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-theses-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

Nancy-Université
Université Henri Poincaré

Département de formation doctorale en informatique **École doctorale IAEM Lorraine**
UFR STMIA

Aide en ligne adaptative et assistants conversationnels animés : mise en œuvre et évaluation ergonomique

THÈSE

présentée et soutenue publiquement le 09 Octobre 2007

pour l'obtention du

Doctorat de l'université Henri Poincaré – Nancy 1

(spécialité informatique)

par

Jérôme Simonin

Composition du jury

Président :	Mme Claire Gardent, Directrice de Recherche, LORIA-CNRS
Rapporteurs :	Mme Catherine Pélachaud, Professeur, Université de Paris 8 M. Jean-Paul Sansonnet, Directeur de Recherche, LIMSI-CNRS
Examinateurs :	M. Jean Vanderdonckt, Professeur, Université Catholique de Louvain-la-Neuve M. René Amalberti, Professeur et Directeur de Recherche, IMASSA M. Joseph Mariani, Directeur de Recherche, LIMSI-CNRS Mme Noëlle Carbonell, Professeur, Université Henri Poincaré Nancy 1

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Loria

Mis en page avec la classe thloria.

Vous ne pouvez pas résoudre un problème avec la même sorte de pensée que celle qui a créé le problème. [A. Einstein]

A mes parents et à Aurél.

Remerciements

Je tiens à exprimer mes plus sincères remerciements à Noëlle Carbonell pour m'avoir accueilli au sein de l'équipe MErLIn et m'avoir conduit jusqu'au terme de ce travail. Durant ces quatre années de travail commun, j'ai pu apprécié ses qualités scientifiques mais surtout ses qualités humaines qui m'ont apporté conseils et soutien tout au long de cette aventure.

J'aimerais remercier également tous les membres ou anciens membres de MErLIn avec qui j'ai pris grand plaisir à travailler, Antonio Capobianco, Suzanne Kieffer, Daniel Gepner, Maryus Hategan, Olivier Christmann, Luo Zaholu, Matjaz Divjak. Sans oublier les stagiaires, avec qui nous avons conçus des expérimentations, Sylvain Houillons, Mickael Hauwy, Baptiste Payan, Pierre Freund, Laurent Prost, Sébastien Weber et Alexandre Roy. Un merci tout particulier à mes collègues de bureau Maryus et Olivier qui sont devenus de véritables amis !

Je tiens également à remercier chaleureusement les membres du jury, Catherine Pélachaud, Jean-Paul Sansonnet, Claire Gardent, Jean Vanderdonckt, René Amalberti, Joseph Mariani, Didier Bazalgette, pour m'avoir fait l'honneur d'accepter de participer à l'évaluation de ce travail, merci beaucoup pour votre attention et vos commentaires.

Merci aussi aux personnes, que j'ai rencontrées durant cette thèse et qui m'ont beaucoup appris. Merci à Yannick Parmentier, Sébastien Hinderer, Eric Kow, Jospeh Roumier, Jospeh Razik, Emmanuel Didiot... Pardon à tous ceux que j'oublie.

Pour la technique, une dédicace spéciale à notre commercial d'immersion Stéphane qui m'a fait découvrir l'eye-tracking et un merci à France Télécom de nous avoir fourni l'ACA nommé Eloona que j'ai utilisé dans ma thèse. Ainsi, merci à l'équipe de FT R & D, Gérard Poulain, Gaspard Breton, David Caillère et Marc Mersiol.

Je tiens aussi à remercier la DGA de m'avoir financé et j'espère que mes travaux leur seront utiles.

Merci aux membres de l'UFR STMIA pour m'avoir permis de dispenser des enseignements et d'avoir des échanges enrichissants, et tout spécialement, Jean-Pierre Jacquot, Martine Cadot, Monique Grandbastien, Frédéric Suter, Joseph Di-Martino, Loutfi Soufi, et Brigitte Wrobel-Dautcourt.

Enfin, j'aimerais remercier ma famille qui m'a encouragé. En particulier, un grand merci à Aurél, à qui je dédie ce document.

Table des matières

Chapitre 1 Présentation générale **1**

Chapitre 2 Motivations **5**

2.1 Agents Conversationnels Animés 5

2.1.1 Apparence et fonctionnalités des ACAs 5

2.1.2 Évaluation de l'apport des ACAs à l'interaction Homme-Machine 11

2.2 Interfaces adaptatives 17

2.2.1 Définitions et classifications 19

2.2.2 Raffinement de la classification générale 21

2.2.3 L'existant 24

2.3 Sujet de recherche 28

Chapitre 3 Étude préliminaire **31**

3.1 Motivations, Objectifs - Systèmes d'aide 32

3.1.1 Contexte 32

3.1.2 Aide en ligne : Définition 33

3.1.3 Comportements exploratoires 33

3.1.4 Conclusion 34

3.2 Méthode - Protocole expérimental 34

3.2.1 Présentation globale du protocole expérimental 34

3.2.2 Choix de conception, mise en œuvre et déroulement 35

3.3 Analyses - Résultats 42

3.3.1 Critères 43

3.3.2 Analyse des données expérimentales 43
3.3.3 Analyse des données post-expérimentales 47
3.4 Résumé 49
3.5 Impacts sur le programme de recherche 50

Chapitre 4 Démarche, méthode et développement logiciel **53**

4.1 Tâche expérimentale 53
4.2 Démarche - Méthode 54
4.2.1 Conception du système d'aide 54
4.2.2 Forme de multimodalité retenue 57
4.2.3 Trace des interactions 57
4.2.4 Déroulement d'une passation 58
4.3 Développement logiciel 58
4.3.1 Plate-forme logicielle :
Système d'aide, Magicien d'Oz et Enregistrement des interactions 59
4.3.2 Rejeu et annotation des données sauvegardées par la plate-forme 64
4.4 Conclusion 68

Chapitre 5 ACAs : Eloona en situation d'aide **69**

5.1 Contexte - Généralités 69
5.1.1 Rappel des objectifs 69
5.1.2 Choix de l'ACA 70
5.2 Protocole expérimental - Définition 70
5.2.1 Cadre général 70
5.2.2 Plan expérimental - Aperçu global 71
5.2.3 Hypothèses de travail - Mesures 73
5.2.4 Sujets 75
5.3 Protocole expérimental - Réalisation 76
5.3.1 Animation d'Eloona 76
5.3.2 Conception - Activation des messages d'aide 82

5.4 Analyse des données 86
5.4.1 Profils des sujets - Questionnaire initial 86
5.4.2 Préférences des sujets : Présence ou absence d'un agent conversationnel . 89
5.4.3 Apport potentiel d'un ACA à l'interaction avec une aide en ligne 93
5.4.4 Facteurs à l'origine des préférences - Hypothèses 101
5.5 Conclusion et discussion 127

Chapitre 6 Interfaces adaptatives : adaptation aux connaissances 129

6.1 Contexte - Généralités 130
6.1.1 Rappel des objectifs 130
6.1.2 Choix des modalités d'expression des informations d'aide 130
6.2 Protocole expérimental - Définition 131
6.2.1 Plan expérimental - Aperçu global 131
6.2.2 Hypothèses de travail - Mesures 131
6.2.3 Sujets 132
6.3 Protocole expérimental - Réalisation 132
6.3.1 Messages d'aide - Interface et conception 133
6.3.2 Stratégie d'adaptation : Rôle du compère 135
6.4 Analyse des données 137
6.4.1 Profils des sujets - Questionnaire initial 137
6.4.2 Expression des préférences : aide adaptative versus aides classiques 137
6.4.3 Avantages et limites du système d'aide expérimenté 143
6.5 Conclusion 151

Chapitre 7 Conclusions et perspectives 153

Annexe A Étude préliminaire - compléments 161

Annexe B Études 2 et 3 : Détails du déroulement d'une passation 165

Annexe C ACA de FT R & D : Eloona 213

Bibliographie **225**

Chapitre 1

Présentation générale

Ce travail de thèse a été réalisé au LORIA[1] (Laboratoire Lorrain de recherche en Informatique et ses Applications), au sein de l'équipe MerLIn, et financé par la DGA (Délégation Générale pour l'Armement). Le projet MerLIn s'articule actuellement autour de plusieurs axes de recherche, et s'intéresse plus particulièrement à l'ergonomie du Web, à l'interaction multimodale, à l'aide en ligne, dans une démarche qui allie méthodes ergonomiques et techniques d'intelligence artificielle. Le domaine d'application choisi pour ce travail est l'aide en ligne à l'utilisation des logiciels grand public. La thématique générale de la thèse consiste à mettre en œuvre et à évaluer l'apport d'*Agents Conversationnels Animés* (ACAs) et d'*Interfaces dites "Adaptatives"* à l'aide en ligne en étudiant les comportements et réactions d'utilisateurs.

Depuis l'apparition des premiers ordinateurs, l'interface présente le système à l'utilisateur. Son rôle, tel qu'il est défini dans [Nanard, 1990] est "d'établir la communication entre deux mondes, celui de l'homme et celui de la machine, et d'assurer la correspondance des concepts propres à chaque partenaire, en s'appuyant sur un ensemble de représentations qui doivent avoir un double aspect : formel pour la machine et perceptible pour l'homme". Il résulte de cette définition que l'interface est le lieu où s'opère la communication entre l'utilisateur et le système informatique, sachant que le premier utilise le second comme aide à son activité. Cela dit, compte tenu des récents développements en matière de technologies de l'information et de la communication, les utilisateurs de logiciels (novices ou experts) sont désormais confrontés à des logiciels qui offrent de nombreuses fonctionnalités et permettent des tâches extrêmement variées. La contrepartie de cette richesse est la difficulté d'apprentissage pour un utilisateur novice, et le risque de n'utiliser que partiellement les possibilités du logiciel (novice comme expert).

En outre, on observe que les interfaces n'ont pas toujours évolué en parallèle avec les logiciels, et qu'un fossé entre l'utilisateur et le logiciel [Grudin, 1993] s'est créé. Ceci s'explique historiquement par le fait que l'intérêt pour l'utilisateur est venu bien après celui pour le traitement des problèmes de calcul. Désormais de nouvelles techniques ont été testées afin d'aider les utilisateurs à faire face à ces nouvelles fonctionnalités et à la surcharge d'informations. Ainsi, adapter les interfaces homme-machine logicielles au profil de l'utilisateur courant et au contexte dans lequel se situe son interaction devient une nécessité impérieuse aujourd'hui où le développement

[1] http ://www.loria.fr

de la Société de l'Information accroît considérablement la diversité des utilisateurs et celle des situations d'utilisation. En effet, si l'objectif est de promouvoir une société où l'informatique accompagne dans leurs activités quotidiennes, mobiles ou non, tous les citoyens, quelles que soient leurs capacités cognitives et physiques, leurs connaissances et compétences, leurs motivations et intérêts, il est impératif d'accroître la flexibilité des interfaces utilisateur actuelles, c'est-à-dire leur capacité à évoluer au cours de l'interaction en fonction du profil de l'utilisateur courant et des contraintes qu'induit sur ses capacités perceptives, motrices et/ou cognitives l'environnement dans lequel se déroule l'interaction.

Ainsi le monde de l'information dans lequel nous vivons actuellement, doit apporter à l'utilisateur l'information dont il a besoin quand il en a besoin et de manière appropriée. Ceci représente un véritable challenge et nécessite diverses adaptations des interfaces actuelles.

L'objectif de la thèse est donc d'évaluer dans un contexte d'assistance en ligne (utilisation d'un logiciel grand public inconnu) la réaction des utilisateurs face à des interfaces innovantes :

- **qui intègrent des personnages virtuels interactifs, les Agents Conversationnels Animés (ACAs)** ;
- **ou qui évoluent sans prévenir, de manière autonome, les interfaces dites "Adaptatives"**.

Bien que de multiples prototypes d'ACAs aient été développés dans les laboratoires de recherche et les centres de R&D, rares sont les applications destinées au grand public dont l'interface utilisateur graphique, fondée sur le paradigme toujours prévalent de la manipulation directe (Shneiderman, 1983), intègre un ACA. Cette étude est motivée, d'une part, par la volonté de déterminer la contribution potentielle de la présence d'un ACA à l'efficacité et à la convivialité de l'interaction homme-machine dans des situations d'assistance. D'autre part, la représentation du système par des ACAs de forme humanoïde laisse supposer que le système possède des capacités d'intelligence humaine. Or une des formes de cette intelligence est l'adaptation au contexte / environnement et, en situation de communication, à l'utilisateur tel qu'on le perçoit. Les ACAs peuvent-ils être vus, entre autres, comme une forme de notification implicite de l'adaptation du système à l'intention de l'utilisateur ?

Concernant l'adaptation des interfaces, on observe tout d'abord que la raison de l'adaptation peut être de nature diverses et varie suivant le contexte d'utilisation. Deux grandes catégories d'interfaces existent : les interfaces dites "Adaptables" que l'utilisateur personnalise selon ses préférences (Netscape peut être configuré suivant les préférences de l'utilisateur) et les interfaces dites "Adaptatives" qui évoluent de manière autonome en fonction, par exemple, des connaissances de l'utilisateur (l'assistant Microsoft propose de l'aide suivant les difficultés rencontrées par l'utilisateur). Nous choisissons d'étudier les interfaces adaptatives avec une approche centrée utilisateur. Toutefois, ce thème de recherche appartient à un domaine difficile en raison de sa jeunesse, de la multiplicité des problèmes à résoudre et de la diversité des capacités physiques et cognitives des utilisateurs. On note cependant l'existence de nombreux prototypes d'interfaces adaptatives et de nombreux travaux sur l'élaboration de modèles dynamiques de l'utilisateur courant dans des domaines d'application variés : navigation sur le Web, systèmes de recommandations, environnements d'apprentissage et d'enseignement. Un constat majeur (basé sur la

taxonomie de [Jameson, 2003]) est qu'il n'existe pas ou peu d'études ergonomiques de l'apport de ces techniques à l'efficacité et à l'utilisabilité de l'interaction Homme-Machine, alors que les systèmes adaptatifs vont à l'encontre d'un des principes majeurs de l'ergonomie des logiciels classique : les réactions du système doivent être prévisibles [Schneiderman, 2000]. Il faut évaluer la réaction des utilisateurs face à des interfaces adaptatives qui évoluent sans prévenir et proposer des solutions qui permettent de concilier les principes traditionnels de l'ergonomie des logiciels et l'innovation qu'apporte l'adaptivité. L'évolution de l'interface en dehors du contrôle de l'utilisateur est-elle utile et, si oui, les interfaces adaptatives sont-elles utilisables au sens de [Nielsen, 1993] ?

Le programme de recherche expérimental réalisé comporte trois études ergonomiques. Le domaine d'application choisi est l'aide en ligne à l'utilisation de logiciels grand public (situation d'apprentissage).
Dans chaque étude, la tâche expérimentale proposée aux sujets était de concevoir de petites animations sous Flash, logiciel qui leur était inconnu.

Le travail se présente comme suit. Le chapitre 2 décrit les motivations des études portant sur les ACAs ainsi que sur les interfaces adaptatives. Nous décrivons dans un premier temps les travaux de recherche portant sur l'évaluation de l'apport des ACAs à la qualité ergonomique de l'Interaction Homme-Machine (IHM). Cette description comporte une revue des ACAs existants ainsi qu'une présentation des principales taxonomies élaborées pour caractériser et mesurer leur contribution potentielle. Dans un second temps, une revue des travaux de recherche récents consacrés à la mise en œuvre en IHM du concept d'adaptation dynamique à l'utilisateur est présentée. Elle s'accompagne d'une classification des recherches sur les systèmes adaptatifs tenant compte de la nature des capacités et comportements de l'utilisateur.

Le chapitre 3 décrit l'étude préliminaire que nous avons menée et qui porte sur les apports de la multimodalité à la mise en œuvre de consignes dans des documents procéduraux. Lors de l'expérimentation, 8 sujets devaient réaliser une animation sous Flash aidés de messages d'aide procéduraux. Ainsi, deux types de conception multimodale des messages d'aide en ligne ont été testés :

- version multimodale intégrant du Texte (informations textuelles sur l'action à réaliser) et du Graphique (copie d'écran associée au message textuel) ;
- et une version multimodale intégrant de l'Oral (message vocal enregistré que l'utilisateur peut écouter via un bouton) et du Graphique (copie d'écran associée au message vocal).

Les constats et résultats de cette étude préliminaire ont permis de construire la démarche adoptée pour réaliser les deux études ultérieures, ainsi que les choix méthodologiques sur lesquels se fonde leur conception. Le chapitre 4 décrit la mise en œuvre des choix méthodologiques communs à ces deux études, ainsi que le développement d'une plate-forme logicielle qui implémente la technique du magicien d'Oz afin de faciliter la réalisation des interfaces utilisateur expérimentées. Ici, l'emploi du magicien d'Oz permet de simuler l'adaptation pour les interfaces, en effet cela permet de simuler l'intelligence de l'ACA ou d'adapter l'interface en fonction des connaissances de l'utilisateur.

Ensuite, nous décrivons dans les chapitres 5 et 6 les deux études qui ont été menées à la suite de l'étude préliminaire. Pour étudier l'apport des ACAs à l'interaction avec l'aide en ligne (voir le chapitre 5), nous avons utilisé dans le cadre d'un contrat de recherche externalisée avec France Télécom, l'Agent Conversationnel Animé "Eloona", développé et mis à notre disposition par FT R & D. 22 étudiants, informaticiens ou non, ont interagi avec cet agent pour se familiariser avec un logiciel grand public qu'ils découvraient. Les traces de leur interaction avec Eloona, l'enregistrement de leurs mouvements oculaires, un questionnaire post-expérimentation, ainsi qu'un entretien de debriefing sont les données qui nous ont permis d'évaluer les réactions d'utilisateurs potentiels à la présence d'un tel agent. La dernière étude, qui porte sur les interfaces adaptatives, utilise le même protocole expérimental ; toutefois, les mouvements oculaires des participants n'ont pas été enregistrés car la technique d'interaction à évaluer ne modifie pas les aspects visuels de l'interaction. Les questionnaires ainsi que les entretiens de débriefing sont centrés sur les réactions des utilisateurs face à des interfaces qui évoluent de manière autonome. Cette étude est décrite dans le chapitre 6.

Le dernier chapitre présente d'une part, les conclusions relatives aux deux études menées et d'autre part les perspectives de recherche complémentaires à effectuer sur les corpus de données recueillis qui n'ont pas peu être exploités totalement dans ce travail.

Chapitre 2

Motivations

Ce chapitre présente le cadre dans lequel s'inscrit notre recherche. Plus précisément, étant donné que nous désirons améliorer l'interaction homme-machine à l'aide d'ACAs et de techniques d'adaptation dynamique au profil de l'utilisateur, nous allons présenter les principales avancées réalisées dans ces deux directions de recherche. Enfin, nous préciserons les objectifs ainsi que l'intérêt du sujet de recherche abordé pour conclure ce chapitre.

2.1 Agents Conversationnels Animés

Voulant évaluer l'apport des Agents Conversationnels Animés à la qualité ergonomique de l'Interaction Homme-machine, l'état de l'art suivant a été réalisé. Il se structure en deux axes :

- D'abord, une revue sommaire des ACAs existants, classés en fonction de leur apparence, de leurs fonctionnalités et des domaines d'application pour lesquels ils ont été conçus. Ces domaines sont également ceux qui ont été choisis pour étudier l'apport de certains prototypes à l'interaction.
- Puis, une présentation des principales taxonomies élaborées pour caractériser et mesurer leur contribution effective à l'interaction.

2.1.1 Apparence et fonctionnalités des ACAs

Malgré les nombreux prototypes d'ACAs que l'on recense, peu sont destinés à être intégrés dans les applications grand public. Hormis le "Compagnon" introduit en 1998 dans les applications Windows pour personnifier l'aide en ligne, on peut citer Angela, ACA générique conçu par la société américaine 'As an Angel'. Angela est un visage féminin réaliste, doté d'expressions faciales et de capacités de dialogue oral évoluées. Les ACAs francophones, par exemple Semantia de la société française de même nom, ont des capacités plus limitées : l'ACA de Semantia dialogue par écrit avec l'utilisateur. Cantoche est l'une des rares sociétés internationales capables

de produire des ACAs génériques d'apparence variée et dotés de capacités évoluées de dialogue oral multilingue (dont le français).[2]

Toutefois, on note ces dernières années l'apparition de nombreux agents virtuels animés, non communicatifs mais personnalisables sur le Web ; voir, par exemple, le site Yahoo Messenger où l'on peut construire (en choisissant sa forme de visage, couleur de cheveux, etc) selon ses désirs un agent qui sera présent à chaque consultation de mails ; voir également, dans le domaine des consoles de jeu, la Wii de Nintendo où chaque joueur peut s'identifier à la console par le biais d'un personnage qu'il aura au préalable, comme sur le site Yahoo, personnalisé à l'aide d'options encore plus variées. L'agent permet ainsi au joueur de s'identifier à chaque utilisation de la console, mais aussi d'être représenté lors d'une session multijoueur sur Internet.

On constate donc l'émergence d'une utilisation des ACAs par les éditeurs de logiciels et un usage qui commence à se répandre d'agents virtuels non communicatifs. La prudence des éditeurs de logiciels peut s'expliquer par les résultats mitigés des rares évaluations ergonomiques publiées jusqu'à présent et par les conclusions issues de l'observation de l'usage effectif des ACAs intégrés à des applications largement diffusées. En effet, si les jugements subjectifs des participants aux études expérimentales ou empiriques réalisées dès la fin des années 90 sont généralement très positifs, la contribution des ACAs à l'efficacité de l'interaction dans les situations choisies fait l'objet de controverses. Quant au Compagnon Microsoft, sa présence irrite rapidement la plupart des utilisateurs qui très rapidement le font disparaître définitivement. Concernant les jugements subjectifs, la contradiction entre les résultats des études ergonomiques et ceux fournis par l'observation des usages n'est qu'apparente ; elle tient très certainement au fait que les travaux de recherche portent sur une période d'utilisation très brève par rapport à la durée d'utilisation que permet l'observation des usages.

Nous présentons dans la suite de cette section une classification des ACAs en fonction de leur apparence, de leurs compétences et des domaines d'application dans lesquels on envisage de les faire intervenir ou dans lesquels ils ont été expérimentés. La classification est illustrée d'exemples de prototypes réalisés dans les laboratoires de recherche. Sont exclus de cet aperçu les produits commercialisés.

Apparence

Généralement les ACAs sont des représentations graphiques 2D ou 3D animées. Cependant, des tentatives ont été effectuées pour doter des robots de capacités conversationnelles. Par exemple, Merl, robot réalisé au Centre de recherche de Mitsubishi, se présente sous la forme d'un pingouin doté de capacités d'expression multimodale et de dialogue avec l'utilisateur. Merl dispose d'un modèle évolué de dialogue coopératif en langue naturelle, Collagen [Lesh *et al.*, 2004] ; en outre, il est capable de combiner expression orale et geste de désignation pour communiquer avec l'utilisateur. Son rôle principal est l'accueil de visiteurs (musées, laboratoires) ou de clients (grands magasins) ; voir par exemple [Sidner et Lee, 2005].

[2] On trouvera des informations complémentaires sur ces sociétés aux URL suivants (cités dans l'ordre où ces sociétés sont évoquées dans le texte) : http ://www.asanangel.com, semantia.com, cantoche.com/.

Cependant, la plupart des ACAs sont représentés par de simples graphiques 2D ou 3D animés. On peut distinguer plusieurs familles d'ACAs graphiques en fonction des objectifs des concepteurs :

- **Donner à l'interaction une dimension affective et coopérative**, en reproduisant les situations de communication humaine face-à-face, donc doter l'ACA d'une apparence humaine aussi réaliste que possible. Le plus souvent, ces ACAs sont des représentations 3D de visages humains capables de manifester des émotions variées, telles les six expressions faciales prototypiques proposées par Ekman (1992). L'une des premières tentatives de réalisation d'un visage animé d'expressions faciales réalistes est celle de Pélachaud [Pélachaud *et al.*, 1996]. Le buste et les bras sont parfois inclus dans la représentation en vue de permettre l'utilisation de désignations gestuelles manuelles par l'ACA dans un contexte d'expression multimodale parole + geste, ou celle d'un langage de communication par signes à l'intention des mal-entendants. Par exemple, l'agent pédagogique Steve [Shaw *et al.*, 1999] utilise le geste de désignation pour décrire les différents constituants et le fonctionnement d'un dispositif technique complexe ; quant à l'agent développé et commercialisé par Vcom3D[3], il utilise buste, bras et mains pour transcrire un texte en langage des signes ASL. Les concepteurs d'ACAs font également appel aux gestes des bras et des mains pour faciliter à l'utilisateur l'interprétation des expressions faciales de l'ACA en enrichissant et diversifiant leur excentricité ; voir, par exemple, le tuteur Adele qui assiste les apprenants dans le domaine du diagnostic médical [Rickel et Johnson, 1999], et les travaux présentés dans [Gunes *et al.*, 2004] qui portent sur la modélisation des expressions faciales et des gestes des membres supérieurs exprimant des émotions. Une approche originale [Kopp *et al.*, 2004] consiste à tenter de reproduire les gestes des mains auxquels les interlocuteurs humains font appel, soit pour simuler un élément du monde réel visible (gestes iconiques), soit pour ponctuer et structurer le discours en redondance avec certains effets prosodiques (voir [Rossi, 1993]) qui facilitent la segmentation, par l'auditeur, du continuum sonore en unités syntagmatiques [Pagel *et al.*, 1996] ou la perception de la structure discursive et thématique du message (e.g., l'accent de focalisation/emphase).

- **Accroître l'intelligibilité des messages oraux**, notamment ceux du système, et donner à l'interaction une dimension didactique, en proposant à l'utilisateur des représentations bimodales, visuelles et sonores, des énoncés produits par le système et en lui facilitant l'apprentissage des sons d'une langue. Les visages humains, lorsqu'ils sont dotés de capacités d'expression/communication orale, peuvent également être utilisés pour faciliter la compréhension de messages oraux par les mal-entendants. En effet, il est possible actuellement de reproduire avec précision les mouvements des articulateurs des sons de parole et de les synchroniser avec l'énoncé du message. Ces mouvements, en particulier ceux des lèvres et des maxillaires (que l'on génère à partir d'un modèle musculaire) facilitent aux mal-entendants la compréhension des messages oraux du système en leur offrant une réalisation bimodale, sonore et visuelle, des messages. Une telle représentation leur permet de compenser leur déficit auditif en faisant appel à leur perception visuelle ; voir, par exemple l'utilisation d'une tête parlante pour faciliter aux mal-entendants les échanges téléphoniques [Beskow *et al.*, 2004]. Plus généralement, la possibilité, pour l'auditeur, de voir le visage du locuteur

[3] http ://www.vcom3d.com/

[Massaro et Stork, 1998]). Les têtes parlantes peuvent également être utilisées pour l'apprentissage de la production des sons d'une langue par les sourds et les mal-entendants, grâce aux vues de face animées du visage qui montrent les mouvements des lèvres et des maxillaires pendant l'articulation des sons d'un énoncé et la présentation de coupes sagittales animées du conduit vocal montrant l'activité des principaux articulateurs, langue et mâchoires notamment. Massaro est l'un des pionniers dans ce domaine [Massaro *et al.*, 1998].

– **Donner à l'interaction une dimension ludique ou attractive**. Lorsque la dimension affective de l'interaction n'est pas un objectif prioritaire, les ACAs peuvent être non seulement des personnages humains mais aussi des animaux ou des objets animés. Le trombone (ou paper clip), l'une des nombreuses incarnations du Compagnon Microsoft, illustre cette dernière catégorie. Dans ce contexte, les ACAs sont souvent représentés de pied en cap ; le trombone de Microsoft dépourvu de corps et de membres fait figure d'exception. Ils ont donc la possibilité de se déplacer au sein de l'univers virtuel affiché à l'écran, ce qui permet à l'utilisateur de s'identifier à eux, lors de visites virtuelles de musées par exemple. Ils peuvent également désigner gestuellement des objets affichés au cours de la présentation d'informations multimédias (dans le contexte d'applications variées), ou lors de la description de produits (voir les applications de téléachat et les "recommending systems"). Ils peuvent montrer, écrire et dessiner dans le cadre d'applications pédagogiques où ils jouent un rôle de tuteur. Le réalisme anthropomorphique n'étant plus nécessaire, les ACAs de cette famille sont souvent représentés par des objets graphiques 2D animés, caricatures d'êtres humains ou d'animaux, ou produits de l'imagination du concepteur ; leur dimension ludique et attractive est analogue à celle des personnages de bandes dessinées ou de dessins animés. Néanmoins, des facultés humaines d'expression/communication verbale et non verbale sont un atout important, voire indispensable, pour assurer la dimension ludique de l'interaction. Parmi les premiers prototypes issus des recherches menées dans ce domaine, on peut citer la PPP Persona présentée dans [André *et al.*, 1998]. Cet ACA, représenté de pied en cap, peut prendre différentes apparences physiques (homme ou femme) ; il tient à la main une baguette qu'il utilise pour désigner des objets affichés à l'écran. Il a été expérimenté pour la présentation d'informations multimédias (e.g., description d'un dispositif technique, présentation de produits commerciaux et assistance à la navigation de consommateurs potentiels) dans le contexte de différentes applications Web de télécommerce. On peut citer également la conseillère en immobilier Rea [Cassell *et al.*, 2000] dont l'apparence physique est réduite à un buste, ou le visage féminin Greta utilisé pour jouer un rôle de médecin [Pélachaud *et al.*, 2002] ou encore, dans le domaine de l'apprentissage, l'agent pédagogique spécialiste de biologie végétale, Herman the Bug, insecte représenté de pied en cap [Lester *et al.*, 1997] et les ACAs à apparence humaine Lea, Marco et Julien, représentés eux-aussi aussi de pied en cap, dont le rôle est de présenter interactivement le mode d'utilisation de dispositifs techniques simples [Buisine *et al.*, 2004]. Enfin, à noter une tendance récente, encore marginale, à multiplier les ACAs dans des situations à vocation ludique ou pédagogique. Par exemple, Agneta et Frida, représentées de pied en cap sous forme de dessins animés, commentent entre elles les pages Web que consulte l'utilisateur

[Höök *et al.*, 1999] ; ce ne sont pas véritablement des ACAs car elles dialoguent uniquement entre elles, mais elles sont capables de "lire" et "comprendre" le contenu des pages Web que l'utilisateur parcourt et donc de lui en faciliter la lecture. On peut mentionner également le recours à des poissons animés pour familiariser des élèves de 7 à 10 ans avec la biologie marine [Darves *et al.*, 2002] ; [Darves et Oviatt, 2004]. Dans un domaine d'application différent, on peut citer le projet d'utilisation de plusieurs ACAs pour présenter des produits dans un hall de démonstration virtuel (ou eShowroom). Cette application potentielle est décrite brièvement dans [Rist *et al.*, 2003] qui présente une plate-forme logicielle susceptible d'assurer l'animation de plusieurs ACAs.

