 52 : Diagrammes récapitulatifs pour les prépositions

DUSSERE Léa

Compréhension morphosyntaxique dans une situation de communication audiovisuelle : incidence du choix d'un avatar parlant auprès d'enfants présentant une surdité

Résumé : La technologie Dynalips, élaborée par l'équipe MULTISPEECH d'Inria-Loria Grand Est, a permis la création d'avatars parlants reproduisant les mouvements labiaux de la parole. Cette technologie pourrait permettre le développement d'outils de communication à destination de personnes atteintes de troubles du langage oral. C'est dans ce but que cette étude est née. Dans le cadre de mon mémoire de fin d'étude, j'ai étudié l'intelligibilité de l'avatar, en évaluant la compréhension d'enfants atteints de surdité et en la comparant avec une modalité audiovisuelle optimale : un locuteur humain et une modalité strictement auditive. Les résultats ont montré que l'avatar apportait des indices visuels supplémentaires à la modalité audio, permettant une compréhension de phrases plus rapide et meilleure. Nous avons aussi observé que la modalité « avatar » avait été la plus appréciée par les enfants. Le peu de sujets rencontrés en raison de la crise sanitaire de 2020 n'ont néanmoins pas permis de rassembler des résultats statistiquement significatifs mais encourage à poursuivre l'étude.

Mots-clés : avatar, surdité, compréhension, morphologie, syntaxe

Morphosyntactic understanding in an audiovisual communication situation: incidence of the choice of an avatar speaking with children with deafness

Abstract: The Dynalips technology, developed by the MU LTISPEECH team at Inria-Loria Grand Est, permit the creation of a speaking avatars reproducing the labial movements of speech, called visemes. This technology could allow the development of communication tools for people with oral language disorders. It is for this purpose that this study was born. As part of my final dissertation, I studied the intelligibility of the avatar, assessing the understanding of children with deafness and comparing it with an optimal audio-visual modality: a human speaker and a strictly auditory modality. The results showed that the avatar provided additional visual clues to the audio modality, allowing for faster and better understanding of sentences. We also observed that the "avatar" modality was most appreciated by the children. The few subjects encountered due to the health crisis of 2020 did not permit to collect statistically significant results but encourages further study.

Keywords: avatar, deafness, comprehension, morphology, syntax