Capacités, rôles et domaines d'utilisation

Capacités

Si les représentations et capacités physiques des ACAs présentent une grande diversité, du visage isolé doté d'expressivité à l'incarnation complète, si les éléments de la réalité dont ils sont issus varient aussi dans de larges proportions, des êtres vivants, personnages humains ou humanoïdes, aux objets (e.g., le trombone Microsoft), les capacités conversationnelles des ACAs développés récemment dans les laboratoires de recherche sont semblables. Ces ACAs sont doués de parole, font preuve d'une bonne maîtrise de l'expression en langue naturelle, y compris concernant l'utilisation de l'anaphore et des déictiques [Lester *et al.*, 1999], et sont capables de conduire des dialogues relativement complexes avec l'utilisateur. Ces dialogues peuvent être considérés comme "intelligents" grâce à la mise en œuvre de techniques de planification évoluées ([André *et al.*, 1999] ; [André *et al.*, 2000] ; [Rist *et al.*, 2003]), et jugés "naturels" grâce à un contrôle flexible des initiatives [Smith et Gordon, 1997] et à la mise en œuvre de la notion d'engagement [Sidner *et al.*, 2004]. Quant à leur expressivité, elle varie essentiellement en fonction du réalisme de leur représentation (3D versus dessin 2D), de leur apparence (visage versus buste ou bien tête, corps et membres) et de leur rôle dans l'interaction (assistant ou présentateur d'informations versus tête parlante).

Rôles

Concernant les rôles et fonctions des ACAs dans l'interaction, ils sont relativement peu diversifiés. On peut regrouper les fonctions des agents dans trois classes principales :

- Présentation interactive d'informations multimédias (voir, par exemple, la PPP Persona dans la section intitulée 'Apparence').
- Assistance à des activités de l'utilisateur plus ou moins complexes, telles que navigation sur le Web, recherche d'informations, prise de décision, apprentissage, notamment procédural (voir les agents pédagogiques et tuteurs).
- Interaction avec des acteurs humains dans des environnement virtuels (e.g., jeux, conception participative).

Domaines d'utilisation

Les domaines d'application des ACAs, encore à venir, sont relativement variés. Les ACAs présentés précédemment ont pour la plupart d'entre eux été intégrés à des prototypes ou des démonstrateurs qui illustrent le vaste champ des applications potentielles proposées par les chercheurs.

Les `environnements d'apprentissage` comptent parmi les premiers domaines applicatifs envisagés par les chercheurs concepteurs d'ACAs. Dès la fin des années 90, des prototypes d'agents pédagogiques ont été développés pour démontrer l'apport des ACAs aux logiciels interactifs d'assistance à l'apprentissage et à l'enseignement. C'est aussi à cette époque que sont apparues les premières classifications portant sur les ACAs susceptibles d'intervenir de façon bénéfique dans les environnements d'apprentissage humain. Celle de [Baylor, 1999] ne comporte que deux classes : d'une part, les tuteurs personnifiés qui interagissent avec les apprenants, d'autre part, les outils ou agents cognitifs classiques des systèmes multi-agents cognitifs. Cette classification a été rapidement enrichie. On distingue actuellement trois profils d'ACAs pédagogiques : les précepteurs/guides, véritables professeurs, les compagnons d'apprentissage qui assistent l'apprenant et les assistants personnels destinés à l'apprenant ou au formateur [Chou *et al.*, 2003]. La classification proposée dans [Payr, 2003] distingue les assistants personnels qui offrent une assistance en ligne à l'utilisation de logiciels, des agents pédagogiques.

L'`aide en ligne` constitue un autre domaine d'application privilégié des ACAs, envisagé également dès la fin des années 90 (voir le Compagnon Microsoft). Ces ACAs doivent être disponibles/accessibles en permanence puisque leur rôle principal est de fournir une assistance à l'utilisation des logiciels grand public. Ils doivent donc être capables de prendre différentes formes d'apparence en fonction de l'espace d'affichage disponible, de façon à adapter leur présence "physique" à la diversité des équipements (PDA, station de travail, téléphone portable, etc.) de l'utilisateur mobile, profil d'utilisateur qui se généralise rapidement ; voir [Berner et Rieger, 2003] pour la mise en œuvre de ce type d'adaptation.

Le `télé-commerce` et, plus généralement, l'assistance à la navigation sur le Web, la présentation ou le commentaire d'informations multimédias, le conseil personnalisé sont des domaines qui ont également motivé la réalisation de nombreux démonstrateurs dans les laboratoires de recherche. En particulier, nombreux sont les prototypes de conseil aux consommateurs (i.e., "recommending systems") dans lesquels ont été intégrés des ACAs. A titre d'exemple, voir dans [Semeraro *et al.*, 2003] l'utilisation d'un ACA, Susanna, en tant que libraire sur un site Web de télé-achat. Voir également [Guinn et Hubal, 2004] qui porte sur l'évaluation de kiosques d'information faisant intervenir un ACA pour présenter l'information. Plus généralement, on peut envisager d'utiliser les ACAs, en tant que "hôtesses d'accueil" ou guides, dans de nombreuses applications : visites virtuelles de musées, sites et services Web, etc.

Les `jeux`, en revanche, bien qu'ils constituent un domaine d'application potentiel majeur, n'ont pas bénéficié des travaux de recherche menés dans les laboratoires de recherche sur les ACAs. Cet état de fait résulte vraisemblablement des objectifs différents qui président, d'une part, à la création des personnages qui interviennent dans les jeux et, d'autre part, à la conception des ACAs. Pour mémoire, les ACAs sont conçus pour satisfaire l'un ou l'autre des deux objectifs suivants (si l'on exclut les têtes parlantes) : donner une dimension affective et/ou un

caractère ludique/attractif à l'interaction avec les logiciels courants. Cependant, la dimension ludique n'est pas une fin en soi pour les concepteurs d'ACAs, elle sert d'autres objectifs, entre autres : faciliter l'apprentissage (pour les ACAs pédagogiques) ou promouvoir un produit (pour les ACAS des sites de télé-commerce), ou encore susciter l'intérêt de l'utilisateur et lui faciliter l'accès à l'information (pour les ACAs présentateurs d'informations). Il n'est donc pas étonnant qu'aucun prototype ou démonstrateur issu des recherches sur les ACAs n'aborde le domaine des jeux, même si certains ACAs, comme Merl, Agneta et Frida, visent à introduire une dimension ludique dans l'interaction. Voir également les applications envisagées pour la plate-forme de développement d'ACAs présentée dans [Eliëns *et al.*, 2002]. Seuls les ACAs destinés aux enfants, tels les poissons parlants de l'application pédagogique présentée dans [Darves *et al.*, 2002], peuvent être rapprochés des personnages, partenaires ou adversaires, avec lesquels interagissent les utilisateurs des jeux actuels sur ordinateur personnel ou console, ou sur Internet ; cependant, les premiers ont un rôle éducatif que n'ont pas les seconds. A noter quelques initiatives de la communauté scientifique en direction de jeux interactifs de création (narration) et de jeux permettant aux enfants comme aux adultes de libérer leurs émotions et d'exprimer leur affectivité. On peut citer le dragon QuiQui (dessin animé) que les enfants utilisateurs contrôlent gestuellement et oralement [Hämäläinen et Höysniemi, 2003] ou la poupée SenToy qui permet à l'utilisateur de projeter ses émotions en la manipulant [Paiva *et al.*, 2003].

La possibilité, grâce à la plate-forme MIAU [Rist *et al.*, 2003], d'animer plusieurs ACAs en parallèle et de leur faire jouer des rôles différents permet d'envisager de nouvelles situations d'interaction et donc de nouveaux domaines d'application encore inexplorés. En particulier, il est possible de donner l'impression à l'utilisateur qu'il est dans un théâtre virtuel peuplé d'ACAs qu'il peut contrôler en tant que metteur en scène, ou avec lesquels il peut interagir et collaborer s'il préfère se donner un rôle d'acteur partenaire. Collecticiels, logiciels de téléconférence, forums et "chats" sur Internet pourraient mettre à profit les possibilités techniques offertes par de telles plate-formes logicielles.

Pour conclure, les domaines potentiels d'utilisation des ACAs ne sont actuellement que partiellement explorés et exploités puisque, d'après [Ruttkay *et al.*, 2002], les rôles principaux qu'il est possible d'assigner aux ACAs, donc les principaux domaines d'utilisation envisageables pour ces agents, se limitent à la présentation d'informations, le renseignement, la vente, le divertissement et l'enseignement.

2.1.2 Évaluation de l'apport des ACAs à l'interaction Homme-Machine

Les premières tentatives d'évaluation ergonomique de l'apport effectif de l'intégration d'ACAs aux interfaces homme-machine en termes d'utilité, d'utilisabilité et de jugements subjectifs d'utilisateurs potentiels, datent de la fin des années 90 ; voir par exemple l'évaluation, par des enfants souffrant d'un déficit auditif, de la tête parlante Baldi [Massaro *et al.*, 1998] ou celle de l'agent pédagogique Herman the Bug [Lester *et al.*, 1997]. L'apparition de taxonomies proposant des critères d'évaluation spécifiques à prendre en compte dans l'évaluation ergonomique d'ACAs est plus tardive ; elle se situe au début des années 2000. Nous présentons les principales d'entre elles avant de décrire les résultats et conclusions des études majeures réalisées jusqu'à présent sur l'apport des ACAs à la qualité ergonomique de l'interaction.

Critères d'évaluation des ACAs

Les principales taxonomies observées mettent l'accent sur des aspects spécifiques des ACAs qu'elles tentent de recenser et pour lesquels il est souvent proposé des méthodes d'évaluations. Une des plus complètes est celle proposée par [Ruttkay *et al.*, 2002]. Ici, l'objectif visé par les auteurs est de définir des critères ainsi que des mesures permettant de comparer l'apparence et les fonctionnalités d'ACAs opérationnels selon quatre points de vue : conception, utilisabilité, utilisation pratique et perception des utilisateurs (ou appréciation subjective des utilisateurs).

Caractéristiques de l'apparence

- Sa représentation
  - Forme d'incarnation : personne (célébrité, instance d'une catégorie/classe, personnage imaginaire), être vivant ou objet ; êtres vivants non humains et objets peuvent avoir un comportement anthromorphique ou non.
  - Réalisme : représentation réaliste, artistique ou caricaturale.
  - Détails physiques : tête ; tête + cou, buste ; représentation de la tête aux pieds ; niveau de détail du visage, des mains.

- Ses moyens de communication : Texte (sous forme de bulles) ou parole. Les caractéristiques de la parole sont importantes à étudier car elle influencent sensiblement la perception de l'ACA par les utilisateurs et influent sur l'intelligibilité des messages oraux :
  - parole enregistrée ou voix de synthèse ;
  - naturel (articulation et prosodie) des énoncés et expressivité (mise en relief prosodique, expression d'une personnalité/individualité, d'émotions) ;
  - spontanéité (reprises, hésitations, exclamations, rire).

- Ses expression faciales
  - Synchronisation des mouvements des lèvres avec les sons prononcés, marques d'emphase, de ponctuation, d'annonce de fin de tour de parole (pour les ACAs communicatifs) ;
  - Expressions faciales dans les moments d'inactivité (e.g., approbation, désapprobation) ;
  - Mouvements des yeux, variation du diamètre de la pupille, mouvements de la tête (e.g., hochements) ;
  - Répétition à l'identique ou non du même jeu d'expressions ;
  - Changement de couleur ; transpiration ;
  - Réalisme des expressions faciales (modélisation des expression d'une personne réelle ou d'une catégorie de personnes, versus expressions "génériques", conventionnelles ou caricaturales).

- Ses mains : Utilisation des mains pour structurer et ponctuer le discours ; geste de désignation ; geste iconique ; expression gestuelle des émotions ; geste ergotique (démonstration d'actions).

- Son corps
  - Postures du corps en accord avec le discours, avec les mouvements des mains ;

- Expression d'états physiques, émotionnels, cognitifs ;
- Capacité et mode de déplacement ;
- Mouvement répétitifs ou variés, typiques d'une personne particulière ou d'un groupe.

Caractéristiques des fonctionnalités

D'après les auteurs de [Ruttkay *et al.*, 2002], les capacités mentales qui entrent dans la composition de la personnalité des ACAs se réduisent à l'aptitude à manifester des émotions et à poursuivre des buts/intentions pré-définis ou évolutifs en fonction du contexte de l'interaction (e.g., les ACAs vendeurs). Concernant les émotions, celles-ci peuvent être déclenchées par des facteurs internes ; c'est le cas, par exemple, du couplage pré-défini, donc statique, de l'état émotionnel de l'ACA avec la nature des informations incluses dans le message qu'il délivre à l'utilisateur). Alternativement, elles peuvent être externes ; c'est le cas des réactions du système générées dynamiquement à la suite d'une intervention de l'utilisateur. Les autres facultés/aptitudes mentales susceptibles de caractériser les ACAs ne sont pas mentionnées ici car elles sont communes à l'ensemble des agents conversationnels et assistants "intelligents". Néanmoins, elles doivent être prises en compte lors de la conception ou de l'évaluation d'un ACA. A titre d'exemple, voir la PPP Persona décrite dans [André *et al.*, 1998].

Contrôle et interactivité

Concernant cet aspect, la taxonomie de [Ruttkay *et al.*, 2002] est sommaire car elle passe sous silence les caractéristiques communes à l'ensemble des agents conversationnels, qu'ils soient incarnés ou non. Trois aspects seulement sont évoqués :

- Le mode de réaction de l'ACA : Il peut être pré-défini et planifié off-line, ou généré dynamiquement au cours de l'interaction. Lorsqu'il est pré-défini, se pose le problème du langage d'expression du contrôle de l'ACA.
- Le contrôle de l'ACA : Il est à la charge, soit de l'application qui l'utilise (e.g., ACA présentateur d'informations multimédias), soit de l'utilisateur (cas des avatars des forums virtuels), soit des deux (ACAs des environnements d'apprentissage). Le contrôle d'un ACA peut être soit statique soit dynamique, selon que le comportement de cet agent est capable d'évoluer (contrôle dynamique) ou non (contrôle statique) sous l'influence de l'état courant du dialogue, des réactions et de l'état émotionnel de l'utilisateur. Le contrôle peut ou non permettre le traitement des anomalies, par exemple les réactions inappropriées ou l'absence de réaction de la part de l'utilisateur, la gestion des tours de parole, etc.
- La sensibilité de l'ACA au contexte : la plupart des ACAs actuels ignorent le contexte dans lequel ils interviennent, qu'il s'agisse de l'utilisateur avec lequel ils interagissent ou de l'environnement souvent changeant dans lequel se situe l'interaction. Cependant, pour que les comportements des ACAs soient effectivement coopératifs et réactifs, il est nécessaire de doter ces agents, d'une part, de capacités d'appréhension et d'interprétation des réactions de l'utilisateur (état courant et évolution des connaissances, intentions, état émotionnel de celui-ci) et, d'autre part, de capacités de perception de l'environnement dans lequel se déroule l'interaction. La mise en œuvre de ces capacités impose de doter l'agent de capteurs et de mécanismes d'interprétation lui permettant de prendre en compte dans son

comportement les données issues de ces capteurs ; voir, concernant ce second point, les travaux de recherche portant sur la mise en œuvre des concepts de "awareness", "ambient intelligence", "smart rooms" dont on trouvera un aperçu dans [Carbonell, 2005]. A noter que les recherches sont moins avancées dans cette direction que dans les précédentes.

Des propriétés spécifiques des ACAs et des objectifs visés par leurs concepteurs tels qu'ils apparaissent aujourd'hui, il est possible de déduire des critères et des méthodes d'évaluation appropriés.

Critères d'évaluation des ACAs

Il est difficile d'établir des critères d'évaluation indépendants du contexte applicatif dans lequel les ACAs sont utilisés, et donc du rôle qu'ils sont censés jouer dans l'interaction : accueil, présentation d'informations, assistance pédagogique, vente, loisirs [Ruttkay *et al.*, 2002]. Cependant, ce sont les objectifs de l'évaluation, c'est-à-dire la dimension de l'apport des ACAs à l'interaction que l'on souhaite mesurer, qui influencent le plus fortement, voire déterminent, les critères utilisés. Deux dimensions principales ont été abordées jusqu'à présent par les taxonomies proposées et les évaluations réalisées. Ce sont, d'une part, les effets induits par l'adjonction d'une apparence "physique" aux agents conversationnels, d'autre part, les effets de la dimension affective et ludique introduite dans la relation entre l'utilisateur et le système, grâce à la présence d'une représentation visuelle à caractère anthropomorphique du système alors que, jusqu'à présent, le système n'a été que rarement représenté dans l'affichage et qu'il a toujours été considéré comme une machine, un outil, un artefact technologique.

Utilisabilité des ACAs

`Facilité d'apprentissage`

Il s'agit de la facilité/difficulté à se familiariser avec l'ACA, à apprendre comment communiquer et interagir avec lui. En d'autres termes, l'interaction avec l'ACA est-elle facile à apprendre ? Est-elle naturelle ? Les limites des capacités cognitives et communicationnelles de l'ACA sont-elles faciles à percevoir ?

`Efficacité`

L'efficacité considérée ici porte sur la réalisation des tâches interactives que permet l'application. Sont pris en compte : la rapidité d'exécution, la correction du résultat (i.e., les taux d'échecs/réussites, les erreurs) et le coût de réalisation (en termes de charge de travail physique, mental ou cognitif). Les aspects considérés sont le temps d'exécution des tâches, la pertinence du dialogue oral avec l'ACA, la qualité de la réalisation des tâches. L'évaluation peut se faire en comparant les performances obtenues par des utilisateurs volontaires lors de l'exécution de tâches identiques, avec ou sans l'assistance de l'ACA à évaluer.

`Erreurs`

L'analyse des erreurs commises par l'utilisateur au cours de l'interaction porte à la fois sur leur nature et leur volume.

Perception des ACAs par les utilisateurs

Concernant la perception que les utilisateurs ont d'un ACA, les critères d'évaluation sont plus nombreux. Certains sont relativement généraux, dans la mesure où ils sont applicables à l'ensemble des agents conversationnels, animés ou non. La portée des autres, en revanche, est plus restreinte et se limite aux agents dotés d'une apparence "physique".

`Qualité de l'assistance et de la coopération`

Ce critère porte principalement sur les jugements subjectifs portés par des utilisateurs potentiels sur l'efficacité de l'assistance fournie par l'ACA en vue d'améliorer la communication de l'utilisateur avec le système, d'accroître l'efficacité de l'activité de l'utilisateur et de lui faciliter la réalisation de ses buts et intentions. Il vise essentiellement à évaluer le degré de contrôle de l'utilisateur sur l'interaction avec l'application, c'est-à-dire à déterminer si celui-ci se sent souvent perdu ou s'il pense maîtriser l'interaction avec l'ACA et l'application. La portée de ce critère est très générale : il peut être utilisé pour évaluer l'interaction avec un système coopératif quelconque.

`Satisfaction de l'utilisateur`

Au cas particulier, ce critère très général sera utilisé pour mesurer les réactions affectives d'utilisateurs à la présence d'un ACA, évaluer la capacité de séduction de ce dernier et l'attrait de sa personnalité. En d'autres termes, l'objectif visé lors de la mise en œuvre de ce critère est de déterminer si l'utilisateur souhaiterait ou non continuer à utiliser l'ACA plutôt q'une interface classique. Cet objectif est difficile à atteindre, en raison de la faible durée des expérimentations à finalité ergonomique, comparée à celle d'une utilisation en situation réelle. En effet, l'attrait exercé par l'ACA peut être dû uniquement à la nouveauté du type d'interaction homme-machine qu'il permet. Cet attrait peut donc disparaître rapidement dans un contexte réel d'usage intensif et/ou prolongé.

`Crédibilité (believability)`

C'est le critère majeur utilisé pour évaluer les ACAs. Un ACA est crédible s'il agit conformément aux attentes de l'utilisateur. Cette définition est difficilement exploitable telle quelle. Elle doit être précisée en fonction du contexte applicatif, du rôle assigné à l'ACA et de l'apparence "physique" de celui-ci. La crédibilité d'un ACA est souvent assimilée au réalisme de son apparence, surtout si celle ci est anthropomorphique (e.g., visage humain, homme/femme représentés en buste ou de pied en cap). Ce qui est une erreur, car les utilisateurs fondent leur jugement de la crédibilité d'un ACA non seulement sur son apparence mais aussi sur son comportement conversationnel. Déterminer le poids respectif de chacun de ces deux facteurs dans la formation du jugement subjectif de l'utilisateur est une question de recherche cruciale pour le développement des ACAs dans la Société de l'Information encore à venir. D'après les auteurs de [Ruttkay *et al.*, 2002], il a été montré que plus l'apparence d'un ACA est réaliste, plus les exigences des utilisateurs pour le juger crédible sont élevées.

`Confiance`

Les utilisateurs accorderont leur confiance à un ACA dans une situation donnée, si celui-ci se montre coopératif, compétent, honnête et prédictible dans cette situation. Le degré de confiance suscité par un ACA est déterminé en grande partie par les informations qu'il communique à l'utilisateur, donc par sa compétence dans le domaine de connaissances ou de savoir-faire couvert

par l'application ou le service considéré. La confiance qui lui est accordée dépend également de sa crédibilité et de facteurs tels que sa personnalité, son apparence, ses capacités de communication.

`Attractivité (engagement)`
L'attraction qu'est susceptible d'exercer un ACA sur l'utilisateur est un aspect de l'évaluation de l'ACA à ne pas négliger bien qu'elle soit difficile à mesurer. L'attractivité d'un ACA tient à la pertinence de ses interactions avec l'utilisateur, l'esthétique de son aspect et de ses mouvements, sa personnalité. Elle n'est pas indépendante de sa crédibilité.

Techniques d'évaluation des ACAs

Les techniques ergonomiques usuelles peuvent être adoptées pour évaluer l'apport des ACAs à l'interaction, moyennant des aménagements mineurs mentionnés dans l'article de [Ruttkay *et al.*, 2002]. Nous passons en revue ci-après les principales techniques de recueil de données sur les réactions et attitudes d'utilisateurs futurs aux ACAs.

`Observation et étude expérimentale`
L'observation n'appelle aucune remarque particulière, sinon la nécessité de définir des scénarios réalistes dans les différents domaines d'application des ACAs, car de tels scénarios font encore défaut aujourd'hui. Quant aux études expérimentales, elles sont particulièrement utiles ici pour tester la pertinence des différents choix de conception possibles lors de la création d'un ACA, en particulier pour évaluer systématiquement les effets de chaque paramètre considéré isolément sur l'interaction. A noter que de telles études expérimentales sont également susceptibles de fournir des informations utiles sur la communication non verbale humaine en situation de co-présence.

`Tests de performance (benchmarks) et tests comparatifs`
Des tests de ce type seraient très utiles pour évaluer et comparer des prototypes existants, globalement ou par rapport à une dimension particulière. Malheureusement, il n'existe pas encore à l'heure actuelle de tests standard, acceptés par l'ensemble de la communauté scientifique du domaine, permettant d'effectuer de tels tests. L'élaboration des scénarios de test nécessaires est une entreprise de longue haleine dans la mesure où une batterie de tests spécifique doit être construite pour chacun des divers domaines d'application des ACAs.

`Questionnaires et entretiens (interviews)`
Les questionnaires sont utiles pour recueillir les jugements subjectifs, les préférences, les compétences ou l'expérience d'utilisateurs potentiels. Si l'on souhaite toucher une vaste population, on utilisera exclusivement des questions fermées ou à choix multiple, et on diffusera le questionnaire sur Internet (surveys). S'il s'agit de recueillir les opinions des participants à une étude empirique ou expérimentale, on combinera questions fermées (e.g., échelles de Likert ou différenciateurs sémantiques) et ouvertes. Dans ce contexte, les entretiens permettent de recueillir les informations les plus riches à l'issue d'une étude. Mais leur coût élevé limite leur utilisation. Questionnaires et entretiens sont les principales techniques qui permettent d'obtenir des informations sur la manière dont la présence des ACAs et celle d'un ACA spécifique sont perçues par des utilisateurs potentiels. Les mesures objectives, indicateurs physiologiques des émotions de l'utilisateur ou suivi de l'évolution de son attention visuelle, donc de son intérêt, à l'aide d'un oculomètre, sont lourdes à mettre en œuvre, ce qui limite le nombre d'utilisateurs qu'il est possible de soumettre

à ces mesures ; en outre, si elles sont plus fiables que les témoignages verbaux, elles sont moins précises et moins riches en informations.

`Focus groups`

Les réunions de petits groupes d'utilisateurs sont très utiles au cas particulier car elles permettent de recueillir les impressions et les préférences d'utilisateurs. Toutefois, le succès de cette technique, en particulier la fiabilité des informations obtenues, dépend dans une large mesure de l'animation du groupe.

`Trace des usages`

L'avantage de cette technique est qu'elle permet d'observer les comportements d'utilisateurs dans la durée. Son utilisation permettrait donc d'éviter les biais induits par l'effet de nouveauté sur les comportements et jugements subjectifs d'utilisateurs. Cependant, sa mise en œuvre nécessite un prototype opérationnel, ce qui interdit son utilisation dans le cas d'applications complexes comme celles où interviennent les ACAs. En outre, les informations fournies se limitent à une suite très volumineuse d'interactions élémentaires qu'il faut analyser et interpréter en grande partie manuellement pour obtenir des informations utiles sur les comportements et les jugements subjectifs des utilisateurs. La tâche peut s'avérer infaisable en raison du volume des données de base enregistrées et de leur atomicité, même si l'on effectue un pré-traitement informatique sur elles.

La section suivante décrit brièvement les recherches menées jusqu'à présent sur l'adaptation dynamique de l'interaction homme-machine au profil de l'utilisateur courant et à son activité.

2.2 Interfaces adaptatives

Le second objectif de notre recherche est de contribuer aux efforts entrepris depuis une quinzaine d'années pour améliorer la flexibilité de l'interaction homme-machine. Si nous effectuons un bref historique de l'évolution des interfaces utilisateur, nous pouvons distinguer quatre grandes périodes dans l'évolution de celles-ci [Dam, 1997] :

- *Années 50-60 : Interfaces à cartes perforées.* Elles correspondent à la naissance de l'informatique, cette période utilise des ordinateurs dépourvus d'interfaces d'interaction. En effet, l'utilisateur introduit une carte perforée en entrée, puis récupère une impression ou listing en sortie. L'interaction, dans ce cas est réellement nulle puisque le dialogue se résume à une action de l'utilisateur.

- *Années 60-80 : Interfaces de commande.* Apparition des premiers systèmes informatiques qui permettent une véritable interaction utilisateur-système grâce à un langage de commande relativement complexe. Les échanges se font via un écran alphanumérique et un clavier. L'interaction évolue, mais ne reste accessible qu'aux informaticiens et ne peut s'étendre au grand public.

- *Années 80-90 : Interfaces graphiques et manipulation directe.* Au début des années 80 se développe le concept appelé "Manipulation directe" qui a contribué à la popularisation des ordinateurs personnels de part sa convivialité très supérieure à celle des langages de

commande classiques destinés aux professionnels de l'informatique [Shneiderman, 1983] ; [Shneiderman, 1993]. Ce paradigme basé sur l'interaction gestuelle avec des objets graphiques a suscité l'espoir d'atteindre la "transparence" c'est-à-dire la possibilité d'offrir à tout utilisateur, quelles que soient ses compétences en informatique, une interface homme-machine qui lui permette d'acquérir rapidement et sans effort la maîtrise de tout logiciel inconnu de lui. Le principal but était d'obtenir une maîtrise rapide, voire immédiate, de l'utilisation d'un logiciel, grâce à la simplicité de mise en œuvre des fonctions de celui-ci. Ce qui a permis de rendre les logiciels d'application effectivement accessibles aux non-spécialistes en éliminant les difficultés d'utilisation d'ordre syntaxique. Cependant ce modèle a rapidement montré ses limites au niveau sémantique essentiellement [Capobianco, 2002b], bien que ce paradigme ait représenté de toute évidence une grande avancée dans le domaine de l'"utilisabilité", au sens de [Nielsen, 1993], des systèmes informatiques, en démontrant la nécessité du développement de supports d'aide à l'utilisateur. Cette nécessité peut expliquer l'apparition et le développement de nouvelles techniques d'interaction destinées à pallier les limites de ce paradigme, et la réalisation de systèmes d'aide en ligne conçus comme un moyen de contribuer à la transparence visée par la manipulation directe.

– *Années 90-... : Adaptation des interfaces.* La conception d'interfaces adaptatives représente notre second axe de recherche. D'après [Gentner et Nielsen, 1996], le contrôle de l'interaction ne doit pas être donné entièrement aux utilisateurs. Pour rendre l'interaction plus naturelle, il faut que le système devienne intelligent. De nouvelles techniques ont été proposées et testées en vue d'aider les utilisateurs à faire face à la complexité croissante des logiciels grand public. Le développement de systèmes adaptatifs semble être une solution intéressante à explorer pour faire évoluer le comportement des interfaces afin de satisfaire au mieux les besoins et les attentes de chaque utilisateur.

Aujourd'hui, on considère qu'une interface, aussi transparente et intuitive soit-elle, ne suffit pas pour résoudre les difficultés d'utilisation [Ringle et Halstead-Nussloch, 1989] contrairement aux espoirs fondés au début des années 80. La recherche sur l'adaptation au profil de l'utilisateur a débuté vers la fin des années quatre-vingt, qui furent riches en découvertes et en publications. L'Intelligence Artificielle, en quête de développer l'automatisation de l'adaptation, est ainsi la source du défi actuel qui est de réaliser des interfaces intelligentes, capables d'évoluer au cours de l'interaction de façon autonome. De plus, l'utilisabilité exige que les logiciels via leurs interfaces s'adaptent à un ensemble d'individus [Shneiderman, 1999]. Toutefois l'utilisation de l'ordinateur diffère d'un individu à l'autre, car les intérêts, les niveaux de qualification et les buts d'utilisation indivuels sont très variés. Généralement, les logiciels ne sont pas développés pour un utilisateur unique ni pour une tâche particulière, bien définie, même si les progrès réalisés récemment en interaction homme-machine ont rendu l'ordinateur accessible a un large éventail d'utilisateurs. Dorénavant, les logiciels peuvent effectuer une multitude de tâches, ils sont donc destinés à une large communauté d'utilisateurs. Ceci explique en partie le problème de l'adaptation rencontré depuis la naissance de l'informatique grand public. De plus, comme le souligne [Salvendy *et al.*, 2002], il est réellement difficile de faire anticiper par la machine les souhaits et intentions de l'utilisateur, ses capacités, ses préférences ou encore sa personnalité. Il faut donc un système

flexible, qui donne toute liberté à l'utilisateur, améliore les relations entre l'utilisateur, la tâche et les caractéristiques du système, et augmente l'efficacité de celui-ci [Oppermann, 2002]. Les systèmes adaptatifs vont donc s'efforcer de prendre en compte les réactions de l'utilisateur et modifier la nature et la présentation des informations de l'interface en fonction du comportement de celui-ci.

Depuis le début des années quatre vingt dix, les travaux de recherche sur le concept d'adaptation se sont développés dans un nombre croissant de domaines d'application. La conception de systèmes interactifs doués de capacités d'adaptation, en particulier d'auto-adaptation, est un domaine de recherche actuellement en plein essor qui fait appel à des compétences pluridisciplinaires variées : psychologie, intelligence artificielle et génie logiciel, principalement. La difficulté majeure sur laquelle se centrent les efforts actuels est la construction de modèles pertinents et efficaces, qui permettent une adaptation dynamique du comportement de l'interface portant essentiellement sur le contenu, la présentation et les modalités des échanges d'informations entre l'utilisateur et le système. Après quelques définitions et la proposition d'une taxonomie qui permette de mettre en perspective les recherches actuelles et les avancées récentes sur les interfaces adaptatives, nous présentons quelques exemples représentatifs de ces travaux. Compte tenu du foisonnement des recherches sur la flexibilité des interfaces utilisateur et de l'ampleur du champ couvert par ce concept, nous avons limité la portée de cette étude aux interfaces utilisateur dont le comportement évolue au cours de l'interaction en fonction du profil de l'utilisateur et de ses évolutions telles que ces interfaces les "perçoivent". Pour la même raison, l'adaptation de l'interaction au contexte d'utilisation n'est pas abordée ici. On trouvera dans [Bobillier-Chaumon *et al.*, 2005] un panorama des recherches auxquelles a donné lieu la mise en œuvre du concept de flexibilité/adaptation en interaction homme-machine (IHM).

2.2.1 Définitions et classifications

Les systèmes adaptatifs auxquels nous nous intéressons visent à personnaliser un logiciel en fonction de l'utilisateur courant. Ce domaine de recherche a donné lieu à de nombreux travaux dont certains contribuent à caractériser la diversité des formes que peut prendre l'adaptation, les différents niveaux du dialogue entre l'homme et la machine où elle peut intervenir, ou encore la variété de ses objectifs. Les travaux précurseurs les plus connus, sont ceux d'[Edmonds, 1987], de [Malinoski *et al.*, 1992] et de [Dieterich *et al.*, 1993]. Ces auteurs nous permettent d'avoir une vue d'ensemble des possibilités offertes par l'adaptation, d'acquérir un cadre conceptuel de référence, nécessaire pour aborder ce difficile sujet.

Plus récemment, avec A. Jameson [Jameson, 2003] nous distinguons deux types principaux d'adaptation, en fonction du but et de la finalité visés : adaptation au profil de l'utilisateur courant ou adaptation à l'environnement dans lequel se situe l'interaction. On utilise fréquemment le terme plasticité pour désigner cette seconde forme d'adaptation qui n'est pas étudié ici ; voir par exemple [Thévenin *et al.*, 2003]. La mise en œuvre du concept de plasticité vise essentiellement à résoudre les problèmes posés par la mobilité de l'utilisateur qui conduit celui-ci à utiliser des dispositifs d'interaction variés aux capacités d'échanges différentes de celles des stations de travail classiques (e.g., PDAs, ordinateurs de poche, téléphones cellulaires, ordinateurs portables au sens de 'wearable', etc.). La plasticité mise en œuvre dans la majorité des solutions proposées porte

essentiellement sur l'adaptation des modalités d'interaction aux ressources limitées offertes par les différents équipements et aux contraintes qu'impose la mobilité sur les capacités perceptives et motrices de l'utilisateur ; voir, par exemple, [Calvary *et al.*, 2005].

L'adaptation à l'utilisateur courant, appelée aussi **personnalisation** de l'interface, peut s'effectuer de deux manières différentes. Si l'utilisateur a la possibilité d'adapter l'interaction à ses préférences personnelles, l'interface est dite configurable (customizable) ou **adaptable**. Si le système est capable d'adapter son comportement aux besoins, capacités et préférences de l'utilisateur courant pendant l'interaction, grâce à ses capacités de perception et d'interprétation de l'interaction et de son contexte, l'interface est dite **adaptative**.

Interfaces adaptables : adaptation statique

Depuis le paradigme de la "manipulation directe", l'évolution des interfaces tend de plus en plus à se rapprocher de l'utilisateur, afin de lui donner un maximum de contrôle sur son interaction avec l'ordinateur. Un système est dit adaptable, "s'il fournit à l'utilisateur les outils qui lui permettent de modifier les caractéristiques du système" [Oppermann, 2002]. L'interface peut être ajustée à l'initiative de l'utilisateur. Il peut alors décider d'adapter l'interface selon ses besoins en vue de pourvoir réaliser l'action désirée plus facilement. Le système n'est basé sur aucun algorithme intelligent, il est passif au sens où c'est l'utilisateur qui prend les décisions. Ce type d'interface est également désigné par le terme d'interface paramétrable ou configurable, puisque l'utilisateur peut ajuster à n'importe quel instant des paramètres de l'interaction.

Interfaces adaptatives : adaptation dynamique

Une interface peut être non seulement adaptable mais également "adaptative" [Dufresne, 1992]. Une interface adaptative est une interface qui s'adapte d'elle-même à l'usager sans que celui-ci n'ait à intervenir pour la paramétrer. Ce point de vue correspond à celui d'[Oppermann, 2002], pour qui un système est dit adaptatif, "s'il est capable de modifier ses propres caractéristiques en fonction des besoins de l'utilisateur". Ce point de vue est partagé par [Totterdell *et al.*, 1987], pour qui un système adaptatif est un système qui possède un mécanisme capable de changer l'état du système en réponse à l'évolution de l'environnement, dans l'objectif de rendre le système en meilleure adéquation avec cet environnement que s'il était resté dans son état initial. Toutefois, d'autres chercheurs se sont penchés sur une définition centrée exclusivement sur l'adaptation à l'utilisateur. [Langley et Fehling, 1998] considère une interface adaptative comme un système d'apprentissage ; c'est un logiciel qui améliore les capacités d'interaction avec l'utilisateur tout en construisant un modèle de l'utilisateur en fonction de l'expérience que celui-ci possède de l'utilisation du logiciel. [Jameson, 2006] ajoute une caractéristique importante à la définition d'[Oppermann, 2002]. Cette caractéristique prend en compte le comportement du système. Pour lui, "un système adaptatif est un système interactif qui adapte son comportement à chaque individu sur la base d'inférences non triviales, élaborées à partir des informations venant de l'utilisateur". Ainsi, le système modifie son comportement en fonction de celui de l'utilisateur par le biais d'informations que le système collectera grâce à l'interaction. D'un point de vue également centré sur l'utilisateur, [Salvendy *et al.*, 2002] propose cette définition : "une interface adaptative autonome adapte son affichage et les actions possibles aux buts et aux capacités de l'utilisateur en recueillant et exploitant des informations sur l'utilisateur, la tâche en cours de réalisation et la situation courante". Cette proposition de [Salvendy *et al.*, 2002] a donné lieu

à une taxonomie basée sur l'utilisateur, la situation et la tâche courantes. Dans un contexte d'adaptation dynamique, l'initiative des évolutions du comportement de l'interface utilisateur est souvent entièrement à la charge du système ; elle est alors totalement transparente pour l'utilisateur. Par exemple, les " Smart Menus ", qui ont fait leur apparition dans les applications commercialisées destinées au grand public (voir Windows 2000), masquent les items rarement (ou jamais) sélectionnés par l'utilisateur courant qui n'a d'autre contrôle sur cette option de personnalisation que de la désactiver. Plus rarement, des possibilités de contrôler l'évolution de l'interface sont offertes à celui-ci, le partage des initiatives entre le système et lui variant dans des proportions importantes d'un logiciel à l'autre. Par exemple, les résultats d'un filtrage des dépêches de l'AFP (diffusées sur le Web) en fonction des intérêts de l'utilisateur, peuvent être présentés sous la forme d'une liste ordonnée des titres de la totalité des dépêches ; l'algorithme de filtrage peut aussi prendre l'initiative de supprimer certaines d'entre elles (e.g., en appliquant un critère quantitatif de redondance), sans en informer l'utilisateur (transparence).

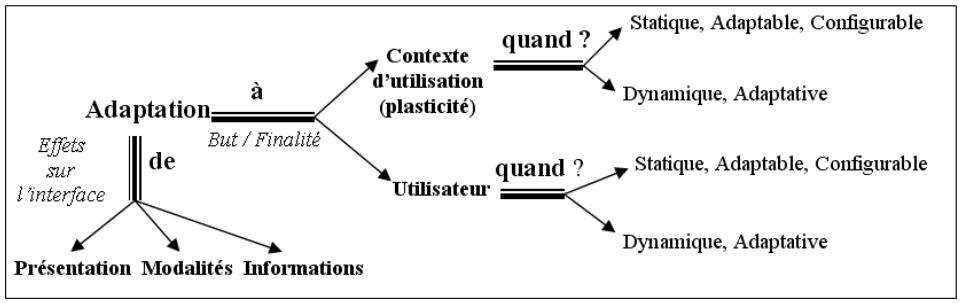


FIG. 2.1 – Classification générale des formes de mise en œuvre du concept d'adaptation en IHM. Voir [Simonin, Carbonell, 2004].

La figure 2.1 résume la classification générale issue d'A. Jameson. On trouvera dans [Vanderdonckt *et al.*, 2005] la présentation d'une taxonomie plus vaste qui couvre, outre l'adaptation statique ou dynamique au profil de l'utilisateur, celle à la situation d'interaction et au contexte d'utilisation. L'objectif des auteurs est de recenser, à l'intention des concepteurs d'interfaces adaptables, adaptatives ou plastiques, un panorama détaillé des choix de conception que permettent les avancées scientifiques et technologiques récentes.

2.2.2 Raffinement de la classification générale

Les dimensions de la classification présentée dans la figure 2.1 ne rendent pas compte entièrement de la diversité des directions de recherche explorées jusqu'à présent concernant l'adaptation de l'interaction au profil de l'utilisateur. Elles doivent être complétées. Nous précisons ici ces dimensions complémentaires :

- La nature des caractéristiques de l'utilisateur ciblées par l'adaptation.
- La répartition de l'initiative et du contrôle de l'adaptation entre celui-ci et le système.
- Le rôle de l'adaptation, c'est-à-dire la nature de l'activité de l'utilisateur qu'elle tente de faciliter et de rendre plus efficace et donc, indirectement, le type d'application considéré.
- Les différentes techniques d'adaptation.

Diversité des constituants du profil et du modèle utilisateur

Les dimensions de l'individualisation de l'interaction que permettent les modèles de l'utilisateur, statiques ou dynamiques, élaborés récemment sont nombreuses et d'une diversité remarquable. Les composantes suivantes de la personnalité de l'utilisateur courant ont fait l'objet de nombreuses modélisations :

- Ses *connaissances*, ses *compétences* et son *expérience*, ses *croyances*.
- Ses *intérêts*, ses *préférences*.
- Ses *habitudes*, par exemple, les séquences d'actions qu'il effectue de façon routinière sur l'interface. Cette dimension ne peut donner lieu qu'à des modèles dynamiques puisqu'il s'agit d'une adaptation au comportement de l'utilisateur.
- Ses capacités physiques. Le handicap physique est un domaine de recherche à part entière qui privilégie la mise en œuvre du concept d'adaptabilité, compte tenu de la stabilité temporelle à court terme de la majorité des déficits perceptifs et moteurs. Une des réalisations les plus abouties est le prototype développé dans le cadre du projet européen AVANTI [Stephanidis, 2001]. Cette dimension ne sera plus évoquée dans la suite du chapitre puisqu'elle privilégie la conception d'interfaces utilisateur adaptables plutôt qu'adaptatives.
- Ses *intentions* et *buts* courants. Pour offrir à l'utilisateur une assistance efficace à la réalisation de ses intentions, il faut que le système soit capable de détecter de façon fiable les intentions de l'utilisateur sans que celui-ci ait besoin de les préciser au système, donc à partir des seules informations contextuelles disponibles. Modéliser dynamiquement les intentions de l'utilisateur courant représente un objectif ambitieux, extrêmement difficile à atteindre pour l'instant, sauf dans des contextes applicatifs particuliers ; ce qui peut expliquer pourquoi seules quelques recherches à caractère préliminaire et de portée limitée aient été menées dans cette direction jusqu'à présent. Voir par exemple [Conati *et al.*, 1997] qui présente l'utilisation d'un réseau bayésien pour identifier et évaluer la stratégie de résolution de problème mise en œuvre par l'apprenant à partir des actions que celui-ci effectue sur l'interface du tuteur intelligent ANDES ; le domaine de connaissances considéré est la physique newtonienne.
- Ses *états psychologiques*, en particulier ses émotions (anxiété, stress, plaisir, etc.). Les recherches menées dans cette direction sont encore embryonnaires, mais elles sont susceptibles de se développer rapidement sous l'impulsion des travaux centrés sur l'apport à l'interaction d'agents conversationnels animés, ou ACAs, (voir, par exemple, [Bevacqua et Pelachaud, 2004], et le soutien institutionnel apporté aux recherches sur l'interaction " affective " ; voir, notamment les objectifs du réseau européen d'excellence HUMAINE qui compte plusieurs équipes possédant une solide expérience dans le domaine des ACAs [HUMAINE, 2003].

Cette liste n'est pas exhaustive. On pourrait y inclure, par exemple, le rôle professionnel de l'utilisateur, son rôle social, etc. Mais ces dimensions représentent des caractéristiques individuelles relativement stables temporellement. Elles relèvent donc d'une adaptation de l'interaction statique plutôt que dynamique. A noter que les deux premières dimensions mentionnées ci-dessus ont suscité les modélisations statiques et dynamiques les plus nombreuses.A noter également que le choix de la composante du profil utilisateur sur laquelle porte la personnalisation n'est pas indépendant du domaine d'application envisagé. Par exemple, les aides en ligne, les didacticiels

et les Environnements Interactifs d'Apprentissage Assisté par Ordinateur, qui tentent d'adapter dynamiquement leur comportement au profil cognitif du novice ou de l'apprenant, se focalisent sur la modélisation des connaissances de l'utilisateur courant, de ses compétences spécifiques ou de son expérience personnelle. En revanche, les navigateurs Internet adaptatifs tentent le plus souvent de cerner les intérêts, goûts et préférences de chaque utilisateur.

Répartition du contrôle de l'adaptation entre l'utilisateur et le système

L'initiative et le contrôle de l'évolution de l'interface peuvent être du ressort exclusif de l'utilisateur. C'est le cas de la plupart des interfaces adaptables qui offrent à l'utilisateur un ensemble d'options et de possibilités de paramétrage sous différentes formes, par exemple : fichiers texte de configuration associés aux utilitaires des premiers systèmes d'exploitation (e.g., Unix), menus 'Affichage' et 'Fenêtre' des logiciels destinés au grand public (e.g., Word), ou item 'Préférences' de nombreux navigateurs (e.g., Netscape). Le contrôle de l'évolution du comportement de l'interface peut également être assuré entièrement par le système. Ainsi, certains prototypes de recherche prennent-ils en charge la sélection du profil utilisateur au sein d'un ensemble plus ou moins riche de stéréotypes prédéfinis, en vue d'assurer la transparence de l'adaptation. Cette solution a été souvent adoptée, notamment, pour les applications destinées au grand public, telles que l'assistance à la recherche d'informations et la navigation sur Internet ou les logiciels éducatifs, car cette catégorie d'utilisateurs est censée rechercher en priorité et privilégier la transparence de l'interaction. Par exemple, c'est cette solution que les premiers prototypes de filtrage des informations sur Internet en fonction des intérêts de l'utilisateur courant ont utilisée [Lieberman et Maulsby, 1996]. Les environnements expérimentaux d'apprentissage adaptatifs des années 90 l'ont également adoptée pour la mise en œuvre de modèles complexes des connaissances de l'apprenant ; voir, par exemple, le modèle décrit dans [Paiva et Self, 1994] et [Paiva *et al.*, 1994]. A noter que la transparence de l'évolution du comportement du système a été également choisie pour les versions récentes de Pack Office ; voir les "Smart Menus" évoqués précédemment.

Rôle de l'adaptation

La finalité de l'adaptation dynamique au profil de l'utilisateur courant est généralement d'assister celui-ci dans la réalisation des objectifs et des activités qui motivent son interaction avec le logiciel adaptatif. Un modèle centré sur les intérêts et les préférences de l'utilisateur permet de fournir à celui-ci une assistance explicite ou implicite à des activités de navigation et de recherche d'informations. Les systèmes "conseillers" (recommending systems), par exemple, assistent explicitement l'utilisateur dans ses prises de décisions en lui proposant des solutions et en argumentant leur bien-fondé ([Ardissono et Goy, 2000] ; [Fink et Kobsa, 2000]). En revanche, les systèmes de filtrage des informations diffusées sur Internet, qui sélectionnent les informations présentées à l'utilisateur en fonction de ses intérêts, l'assistent implicitement dans sa recherche d'informations [Ardissono *et al.*, 2001]. Un modèle de l'utilisateur capable d'évaluer les connaissances du novice ou de l'apprenant et de suivre leur évolution au cours de l'interaction permet de fournir aux utilisateurs d'un didacticiel, des informations pertinentes au bon moment et dans

une présentation qui lui en facilitera la compréhension et l'assimilation ([Mitchell *et al.*, 1994] ; [Brajnik et Tasso, 1994]).

Techniques d'adaptation

Concernant l'acquisition et l'exploitation des informations sur l'utilisateur courant, de nombreuses techniques ont été utilisées dans le cadre de différentes études :

- Réseau de neurones dans le projet Lumiere [Horvitz *et al.*, 1998],
- Réseau de Petri [Medina-Medina *et al.*, 2002],
- Apprentissage automatique [Langley, 1992],
- Multi-Agent dans Amebica [Vaudry, 2002],
- Inférence et planification [Berthomé-Montoy et Fouet, 1995].

Ces techniques permettent de construire un modèle de l'utilisateur et un modèle de la tâche. La majorité des prototypes existants utilisent une approche de type apprentissage automatique (Machine Learning). On trouvera une description détaillée de ces techniques dans [Jameson, 2006] et [Ruvini, 2000].

2.2.3 L'existant

Nous décrivons dans cette section quelques exemples représentatifs de mise en œuvre du concept d'adaptation dynamique au profil de l'utilisateur, plutôt que de survoler l'ensemble des travaux publiés. Les recherches dans ce domaine ont souvent donné lieu au développement de systèmes adaptatifs plus ou moins aboutis, du démonstrateur ou du prototype de laboratoire au système opérationnel testé dans des situations réelles d'utilisation prolongée. La présentation suit l'ordre de la classification proposée dans la section précédente.

Diversité des composantes du profil utilisateur visées par l'adaptation

Activités de routine et habitudes

Les recherches dans ce domaine, relativement anciennes, se sont orientées dans deux directions principales. L'une vise à accroître l'efficacité de l'interaction en déchargeant l'utilisateur de tâches routinières comme la gestion des messages [Maes, 1994] ou le suivi de planning/agenda [Mitchell *et al.*, 1994]. Généralement, l'utilisateur garde un contrôle sur la prise en charge de ces activités par le système. Par exemple, SwiftFile, logiciel téléchargeable sur le site de recherche d'IBM, assiste l'utilisateur dans le rangement des messages qu'il reçoit, en lui proposant, pour chaque nouveau message, trois dossiers qu'il considère comme les plus pertinents. La sélection de ces dossiers s'appuie sur une analyse statistique du contenu des messages et des pratiques de l'utilisateur [Segal et Kephart, 1999]. L'autre direction de recherche porte sur l'adaptation de la présentation des informations, principalement des menus, aux habitudes de travail de chaque utilisateur. Outre les "Smart Menus", on peut citer les "Split Menus" proposés dans [Sears et Shneiderman, 1994] qui mettent en œuvre une stratégie différente. Cette stratégie consiste, lors de la sélection d'un menu, à afficher en début de liste (i.e., en haut de l'écran, sous l'en-tête du menu), les items les plus fréquemment sélectionnés par l'utilisateur, ce qui, d'après l'évaluation

réalisée par les auteurs, accroît les performances (temps de sélection) et la satisfaction des utilisateurs ; comme pour les "Smart Menus", il s'agit d'une option de l'interface que l'utilisateur ne peut qu'activer ou désactiver.

Capacités cognitives, connaissances, compétences et expérience

L'adaptation de l'interaction aux connaissances et compétences de l'utilisateur courant a suscité de nombreux travaux de recherche. Elle représente l'une des mises en œuvre les plus étudiées en IHM du concept d'adaptation dynamique. Le projet Lumiere [Horvitz *et al.*, 1998] mérite une attention particulière en raison de l'influence que ses résultats ont exercée sur l'évolution des logiciels Microsoft destinés au grand public. Une version réduite (en termes de fonctionnalités) du prototype développé dans le cadre de ce projet a été intégrée aux aides en ligne des logiciels de Pack Office à la fin des années 90. Ces aides en ligne prennent l'initiative de proposer à l'utilisateur une ou plusieurs rubriques thématiques d'aide qu'elles jugent susceptibles de lui être utiles sur la base d'une analyse des actions récentes qu'il a effectuées sur le logiciel et de l'action en cours d'exécution ; pour accéder aux informations d'aide correspondant à ces rubriques, il suffit de sélectionner l'une d'entre elles. Cette stratégie est discutable dans la mesure où les interventions "spontanées" de l'aide risquent d'interférer avec l'activité principale de l'utilisateur, à savoir l'interaction avec le logiciel, et donc d'être perçues par lui comme des interruptions intempestives, voire irritantes, de cette activité. En outre, anticiper à bon escient les besoins d'informations d'aide de l'utilisateur requiert des compétences qui sont aujourd'hui encore hors de la portée des systèmes dits "intelligents". Enfin, au cas particulier, le comportement et les performances de la version implantée dans Pack Office sont trop limitées par rapport à celles du prototype issu du projet Lumiere pour que l'on puisse se fonder sur l'analyse des réactions des utilisateurs de Pack Office pour déterminer l'utilité et l'utilisabilité (usability) d'une aide en ligne qui tente d'anticiper les besoins en informations d'aide de l'utilisateur novice. On trouvera à la fin de cette section un exemple plus classique, représentatif des environnements adaptatifs d'enseignement et des tuteurs "intelligents" qui mettent à profit un suivi de l'évolution des connaissances de l'apprenant au cours de l'interaction pour améliorer l'efficacité pédagogique du système.

Etats émotionnels

Dans le contexte des travaux sur l'interaction homme-machine affective, l'objectif qui sous-tend le plus souvent la modélisation de l'état émotionnel courant de l'utilisateur est moins d'adapter le comportement du système à cet état et à ses fluctuations au cours de l'interaction que d'agir sur lui s'il nuit à la réalisation des buts de l'utilisateur.
Les travaux de recherche présentés dans [Schäfer et Weyrath, 1997] s'inscrivent dans une perspective d'adaptation dynamique du comportement d'un système d'assistance à l'état émotionnel de l'utilisateur, en particulier aux effets de cet état sur ses capacités cognitives. READY (Resource-Adaptive Dialog System) est un prototype qui tente d'évaluer le degré de stress et d'anxiété des auteurs d'appels à un service téléphonique d'assistance aux urgences (incendies) afin, d'une part, d'adapter le contenu et la présentation des messages oraux à l'altération temporaire des capacités cognitives des appelants (i.e., mémorisation, élocution et maîtrise de la situation) due aux effets négatifs de ces émotions et d'une intoxication éventuelle et, d'autre part, de tenir compte dans la reconnaissance de leurs énoncés de la dégradation de leur élocution. La détection de l'état émo-

tionnel de l'appelant utilise un réseau bayésien dynamique pour analyser son rythme respiratoire, son élocution et les bruits de fond. Le comportement du système met à profit les connaissances issues de l'étude d'un corpus de dialogues enregistrés dans un centre d'appels téléphoniques des pompiers.
Des travaux plus récents visent non pas à adapter le comportement du système à l'état émotionnel de l'utilisateur courant mais à agir sur cet état. Ainsi, [Prendinger *et al.*, 2005] décrit une étude où le comportement empathique d'un agent conversationnel animé réduit significativement l'anxiété de l'utilisateur pendant un entretien d'embauche simulé. L'objectif est d'entraîner celui-ci à contrôler ses émotions négatives au cours de ce type d'entretien par un apprentissage (conditionnement) de type behaviouriste. La détection de l'état émotionnel (arousal) de l'utilisateur est effectuée au moyen de capteurs physiologiques du rythme cardiaque (électromyographie) et de la conductance de la peau (capteur galvanique).

Répartition de l'initiative et du contrôle de l'adaptation - Rôle de l'adaptation

Nous avons rassemblé dans une même section ces deux dimensions car elles sont liées au domaine d'application envisagé. Les recherches sur l'adaptation dynamique à l'utilisateur privilégient les classes d'applications grand public suivantes qui se prêtent naturellement à une personnalisation de l'interaction :

- Les systèmes et agents "conseillers" (recommending systems), destinés souvent à des sites Web commerciaux.
- Les systèmes d'assistance personnalisée à la recherche d'informations sur Internet.
- Les systèmes à vocation didactique : environnements d'enseignement à distance, didacticiels, tutoriels et aides en ligne.

Les exemples présentés dans la suite de cette section ont été regroupés en fonction du domaine applicatif ciblé.

Les systèmes conseillers

En raison de leur vocation, les systèmes conseillers détiennent l'initiative et le contrôle de l'adaptation puisque, par définition, leur rôle est d'assister l'utilisateur dans un choix, une prise de décision. Ils se distinguent les uns des autres par la nature des informations qu'ils utilisent pour modéliser les intérêts et préférences de l'utilisateur et par la méthode qu'ils appliquent pour acquérir ces informations. Par exemple, MovieCentral, site de Entertainment Inc. qui fournit des recommandations personnalisées sur les films à l'affiche, initialise le modèle de l'utilisateur courant U à partir de données explicites, à savoir l'évaluation par celui-ci de dix films (à l'aide de notes comprises entre 1 et 10). Le système, qui applique une approche dite "collaborative", identifie, dans sa base des profils utilisateur, le groupe d'utilisateurs dont les vecteurs de notes sont les plus proches de celui de U. Ce qui permet au système de prédire, en réponse à une requête de U sur un film donné, si ce film lui plaira ou non. Certains sites imposent à U d'expliciter ses critères d'appréciation plutôt que d'évaluer un ensemble limité de produits. D'autres encore construisent le profil de l'utilisateur de façon incrémentale. C'est le cas du prototype présenté dans [Burke *et al.*, 1997]. Le système S propose à U un produit, un restaurant au cas particulier, et invite celui-ci à critiquer et évaluer qualitativement cette proposition initiale. S fait ensuite

une nouvelle proposition qui tient compte des critiques de U, et répète cette procédure jusqu'à ce que U ne formule plus de critiques sur la proposition courante. Cette démarche itérative permet à l'utilisateur de raffiner progressivement et de façon naturelle ses critères de choix, et au système d'obtenir au terme du raffinement un profil utilisateur pertinent et précis sans imposer à celui-ci un questionnaire initial artificiel et fastidieux.

L'assistance à la navigation et à la recherche d'informations sur Internet

Le rôle principal des systèmes adaptatifs qui visent ce domaine d'application consiste à filtrer la masse d'informations disponibles sur la Toile en fonction des intérêts de l'utilisateur courant. Si le système de personnalisation ne présente à l'utilisateur que les informations qui sont censées l'intéresser, l'adaptation est transparente pour celui-ci, il ne peut donc la contrôler. En revanche, si le système se borne à ordonner les résultats de la recherche en fonction de ce qu'il suppose être les intérêts de l'utilisateur, celui-ci peut contrôler l'adaptation et, par ses réactions, contribuer à améliorer la pertinence du modèle sur lequel s'appuie la personnalisation. Parmi les travaux de recherche publiés sur le filtrage adaptatif d'informations, l'une des études les plus intéressantes est celle décrite dans [Billsus et Pazzani, 2000]. C'est une recherche parmi les plus abouties puisqu'elle a donné lieu à un prototype qui a fait l'objet d'une évaluation à grande échelle en situation d'utilisation réelle (recueil des traces de 2000 utilisateurs). C'est aussi l'une des plus évoluées puisqu'elle met en œuvre un modèle à court terme et un modèle à long terme des intérêts de l'utilisateur, et utilise un feedback explicite et implicite fourni par celui-ci pour la mise à jour de ces deux modèles. Le problème de la plasticité est également abordé puisque deux versions du prototype ont été réalisées, l'une pour le Web, l'autre pour PDAs et téléphones portables. Les auteurs ont réalisé la conception, le déploiement et l'évaluation d'un système interactif S d'accès aux actualités (i.e., les news). La version Web de S fournit, en réponse à une requête de l'utilisateur U, exprimée sous la forme d'une expression en langue naturelle ou d'une suite de mots-clés, la liste des titres des bulletins d'information susceptibles de satisfaire la requête, ordonnés en fonction du profil de U ; S affiche également, pour chaque titre, son score de prédiction de l'intérêt de U. La sélection d'un titre dans cette liste entraîne l'affichage du bulletin d'information correspondant. U peut alors, s'il le souhaite, fournir un jugement explicite sur l'information affichée en sélectionnant l'un des items suivants : "Intéressant", "Inintéressant", "Encore", "Déjà connu", "Demande d'explication". S'il sélectionne le dernier item, S répond en affichant : (i) son score de prédiction pour cette information, accompagné d'une phrase d'explication en langue naturelle, (ii) le nombre d'autres utilisateurs qui l'ont évaluée et (iii) leur intérêt pour l'information considérée. L'algorithme d'apprentissage qui construit le modèle global d'un utilisateur et le met à jour utilise les jugements explicites de U ("Intéressant", "Inintéressant"), les titres non sélectionnés par lui, les bulletins qu'il a lus (temps passé). Pour prédire l'intérêt qu'un nouveau document est susceptible de provoquer de la part de U, l'algorithme utilise une mesure de similitude entre ce document et ceux qui ont suscité l'intérêt de U. A partir de ces informations, S construit deux modèles complémentaires de U : un modèle à court terme qui tient compte de l'intérêt de U pour les sujets et événements récents de l'actualité, et un modèle à long terme centré sur les préférences générales de U indépendantes de ses intérêts conjoncturels. Le calcul du score à court terme d'un nouveau document utilise l'algorithme " du plus proche voisin ", et l'évaluation de son score à long terme un classificateur bayésien. Ces deux scores sont combinés dans un algorithme global multi-stratégies qui définit, sous la forme d'un score,

la prédiction de S concernant l'intérêt global de U pour le document considéré.
Les principales avancées récentes portent sur le filtrage collaboratif, c'est-à-dire l'utilisation des profils d'autres utilisateurs pour compléter et préciser le profil d'un nouvel utilisateur (voir, par exemple, [Kleinberg et Sandler, 2004]). Le profil initial du nouvel utilisateur peut être établi, par exemple, à partir des réponses à un questionnaire initial [Ardissono *et al.*, 2001].

Environnements d'enseignement

Les environnements d'assistance à l'enseignement à distance suscitent depuis plusieurs décennies une activité de recherche intense et constituent un domaine à part entière qui accorde une place importante aux travaux sur la modélisation de l'apprenant et la personnalisation des apports de connaissances, des exercices et de l'évaluation. On observe, en particulier, une grande diversité dans :

- les stratégies de contrôle de l'apprentissage, qui varient d'un guidage strict du parcours de l'apprenant à une liberté plus ou moins grande d'exploration des unités.
- l'aspect de l'interaction sur lequel porte l'adaptation dynamique, navigation dans les unités d'enseignement, choix des connaissances et concepts nouveaux présentés, et modulation de leur présentation, du contenu des suggestions et retours système, ou encore de la stratégie d'intervention du système " tuteur ".
- les caractéristiques de l'apprenant sur lesquelles portent la modélisation et l'adaptation : connaissance des contenus de l'enseignement, style d'apprentissage, motivation, activité en cours.

Le prototype ELM-ART [Weber et Specht, 1997], bien que relativement ancien, est intéressant ici, car il offre à l'apprenant un guidage flexible pour l'apprentissage de LISP. Ce guidage tient compte des compétences de celui-ci et des pré-requis des unités de cours. Il lui propose deux formes de recommandations. L'une est une liste d'unités associées chacune à une icône dont la couleur exprime la force de la recommandation ; l'apprenant dispose donc d'une relative liberté d'exploration. L'autre recommande une seule unité ; elle constitue un guidage strict de l'apprenant.

2.3 Sujet de recherche

D'après l'état de l'art réalisé dans les sections précédentes, on observe que les ACAs et les interfaces adaptatives sont des thèmes de recherche en plein essor. Les problématiques encore peu abordées et les progrès techniques récents à explorer sont, respectivement :

1. Comprendre et évaluer l'influence de la présence des ACAs sur l'interaction,
2. Répondre aux nombreuses questions d'ordre ergonomique que soulèvent les interfaces adaptatives quant à leur apport effectif à l'interaction.

D'une part, les études portant sur l'évaluation de l'incidence effective des ACAs sur l'interaction n'utilisent généralement que des techniques de recueil des jugements subjectifs d'utilisateurs potentiels, à savoir, questionnaires, interviews, 'focus groups'. En effet, le recours, à l'observation des comportements, et à l'analyse des performances et de l'usage (trace des interactions) est plus

rare comme le souligne [Ruttkay *et al.*, 2002]. A noter que les techniques de recueil des jugements subjectifs sont utilisées principalement pour mesurer les effets de la présence des ACAs non seulement sur la satisfaction d'utilisateurs futurs mais aussi sur l'efficacité de l'interaction, des capacités d'assistance et de communication orale ou multimodale des ACAs. Ce qui est discutable en raison de l'écart fréquent entre les effets perçus, ressentis et les effets réels. Nous nous intéressons donc à l'observation et l'analyse de l'apport effectif de la présence d'un ACA, en tant que variable libre, à l'efficacité et l'utilisabilité de l'interaction : présence/absence d'un ACA dans une situation réaliste, à savoir l'assistance à l'utilisation d'un logiciel grand public.

D'autre part, comme l'utilisation et l'étude ergonomique des systèmes d'interaction adaptatifs suscitent de nombreuses difficultés, nous nous proposons de tenter de répondre aux questions suivantes : Quel contrôle donner à l'utilisateur sur l'évolution du comportement du système ? La transparence de l'adaptation est-elle souhaitable ? Ne vaut-il pas mieux notifier l'utilisateur de l'évolution spontanée du comportement du système au cours de l'interaction et, si oui, sous quelle forme ? Ces questions, ainsi que les remarques des participants à l'étude préliminaire (décrite dans le chapitre suivant) font que nous avons choisi d'étudier les réactions d'utilisateur face une interface qui s'adapte de manière autonome aux connaissances et compétences de l'utilisateur novice et à leur évolution au cours de l'interaction.

Les éléments de réponse à ces questions fournis par nos travaux permettront d'offrir aux concepteurs des guides et des recommandations utiles pour la conception d'interfaces utilisateur capables d'accompagner le grand public dans la prise en main d'un logiciel courant grâce à la présence d'un ACA ou d'un mécanisme d'adaptation dynamique à l'utilisateur courant. Cela permettra, d'une part, d'accroître les capacités d'assistance et de coopération des systèmes interactifs et, d'autre part, de les rendre plus attractives et plus agréables à utiliser, voire d'humaniser l'interaction.

Chapitre 3

Étude préliminaire

Une première étude expérimentale menée à titre exploratoire a pour objectif d'évaluer l'efficacité et la convivialité de messages d'aide en ligne intégrant la parole, graphique et texte. Cette expérimentation fait intervenir des utilisateurs potentiels novices (8 sujets) ayant pour consigne de réaliser une animation sous le logiciel Flash qui leur est inconnu. Durant l'expérimentation, les sujets disposent d'une aide en ligne générant des messages d'aide procéduraux de deux types :

- une version multimodale intégrant du Texte (message textuel de l'action à réaliser) et du Graphique (copie d'écran associée au message textuel) - version "T+G" ;
- une version multimodale intégrant de l'Oral (message vocal enregistré que l'utilisateur peut écouter via un bouton) et du Graphique (copie d'écran associée au message vocal) - version "O+G".

L'analyse des données expérimentales s'appuie sur des mesures classiques (performances, questionnaires) ainsi que sur l'interprétation des mouvements oculaires des participants recueillis par un dispositif de suivi du regard. Tous les sujets ont expérimenté les deux systèmes d'aide ; l'ordre de passation a été contrebalancé.

L'analyse des résultats a montré que la condition "T+G" se révèle plus efficace que la condition "O+G", concernant les temps de réalisation des tâches complexes. En effet, si les tâches sont relativement courtes et simples, les performances des deux conditions sont identiques. Toutefois, l'analyse des taux de réussite a montré que la condition "O+G" est plus efficace. Ceci peut s'expliquer éventuellement par le fait que les sujets, dans cette condition pouvaient se concentrer davantage sur la réalisation des animations ; ils n'avaient pas à interrompre leur activité pour consulter l'aide (ce que l'analyse des traces des mouvements oculaires a montré).
Ainsi, des messages d'aide oraux permettent-ils d'interagir avec le logiciel en même temps qu'on les écoute, alors que des informations procédurales sous forme textuelle imposent une interruption de l'activité. En revanche, la présentation textuelle facilite la recherche d'une information ponctuelle, tandis qu'il faut réécouter entièrement un ou plusieurs messages oraux pour trouver l'information précise dont on a besoin.
Les résultats de cette étude préliminaire suggèrent que, pour faciliter l'accès aux informations d'aide, deux directions de recherche principales se dégagent :

- améliorer la conception et la manière de présenter les messages oraux, par exemple grâce à la présence d'un ACA anthropomorphe doté de capacités d'expression orale (voir l'étude présentée dans le chapitre suivant) ;
- créer des messages adaptatifs qui tiennent compte des connaissances acquises par l'utilisateur courant au fur et à mesure de l'interaction et de son profil cognitif (voir le chapitre consacré à l'évaluation d'un système d'aide adaptatif).

Après avoir présenté les motivations et objectifs de cette étude basée sur les systèmes d'aide, nous décrivons le protocole expérimental ainsi que le matériel utilisé au cours de l'expérimentation. Les résultats issus de l'expérience sont ensuite commentés.

3.1 Motivations, Objectifs - Systèmes d'aide

3.1.1 Contexte

Dans le domaine de l'interaction homme-machine, les raisons justifiant la nécessité d'une assistance efficace à l'utilisateur sont nombreuses. Elles comprennent le développement de nouvelles formes d'interaction, l'apparition de modalités nouvelles (aussi bien en entrée qu'en sortie) et la diversité des tâches que peut réaliser l'utilisateur avec le système. Ceci donne lieu à une grande complexité et une variété très riche des tâches que l'utilisateur peut accomplir. Néanmoins ceci impose aux utilisateurs un apprentissage constant pour acquérir une connaissance minimale et une pratique efficace d'un nouveau logiciel.

Notre travail porte principalement sur les interfaces graphiques de type WYSIWIG (Waht You See Is What You Get), dans la mesure où celles-ci se sont imposées sur le marché des systèmes grand public au cours des dernières années. Cependant, on peut noter que les systèmes d'aide proposés aux utilisateurs restent, malgré une utilité et une demande évidentes, largement insuffisants pour satisfaire les besoins des utilisateurs visés. Ceci nécessite donc un important travail de recherche afin de proposer aux utilisateurs (tous niveaux confondus) de systèmes grand public, des supports d'aide conviviaux et efficaces.

La nécessité d'une assistance à l'utilisation d'un logiciel tient également à la limitation des systèmes informatiques, limitation liée d'une part aux modalités d'interaction souvent mal combinées ou sous-utilisées et à une mauvaise utilisation des fonctions du logiciel d'autre part. En effet, un logiciel 'ergonomique', d'utilisation intuitive, peut assurer à l'utilisateur une maîtrise partielle de l'utilisation de celui-ci sans qu'il soit nécessaire de recourir à une formation ou à un apprentissage spécifique. Cependant, s'il désire acquérir une maîtrise complète du logiciel, cela lui sera impossible sans support d'aide. Au début des années 80, se développe le concept appelé "manipulation directe" qui a contribué à la popularisation des ordinateurs personnels de par sa convivialité par rapport aux langages de commandes classiques destinés aux professionnels de l'informatique [Shneiderman, 1983] ; [Shneiderman, 1993]. Ce paradigme basé sur l'interaction gestuelle et graphique a suscité l'espoir d'atteindre la "transparence" c'est-à-dire la possibilité d'offrir à tout utilisateur, quelles que soient ses compétences en informatique, une interface homme-machine qui lui permette d'acquérir rapidement et sans effort la maîtrise de tout

logiciel lui étant inconnu. Le principal but était d'obtenir une maîtrise rapide, voire immédiate, de l'utilisation d'un nouveau logiciel, grâce à la simplicité de mise en œuvre de ses fonctions. On espérait ainsi rendre les logiciels d'application accessibles aux non-spécialistes en éliminant les difficultés d'utilisation d'ordre technique. Cependant, ce modèle a rapidement montré ses limites, techniques et sémantiques [Capobianco, 2002b], bien que ce paradigme ait représenté de toute évidence une grande avancée dans le domaine de l'"utilisabilité", au sens de [Nielsen, 1993] des systèmes informatiques en démontrant indirectement la nécessité du développement de supports d'aide à l'utilisateur. Cette nécessité a permis le développement et l'apparition de nouvelles technologies de l'information pour palier aux limites de la manipulation directe et la réalisation de systèmes d'aide efficaces, capables de contribuer à la transparence visée par la manipulation directe.

3.1.2 Aide en ligne : Définition

Aujourd'hui, on considère qu'une interface aussi transparente et intuitive soit-elle, ne suffit pas pour résoudre les difficultés d'utilisation des novices [Ringle et Halstead-Nussloch, 1989] contrairement aux espoirs fondés au début des années 80. La conception des systèmes d'aide en ligne à l'utilisation de logiciels pose de réels problèmes, d'où l'étude proposée qui consiste à préparer une expérimentation visant à évaluer l'efficacité et la convivialité de messages d'aide en ligne intégrant parole ou texte, et graphique. Nous avons, dans un premier temps, évalué les problèmes spécifiques que pose la conception d'un système d'aide en ligne. En effet, le système doit permettre à l'utilisateur d'accéder aux informations qui lui sont nécessaires, tout en tenant compte des problèmes posés par le manque de connaissances et la propension de ce dernier à effectuer des actions contraires à ses intentions et à la réalisation de ses buts, par ignorance essentiellement.

Notre système devra donc :

- Assister le novice dans l'utilisation du logiciel ;
- Lui faciliter l'acquisition des savoirs-faire nécessaires pour maîtriser l'utilisation du logiciel sans perturber la réalisation des tâches qui motivent son recours au logiciel ;
- Favoriser un apprentissage rapide, facile et durable en les ancrant dans l'activité de l'utilisateur et en mettant à profit les possibilités d'interaction avec lui [Carroll et Mack, 1984].

Concernant le dernier point, notre expérimentation vise à explorer l'utilisation de l'oral comme modalité en sortie. En effet, les modalités usuelles, texte et graphique, sont généralement utilisées dans les systèmes d'aide en ligne statiques pour résoudre les problèmes d'accès à l'information rencontrés par les utilisateurs. La parole a été peu utilisée jusqu'à présent comme modalité d'expression des messages d'aide alors qu'elle évite à l'utilisateur novice d'interrompre son interaction avec le logiciel, ce qui est le cas lorsqu'il doit lire des messages textuels.

3.1.3 Comportements exploratoires

Outre les problèmes de conception des systèmes d'aide en ligne et malgré les progrès réalisés grâce au paradigme de la manipulation directe dans le domaine de l'utilisabilité, on recense des difficultés d'utilisation dues au manque de connaissance des novices, et un coût cognitif supplé-

mentaire induit par le recours à l'aide qui peut s'avérer non négligeable [Kearsley *et al.*, 1988]. En effet, lorsque l'utilisateur fait appel à l'aide pour sortir d'une impasse, il doit interrompre la tâche en cours et garder en mémoire les informations contextuelles nécessaires à la reprise de son exécution. Les problèmes liés aux utilisateurs se répartissent en deux catégories : problèmes d'utilisabilité et problèmes de lisibilité.

Comment résoudre ces problèmes ? Pour tenter de les résoudre, il faut vraisemblablement les considérer individuellement et entreprendre des études spécifiques à leur égard. En effet, les caractéristiques individuelles des utilisateurs exercent une influence sur la nature des problèmes d'utilisation qu'ils rencontrent au cours de l'interaction. Il faut que l'aide agisse au niveau du fonctionnement du système utilisé et permette à l'utilisateur d'obtenir les connaissances nécessaires à la réalisation des tâches qu'il désire accomplir. Ainsi, ceci nous permet de distinguer deux catégories principales d'utilisateurs : les experts et les novices. Le comportement des experts se caractérise par la recherche de l'efficacité, et le comportement des utilisateurs novices se caractérise par un double paradoxe [Capobianco, 2002b] qui explique pourquoi l'exploration d'un nouveau logiciel représente pour la plupart d'entre eux une activité marginale et que l'optimisation de l'interaction entre rarement dans leurs préoccupations. Ce comportement peut expliquer pourquoi, selon [Duffy *et al.*, 1992], ces utilisateurs développent rarement une véritable maîtrise des logiciels qu'ils utilisent [Duffy *et al.*, 1989].

3.1.4 Conclusion

En conséquence, la réalisation d'un système d'aide ne se limite pas à la présentation d'informations pertinentes. Il importe de définir des modes d'accès à l'information efficaces, d'employer des modalités adéquates et d'un coût cognitif minimum, de façon à interférer le moins possible avec l'activité du novice. Si l'accès à l'aide est trop coûteux, l'utilisateur limitera ses consultations ; s'il est inefficace, il renoncera à utiliser le système d'aide. Ainsi, l'efficacité et la qualité de l'aide dépendent de deux facteurs : la pertinence des méthodes d'accès à l'information proposées, l'intelligibilité de l'information donnée. Ceci est donc le défi de notre sujet, pour éviter tout rejet du système pour l'utilisateur.

3.2 Méthode - Protocole expérimental

3.2.1 Présentation globale du protocole expérimental

L'expérience dont il est question ici a pour objectif d'évaluer l'usage de la multimodalité parole+graphique (en sortie) pour la présentation d'informations procédurales. Dans ce but, deux combinaisons de modalités sont exploitées pour effectuer un premier ensemble d'analyses comparatives dans un cadre d'étude prospective exploratoire. Lors de l'expérience les sujets ont pour objectif de réaliser un ensemble d'animations en utilisant le logiciel graphique Macromedia Flash 5. Ce logiciel combine une interface graphique de type WYSIWYG et un langage de script puissant permettant de réaliser non seulement des animations graphiques pouvant être mises en

ligne, mais également des pages web interactives incluant des fonctions avancées, formulaires ou pages dynamiques.

Les sujets sont confrontés successivement aux deux types de présentation multimodale, texte+graphique (T+G) et oral+graphique (O+G) dans un ordre de passation contrebalancé (cf. Annexe A). Le protocole expérimental comporte trois phases principales :

- Phase 1 : **Questionnaire et présentation de Flash** (environ 30 mn) L'expérience commence par un questionnaire qui nous sert à établir le profil des sujets (connaissances générales et expérience informatique principalement). Ensuite, une présentation de Flash sous la forme d'une quinzaine de diapositives PowerPoint leur est proposée afin qu'ils acquièrent les concepts de ce logiciel.

- Phase 2 : **Expérimentation** (environ 30 à 45 mn) Il s'agit pour les sujets de réaliser un ensemble de tâches, constituant un scénario d'utilisation possible du logiciel Macromédia Flash 5, à l'aide d'un ensemble de messages procéduraux d'assistance à la réalisation des animations pré-définies. Cette première partie de l'expérience a notamment pour objet de nous permettre d'étudier les stratégies visuelles d'exploration des documents, ainsi que les stratégies d'application des consignes adoptées par les sujets, en fonction des deux formes de présentation des informations d'aide.

- Phase 3 : **Post-test, questionnaire, entretien** (25 mn environ) Pour étudier l'impact des modalités employées sur l'apprentissage de l'utilisation du logiciel, nous distinguons deux types d'apprentissage :
  - Un apprentissage procédural (le savoir-faire) : pour cela, il est demandé aux sujets de réaliser un second scénario expérimental grâce au logiciel Flash. Cette seconde manipulation se fait sans recours à une aide en ligne ou à d'autres documents procéduraux, simplement sur la base des savoir-faire acquis au cours de la première manipulation. Les performances des sujets sont ensuite analysées pour pouvoir évaluer l'expérience acquise. Le scénario proposé reprend des notions vues au cours de la première manipulation et des notions voisines.
  - Un apprentissage conceptuel ou sémantique (les connaissances) : pour cela, nous avons recours à la méthode du questionnaire. Les sujets répondent à un certain nombre de questions qui ont pour objectif d'identifier l'ensemble des connaissances qui ont été conceptualisées au cours de la manipulation de Flash (cf. la phase 2 de l'expérimentation).

3.2.2 Choix de conception, mise en œuvre et déroulement

Ce qui va suivre détaille chaque phase de l'expérimentation en indiquant les choix effectués et les développements logiciels réalisés. Avant de présenter chacune d'entre elles, nous vous invitons à lire un résumé des informations que nous donnons oralement au sujet pour présenter l'expérimentation.

Nous informons le sujet de la sorte : "Vous allez participer à une expérimentation dont l'objectif est d'évaluer deux types de présentation multimodale d'informations procédurales. Pour

cela, nous allons vous demander de réaliser plusieurs tâches de création d'animations en utilisant le logiciel Flash. Pour vous aider à réaliser ces tâches, vous disposerez d'informations d'aide présentées sous deux formes différentes, à savoir, dans un ordre défini aléatoirement :

- des messages associant texte et graphique ;
- des messages associant parole et graphique.

Pour nous permettre de déterminer avec précision l'efficacité de ces deux formes de présentation des informations d'aide, vous serez équipé d'un dispositif mobile de suivi du regard. Ce dispositif, ou oculomètre (eye-tracker), est monté sur un casque."

Phase 1 : Questionnaire et présentation de Flash (environ 30 mn)

Questionnaire (environ 15 mn)

Tous les travaux sur l'apprentissage font ressortir clairement que le niveau d'expertise initial est capital dans l'efficience différentielle de tout apprentissage [Barth, 1993], cité par [Capobianco, 2002b]. De la même manière, les travaux traitant de la conception de documents procéduraux font apparaître les connaissances initiales des sujets comme une dimension primordiale lors de l'utilisation d'informations procédurales : le niveau d'expertise des sujets influence de façon significative les taux de réussite dans les tâches demandées, ainsi que la compréhension du document et les connaissances inférées. Aussi, dans le contexte de notre expérience, où les aspects différentiels des variables interindividuelles ne sont pas l'objet de l'étude, il convient de viser à la plus grande homogénéité possible des groupes expérimentaux. Cela revient d'abord à vérifier au maximum l'équivalence de leur niveau d'expertise en informatique et dans l'utilisation des outils graphiques standard (le logiciel Macromedia Flash 5 faisant appel à de nombreuses notions issues de l'infographie). Éventuellement, on peut choisir, parmi les sujets volontaires, ceux présentant les plus grandes similitudes du point de vue des aptitudes cognitives, de façon à restreindre au mieux les variables externes pouvant influencer l'impact des modalités de présentation. Cette option ayant été retenue, nous avons élaboré un dispositif d'évaluation permettant de contrôler ces variables interindividuelles. Nous avons donc prospecté un large échantillon de sujets potentiels. Seuls les sujets respectant les critères d'inclusion définis ci-dessus ont été retenus pour participer à l'étude.

De manière précise, un questionnaire a été rédigé. Il comporte, outre une introduction à l'objectif et/ou aux modalités de déroulement de l'expérience (voir l'extrait présenté au début de la section), un ensemble de questions permettant de fixer les informations nécessaires pour retenir ou rejeter le sujet en toute connaissance de cause. Sa structure est la suivante :

1. Présentation de l'expérimentation
2. Identification du sujet (Nom, Prénom, Age, Téléphone, Email, Numéro d'identification...)
3. Définitions des critères d'inclusion et d'exclusion
   - Familiarité avec les outil informatiques : Ici figurent les questions visant à exclure les sujets dépourvus des compétences informatiques minimales requises, et ceux ayant une trop grande familiarité avec les outils informatiques graphiques, ou tout simplement avec Macromédia Flash 5.

- Le niveau de connaissances générales et informatiques : Il semble qu'un niveau d'études équivalent, dans des domaines et disciplines technologiques, peut garantir une certaine homogénéité du point de vue des capacités cognitives et des compétences des sujets. Nous avons choisi une catégorie d'utilisateurs actuellement négligée par les concepteurs au profit du grand public ; il s'agit de celle des "experts" selon la classification des utilisateurs proposée par Shneiderman dans [Shneiderman, 1987].
- Sujets ayant déjà des connaissances en programmation et langage de script.
- Sujets n'ayant pas d'expérience des outils graphiques.
- L'âge : (donné par l'identification du sujet) : celui-ci pouvant avoir une influence au regard de la plasticité des processus cognitifs et de la motivation.
- Des sujets équivalent d'un point de vue cognitif : On peut également essayer de retenir un échantillon de sujets homogènes du point de vue des aptitudes cognitives pouvant être mises en œuvre dans l'utilisation de logiciels. Il s'agit alors de repérer ces aptitudes cognitives.

Dans l'utilisation de logiciels de type WYSIWYG, plusieurs dimensions cognitives peuvent être considérées : la mémoire de travail (les situations d'apprentissage mettant nécessairement en jeu des mécanismes de sélection et de mémorisation des informations extraites de l'environnement), la mémoire procédurale, la coordination visuo-motrice, l'orientation spatiale, le traitement de l'information symbolique (qu'il s'agisse de symboles verbaux ou iconiques, ces derniers s'avérant particulièrement importants), la planification des actions (définie par [Hoc, 1987], cité par [Capobianco, 2002b] comme la construction et/ou l'utilisation de représentations anticipatrices hiérarchisées pour guider l'activité permettant le contrôle selon [Richard, 1998], cité par [Capobianco, 2002b]), le raisonnement (correspondant à la production d'inférences), l'abstraction, etc...

Dans le cadre d'une étude exploratoire, nous n'avons pas tenu compte de ces facteurs de variabilité cognitive interindividuelle dont la mise en œuvre aurait nécessité le recours à des tests psychologiques lourds à réaliser.

4. Recueillir un certain nombre de renseignements complémentaires. Il peut également être utile de disposer d'informations complémentaires concernant la motivation des sujets. Nous avons retenu huit élèves-ingénieurs et étudiants du DEA Informatique (i.e., niveau Bac + 4 ou 5).

Présentation de Flash (environ 30 mn)

Pour une homogénéisation des connaissances, un tutoriel sous PowerPoint a été réalisé. Ce tutoriel aborde les différents concepts du logiciel à utiliser dont le sujet a besoin au cours de l'expérimentation. Ainsi tous les sujets disposent des connaissances de manipulation de base et des concepts fondamentaux du système (concepts du fonctionnement, noms des éléments de l'interface graphique, etc...), nécessaires pour réaliser les animations demandées. Les connaissances procédurales permettant la réalisation des animations seront fournies au cours de l'expérience. Elles constituent le contenu des messages d'aide.

Phase 2 : Expérimentation (environ 30 à 45 mn)

Calibrage (environ 10 mn, le temps de cette phase n'est pas fixe et varie selon les sujets)

A la suite de la présentation multimédia des concepts de base de Flash, le sujet entre dans la phase expérimentale proprement dite. Cette phase consiste d'abord à installer et à préparer le sujet. Ayant à notre disposition un système permettant d'enregistrer les mouvements oculaires (oculomètre modèle 501 d'ASL), il convient d'installer ce dispositif fixé à un casque, sur la tête de l'utilisateur.



FIG. 3.1 – Eye-tracker modèle 501 (cf. Annexe A).



FIG. 3.2 – Eye-head modèle 501.

En fait, une source infrarouge illumine l'oeil dont l'image est réfléchie sur un miroir translucide, le reflet est enregistré par une caméra fixée sur le casque. A noter la présence d'un capteur de tête magnétique (head-tracker) incorporé au casque, qui permet d'enregistrer les mouvements de la tête. C'est donc à partir de la détection de la pupille, du reflet cornéen et de l'intégration des données en provenance du capteur de tête, qu'un logiciel (fourni par le constructeur, e5win sous Windows) délivre la position du regard sur l'écran. Toutefois, avant d'acquérir les données, il est nécessaire de calibrer l'oculomètre, une fois celui-ci installé sur la tête de l'utilisateur. La procédure de calibrage consiste à faire regarder une mire à 9 points sur l'écran (mire couvrant au minimum 80% de l'écran). En faisant regarder successivement les points de la mire de 1 à 9, on acquériert pour chaque point l'orientation de l'oeil, ce qui permet d'établir la correspondance entre l'orientation de l'oeil et le point de regard.

Manipulation et conception (environ 20 mn)

La deuxième partie de l'expérimentation porte sur la manipulation de Flash (logiciel que j'ai sélectionné en fonction des objectifs de l'expérience) proprement dite. L'affichage (figure 3.3) est organisé comme suit : à gauche se trouve l'espace de travail Flash, à droite se trouve l'espace des messages d'aide en ligne.

L'évaluation des stratégies d'aide a été réalisée grâce à la comparaison entre une situation "T+G" et une situation "O+G" que l'on contrebalance entre deux groupes comprenant 4 sujets chacun. Le contenu informationnel des messages et les illustrations graphiques (captures

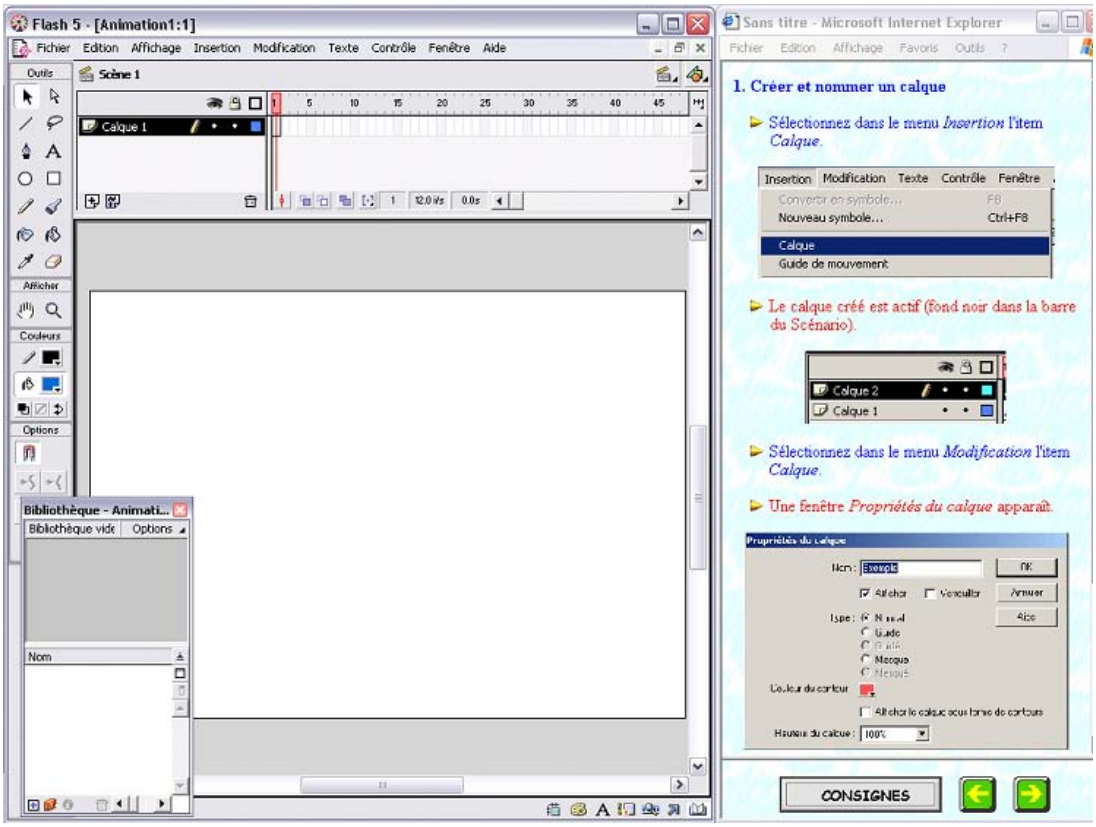


FIG. 3.3 – Interface utilisateur.

d'écrans) sont identiques dans les deux conditions, la seule différence étant la modalité d'expression des informations procédurales, lues oralement et enregistrées dans des fichiers audio (condition "O+G"), ou présentées sous forme textuelle au dessus des illustrations graphiques (condition "T+G"). Pour obtenir la meilleure équivalence possible entre modalités orale et textuelle, nécessaire pour assurer la validité de l'analyse comparative, nous avons tenu compte des points suivants :

- Obtention de la meilleure qualité audio possible ;
- Découpage des messages favorisant des messages audio courts ;
- Contenu identique des messages écrits et des messages oraux.

Les messages d'aide en ligne sont réalisés en HTML de façon à être lus facilement par n'importe quel navigateur Web. Afin de les construire nous avons suivi les principes de [Mayer, 1989] pour l'élaboration de documents explicatifs multimédias et les conseils de [Jamet et Heurley, 2002] pour la condition "O+G". Ces principes, résumés dans le tableau 3.1, ont servi de guide pour la conception des messages.

Les différentes zones des messages sont séparées comme on peut le voir sur la figure 3.4. Pour la condition "O+G" une applet a été réalisée qui crée un bouton et qui lit un son ".au" lorsque l'on clique sur celui-ci.

La navigation dans un message textuel se fait par l'ascenseur à droite et la navigation entre les différents messages se fait par les deux boutons flèches situés en bas de la fenêtre d'aide.

Principe	**Effet sur l'apprentissage et la compréhension**
Principe multimédia	Effet positif de la présence d'illustrations
Principe de la contiguïté spatiale	Effet positif de la proximité physique des sources d'informations visuelles
Principe de la contiguïté temporelle	Effet positif de la proximité temporelle des sources d'informations visuelles et orales
Principe de modalité	Effet positif de l'utilisation de la modalité orale pour expliquer une source d'information visuelle
Principe de cohérence	Effet positif de la suppression d'informations non pertinentes pour l'apprentissage
Principe de redondance	L'utilisation d'informations redondantes dans des modalités différentes peut avoir un effet négatif sur l'apprentissage
Principe des différences interindividuelles	Les principes évoqués précédemment sont liés à des aptitudes qui varient d'un individu à l'autre

TAB. 3.1 – Principes de Mayer.

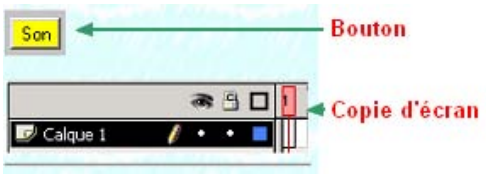


FIG. 3.4 – Exemple de message dans la condition "O+G".

Le bouton consigne, lui, ouvre une fenêtre occupant tout l'écran qui renseigne l'utilisateur sur l'animation à réaliser, les procédures qu'il doit effectuer et l'affichage du résultat qu'il doit obtenir.

On obtient les deux conditions suivantes pour les messages d'aide :

1. Condition textuelle : Chaque message textuel est illustré par une copie d'écran placée sous le message ; en bleu, les actions, en rouge le résultat attendu.
2. Condition orale : Chaque message est divisé horizontalement en zones (séparées par un trait horizontal noir). Chaque zone comprend une ou plusieurs copies d'écrans illustrant les informations procédurales (action + résultat) contenues dans les messages oraux. Pour "jouer" ces messages, il faut cliquer sur le bouton placé à gauche en début de zone, au dessus des copies d'écrans.

En ce qui concerne la manipulation elle-même, les sujets avaient à réaliser deux animations, la première faisant déplacer un ver dans un décor déjà créé, la deuxième consistant à créer un décor, puis à faire se déplacer une abeille suivant une trajectoire.

Pour permettre la réalisation de ces animations, un découpage a été réalisé en tâches elles-mêmes subdivisés en sous-tâches ou étapes. Ces notions de tâche, sous-tâche et étape sont intégrées à l'aide : lorsque l'on appuie sur le bouton "Consignes", deux fenêtres recouvrent la totalité

de l'écran, celle de droite reproduit l'animation à réaliser, celle de gauche décrit brièvement les différentes étapes/tâches à effectuer ; une couleur différente permet à l'utilisateur d'identifier la tâche et l'étape qu'il est en train de réaliser. Ainsi, le scénario 1 se décompose en trois tâches et le scénario 2 en quatre tâches.

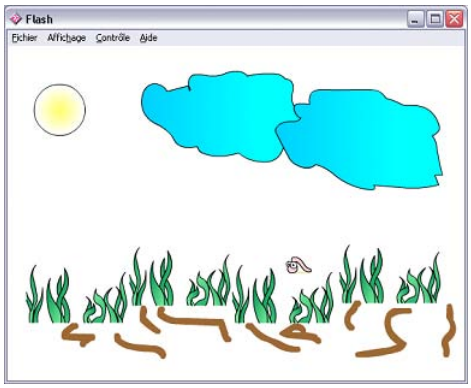


FIG. 3.5 – Résultat du scénario 1.



FIG. 3.6 – Résultat du scénario 2.

Parmi les huit sujets qui ont participé à l'expérimentation, quatre ont effectué le scénario n°1 en condition "T+G" et le scénario n°2 en condition "O+G". Quant aux quatre autres, ils ont d'abord expérimenté la condition "O+G" (scénario n°1) puis la condition "T+G" (scénario n°2).

Afin de stocker les données expérimentales, j'ai réalisé deux logiciels qui tournent en tâche fond sur la station de travail :

1. Le premier permet de stocker les données oculaires provenant de l'oculomètre via la liaison série. La sauvegarde se fait dans un fichier texte nommé "eye.txt". On stocke d'abord la date puis, à la suite, les différentes données fournies, pour chaque échantillon, par l'oculomètre à une fréquence de 60 Hz. Les données reçues sont décodées et stockées de la sorte : diamètre de la pupille, position horizontale en pourcentage par rapport à la dimension de l'écran et position verticale en pourcentage par rapport à la dimension de l'écran.
2. Le deuxième permet d'effectuer des captures d'écran des deux espaces de travail présentés précédemment (application écrite en Java). Ces captures sont sauvegardées dans un répertoire et compressées afin de minimiser la place sur le disque dur (format jpeg). Une capture d'écran est réalisée dès que l'affichage change, en d'autres termes, dès que l'utilisateur réalise une action.

Phase 3 : Post-test, questionnaire, entretien

Post-test (environ 15 mn)

Pour mesurer l'efficacité relative des deux types de message d'aide, il est ensuite demandé aux sujets de réaliser deux scénarios qui utilisent les notions présentées dans le tutoriel initial et qui reprennent les notions utilisées (au moins pour la plupart d'entre elles) lors de la manipulation/utilisation du logiciel. Au cours de ce test, les sujets ne sont plus munis de l'eye-tracker. Ils n'ont aucune possibilité d'accès à une quelconque aide en ligne. De plus, ils sont totalement autonomes dans la mesure où on ne leur fournit aucune consigne ni aucune indication sur la

procédure à suivre pour réaliser l'animation choisie. Nous avons fixé le temps de réalisation à 15 mn pour chacun des deux scénarios afin de limiter les activités de type exploratoire ou le recours à des approches de type "essai-erreur". Pour ce post-test, deux scénarios ont été conçus (cf. Annexe A). Le premier utilise l'interpolation de mouvement suivant une trajectoire. L'objet à animer est un moustique et, le décor est une plage. Le deuxième utilise la notion de calque masque et de calque masqué. On fait se déplacer un cercle sur un décor et seul l'intérieur de cercle affiche le contenu du décor. La notion de calque masqué est présentée lors de la familiarisation avec Flash mais ne fait l'objet d'aucune mise en œuvre pendant la phase de manipulation du logiciel. Ce scénario était destiné à tester une des dimensions du processus d'apprentissage : la capacité à mettre en œuvre de façon procédurale des connaissances conceptuelles.

Les données suivantes ont été retenues à des fins d'analyse : l'ordre de passage du sujet (T+G puis O+G, ou le contraire) et les taux de réussite aux tâches (comptage du nombre de tâches prescrites réalisées avec succès, et du temps de réalisation des tâches et des scénarios).

Pour stocker ces données post-expérimentales, nous faisons appel au même logiciel que celui utilisé lors de l'expérimentation pour effectuer les captures d'écran, et conserver la trace de l'activité des sujets.

Questionnaire (environ 5 mn)

Afin de connaître les impressions et jugements sur les deux types de présentation des messages d'aide, un questionnaire a été réalisé. Il comporte trois parties : une partie par scénario et une troisième plus générale, concernant les impressions des sujets, leurs observations, leurs réactions ... Ce questionnaire a permis de comparer les préférences des sujets vis-à-vis des deux formes de présentation de l'aide et d'obtenir des informations sur les améliorations éventuelles des modes de présentation des informations procédurales, souhaitées par les sujets. Nous avons pris soin de préciser aux sujets, avant de leur donner le questionnaire, que l'évaluation portait non pas sur leurs performances, mais bien sur celles des formes de présentation des informations d'aide, pour ne pas provoquer de stress lié à l'évaluation.

Entretien (environ 10 mn)

Au cours de cet entretien, nous avons recueilli leurs impressions concernant l'expérimentation. Il s'agit d'un bref "débriefing" ou entretien semi guidé. Les principaux points abordés portent sur les préférences individuelles des sujets (condition "O+G" versus "T+G"), les problèmes spécifiques rencontrés par chacun d'eux dans la réalisation des animations, ainsi que leurs impressions et suggestions éventuelles.

3.3 Analyses - Résultats

Nous présentons les critères que nous avons retenus pour évaluer les performances des sujets puis nous évaluons l'influence du type d'aide utilisé sur leurs performances.

3.3.1 Critères

Un système d'aide à l'utilisateur a pour objectif de permettre aux utilisateurs novices d'atteindre leurs buts (i.e, de réaliser les objectifs qui motivent leurs recours au logiciel). Les performances des sujets ont été évaluées en fonction du taux de réussite observé lors de la réalisation des tâches. Nous avons relevé au cours de l'expérience de nombreuses erreurs causées par une compréhension ou une représentation erronée de la tâche demandée. Ainsi, dans la mesure où la tâche demandée n'a pas été réalisée, ces cas ont été comptabilisés comme des échecs de la part des sujets.

3.3.2 Analyse des données expérimentales

Nous abordons cette partie en quatre points successifs :
- Présentation des groupes : répartition des sujets en fonction des situations expérimentales,
- Résultats issus de l'analyse des temps de réalisation,
- Résultats issus de l'analyse des taux de réussite,
- Résultats issus de l'analyse des données propres à chaque sujet (fixations et diamètre pupille).

Présentation des groupes

Les 8 sujets volontaires ont été aléatoirement répartis en deux groupes. De la même manière, les scénarios ont été répartis en deux groupes, le premier scénario étant associé à un système d'aide, et le deuxième scénario à l'autre système. De la sorte, chaque utilisateur a pu tester les deux systèmes d'aide, selon l'ordre suivant :

	Groupe 1	Groupe 2	
Scénario 1	"T+G"	"O+G"	Groupe 1 : Sujets 1, 2, 3, 4
Scénario 2	"O+G"	"T+G"	Groupe 2 : Sujets 5, 6, 7, 8

TAB. 3.2 – Répartition des sujets en fonction des situations expérimentales.

Le premier groupe (Groupe 1) a eu pour support d'aide d'abord le système "Texte+Graphique" puis le système "Oral+Graphique". Le second groupe (Groupe 2) a eu pour support d'aide d'abord le système "Oral+Graphique" puis le système "Texte+Graphique".

Temps de réalisation

Les temps de réalisation des scénarios sont donnés en minutes.

On considère que ces temps représentent principalement le temps consacré par chaque sujet :
- à la réalisation de chaque scénario,
- à la réalisation de chaque tâche,
- à l'observation visuelle de chacune des zones de l'affichage définies lors de l'analyse des données oculométriques. concrètement, nous avons défini les zones suivantes : Flash, Aide

ou Consignes et, à l'intérieur de ces zones, plusieurs sous-zones. Pour Flash, les sous-zones sont : le menu, la barre des outils, le scénario, la scène et le reste. Pour l'Aide, les sous-zones sont : le contenu du message d'aide et la barre des boutons.

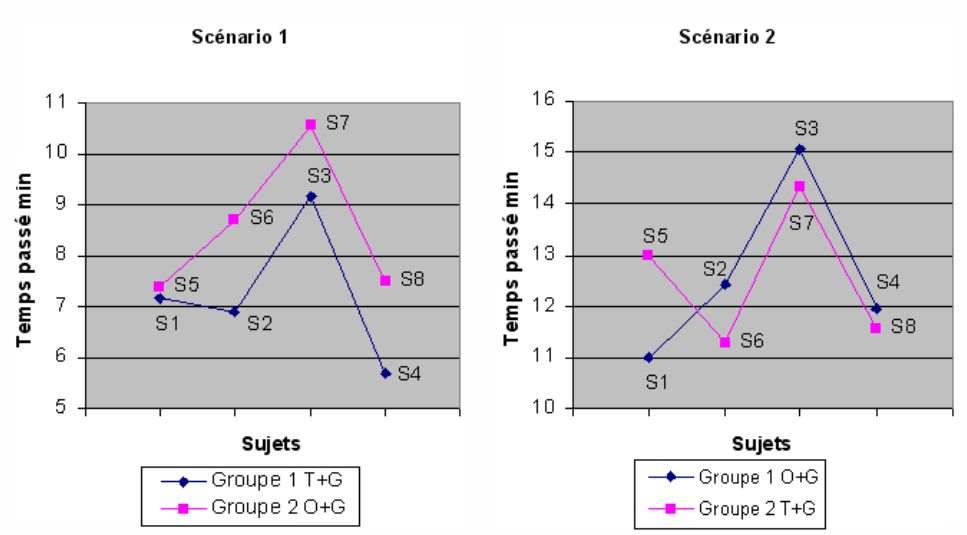


FIG. 3.7 – Temps passé par les sujets à la réalisation de chaque scénario (Temps en min).

De manière générale, on observe que les temps obtenus avec le système "T+G" sont plus courts que ceux obtenus avec le système "O+G" (voir figure 3.7).

Ainsi, si l'on établit une moyenne sur les deux scénarios, on s'aperçoit que les sujets passent moins de temps à la réalisation des tâches avec le système "T+G". Ceci peut être dû au fait que la parole introduit une forte charge cognitive et/ou au fait que, lorsque l'on cherche une information procédurale spécifique ou ponctuelle, il suffit de survoler la présentation textuelle pour la trouver, alors qu'il faut réécouter le message dans sa totalité.

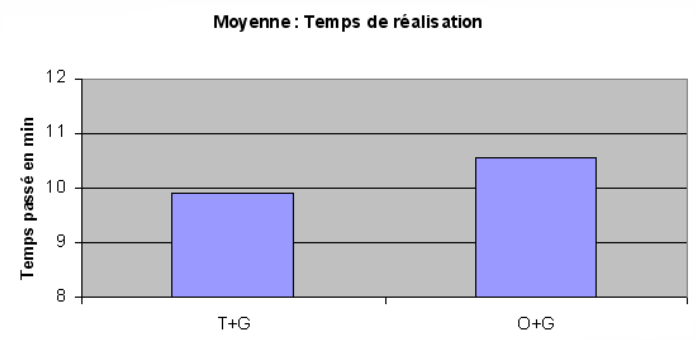


FIG. 3.8 – Temps moyen (en min.) passé dans chaque condition (T+G et O+G) pour les deux groupes confondus.

Pour tenter de comprendre ces temps de réalisation et raffiner l'analyse, nous avons étudié le temps passé, respectivement, dans la zone de Flash et dans la zone de l'aide. Nous présentons

uniquement les moyennes globales des deux groupes confondus pour les deux conditions afin de ne pas surcharger cette partie (voir la figure 3.8). Comme on peut le constater, le temps de réalisation du scénario dans la condition "T+G" est supérieur à celui de la condition "O+G". Ceci confirme ce que les sujets nous ont signalé dans l'entretien de debriefing : *"L'avantage de l'oral, c'est de permettre d'écouter en faisant des actions. (Sujet n°3)"* ou *"L'oral est répétitif. A force ça devient verbeux. Les écrans peuvent suffire ; c'est faisable mais c'est peut-être dur (Sujet n°6)"*.

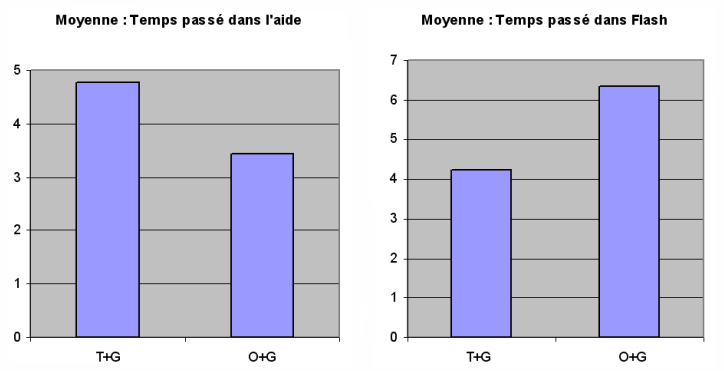


FIG. 3.9 – Données oculométriques : Temps moyen (en min.) passé dans chaque zone, aide et Flash, tous sujets confondus.

Quant aux résultats présentés dans la figure 3.9, ils suggèrent que les sujets accordent moins d'attention aux messages d'aide oraux qu'aux messages textuels, soit parce qu'ils mettent en œuvre les informations procédurales des messages oraux au fur et à mesure, soit parce qu'ils utilisent surtout les copies d'écran et répugnent à rejouer les messages sonores. La seconde hypothèse semble la plus plausible, compte tenu des temps de réalisation supérieurs observés avec le système d'aide "O+G" (voir la figure 3.8).
Enfin, il nous a paru intéressant de donner les temps moyens de réalisation des tâches (en minutes) pour les raisons suivantes : ces données indiquent, d'une part, la condition la mieux adaptée à chaque tâche et, d'autre part, la difficulté et la longueur relatives de chacune d'elles indépendamment de la condition. Encore une fois, on observe que majoritairement la condition "T+G" est plus rapide que la condition "O+G". Cependant pour les tâches très courtes comme S1T3 ou S2T4, les performances de l'une ou de l'autre condition semblent équivalentes. De même pour les tâches simples comme S1T1 ou S2T1 (où on demande de renommer ou de créer un calque), on s'aperçoit que les performances sont très voisines. La seule véritable distinction se fait au niveau des tâches où la difficulté est accrue comme les tâches S1T2 ou S2T2 (manipulation du calque guide de mouvement, insertion d'image clé, interpolation de mouvement). On notera que pour S1T2, "T+G" conduit à des performances meilleures que "O+G", tandis que pour S2T2on observe le contraire.

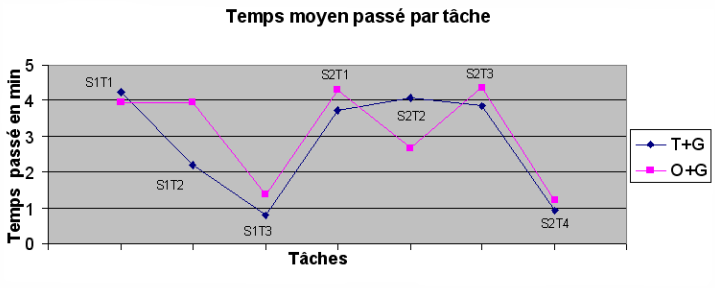


FIG. 3.10 – Temps moyen (en min.) passé dans chacune des tâches tous sujets confondus.

Taux de réussite

L'évaluation du taux de réussite est assez délicate au cas particulier. En effet, les scénarios sont composés de tâches non indépendantes et l'on ne peut demander au sujet de stopper ses actions s'il se trompe dans une tâche, car cela lui donnerait l'impression de n'avoir aucune liberté d'utilisation du logiciel. Pour tenter d'évaluer la réussite de chaque sujet, nous les avons laissé manipuler et lorsqu'ils s'apercevaient qu'ils s'étaient trompés ou qu'ils étaient bloqués l'expérimentateur entrait en action. Il leur expliquait ce qui n'allait pas et les remettait dans de bonnes conditions pour continuer le scénario. Ainsi, pour pouvoir noter leurs travaux, nous avons décidé de noter les deux scénarios et de soustraire des points à chaque erreur commise. On obtient le tableau suivant :

	S1	S2	S3	S4	S5	S6	S7	S8	Moyenne
Scénario 1	100	100	60	60	100	60	80	100	82,5
Scénario 2	100	100	60	80	60	60	100	80	80
1 erreur = -20 %									

TAB. 3.3 – Pourcentages de réussite par scénario et par sujet.

Ceci ce traduit en terme de condition "T+G" et "O+G" par le graphique suivant :

Le tableau 3.11traduit le fait que la condition "O+G" semble plus performante que la condition "T+G", toutefois, comme nous l'avons signalé auparavant ceci n'est qu'une approximation du fait de la difficulté de définir des scores de réussite. De plus, les résultats sont très proches, et pour obtenir des résultats plus fiables et éventuellement réaliser une analyse statistique, il faudrait un nombre très supérieur de sujets.

Données propres à chaque sujet

Dans le but de compléter notre analyse, nous avons comptabilisé le nombre de fixations oculaires d'une durée comprise entre 100 et 250 ms (activité perceptive) et le nombre de fixations supérieur à 250 ms (activité cognitive). On obtient les résultats de la figure 3.12.

Ces résultats suggèrent que, pendant la condition "O+G", la charge cognitive des sujets était plus forte que pendant la condition "T+G", étant donné le nombre de fixations supérieur observé

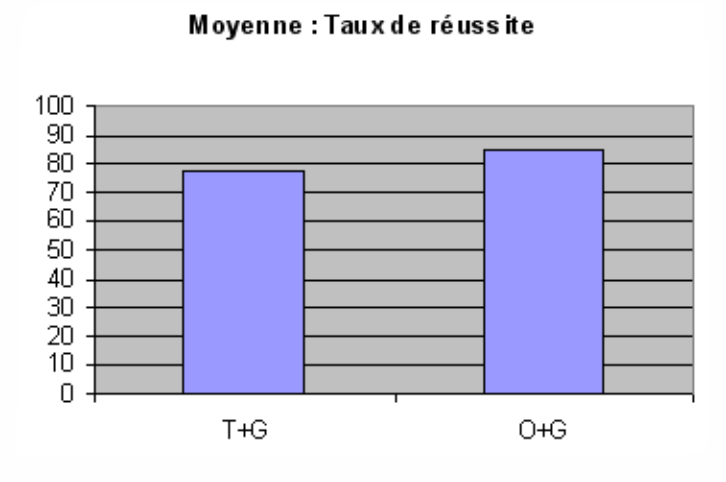


FIG. 3.11 – Pourcentage global de réussite par condition.

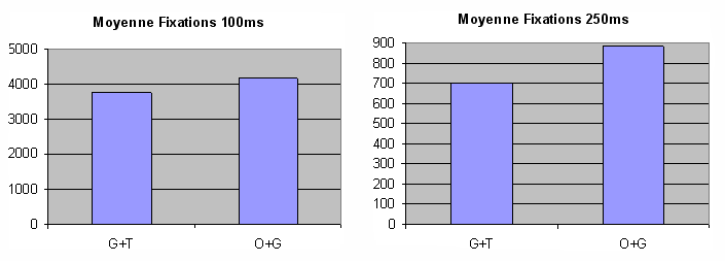


FIG. 3.12 – Nombres moyens de fixations "perceptives" et "cognitives" pendant chaque condition.

pendant la condition "O+G". On peut interpréter cette différence comme une confirmation de l'hypothèse selon laquelle les sujets auraient moins utilisé les messages d'aide oraux que les messages d'aide textuels, ce qui aurait rendu plus difficile et plus longue la réalisation des animations dans le contexte de la condition "O+G".

3.3.3 Analyse des données post-expérimentales

Les données fournies par les questionnaire finals, les deux scénarios du post-test et les entretiens de debriefing permettent de tester l'apprentissage du logiciel et d'analyser les préférences des utilisateurs.

Données du post test

Scénario 1 du post-test

Le scénario 1 du post-test est identique au scénario 2 réalisé pendant l'expérimentation. Les tâches à effectuer sont strictement identiques, hormis le fait que les objets à manipuler sont différents et que les sujets ne disposent d'aucune aide en ligne.

	S1	S2	S3	S4	S5	S6	S7	S8	Moyenne
Scénario 2	100	100	60	80	60	60	100	80	80

TAB. 3.4 – Taux de réussite du scénario 2 réalisé pendant l'expérimentation.

	S1	S2	S3	S4	S5	S6	S7	S8	Moyenne
Scénario 1	100	100	100	100	80	80	60	100	90

TAB. 3.5 – Taux de réussite du scénario 1 réalisé pendant le post-test.

On observe, en comparant les données des tableaux 3.4 et 3.5, que l'apprentissage est plutôt satisfaisant car, pour les mêmes tâches à accomplir, le taux de réussite des sujets a augmenté, sauf pour l'un d'eux, en l'absence de toute aide en ligne. Cela dit, pour être véritablement concluant, le post-test devrait être effectué plusieurs jours après l'expérimentation. L'effet observé peut dénoter une simple mémorisation des procédures plutôt qu'une véritable acquisition de connaissances procédurales, en raison de la proximité temporelle entre l'expérimentation et le post-test.

Scénario 2 du post-test

Le scénario 2 du post-test aborde des concepts non utilisés lors de l'expérimentation, mais décrits dans le tutoriel par un exemple et tels que leur mise en œuvre peut être inférée facilement à partir des deux scénarios effectués pendant l'expérimentation. Les sujets ne disposent d'aucune aide en ligne. Les taux de réussite sont présentés dans le tableau 3.6.

	S1	S2	S3	S4	S5	S6	S7	S8	Moyenne
Scénario 2	80	80	100	80	80	100	80	80	85

TAB. 3.6 – Taux de réussite du scénario 2 réalisé pendant le post test, par sujet.

On observe un taux de réussite relativement élevé pour ce scénario, légèrement supérieur à ceux obtenus pour les scénarios réalisés pendant l'expérimentation (80 et 82,5% respectivement. ce qui indique que l'expérimentation a permis aux sujets de maîtriser les principaux concepts de Flash et de transférer les compétences acquises pendant l'expérimentation pour acquérir et mettre en œuvre des concepts voisins présentés rapidement dans le tutoriel, sans le secours d'aucune aide en ligne.

Données des questionnaires et des entretiens de debriefing

Les informations fournies par l'analyse des questionnaires sont rassemblées dans le tableau 3.7. Ces données sont exprimées en pourcentages par rapport aux réponses que nous ont fournies les sujets qui disposaient d'échelles (à 5 cases) pour exprimer leurs jugements. Pour les catégories "conception messages" et "contenu messages", nous avons moyenné les réponses aux différents critères proposés pour caractériser les messages d'aide.

Les premiers résultats qui ressortent de ce tableau sont que 25% des sujets préfèrent la condition "O+G", 62.5% préfèrent la condition "T+G" et que 12.5% n'ont pas de préférence. On observe également que les sujets trouvent la condition "O+G" appropriée à 75% contre 87.5%

	S1	S2	S3	S4	S5	S6	S7	S8
T+G est approprié	100	100	100	75	75	75	100	75
Conception messages	66	75	66	58	58	58	66	25
Contenu messages	91	100	66	75	75	75	75	75
O+G est approprié	75	50	75	100	75	50	100	75
Conception messages	41	50	66	50	66	58	66	58
Contenu messages	83	91	50	75	75	66	91	75
Préfère T+G	X	X	X			X		X
Préfère O+G				X			X	

TAB. 3.7 – Taux de réussite du scénario 2 réalisé pendant le post-test.

pour la condition "T+G". La conception des messages est jugée de manière similaire pour les deux conditions (57% contre 59%), ceci est aussi le cas pour le contenu des messages (76% contre 79%) ; on observe toutefois un léger avantage pour la condition "T+G".

3.4 Résumé

Nous avons mis en place un protocole expérimental qui nous a permis, lors de cette expérimentation préliminaire, d'obtenir des informations sur la mise en œuvre de systèmes d'aide en ligne multimodaux susceptibles de répondre aux besoins et attentes des utilisateurs novices issus du grand public.

L'analyse des résultats a montré que, globalement, les messages "Texte+Graphique" ne s'avèrent pas vraiment plus efficaces que les messages "Oral+Graphique" même si les temps de réalisation de certaines tâches sont légèrement meilleurs dans la condition "T+G" que dans l'autre condition. En effet, si les tâches sont relativement courtes et non complexes, les performances dans les deux conditions sont semblables. En revanche, pour les tâches plus longues et plus complexes, il apparaît que la condition "T+G" l'emporte nettement sur l'autre en terme de rapidité. On peut avancer deux interprétations complémentaires pour rendre compte de cette observation. La parole pourrait accroître la charge cognitive des sujets ; en outre, les sujets n'étaient pas habitués à la présentation orale des informations d'aide, ils la découvraient. En revanche, l'analyse des taux de réussite suggère que la condition "Oral+Graphique" serait plus efficace que l'autre ; ceci peut s'expliquer éventuellement par le fait que les sujets, dans cette condition, pouvaient se concentrer davantage sur la réalisation des animations ; ils n'avaient pas à l'interrompre pour consulter l'aide. L'analyse des données oculométriques confirme cette interprétation.

Concernant les préférences des sujets, on s'aperçoit que les systèmes d'aide leur ont paru dans l'ensemble bien conçus. Cependant, la majorité d'entre eux ont préféré les messages "T+G". Les débriefings suggèrent des pistes que nous expérimenterons ultérieurement pour palier aux inconvénients de l'oral par rapport au texte. Certes, des consignes orales autorisent à interagir avec le logiciel en même temps qu'on les écoute, alors que des informations procédurales sous

forme textuelle imposent une interruption de l'activité. Mais la présentation textuelle facilite la recherche d'une information ponctuelle, tandis qu'il faut réécouter entièrement un ou plusieurs messages oraux pour trouver l'information précise dont on a besoin. Une solution que nous avons expérimentée dans le cadre de l'étude présentée dans le chapitre 4 consiste à raccourcir les messages oraux en les découpant en sous-messages contenant chacun un seul type d'information d'aide (e.g., procédurale, explicative, didactique), et en associant à chaque type d'information un bouton de couleur différente. A noter que tous les sujets se sont montrés très intéressés par cette expérimentation, et que le port de l'eyetracker ne leur a occasionné aucune gêne ; plusieurs d'entre eux ont même "oublié" le casque.

3.5 Impacts sur le programme de recherche

Pour illustrer cette section, nous commençons par présenter le recueil des impressions des sujets à l'issue de l'expérimentation des deux systèmes d'aide. Pour chaque sujet, nous leur avons demandé ce qu'ils pensaient des modalités utilisées (Texte et Graphique ou Graphique et Oral), ainsi que les améliorations qu'ils souhaiteraient voir apportées aux messages d'aide. Nous citons les remarques les plus représentatives pour les deux thèmes abordés :

- `Jugements sur les modalités :`
  <"Les messages sonores sont trop atomisés (i.e., l'information est trop délayée). C'est fastidieux pour les procédures simples qui se répètent au cours de l'expérimentation, voir par exemple 'Créer un calque'." Ce sujet a "testé" tous les messages sonores pour être sûr de ne pas manquer une information importante. Ceci explique pourquoi il préfère les messages texte + graphique. Toutefois, selon lui, les messages oraux permettent d'éviter les va-et-vient entre Flash et l'aide.>(Sujet 1)

  <Ce sujet préfère Texte+Graphique, "car c'est le mode de présentation familier. L'avantage de l'oral, c'est de permettre d'écouter en faisant des actions. Mais on ne regarde pas les copies d'écran et on a des informations incomplètes. En revanche, avec l'écrit, on lit et on regarde.">(Sujet 3)

  <Oral+graphique : "C'est bien pour la première fois, la première utilisation, quand on découvre le logiciel, quand on sait pas comment faire. On va plus vite, on a l'impression qu'il y a quelqu'un qui vous aide. [...] On peut pas faire toute l'aide à la voix. C'est bien, juste pour les manipulations de base, la première fois. C'est pas pour les utilisateurs avancés" car le texte peut se lire en diagonale en utilisant les mots-clés et facilite la recherche d'informations obtenues antérieurement ou complémentaires (points précis, détail). "La parole est linéaire, l'oeil va plus vite que l'oreille.">(Sujet 4)

  <Pour une première approche l'oral convient. Pour chercher "ce qu'on veut quand on veut" le texte est préférable. Ce sujet a apprécié le passage de l'oral à l'écrit.>(Sujet 5)

<"Il fallait écouter les messages jusqu'au bout. Avec l'écrit les allers-retours sont possibles." On pourrait imaginer "un long fichier audio avec des balises écrites pour se rappeler l'endroit que l'on veut réécouter, par exemple on pourrait représenter le fichier audio par une règle que l'on pourrait faire défiler." Ce sujet éprouve de la difficulté à manipuler en même temps qu'il écoute l'audio car l'audio est trop lent, d'où il n'écoute plus, continue à interagir avec Flash et "se plante".>(Sujet 6)

<"L'oral permet de faire l'action en même temps qu'on écoute." Ce sujet préfère la modalité orale au texte. Il note cependant qu'il a peu regardé les copies d'écran : cela vient peut-être de sa façon de prendre en main un logiciel par exploration.>(Sujet 7)

<"L'avantage de l'oral : on peut exécuter en même temps. Mais le problème c'est que il faut pas être trop en avance, trop en retard. L'oral impose le rythme. Il faut un truc mixte : un oral qui prend en compte les actions de l'utilisateur." Ce sujet donne un exemple : le système suggère de sélectionner un item spécifique dans un menu ; l'utilisateur ne le fait pas ; le système reformule l'énoncé et augmente l'insistance au fur et à mesure des répétitions.>(Sujet 8)

- `Améliorations souhaitées :`
  <Oral+Texte : Quand les messages comportent plusieurs étapes, enchaîner les énoncés en introduisant une pause après chaque étape ; ainsi, l'utilisateur n'est pas obligé de cliquer. On peut aussi, si l'utilisateur juge l'enchaînement trop rapide, lui permettre de régler les pause manuellement. Le contenu des messages d'aide est "approprié à la situation".>(Sujet 1)

  <Ce sujet a dû revenir sur certains messages, "mais c'est purement personnel, je lis trop vite, je veux faire trop vite. Les interfaces sont les mêmes pour tous les logiciels, on pense que c'est pareil mais il y a des petites différences." Suggestion : pour les opérations de base, créer des vidéos (parole + graphique animé) où on voit de A à Z la réalisation d'une opération. "C'est ce qui se passe de plus en plus pour les gros logiciels" (ADOBE, voir par exemple le montage d'une carte mère). Ensuite, Texte+Graphique est préférable.>(Sujet 4)

  <"Les messages étaient précis. Ça m'a suffit.">(Sujet 8)

Certains sujets nous ont ainsi suggéré des solutions pour faciliter la recherche d'informations dans les messages oraux, et remédier aux limites d'accès aux informations imposées par le caractère linéaire inhérent aux présentation orales.
Ainsi, à la suite de cette étude, j'ai centré ma recherche sur le test de nouvelles techniques qui facilitent effectivement l'accès aux informations d'aide, ainsi que leur assimilation. Plus précisément, j'ai expérimenté :

- L'utilisation d'un ACA, dont la présence ajoute une dimension affective d'assistance aux messages oraux, dimension susceptible d'accroître l'intérêt des novices pour l'aide et de leur faciliter la prise en main d'un nouveau logiciel en les rassurant et réduisant leur anxiété.

- La mise en œuvre d'une aide adaptative qui tienne compte des connaissances acquises par l'utilisateur courant au fur et à mesure de l'interaction et, plus généralement, de son profil cognitif. Cette adaptation dynamique du contenu et de la présentation des messages à l'utilisateur est de nature à accroître l'efficacité de l'aide.

Les compétences acquises via cette étude préliminaire nous ont servi de référence pour la conception des deux principales études expérimentales que j'ai réalisées ensuite. Elles ont permis, en particulier, d'améliorer :

- la conception du tutoriel et des messages d'aide,
- les méthodes d'analyse des données oculométriques,
- les scénarios d'utilisation de Flash proposés aux sujets et la conception du post-test.

Chapitre 4

Démarche, méthode et développement logiciel

Nous avons adopté une démarche expérimentale pour étudier l'apport d'Agents Conversationnels Animés (ou ACAs) ainsi que les réactions d'utilisateurs face aux interfaces adaptatives. Ce chapitre retrace les parties communes des démarches expérimentales que nous avons adoptées en vue du développement logiciel nécessaire à la réalisation des deux études envisagées. Ainsi, l'objectif n'est pas de présenter la mise en œuvre du programme expérimental de chaque étude mais de décrire la construction logicielle destinée à répondre aux besoins des expérimentations. A noter que la méthodologie et le protocole expérimental seront détaillés pour chaque étude dans les chapitres suivants.

Nous commençons par décrire la tâche expérimentale choisie, commune aux deux études. Puis, nous décrivons les besoins logiciels et les choix retenus pour chaque étude, notamment la conception du système d'aide, la forme de multimodalité utilisée pour les messages d'assistance, les données à recueillir et le déroulement d'une session. Nous terminons ce chapitre par une présentation plus technique des outils logiciels développés, à savoir, une plate-forme logicielle qui permet d'acquérir la trace des interactions et de simuler la technique du magicien d'Oz, ainsi qu'un logiciel de "rejeu" et d'analyse des données enregistrées.

4.1 Tâche expérimentale

Le domaine d'application retenu pour la thèse étant l'aide en ligne, nous avons choisi l'aide à l'utilisation d'un logiciel grand public nouveau pour l'utilisateur. Ainsi le système d'aide créé intègre, dans une situation d'apprentissage interactif, soit un ACA, soit une technique d'adaptation dynamique au profil de l'utilisateur.

Le logiciel retenu, est Flash, logiciel grand public de création d'animations multimédias. A noter que la version utilisée est plus évoluée que celle de l'étude préliminaire : l'interface utilisateur a été totalement remaniée. Toutefois, nous décidons de poursuivre avec ce logiciel,

en raison de l'expérience acquise via l'étude préliminaire, mais aussi des nombreux concepts originaux qu'offre ce logiciel.

Les participants aux expérimentations se familiarisaient d'abord avec les principaux concepts de Flash (i.e., connaissances "sémantiques" à l'exclusion de toute information procédurale), grâce à un court tutoriel multimédia (sous PowerPoint) qu'ils parcouraient à leur rythme. Ensuite, la tâche expérimentale consistait à réaliser trois scénarios d'animation pré-définis qui faisaient intervenir les principaux outils de création d'animations de Flash, à l'exclusion du langage d'écriture de scripts, jugé trop difficile pour des participants non informaticien. Les scénario 1 et 2 portent sur la construction d'un aquarium ; dans un premier temps l'utilisateur crée le décor (figure 4.1) à l'aide d'objets graphiques appartenant à une bibliothèque, puis il anime les poissons (figure 4.2) grâce à la technique d'interpolation de mouvement que propose Flash. Le scénario 3 (figure 4.3) quant à lui consiste à élaborer l'introduction d'un site Web. Ce scénario ne requiert que quelques connaissances procédurales supplémentaires par rapport aux deux premiers, ce qui permet d'évaluer l'assimilation et la mémorisation à court terme des connaissances procédurales impliquées dans la réalisation des deux premiers scénarios.

La nature des expérimentations envisagées conduit aux contraintes suivantes sur les choix de conception de la plate-forme logicielle.

Implication sur la conception de la plate-forme

La plate-forme logicielle doit présenter une certaine généricité. En particulier, pour être réutilisable, elle doit pouvoir fonctionner avec un logiciel quelconque, application propriétaire ou logiciel grand public.



FIG. 4.1 – Décor Aquarium.



FIG. 4.2 – Animation Poissons.



FIG. 4.3 – Site Web.

4.2 Démarche - Méthode

4.2.1 Conception du système d'aide

Durant l'utilisation de Flash, les utilisateurs étaient accompagnés d'un système d'aide de consultation de messages d'aide, qui exploite la technique du magicien d'Oz. Leurs moyens d'interaction avec le système se limitaient au clavier et à la souris, de façon à réduire les aspects innovants à la présence d'un ACA (première étude) ou à l'adaptation des messages d'aide au pro-

fil de l'utilisateur courant (seconde étude). Ainsi, pour interagir avec le dispositif, les participants disposaient de quatre boutons pour caractériser le type de leur question.

Les questions qu'ils pouvaient poser au système reprennent la taxonomie présentée dans [Capobianco, 2002a] et [Capobianco, 2002b] :

- *Demandes d'informations procédurales*, **questions "Comment ?"**. Les participants disposaient d'un lexique d'actions réalisables[4] sur l'interface de Flash dans lequel il leur suffisait de cliquer pour préciser l'objet de leur demande d'information procédurale.

- *Demandes d'informations "sémantiques"* sur le rôle d'une icône (dans une barre d'outils affichée) ou sur la fonction désignée par un item dans un menu (de la barre principale des menus uniquement), **questions "Quoi ?"**. Pour obtenir ces informations, les participants, après avoir cliqué sur le bouton "Quoi ?", sélectionnaient à la souris l'objet graphique ou l'item dont ils souhaitaient connaître le rôle.

- *Demandes d'explications*, **questions "Pourquoi ?"**. Pour simplifier l'implémentation des réponses à ce type de question qui sont souvent d'une grande complexité, la portée des questions "Pourquoi ?" a été limitée aux requêtes portant sur l'état courant du système qui se réduit, dans le cas d'une interface graphique, à l'état courant de l'affichage :
  - Si le compère qui simule les réactions du système d'aide considère que la dernière action du novice est erronée, c'est-à-dire qu'elle va à l'encontre de la réalisation du but courant du novice tel que le compère l'a perçu, celui-ci annulera la (les) action(s) précédente(s) du novice jusqu'à ce que le système soit à nouveau dans un état compatible avec la réalisation du but courant du novice tel que le compère l'a compris. Il affichera ensuite un message de type "Comment ?" correspondant à la réalisation de ce but.
  - Si le compère n'a pas compris le but du novice, il revient à l'état antérieur le plus récent qui corresponde à une étape de la réalisation du scénario.
  - Si l'état courant du système représente une étape valide de la réalisation du scénario, le compère envoie un message de type "OK".

- *Demandes de confirmation*, **questions "D'accord ?"**. Pour ce type de question également, on a restreint l'objet de la requête à l'état courant du système, ce qui permet de traiter les questions de ce type de la même manière que celles de type "Pourquoi ?".

La présence d'un compère humain permet de simuler l'intelligence de l'ACA ou d'adapter le message d'aide en fonction des connaissances de l'utilisateur. Cette technique permet ainsi de réaliser une aide "intelligente" facilement.

Il faut cependant que le compère dispose d'une assistance logicielle pour réaliser certaines tâches où les temps de réponse humains sont plus lents que ceux d'un système informatique (i.e., la recherche d'informations dans une base de données et, ici, la recherche du message d'aide à activer dans la base des messages). De plus, doter le compère d'une assistance permet également

[4] Les actions réalisables constituent un sous-ensemble cohérent de fonctions/procédures de Flash qui inclut strictement celles nécessaires à la réalisation des deux scénarios. Le lexique associe au vocabulaire de la langue courante, qui est celui que connaît le novice, la terminologie de Flash qui est celle du concepteur.

de stabiliser son comportement, d'éviter que celui-ci ne dérive ou fluctue au cours du temps, sous l'influence de phénomènes d'apprentissage ou sous l'effet de la fatigue (en cas de surcharge cognitive, notamment).

1. "Cas de l'étude utilisant l'ACA". Le compère rend l'ACA intelligent, notamment pour les requêtes "Pourquoi ?" et "D'accord ?".

2. "Cas de l'étude intégrant l'adaptation". Le compère joue d'une part le même rôle que pour l'étude utilisant l'ACA, le système devient alors intelligent. D'autre part, il réalise l'adaptation, c'est lui qui détermine le message d'aide approprié à envoyer à l'utilisateur pour les requêtes "Comment ?" et "Quoi ?". En effet, dans cette étude, il existe trois niveaux de messages d'aide et le compère, en fonction de la maîtrise du logiciel dont fait preuve le novice, choisit le niveau pertinent d'aide à fournir à celui-ci. Ces trois niveaux sont les suivants. Lorsque le novice demande pour la première fois une information, il reçoit un message analogue à ceux que fournissent les aides en ligne classiques. S'il renouvelle ultérieurement sa demande il reçoit, si le compère juge que les informations envoyées lors de la première requête ont été assimilées, des informations complémentaires (question de nature sémantique) ou un bref rappel syntaxique (question d'ordre procédural). Si le compère juge que le novice n'a pas compris les informations transmises à la suite de la requête initiale, il lui envoie un message détaillé, contenant explications supplémentaires, recommandations, etc.

Implications sur la réalisation du système d'aide

Créer un environnement d'interaction pour le système d'aide qui s'exécute parallèlement au logiciel d'application (Flash au cas particulier, c'est-à-dire :

- Une zone d'affichage des messages d'aide.
- Un panneau d'interaction permettant la prise en compte des quatre types de requêtes possibles pour l'utilisateur.

Implémenter la technique du magicien d'Oz, en offrant au compère :

- L'affichage en temps réel du contenu de l'écran de l'utilisateur novice, sur son propre écran, dans un espace dédié.
- La possibilité d'agir sur l'application de Flash pour annuler la ou les dernière(s) action(s) effectuée(s) par le sujet.
- L'affichage, dans un autre espace, de la base des messages parmi lesquels le compère peut choisir celui qu'il active en réponse à la requête de l'utilisateur. Ici l'assistance logicielle, filtre les messages en fonction du type de la question courante et de l'objet sur lequel elle porte ; de cette façon, il ne reste plus au compère qu'à choisir le message de réponse approprié. Par exemple, dans le cadre d'une assistance adaptative, le compère aura simplement à choisir, parmi les trois formes possibles d'un message, celle qui correspond le mieux au niveau du novice, tel qu'il peut le définir à partir de l'observation de ses interactions avec le logiciel d'application.

– Les énoncés et les données oculométriques du sujet, également via le réseau local en prévision d'une utilisation par commande vocale et de l'emploi d'un eye tracker.

4.2.2 Forme de multimodalité retenue

Travaillant en situation d'assistance, le système d'aide permet à l'utilisateur de consulter des messages d'aide multimodaux. Ainsi, dans l'espace dédié au système d'aide, le contenu de la zone d'affichage des messages d'aide varie suivant l'étude désirée :

1. "Cas de l'étude utilisant l'ACA". Comme l'ACA est doté de capacités d'expression orale, nous choisissons d'utiliser la Parole comme modalité de sortie et d'accompagner les énoncés de présentations graphiques constituées de copies d'écrans des affichages du logiciel d'application, Flash au cas particulier.

2. "Cas de l'étude intégrant l'adaptation". L'étude préliminaire (voir le chapitre 3) ayant montré la pertinence de messages d'aide associant texte et graphique, nous avons choisi ces modalités pour l'étude de l'apport d'une adaptation dynamique de l'interaction au profil de l'utilisateur.

Implication sur la présentation des messages d'aide

La zone d'affichage des messages d'aide doit être capable d'accueillir un ACA, du texte, de la parole et du graphique.

4.2.3 Trace des interactions

En vue d'analyser l'apport d'un ACA et celui des interfaces adaptatives à l'interaction homme-machine, il est nécessaire d'enregistrer les interactions des utilisateurs avec le système d'aide mais aussi avec le logiciel Flash. Pour cela, au cours de la réalisation des scénarios, un enregistrement de l'activité du sujet doit avoir lieu. De plus, pour obtenir des informations objectives sur l'intérêt suscité par l'ACA, les mouvements oculaires des participants doivent être recueillis. L'oculomètre utilisé est le modèle ASL 501, identique à celui de l'étude préliminaire.

Implication sur la conception de la plate-forme

Enregistrement des interactions pour être en mesure d'analyser les comportements et réactions des utilisateurs.

4.2.4 Déroulement d'une passation

Les passations étaient individuelles, dans une salle d'expérimentation dédiée, et comprenaient les étapes suivantes (cf.annexe B) :

1. Présentation des objectifs de l'expérience (décalés par rapport aux objectifs réels, et proches des objectifs d'utilisateurs potentiels, à savoir l'apprentissage de Flash en vue d'évaluer une aide en ligne innovante).
2. Questionnaires d'identification sous Adobe (i.e., caractérisation du profil de l'utilisateur).
3. – "Étude ACA" : Présentation de l'expérimentation et des tâches à réaliser par l'ACA nommé "Eloona", puis lecture par le sujet du tutoriel multimédia en dehors de la présence de l'expérimentateur (15 minutes environ).
   – "Étude Adaptation" : Présentation de l'expérimentation et des tâches à réaliser par l'expérimentateur, puis lecture du tutoriel multimédia par le sujet en dehors de la présence de l'expérimentateur (15 minutes environ).
4. Vérification de la bonne compréhension du fonctionnement du système d'aide par l'expérimentateur.
5. Courte manipulation du système en présence de l'expérimentateur.
6. – "Étude ACA" : Installation et calibrage de l'eye-tracker par l'expérimentateur (10 minutes environ).
   – "Étude Adaptation" : Eye tracker, non utilisé.
7. Réalisation des deux premiers scénarios (40 minutes en moyenne) en dehors de la présence de l'expérimentateur (qui devient compère).
8. Mise en place du troisième scénario.
9. Réalisation du troisième scénario (20 minutes en moyenne) en dehors de la présence de l'expérimentateur (qui redevient compère).
10. Remplissage des questionnaires verbal et non verbal d'évaluation (le contenu change suivant l'étude) et réalisation d'un post-test sous Adobe (30 minutes environ).
11. Entretien de debriefing (20 minutes).

La durée totale d'une passation varie, selon les participants, de deux heures à deux heures et demie.

4.3 Développement logiciel

Dans le but de répondre aux besoins mis en évidence dans les sections précédentes, des développements logiciels ont été entrepris qui ont conduit à la réalisation de deux composants :

1. La *plate-forme logicielle*, utilisée pendant l'expérimentation. Elle permet de simuler la technique du magicien d'Oz, de gérer le système d'aide et de sauvegarder les traces des interactions.

2. Un *logiciel de Rejeu* qui "rejoue" de manière synchronisée les traces des interactions sauvegardées lors de l'expérimentation. Ce composant permet également d'annoter les données enregistrées en vue de faciliter leur analyse détaillée.

4.3.1 Plate-forme logicielle : Système d'aide, Magicien d'Oz et Enregistrement des interactions

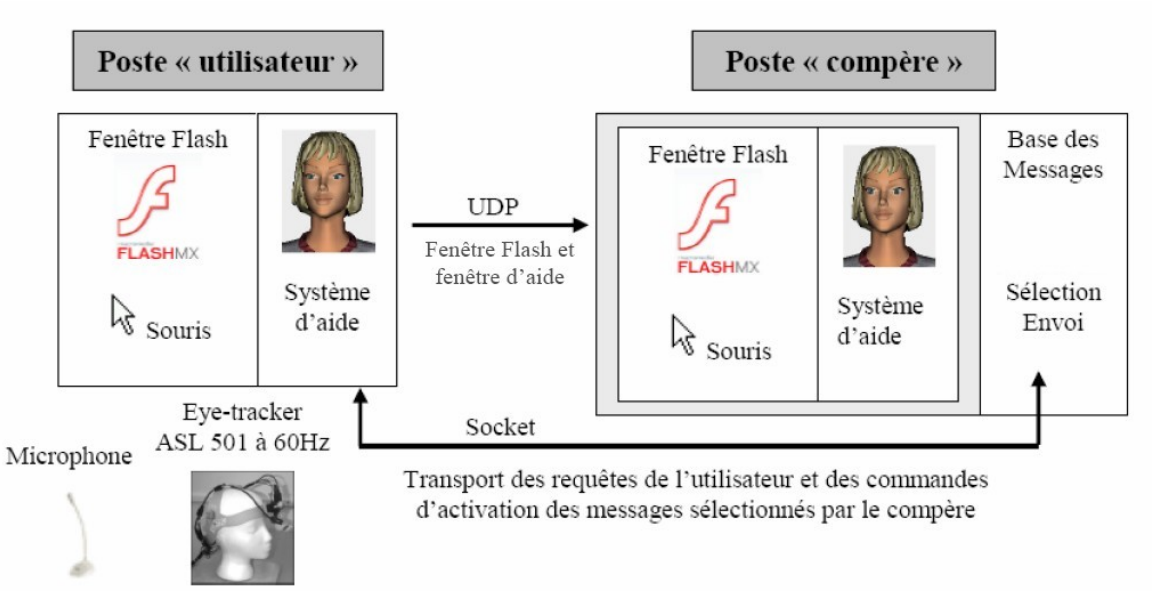


FIG. 4.4 – Plate-forme logicielle : Architecture.

La plate-forme dont la figure 4.4 présente l'architecture globale (de type client-serveur) se compose de deux applications :

- Une application qui s'exécute sur le poste "Utilisateur". Elle assure l'enregistrement de la trace des interactions et communique avec le poste "Compère" notamment pour recevoir les messages d'aide émis par le compère et envoyer à celui-ci la trace des interactions du sujet.

- Une application qui s'exécute sur le poste "Compère". Elle assure la mise en œuvre de la technique du magicien d'Oz. D'une part, elle reçoit la trace des actions de l'utilisateur, filtre les messages d'aide disponibles en fonction des requêtes de l'utilisateur et, d'autre part, elle envoie le message que le compère a sélectionné à partir de la requête de l'utilisateur et l'observation de son activité.

Poste "Utilisateur"

`Configuration - Fenêtre de lancement`

L'application qui s'exécute sur le poste "Utilisateur", se configure via l'interface représentée sur la figure 4.5 ; elle peut aussi s'effectuer par le fichier .ini pour des options plus avancées. En

passant par l'interface de l'application, l'expérimentateur peut choisir les événements et le type de données qu'il désire sauvegarder :

- Événement Clavier : appui sur une touche ;
- Événement système : sélection dans les menus, déplacement de fenêtres ;
- Événement automatique : activé lorsqu'il n'y a pas d'événement utilisateur ou système ; on peut choisir la fréquence de ces événements périodiques ;
- Regard : réception des échantillons sur le port série en provenance de l'eye-tracker.

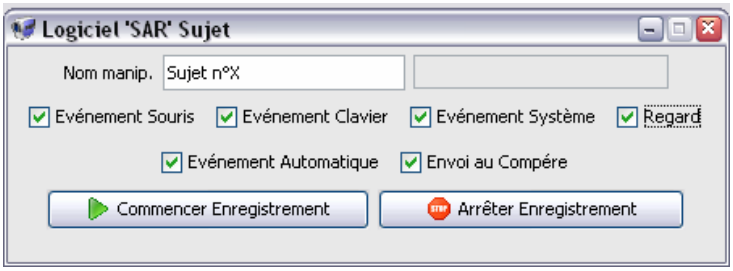


FIG. 4.5 – Interface de lancement de l'application, côté "Utilisateur".

On peut aussi choisir d'établir une connexion avec l'application distante "Compère". Cette connexion permet au compère de recevoir la trace des interactions et, au sujet, de recevoir les messages d'aide. A noter que si l'on ne coche pas cette option, il n'y a pas de connexion établie avec le poste compère mais une sauvegarde des interactions s'effectue en local sur le poste sujet.

Une fois, les options cochées, l'expérimentateur choisit le nom du répertoire où sera stocké l'enregistrement. Enfin, il lance l'enregistrement en cliquant sur le bouton « Commencer Enregistrement ». Le fait de cliquer sur ce bouton lance l'interface du système d'aide dont la configuration a été établie dans le fichier .ini, à savoir :

- Étude utilisant un ACA : Messages multimodaux utilisant la Parole (sur la figure 4.6, on observe des boutons qui permettent de jouer un énoncé) et des graphiques.
- ou Étude des interfaces adaptatives : Messages multimodaux utilisant Texte et Graphique.

`Système d'aide - Interface`

L'interface utilisateur du système d'aide que l'on peut observer sur la figure 4.6 se compose de deux zones principales :

- la zone d'interaction avec le système d'aide (i.e., les boutons d'activation des quatre types de questions), en bas sur la figure.
- et la zone d'affichage des messages d'aide. A noter que nous avons également intégré un dictionnaire que l'utilisateur peut consulter lorsqu'il ne comprend pas un terme technique ou a oublié sa signification.

Les messages d'aide s'affichent sous la forme de pages HTML pour la seconde étude où leur présentation combine Texte et Graphique. Pour la première étude, les affichages comprennent, en outre, un clone de technologie ActiveX fourni par FT R & D. L'intégration de l'ActiveX dans l'application était un objectif difficile à réaliser mais nécessaire pour pouvoir manipuler

dynamiquement le clone depuis le poste Compère par exemple. Une liste des problèmes rencontrés lors de l'intégration de l'ActiveX et des solutions proposées pour résoudre a été retournée à FT R & D.

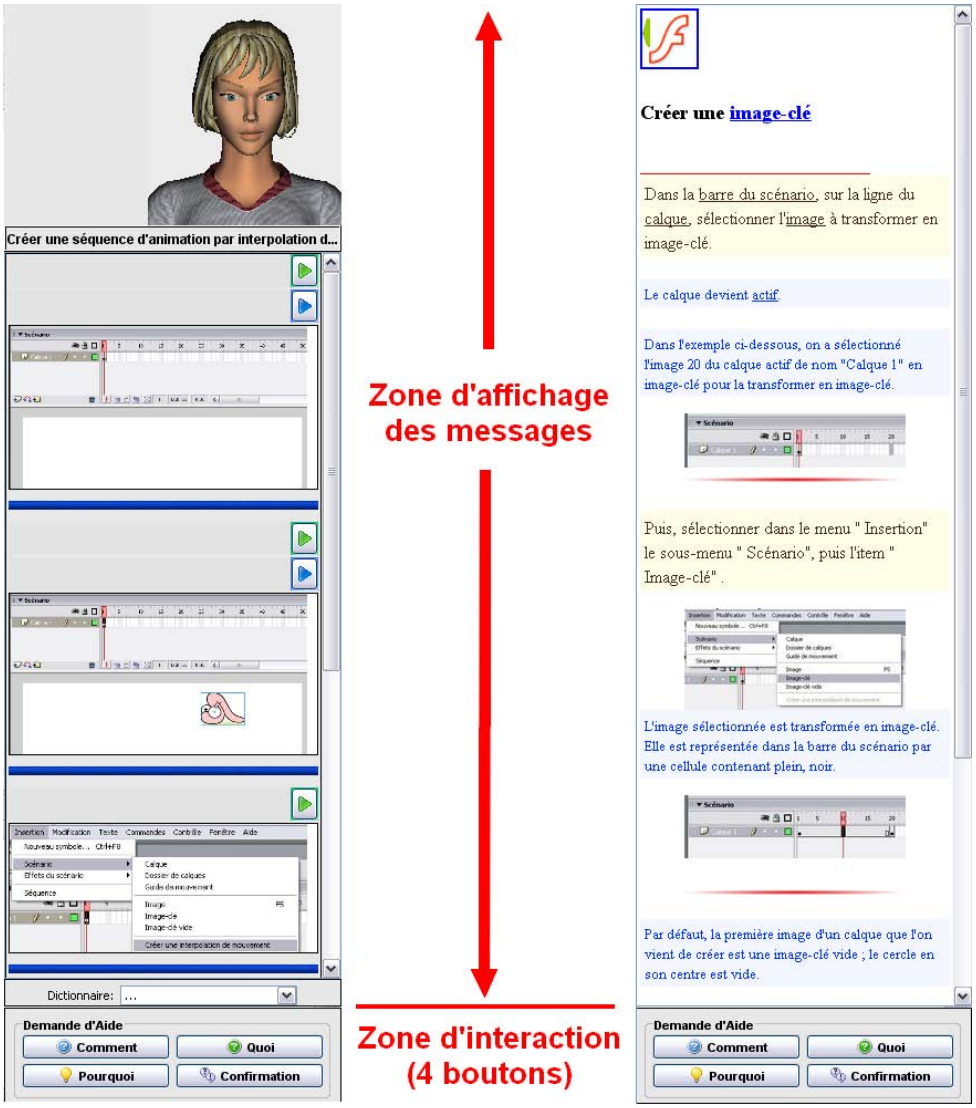


FIG. 4.6 – Interface du système d'aide. A gauche une configuration avec ACA (ACA + Parole + Graphique), à droite la configuration simulant l'aide en ligne adaptative (Texte + Graphique).

Pour mémoire, les quatre boutons visibles dans la zone d'interaction (voir la figure 4.6) permettent à l'utilisateur d'interroger le système, à savoir :

- "Comment ?" : Comment réaliser, par exemple, une interpolation de mouvement ;
- "Quoi ?" : Quelle est la fonction, le rôle, de cette icône/menu ?
- "Pourquoi ?" : Pourquoi ai-je ce résultat, suite à une action,
- "Confirmation ?" : La dernière action effectuée est-elle compatible avec le but courant de l'utilisateur tel qu'il a été compris par le compère ou avec les instructions du message "Comment ?" en cours d'exécution ?

L'implémentation de chaque type de requête ne sera pas explicitée en détail ici étant donné qu'elle varie légèrement, pour certains types de requêtes, d'une étude à l'autre (ACA / Adaptation).

`Données recueillies - Fonctionnement`

Comme il a été mentionné précédemment, l'application gère le système d'aide et réalise en parallèle l'acquisition et la sauvegarde des données représentant l'activité de l'utilisateur. Ainsi, pour l'acquisition, plusieurs 'threads' permettent l'enregistrement des données suivantes :

- **Données oculométriques** : l'application lit sur le port série des trames fournies par l'Eye-Tracker (choix d'une fréquence d'échantillonnage de 60Hz). Le décodage de celles-ci permet d'extraire la position en X et Y du regard à l'écran, la distance entre la tête de l'utilisateur et l'écran, ainsi que le diamètre de pupille. Nous avons choisi de ne pas transmettre ces informations au compère afin d'éviter de le distraire durant sa tâche de suivi de l'activité de l'utilisateur courant. Ces informations sont utilisées par la suite dans le logiciel de "Rejeu", après avoir été traitées par un algorithme de calcul des fixations oculaires (développé au sein de l'équipe). L'étude des fixations permet de déterminer les zones de l'écran, donc des affichages, sur lesquelles se porte successivement l'attention de l'utilisateur. Ces données qui sont datées comme l'ensemble des événements intervenant dans le rejeu sont sauvegardées dans un fichier nommé eye.txt.

- **Activité de l'utilisateur** : Travaillant sous Windows pour pouvoir utiliser Flash, l'application met en place différents 'Hooks' qui interceptent les messages Windows. Chaque événement utilisateur ou système capturé grâce à ces Hooks déclenche une copie d'écran qui lui est associée ainsi que l'enregistrement de la position de la souris au moment de l'événement ; l'ensemble de ces données est sauvegardé. Lorsque l'utilisateur est inactif, un événement automatique paramétrable par l'expérimentateur est déclenché périodiquement. Cet événement est utile, par exemple, lorsque l'affichage comprend une animation ; nous avons utilisé cette fonctionnalité pour sauvegarder une trace représentative des animations de Flash. Les références des événements sont sauvegardées dans un fichier image.txt qui partage le même répertoire que les copies d'écran enregistrées au format .jpg

- **Énoncés verbaux** : Si l'on désire utiliser la parole comme modalité d'entrée, la plate-forme peut également enregistrer les énoncés verbaux formulés par l'utilisateur et les transmettre au poste du compère. Cette fonctionnalité n'a pas été utilisée dans le cadre des deux études expérimentales réalisées, la Parole étant utilisée exclusivement comme modalité de sortie, et uniquement pour la première des deux études.

Poste "Compère"

L'application installée sur le poste "Compère" se configure également via un fichier .ini, mais uniquement pour des options techniques avancées telles que l'adresse IP de la machine "Utilisateur", la résolution de l'affichage ou encore le chemin d'accès aux messages d'aide.

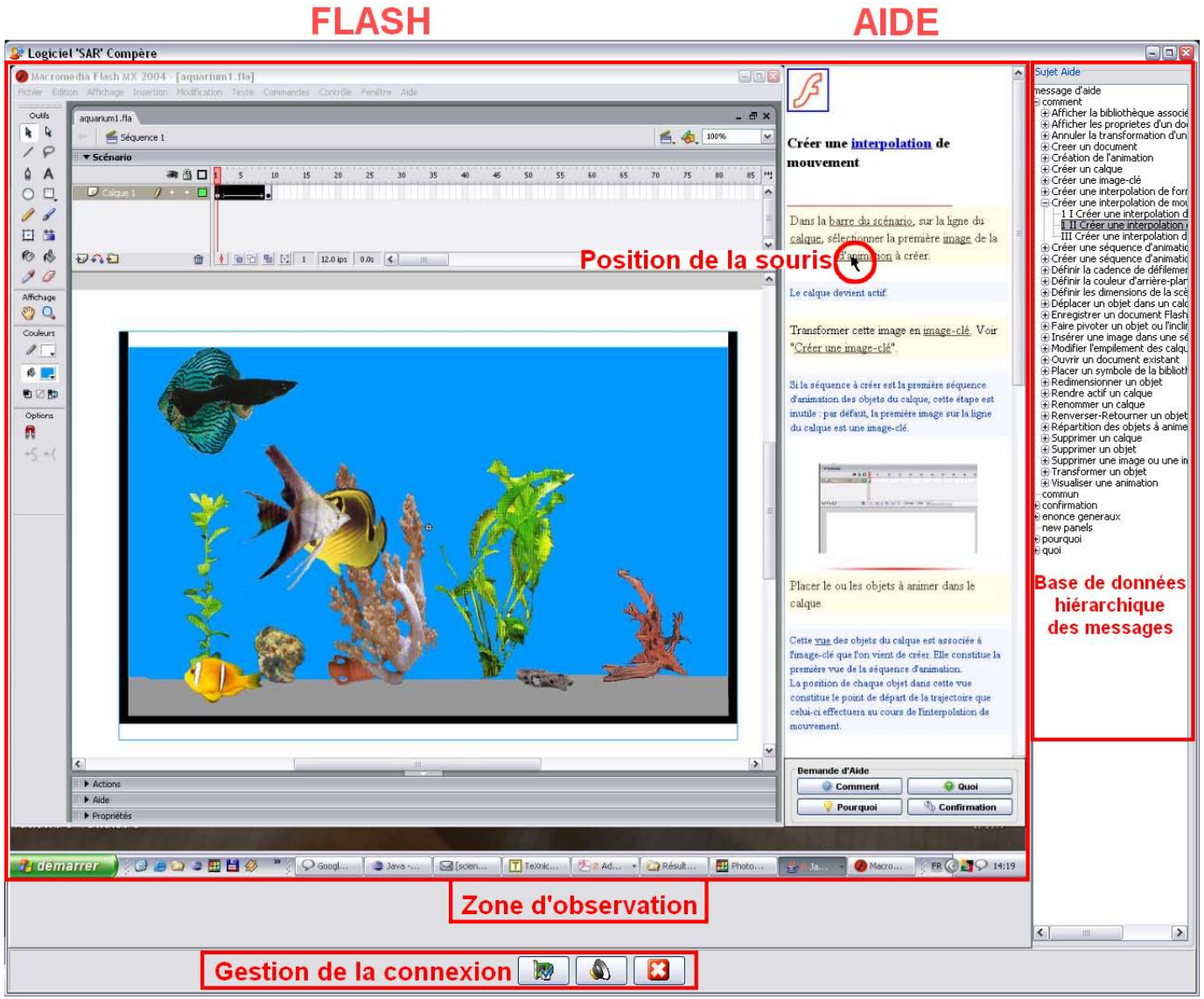


FIG. 4.7 – Poste "Compère".

L'interface de cette application, visible à la figure 4.7, comporte trois parties :

- "Gestion de la connexion" : On positionne l'application en attente de connexion d'un client, à savoir le poste "Utilisateur". Pour mémoire, une fois la communication entre les deux applications ouverte, on dispose d'une socket pour transférer les requêtes de l'utilisateur et les messages d'aide choisis par le compère, ainsi que d'une connexion en UDP pour transférer les copies d'écran et la position de la souris. Le protocole utilisé pour le transport des copies d'écran est UDP, pour des raisons de performance essentiellement étant donné la fréquence assez élevée des événements déclencheurs de copies d'écran. Au cas particulier, UDP offre une fiabilité suffisante, compte tenu de la nature des informations transportées . En revanche, les requêtes de l'utilisateur sont transmises au compère par "sockets", liaison sans perte, car ce sont des séquences d'événements utilisateur.

- "Zone d'observation" : Dans cette zone, apparaissent les affichages du poste "utilisateur" avec, superposée, la position de la souris. Le compère peut ainsi suivre l'activité du sujet.

- "Base de données hiérarchique des messages" : Lorsque le sujet effectue une requête, le message correspondant à celle-ci se sélectionne automatiquement dans la hiérarchie. Pour l'envoyer le compère n'a plus qu'à effectuer un clic droit faisant apparaître un pop-up menu et sélectionner *Parler* dans le cas de l'étude avec ACA ou *Envoyer* dans le cas de l'étude adaptative. C'est également à l'aide de ce menu qu'il peut *pré-consulter* le message sélectionné ou *annuler* la dernière action de l'utilisateur via d'autres items de ce menu.

`Données recueillies`

Côté Compère, nous recueillons uniquement l'historique des requêtes du sujet, accompagnées des messages choisis et envoyés par le compère en réponse à ces requêtes. Tous les enregistrements sont effectués sur la machine "Utilisateur", pour obtenir une datation homogène, nécessaire pour assurer un rejeu correctement synchronisé des traces des interactions.

4.3.2 Rejeu et annotation des données sauvegardées par la plate-forme

Présentation de l'application

L'application de "Rejeu" se charge de rejouer les traces des interactions enregistrées, comprenant les données suivantes synchronisées : copies d'écran, position de la souris et des points de regard (gaze points) ainsi que, le cas échéant, énoncés de l'utilisateur. De plus, comme le nom de l'application le sous-entend, elle possède les mêmes caractéristiques qu'un lecteur multimédia, à savoir, avance, stop, recule, etc.

La figure 4.8 reproduit un affichage extrait du rejeu d'un enregistrement montrant l'environnement de l'utilisateur (Flash et le système d'aide) auquel sont superposées la position courante de la souris et les échantillons de regard (sous forme de petits cercles) acquis pendant la durée d'affichage de la capture écran.

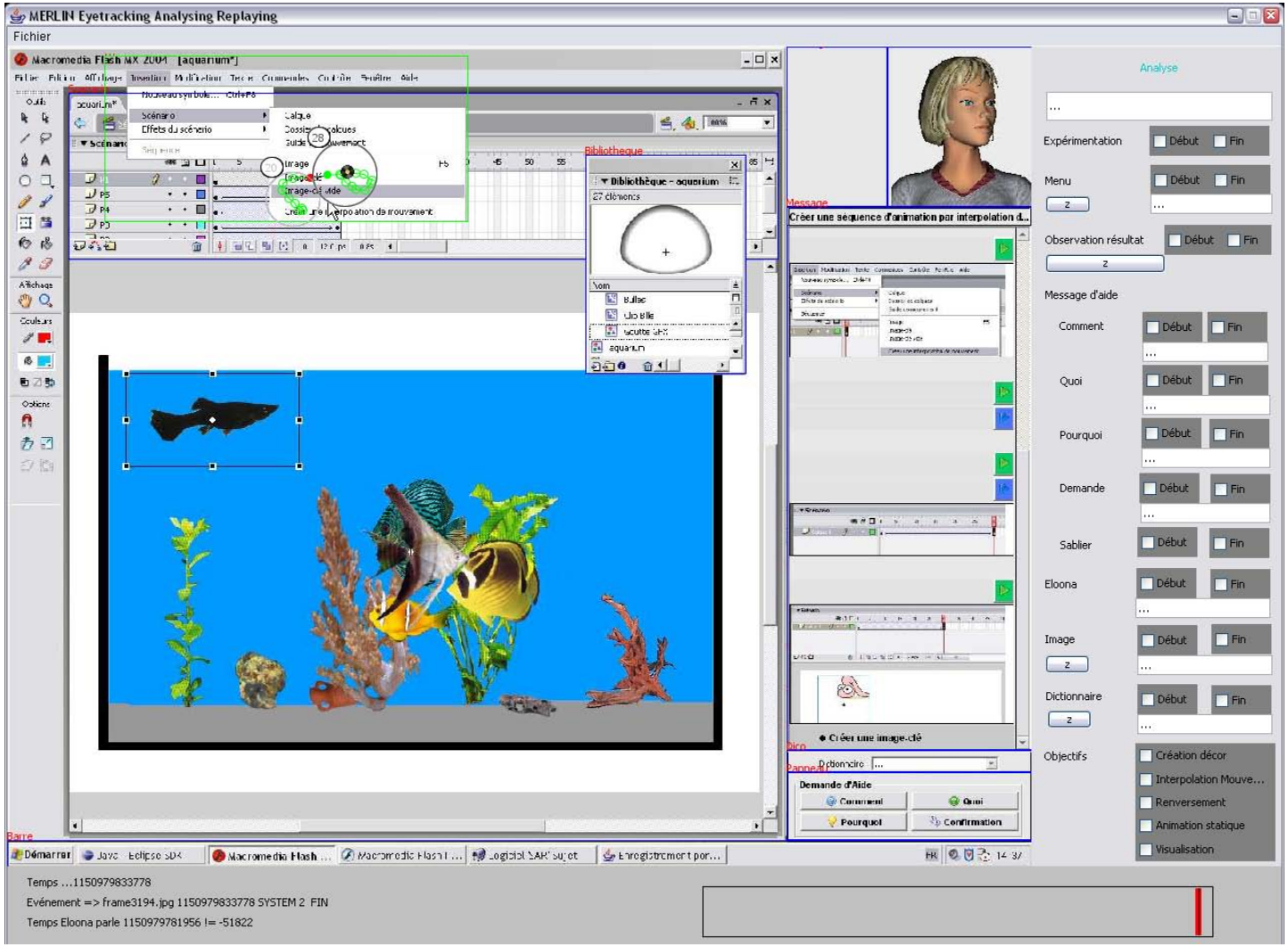


FIG. 4.8 – Application de "Rejeu" et d'annotation. Sur la gauche, affichage extrait du rejeu de la trace d'interactions enregistrées au cours d'une passation. A droite, les menus d'assistance à l'analyse utilisés pour l'annotation. En bas de l'écran, sur la gauche, des informations temporelles sur l'activité de l'utilisateur.

A noter que les échantillons de regard peuvent appartenir ou non à une fixation. Un cercle englobe les échantillons qui composent chaque fixation, et la durée (en nombre d'échantillons, 1 échantillon = 1/60 Hz = 16.66ms) de celle-ci s'affiche à proximité. Un code de couleur permet d'interpréter chaque échantillon :

- Noir : il appartient à une saccade ;
- Vert plein : il correspond à un début de fixation ;
- Vert : il appartient à une fixation ;
- Rouge plein : il correspond à la fin d'une fixation.

Principe de l'annotation

Sur la gauche de l'extrait de rejeu présenté, un panneau de commande est dédié à l'annotation d'une session. Cette zone est configurable suivant l'expérimentation que l'on désire effectuer.

Le principe que nous avons décidé d'adopter est celui de caractériser certaines actions de l'utilisateur par des **marques** qui ont toutes une date de début et de fin ; les marques peuvent porter sur plusieurs actions exécutées en parallèle. Par exemple, l'utilisateur peut lancer l'énoncé d'un message sonore tout en agissant sur les menus du logiciel Flash. En revanche, la manipulation d'un objet dans la scène de Flash ne fera l'objet d'aucune annotation.

Un autre aspect de l'annotation est qu'à chaque couple de marques début-fin délimitant une action de l'utilisateur, il est possible d'associer une **liste de zones** d'attention visuelle que l'expérimentateur dessine ; ces zones entourent les objets graphiques concernés par l'action de l'utilisateur. On peut ainsi déterminer les positions des fixations oculaires pendant la durée de chaque action. Par exemple, lorsque l'utilisateur parcourt un menu, une zone enveloppe ce dernier afin d'identifier les fixations réalisées sur ce menu. A noter également qu'une liste permanente de zones, propre à l'application avec laquelle l'utilisateur interagit (Flash au cas particulier) peut être définie pour recueillir les fixations sur les zones de l'affichage qui restent statiques pendant toute la durée de l'interaction (par exemple, dans notre cas, la zone où le clone est affiché en permanence, ou bien la barre des menus de Flash).

Exemple d'utilisation : L'utilisateur commence par actionner l'item "insertion" de la barre des menus ; l'expérimentateur coche alors la marque de début correspondant à une action sur un menu, puis dessine une zone autour de ce menu sur l'affichage du rejeu ; la zone est matérialisée par un rectangle vert (voir la figure 4.8). Au fur et à mesure que l'utilisateur navigue dans les différents sous-items du menu, l'expérimentateur dessine des zones autour des sous-menus et des items visibles. Enfin, lorsque l'utilisateur sélectionne un item, l'expérimentateur coche la marque de fin correspondant à une action sur un menu et précise la référence de cet item dans la zone de texte relative aux menus.

Pour l'exemple présenté, l'information générée est :

- la durée de parcours du menu : date de fin de marque - date de début de marque ;
- le nom de l'item sélectionné par l'utilisateur, "..." si celui-ci quitte sans choisir d'item ;
- la liste des fixations (position, durée, zone) effectuées pendant le parcours du menu, si l'on utilise l'eye tracker.

Récapitulatif des données disponibles

Plus généralement, on peut regrouper les données issues de l'annotation en deux catégories :

1. Les données sans "Eye tracker", c'est à dire les marques que l'on pose pour segmenter et caractériser l'activité de l'utilisateur.
2. Les données qui exploitent le suivi du regard, c'est-à-dire les zones de l'affichage et les objets graphiques qui mobilisent successivement son attention.

Environnement Utilisateur			Marque	Zone	Données recueillies
Flash	Barre du Scénario			X	C et F
	Bibliothèque			X	C et F
	Barre des Menus		X	X	D, I, C et F
	Scène			X	C et F
Windows	Barre des Tâches			X	C et F
	Observation du résultat à obtenir		X	X	D, C et F
	Environnement d'expérimentation quitté			X	C et F
Système d'aide	ACA et demande d'énoncés		X	X	D, I et F
	Zone d'aide	Comment	X	X	D, I, C et F
		Quoi	X	X	D, I, C et F
		Pourquoi	X	X	D, I, C et F
		Demande	X	X	D, I, C et F
		Sablier	X	X	D, I, C et F
	Zoom d'un graphique		X	X	D, I, C et F
	Dictionnaire		X	X	D, I, C et F
	Panneau d'interaction		X	X	D, I, C et F
Avec : D=Durée de consultation, I=Identifiant de l'action, C=Clic souris et F=Fixations					

TAB. 4.1 – Données issues de l'annotation.

Les données qu'il est possible d'extraire du rejeu de l'enregistrement des interactions sont récapitulées dans le tableau 4.1. Dans la colonne nommée "Marque", une croix indique que l'action ou l'activité correspondante (voir la première colonne) peut être annotée en utilisant l'assistance logicielle évoquée dans l'exemple d'utilisation ci-dessus. L'annotation permet, notamment, de connaître la durée de consultation d'une zone de l'affichage et la nature de chaque type d'action annotée. A noter que cette assistance n'est disponible que pour un sous-ensemble des actions possibles sous Flash ; elle n'a été implémentée que pour les actions et activités qu'il nous a semblé intéressant d'analyser dans le cadre des deux études présentées dans les chapitres suivants, en particulier les interactions des sujets avec le système d'aide. Dans la colonne nommée "Zone" du tableau 4.1, une croix permet de connaître les différentes zones que l'expérimentateur peut dessiner au cours du rejeu de l'enregistrement. Grâce à ces zones, on recueille essentiellement le nombre de clics (utile, par exemple, pour connaître l'activité dans la barre de scénario sous Flash), le nombre et la durée des fixations dans chaque zone.

4.4 Conclusion

L'étude préliminaire, qui a servi de référence pour concevoir la méthodologie des deux études centrées l'une sur l'apport d'un ACA à l'interaction homme-machine, l'autre sur celui de l'adaptation au profil utilisateur, a également orienté la conception des outils logiciels, notamment en montrant les insuffisances (d'ordre ergonomique principalement) des versions initiales des modules d'acquisition-analyse des traces des interactions et du logiciel de "rejeu", ce qui a permis d'y remédier et de disposer actuellement de logiciels d'utilisation facile et efficace. Les modules ajoutés aux logiciels utilisés pour l'étude préliminaire comprennent la simulation de la technique du magicien d'Oz, ainsi que la création de l'interface du système d'aide. Cette interface est paramétrable ; elle permet, en particulier, d'utiliser soit un ACA sous forme d'un ActiveX (e.g., celui de France Télécom) ou des pages Html pour présenter les messages d'aide.

Dans les chapitres suivants, nous verrons qu'une base spécifique de messages a dû être mise en œuvre pour chaque étude. La conception des messages d'aide sera également détaillée. D'un point de vue méthodologique, les deux études dont la présentation va suivre sont conçues selon une approche expérimentale identique (même tâche expérimentale et même schéma de passation), et implémentées en utilisant des protocoles expérimentaux semblables dont l'implémentation met à profit le caractère générique de la plate-forme et des outils logiciels présentés dans ce chapitre et démontre la réduction importante du coût de développement que permet la réutilisation d'outils logiciels génériques. Ces logiciels constituent, dans le cadre de l'étude de la qualité ergonomique d'interfaces innovantes, des outils de prototypage rapide performants ; ils fournissent en outre une assistance efficace à l'analyse des comportements et réactions d'utilisateurs futurs.

Chapitre 5

ACAs : Eloona en situation d'aide

Cette deuxième étude, comme l'étude préliminaire, est conçue dans le cadre d'une approche expérimentale. Elle est centrée sur l'étude de l'apport d'un Agent Conversationnel Animé (ACA) : Eloona en situation d'assistance à l'utilisation du logiciel Flash. Ainsi, le but de l'utilisateur est de réaliser deux animations sous Flash (voir, chapitre 4, la description de la tâche expérimentale utilisée dans l'étude préliminaire), avec l'assistance d'un système d'aide en ligne qui permet de générer deux formes d'aide, une par animation : "Parole + Graphique", versus "Parole + Graphique + ACA". Les utilisateurs ayant testé le système d'aide ont utilisé ces deux formes d'aide dans un ordre contrebalancé.

Dans un premier temps, nous rappelons le contexte général de l'étude, puis, nous présentons la méthodologie utilisée pour concevoir le plan expérimental ainsi que le protocole. Les développements logiciels réalisés ayant été présentés de façon globale dans le chapitre précédent, nous détaillons uniquement les composants spécifiques réalisés pour la mise en œuvre du protocole de cette seconde étude. Enfin, l'analyse des données recueillies lors de l'expérimentation est présentée, suivie de la discussion des résultats obtenus.

5.1 Contexte - Généralités

5.1.1 Rappel des objectifs

L'objectif principal est d'évaluer l'apport d'une représentation graphique animée du système à l'interaction homme-machine, dans une situation d'apprentissage interactif : l'aide en ligne à l'utilisation d'un logiciel grand public nouveau pour l'utilisateur.

Les questions de recherche soulevées sont d'une triple nature :

1. La présence d'un agent conversationnel animé (ou ACA) « avenant » peut-elle accroître la motivation du novice pour l'apprentissage de l'utilisation du logiciel, lui donner confiance en lui pour réussir cet apprentissage et acquérir la maîtrise du logiciel sans assistance, enfin, réduire le stress et le désarroi liés à la prise en main d'un nouveau logiciel ?

2. Si l'ACA permet la visualisation des mouvements articulatoires (en particulier ceux des lèvres) au cours de l'énonciation d'un message d'aide oral, sa présence peut-elle améliorer l'intelligibilité et, par conséquent, la compréhension, l'assimilation et la mémorisation des messages oraux ?
3. Si l'ACA est doté de capacités d'expression gestuelle, la formulation des messages d'aide peut-elle être simplifiée grâce à l'usage, pour désigner les éléments de l'affichage, de déictiques linguistiques associés et concomitants à des désignations gestuelles ? Une désignation multimodale (déictique + geste) des éléments affichés à l'écran est plus efficace, en termes de rapidité et de précision, et plus simple, en terme d'interprétation, que les expressions linguistiques spatiales équivalentes.

L'expérience envisagée vise à déterminer si la présence d'un ACA est susceptible de faciliter l'apprentissage de la maîtrise d'un nouveau logiciel en le rendant plus rapide, plus facile (i.e., exigeant un effort moindre de la part du novice), plus robuste (i.e., moins superficiel et conduisant à une meilleure compréhension du fonctionnement du logiciel, donc durable), et plus agréable subjectivement. La section suivante précise l'ACA choisi parmi les différents types existants, et justifie ce choix.

5.1.2 Choix de l'ACA

Pour étudier les différents types d'apport potentiel des ACAs évoqués ci-dessus, nous avons choisi, parmi les ACAs développés par FT R & D et mis à notre disposition, un visage de femme, Eloona (voir, en annexe C).

Cet ACA comporte un modèle musculaire évolué d'articulation de la parole. Grâce à un éditeur graphique, on peut doter son visage d'expressions variées et utiliser les mouvements de son regard et de sa tête en tant que gestes de désignation des éléments affichés à l'écran. L'intelligibilité de la synthèse vocale est suffisante. Il semble en outre possible de manipuler les paramètres prosodiques pour varier l'expressivité des énoncés. Enfin, le caractère féminin d'Eloona et le graphisme de son apparence physique contribuent de toute évidence à en faire un ACA « avenant » dont la présence dans la situation d'interaction choisie soit propre à rassurer l'utilisateur novice, et les messages d'aide parfaitement crédibles. Le choix d'Eloona satisfait donc l'ensemble des exigences nécessaires concernant le comportement et l'aspect de l'ACA.

5.2 Protocole expérimental - Définition

5.2.1 Cadre général

Le protocole proposé supposait une synthèse vocale de très bonne qualité pour assurer l'intelligibilité des messages, ainsi que la possibilité de contrôler la prosodie d'Eloona pour donner au personnage une dimension conviviale. Les tests effectués sur la synthèse vocale fournie par FT R & D s'étant avérés très décevants, nous avons préféré utiliser des messages pré-enregistrés par une locutrice expérimentée.

Les participants ont effectué une seule session d'une durée de deux heures environ. Deux raisons majeures expliquent cette décision. D'une part, l'utilisation d'un logiciel commercialisé, Flash, pour évaluer l'apport de la présence d'un ACA à l'efficacité et à l'attractivité d'un système d'aide en ligne ne permet pas de contrôler l'usage que les participants sont susceptibles de faire du logiciel entre plusieurs sessions successives distantes dans le temps ; ce qui risque de biaiser l'évaluation de l'apport effectif de l'ACA à l'apprentissage de l'utilisation du logiciel. D'autre part, les applications des ACAs envisagées par FT R & D à court terme concernent des usages ponctuels, donc des populations d'utilisateurs occasionnels plutôt que réguliers : guides d'installation ou d'utilisation de dispositifs grand public simples (e.g., la Livebox), ou systèmes de renseignement. Or, une seule session suffit pour déterminer l'apport des ACAs à l'interaction homme-machine dans ces contextes d'utilisation.

Lors de la définition des expressions faciales d'Eloona, des mouvements de son regard et de sa tête, ainsi que des variations expressives de l'intonation et du rythme de ses énoncés, l'objectif majeur a été d'obtenir des comportements dont l'expressivité et la sémiotique soient interprétables sans ambiguïté par les utilisateurs potentiels, plutôt que des comportements qui copient de façon littérale le comportement humain.

En effet, il a été montré [Mori, 1970] et [MacDorman *et al.*, 2005] que l'effet produit par un avatar humanoïde (human-like) sur l'utilisateur est une fonction complexe qui fait intervenir à la fois l'aspect de l'avatar et son comportement : lorsque la ressemblance de l'avatar avec un être humain croît, on observe que, au delà d'un certain degré de similitude, l'intérêt de l'utilisateur pour l'avatar décroît jusqu'au rejet. Plusieurs explications peuvent être proposées pour expliquer cette chute étrange (ou "uncanny valley"), notamment la distance entre le comportement stéréotypé des avatars humanoïdes actuels et leur apparence physique très convaincante.

5.2.2 Plan expérimental - Aperçu global

La situation d'interaction homme-machine choisie est la consultation d'un système multimodal d'aide en ligne à la découverte et à l'utilisation de Flash, logiciel grand public de création d'animations multimédias. Nous avons analysé et comparé la satisfaction subjective et les performances (variables liées) de participants volontaires dans deux situations d'interaction avec l'aide en ligne qui ne diffèrent l'une de l'autre que par la modalité de présentation des messages d'aide (variable indépendante) :

1. Synthèse vocale simulée et graphique (copies d'écrans de Flash), condition de référence ;
2. Synthèse vocale, présence d'Eloona et graphique.

Le contenu informationnel des messages était identique dans les deux conditions. Les participants ont effectué ces deux conditions dans un ordre de passation contrebalancé, de façon à limiter les effets éventuels de l'ordre de présentation des deux système d'aide sur les jugements subjectifs des sujets et sur leurs performances.

Ainsi, 22 étudiants de premier cycle universitaire ont participé à l'expérience ; 11 d'entre eux ont interagi d'abord avec le système sans ACA, et les 11 autres ont utilisé en premier le système avec ACA. 14 d'entre eux ont été sélectionnés parmi des étudiants engagés dans un cursus à dominante informatique et mathématique, les 8 autres suivent des cursus variés dont

l'informatique en tant que discipline est absente. Tous les participants ont une expérience des jeux vidéos qu'ils utilisent plus ou moins souvent.

Comme on a pu le voir dans le chapitre précédent, la plate-forme développée gère l'enregistrement des interactions de l'utilisateur avec Flash et l'aide, simule la technique du magicien d'Oz et se charge de générer et transmettre au novice les réactions du système d'aide définies par le compère. Nous rappelons ici les besoins conceptuels et matériels ainsi que outils logiciels nécessaires pour réaliser cette seconde étude.

`Enregistrement des mouvements oculaires`

Pour obtenir des informations objectives sur l'intérêt suscité par l'ACA, les mouvements oculaires des participants ont été enregistrés et analysés. L'oculomètre utilisé est un modèle ASL 501 qui offre une bonne précision de détection de la direction du regard (erreur inférieure à 1 degré d'angle visuel). C'est un casque, par conséquent l'utilisateur peut bouger la tête librement sans que le système de suivi ne « perde » le regard ; cette liberté de mouvement est indispensable pour assurer le caractère réaliste (écologique) des deux situations d'interaction homme-machine étudiées. Cependant, elle a pour rançon une certaine fatigue due au poids du casque (de l'ordre de 300 g). Les réponses au questionnaire verbal montrent que la gêne résultante est relativement faible : 3 participants sur 22 seulement ont jugé que le poids du casque les avait fatigués légèrement (2 participants) ou beaucoup (1 participant). L'option d'enregistrement des données oculaires est activée lors du lancement de la plate-forme logicielle.

`Comportements de l'ACA`

Le comportement d'Eloona était statique dans le contexte de cette étude, c'est-à-dire qu'il ne variait pas au cours de l'interaction en fonction, par exemple, des difficultés rencontrées par les participants dans la réalisation des scénarios proposés. Utiliser un système d'aide adaptatif dans les deux conditions n'aurait pas permis d'évaluer les seuls effets de la présence de l'ACA sur l'utilisateur par comparaison avec la condition de référence : les effets de cette innovation (effet de nouveauté principalement) auraient pu risquer d'interférer avec ceux de la présence de l'ACA, voire de les masquer. Toutefois, quelques dérogations à ce principe ont été prévues pour donner une certaine empathie au comportement d'Eloona vis-à-vis du novice. En effet, le compère dispose de quelques messages de sympathie/encouragement destinés aux participants en difficulté, par exemple : « Utilisez les liens hypertexte si vous ne comprenez pas certains mots du message. », ou « C'est bien. Vous pouvez passer à l'étape suivante. ». La dimension empathique de l'expressivité d'Eloona (prosodie, expressions faciales, mouvements du regard et de la tête) a été renforcée pour ces messages.

Pour éviter de même toute confusion entre les effets dus à la simple présence d'une représentation graphique animée du système et ceux dus à un comportement "intelligent" de celui-ci, il fallait également doter l'ACA d'un comportement identique à celui des systèmes d'aide classiques. Ainsi, il n'intervenait que sur demande de l'utilisateur courant, il ne prenait aucune initiative. En outre, l'interaction entre les participants et le système d'aide était limitée à des échanges simples, classiques, de type question-réponse.

`Moyens d'interaction avec le système d'aide proposé`

Les utilisateurs avaient à leur disposition quatre boutons pour interroger le système d'aide et préciser le type de leur question (voir, dans le chapitre 4 la définition des fonctions de ces boutons) ; ils disposaient également d'un lexique pour exprimer l'objet sur lequel elle portait. Leurs moyens d'interaction étaient limités au clavier et à la souris, dispositifs d'interaction classiques de façon à restreindre le nombre de facteurs susceptibles d'agir sur leurs jugements et leurs comportements à la présence/absence d'Eloona. Une étude antérieure [Capobianco et Carbonell, 2003] fournit un argument supplémentaire d'importance en faveur de cette décision. Elle montre en effet que, lorsqu'on autorise les participants à interroger l'aide oralement, sans contrainte d'expression, les effets de cet enrichissement des moyens d'interaction sur la satisfaction subjective des participants sont si importants qu'ils masquent ceux d'autres facteurs (dans le cas de l'étude citée, la comparaison de deux systèmes d'aide en ligne, l'un contextuel, l'autre non).

5.2.3 Hypothèses de travail - Mesures

Hypothèses de travail

La présence d'Eloona (une des deux valeurs de la variable expérimentale sur laquelle est centrée l'expérience) est susceptible d'enrichir l'interaction avec l'utilisateur de deux manières :

- en introduisant une dimension affective (i.e., humanisation) dans la relation utilisateur système, et
- en améliorant l'intelligibilité des messages d'aide oraux grâce à la mise en œuvre de mécanismes multimodaux de perception-compréhension de la parole (i.e., implication de l'audition et de la vision), qui s'avèrent plus efficaces et robustes que les mécanismes exclusivement auditifs.

Par conséquent, l'étude expérimentale envisagée peut fournir des éléments de réponse aux deux questions de recherche suivantes :

H1. La présence d'un ACA est-elle susceptible indirectement d'améliorer de façon significative l'efficacité de l'interaction en accroissant la motivation des participants et leur confiance en eux ? Plus précisément, l'activité considérée combine apprentissage de savoir-faire (tels que la maîtrise des procédures d'utilisation des fonctions de Flash) et acquisition de connaissances, notamment sur le fonctionnement du logiciel ou, en d'autres termes, sur la « sémantique » des fonctions offertes à l'utilisateur. La présence d'un ACA peut-elle améliorer sensiblement l'efficacité de cet apprentissage, c'est-à-dire la rapidité de l'acquisition de la maîtrise du logiciel, le type de connaissances acquises, procédurales ou conceptuelles et la qualité de leur mémorisation ?

H2. Dans un contexte d'interaction peu favorable, tel que la consultation d'une aide en ligne (voir le paradoxe de motivation mis en évidence par Carroll et ses collègues [Carroll *et al.*, 1987] et [Carroll et Rosson, 1987] dès 1987, de même que le succès très mitigé de l'Assistant Microsoft), la présence d'un ACA suscitera-t-elle des réactions positives de la part des utilisateurs potentiels ? Sera-t-elle bien accueillie ? Et surtout, incitera-t-elle les utilisateurs novices à consulter davantage

l'aide en ligne, augmentera-t-elle leur intérêt pour cette forme d'assistance généralement ignorée par les novices ?

Mesures

L'évaluation de l'*efficacité de l'apprentissage* à l'utilisation de Flash dans les deux situations considérées se fonde :

- d'une part, sur un codage et une analyse détaillée des interactions des participants avec Flash, notamment de leurs réussites et échecs dans l'utilisation des fonctions de ce logiciel et la réalisation des animations proposées,
- d'autre part, sur les réponses des participants à un post-test qui a pris la forme d'un questionnaire électronique.

Pour mesurer la rapidité d'un apprentissage, lorsque celui-ci s'effectue principalement par l'action et la pratique, on peut tenter de déterminer l'étendue des savoir-faire et des concepts acquis (i.e., mémorisés) au terme de la réalisation d'un même ensemble d'activités. Dans cette perspective, le post-test établi vise à cerner les connaissances procédurales et « sémantiques » acquises par les participants dans chacune des deux situations, à partir, pour les connaissances procédurales, de la liste a priori des savoir-faire nécessaires à la réalisation des animations proposées et, pour les connaissances « sémantiques », des concepts présentés dans le tutoriel et mis en œuvre au cours de l'interaction avec Flash. L'intérêt principal de ce questionnaire, par rapport à l'analyse des interactions, est de permettre une évaluation précise de la qualité de l'apprentissage grâce à des questions portant sur les concepts acquis. Son contenu sera détaillé ultérieurement, dans la section consacrée à l'analyse des interactions des participants avec Flash.

Concernant la *dimension affective* donnée par l'ACA à l'interaction, elle a été évaluée de façon objective par l'analyse des mouvements oculaires enregistrés pour l'ensemble des participants. Ceux-ci ont rempli en outre deux questionnaires d'évaluation de leurs impressions et jugements subjectifs, l'un verbal et l'autre non verbal, SAM. Les entretiens de debriefing post-expérimentaux, réalisés avec chacun des participants ont permis également d'obtenir des informations sur les réactions subjectives des deux groupes à la présence de l'ACA choisi. Questionnaires et entretiens sont, d'après [Ruttkay *et al.*, 2002], les principales techniques utilisées jusqu'à présent pour étudier la dimension affective introduite dans l'interaction par les ACAs. Cependant, le recours à un questionnaire non verbal est peu courant, à notre connaissance ; il constitue, avec l'utilisation de données oculométriques, une des originalités de cette étude.

Remarque : Aucune contrainte de temps n'a été imposée aux sujets pour la réalisation des animations, car les temps de réalisation des scénarios ne peuvent fournir aucune indication utile sur l'apport de la présence d'un ACA à l'efficacité de l'apprentissage. En outre, la tendance naturelle des novices (voir Carroll et al., 1987) est d'ignorer l'aide et de procéder par essais et erreurs ; l'analyse des questionnaires initiaux (voir plus loin dans ce chapitre) confirme cette tendance. Enfin, la présence d'un ACA peut augmenter le temps de réalisation des scénarios en raison de l'attrait qu'il peut exercer sur l'utilisateur et de la dimension ludique et/ou affective qu'il intro-

duit dans l'interaction. A noter que, pour éviter des passations interminables, l'expérimentateur « dépannait » les sujets inactifs ou bloqués pendant un certain temps dont la durée était laissée à son jugement.

5.2.4 Sujets

D'après [Ruttkay *et al.*, 2002], les principaux critères à prendre en compte lors de la sélection des participants à l'évaluation d'un ACA, notamment lorsque celle-ci porte sur son attractivité, incluent l'âge, le sexe, la culture générale et l'expérience informatique. Le protocole expérimental proposé tient compte de ces critères. Les participants sont tous des étudiants de Licence ; on peut donc considérer que les facteurs âge et culture générale sont contrôlés. En revanche, nous avons diversifié l'expérience informatique des participants en sélectionnant 14 étudiants engagés dans un cursus à dominante informatique, et 8 autres engagés dans des cursus sans enseignement de l'informatique en tant que discipline. Ce choix a rendu impossible le contrôle la variable sexe pour le groupe des informaticiens, la participation féminine aux formations universitaires spécialisées en informatique étant très limitée. Nous n'avons eu en revanche aucune difficulté à équilibrer le nombre d'étudiantes et d'étudiants dans le groupe des « non informaticiens ».

Nous avons décidé, dans le cadre de cette étude, de contrôler l'âge, le niveau de culture générale et, partiellement au moins, le sexe des participants, car ces facteurs nous semblent susceptibles d'exercer une influence moins importante que l'expérience informatique, sur les jugements subjectifs individuels des participants, compte tenu de la situation d'interaction et de l'application choisies. En raison du caractère prospectif de l'étude, nous avons choisi un profil représentatif de ce que seront les compétences d'une fraction importante des utilisateurs grand public dans quelques années, à savoir : un âge compris entre 19 et 22 ans à une exception près, un niveau de culture générale Bac + 2 ou 3, une expérience de la bureautique et d'Internet.

Le choix de participants jeunes offre des avantages supplémentaires. Ce sont des utilisateurs familiers des jeux électroniques mettant en scène des avatars (voir par exemple le succès de Lara Croft) ; par conséquent, on peut s'attendre à ce qu'ils soient moins sensibles à l'effet positif induit par le seul caractère innovant des deux situations expérimentales (i.e., système parlant et système parlant incarné) sur leurs jugements subjectifs. Les jugements qu'ils portent sur ces situations au terme d'une expérimentation limitée (une seule session) risquent d'être proches de ceux qu'ils pourraient formuler au terme d'un usage régulier prolongé. En revanche, les jugements de participants appartenant à des tranches d'âge supérieur risqueraient d'évoluer de façon importante lors du passage d'une utilisation ponctuelle expérimentale à un usage régulier, en raison de l'influence de l'effet de nouveauté sur leurs jugements initiaux. Enfin, classiquement, les sujets jeunes sont considérés a priori comme plus ouverts à l'innovation et plus adaptables à des situations nouvelles. Par conséquent, pour une première approche de l'évaluation d'un mode d'interaction nouveau au terme d'une expérimentation ponctuelle, il est préférable de privilégier cette catégorie d'utilisateurs potentiels, en vue d'obtenir des jugements subjectifs aussi « stabilisés » que possible. Il importe également, dans ce contexte où l'affectivité joue un rôle important, de sélectionner un premier ensemble de participants dont les profils, par rapport aux facteurs d'influence susceptibles d'agir sur leurs jugements, soient semblables entre eux. En effet,

la multiplicité de facteurs susceptibles d'agir sur les comportements humains d'ordre affectif est telle qu'il est nécessaire de contrôler le plus grand nombre possible d'entre eux.

Les résultats les plus significatifs obtenus jusqu'à présent sur la contribution positive des ACAs à l'interaction homme-machine sont ceux d'évaluations ergonomiques ponctuelles, faisant intervenir des groupes culturellement homogènes de sujets jeunes, élèves de l'enseignement primaire [Darves et Oviatt, 2004] ou adolescents élèves du secondaire [Lester *et al.*, 1997]. L'homogénéité des classes d'utilisateurs potentiels ciblées, ainsi que leurs caractéristiques spécifiques, peuvent expliquer les résultats de ces études pionnières.

5.3 Protocole expérimental - Réalisation

Cette section complète le chapitre précédent en précisant les aspects de la mise en œuvre du protocole propres à l'étude décrite dans ce chapitre.

Tout d'abord, nous décrivons la conception du personnage d'Eloona, les principes qui ont guidé la création de sa "personnalité", et leur mise en œuvre pour définir les expressions faciales de cet ACA, ses mouvements de tête, le timbre de sa voix, sa prosodie, le contenu et la longueur de ses énoncés.

Ensuite, nous présentons la conception ainsi que l'activation des messages d'aide dont le contenu couvre l'ensemble des informations nécessaires à la réalisation des scénarios d'animation à réaliser.

5.3.1 Animation d'Eloona

Nous présentons dans un premier temps les principes que nous avons mis en œuvre pour créer la « personnalité » d'Eloona. Nous décrivons ensuite nos choix de conception concernant son expressivité globale, puis la réalisation effective de ses expressions faciales et de ses énoncés.

Principes de création de la "personnalité" d'Eloona

Le premier principe que nous avons mis en œuvre, et le plus important, est que la présence d'un ACA doit être discrète, pour deux raisons au moins. D'une part, les « dialogues » homme-machine sont par essence finalisés. C'est-à-dire que la communication avec un ACA n'est pas une fin en soi pour l'utilisateur ; celui-ci interagit avec le système dans un but précis qui, dans le cas des applications grand public classiques par exemple, est la réalisation d'une activité ou l'obtention d'informations. Dans ce contexte, le rôle principal du système est d'assister l'utilisateur en l'aidant à atteindre les objectifs qui motivent son utilisation du logiciel. Si le système est « incarné » sous la forme d'un ACA, la présence de cet ACA doit donc être aussi discrète que possible pour ne pas interférer avec l'activité principale de l'utilisateur. D'autre part, les techniques d'animation et celles de synthèse vocale ne sont pas suffisamment perfectionnées pour qu'il soit possible à l'heure actuelle de créer des ACAs dont le comportement perceptible soit réaliste. De même, capacités de perception, de raisonnement et de compréhension de ces agents demeurent encore

très limitées. Si l'ACA est mis en évidence, sa présence risque de lasser rapidement l'utilisateur ; le manque de réalisme de son comportement risque même de l'irriter.

Le second principe est que le rôle principal des logiciels, au moins celui des logiciels courants, étant un rôle d'assistance, l'apparence et le comportement de l'ACA représentant le système doivent être en rapport avec cette fonction et le domaine d'application concerné. C'est ce principe qu'enfreint le Compagnon Microsoft qui représente l'aide en ligne des logiciels de Pack Office en raison de son apparence et de son comportement ludiques.

Nous indiquons dans la section suivante comment ces deux principes ont été mis en œuvre dans la création du comportement d'Eloona.

Définition du comportement d'Eloona

En application du second principe, nous avons choisi Eloona parmi les différents ACAs proposés par FT R & D sur la base de l'argumentaire suivant. On peut faire l'hypothèse que les principales qualités attendues d'un assistant à la prise en main d'un logiciel sont un comportement coopératif, capable d'exprimer une certaine empathie, de rassurer et de décontracter l'utilisateur novice. Par conséquent, un visage est préférable à un buste ou à un corps entier, car l'expression de l'empathie fait intervenir principalement le regard ; de même, un visage donne un sentiment de présence plus fort qu'un corps entier, et de ce fait est davantage capable de rassurer, détendre. Enfin, assister, rassurer, détendre sont des fonctions considérées comme féminines par les stéréotypes courants. Le visage d'Eloona est peut-être un peu fade pour un logiciel qui possède une dimension ludique inhérente, puisque Flash est un logiciel de construction d'animations. Mais, il est difficile de donner à la fois un aspect ludique et les qualités mentionnées ci-dessus à l'apparence et au comportement d'un ACA ; en outre, parmi les ACAs proposés par FT R & D, aucun ne possédait l'ensemble des qualités requises. En accord avec le premier principe, nous avons réduit la surface de l'écran occupée par Eloona autant qu'il était possible de le faire sans que ses expressions ne deviennent illisibles.

Nous l'avons également placée en haut à droite de l'écran, au-dessus de la zone des messages ; c'est une zone de l'écran relativement peu utilisée par les logiciels, donc sa présence ne risque pas de gêner l'utilisateur (voir la figure 1). Pour rendre sa présence discrète et donc acceptable, elle n'intervient que sur sollicitation de l'utilisateur et ne le regarde que pendant l'énoncé des messages. Lorsqu'elle est silencieuse, elle ne disparaît pas, mais se borne à détourner son regard de l'utilisateur en tournant la tête ; des apparitions-disparitions répétées risqueraient de fatiguer visuellement l'utilisateur et de le déconcentrer de son activité principale.

Enfin, pour assurer la crédibilité de la « personnalité » d'Eloona et introduire en même temps une source de variété dans son comportement, nous nous sommes efforcés d'assurer une cohérence aussi parfaite que possible entre ses expressions faciales, la prosodie de ses énoncés et le contenu informationnel des messages. Ce souci de cohérence nous a conduits d'abord à classer l'ensemble des messages (plus de 300) en fonction de :

- leur contenu informationnel, caractérisé par le type de la question correspondante ("Comment ?", "Quoi ?", "Pourquoi ?", "D'accord ?" ou requête au dictionnaire),

- l'acte de langage sous-jacent à l'énoncé : (i) apport d'informations, (ii) mise en garde/avertissement (e.g., indication de pré-conditions), (iii) conseils d'utilisation ou recommandations, (iv) explications,
- leur longueur, afin d'introduire un facteur de variété supplémentaire.

Ensuite, pour chacune des classes obtenues, nous avons défini un scénario spécifique d'animation qui précisait à la fois la séquence d'expressions faciales et de mouvements de la tête qu'Eloona devait réaliser, la prosodie globale des énoncés de cette classe et la synchronisation entre l'animation du visage et l'énonciation du message.

A noter que chaque séquence d'animation a été définie manuellement par composition d'expressions élémentaires, réalisées elles aussi manuellement. Ces primitives constituent un sous-ensemble du lexique d'expressions faciales proposé dans (Pélachaud, 2005 [5]) ; on trouvera dans la section suivante la liste des primitives que nous avons utilisées, le scénario général d'animation d'Eloona et, à titre d'exemple, les scénarios d'animation des réponses aux questions de type "Quoi ?" portant sur les icônes de l'interface de Flash. Pour montrer la cohérence entre les expressions faciales et la prosodie des énoncés, nous avons reproduit également la définition des consignes données à la locutrice pour les messages de ce type. Quant à la prosodie des énoncés pré-enregistrés, elle est définie globalement, sur l'ensemble de chaque énoncé ; elle fait intervenir le rythme et l'intonation avec les valeurs qualitatives suivantes que la locutrice s'est entraînée à appliquer de façon constante sur l'ensemble du corpus d'énoncés :

- Rythme : lent, moyen, rapide, varié.
- Intonation : monotone, variée, expressive (accents d'emphase).

En outre, des consignes ont été données à la locutrice pour qu'elle adopte, selon l'énoncé, l'un ou l'autre des styles suivants : factuel (e.g., type bulletin d'informations), didactique, empathique. Ce dernier style désigne un ensemble d'émotions variées en fonction de l'énoncé et du contexte dans lequel il est utilisé.

Réalisation du comportement d'Eloona

Sont présentés dans cette section :

- les primitives d'expression non verbale, expressions faciales et mouvements de la tête, que nous avons utilisées pour animer Eloona,
- le scénario général d'animation, et
- à titre d'exemple, les scénarios d'animation des réponses aux questions de type "Quoi ?" portant sur les icônes de l'interface utilisateur de Flash.

`Primitives d'expression non verbale utilisées`

Les paramètres d'animation que nous avons manipulés sont ceux qui interviennent dans le « lexique » des expressions non verbales proposé par C. Pélachaud :

[5] Il s'agit d'un document non publié transmis par l'auteur : Pélachaud, C. Communicative functions. Décembre 9, 2005, 13 pp. (conçu à partir de documents fournis par I. Poggi et B. De Carolis).

- Orientation de la tête (pour renforcer la direction du regard)
- Mouvements de la tête (hochements négatifs et positifs) ; voir Loredana Cerrato
- Sourcils (froncement, élévation)
- Joues (élévation dans certains sourires)
- Yeux (degré d'ouverture, élévation de la paupière supérieure, direction)
- Bouche (ouverture de la bouche, étirement des lèvres pour le sourire)

Parmi les primitives du lexique proposé par C. Pélachaud, nous avons retenu les combinaisons de paramètres correspondant aux intentions communicatives et/ou expressives suivantes :

- *Certitude* : froncement des sourcils, léger (accord avec l'interlocuteur), marqué (désaccord)
- *Réflexion* : regard vers le haut de côté ou paupières baissées et regard ailleurs que vers l'utilisateur (U)
- *Suggestion* : tête légèrement penchée, légère élévation des sourcils, regard tourné vers U
- *Avertissement* : paupières tendues, léger froncement des sourcils, regard tourné vers U
- *Approbation* : petit hochement de la tête, élévation des sourcils, petit sourire, élévation légère des joues, regard tourné vers U
- *Désaccord* : froncement de sourcils, regard tourné vers U
- *Accord* : petit hochement de tête, froncement des sourcils, regard tourné vers U
- *Conseil* : petit hochement de tête, regard tourné vers U
- *Confirmation* : petit hochement de tête, froncement des sourcils, regard tourné vers U
- *Incitation* : mouvement de tête vers l'avant, sourire, regard tourné vers U
- *Refus* : hochement négatif de la tête, et mouvement de la tête vers l'arrière et le haut
- *Information* : regard tourné vers U
- *Deictique* : direction de la tête et du regard vers un point de l'espace particulier
- *Fin de message (tour de parole)* : direction du regard ailleurs que sur U
- *Joie* : sourire, élévation des joues
- *Surprise* : élévation des sourcils, augmentation du degré d'ouverture de yeux, ouverture de la bouche, chute des mâchoires
- *Contentement* : sourire
- *Soulagement* : petit sourire, ouverture de la bouche
- *Satisfaction* : sourire, élévation des sourcils, hochement de tête

A noter que pour la réalisation des mouvements de la tête, nous nous sommes inspirés des travaux de L. Cerrato [Cerrato, 2002].

`Scénario général d'animation`

Lorsqu'aucun message n'est activé, Eloona a la tête légèrement tournée vers la droite (bord droit de l'écran). L'animation aléatoire définie par FT R & D, que nous avons conservée, lui fait bouger légèrement la tête (vers la gauche, la droite, le haut ou le bas) et cligner des yeux. Lors de l'activation d'un message, les expressions d'Eloona avant et après l'énoncé du message sont contrôlées. Ces deux courtes plages temporelles permettent à l'animation, alors manuelle, de manifester les intentions communicatives et expressives d'Eloona. Le retour à l'inactivité met en jeu l'expression Fin de tour de parole.

`Scénarios d'animation des réponses aux questions "Quoi ?" sur les icônes de Flash`

Les messages : Classification

On dispose de trois critères pour classer les 46 énoncés de réponse aux questions "Quoi ?" portant sur les icônes de l'interface utilisateur de Flash. Ces critères sont les suivants :

- Longueur du message vocal (3 valeurs) : court, moyen ou long
- Intention communicative (2 valeurs) : apport d'information ou conseil-explication
- Illustration du message par une copie d'écran (2 valeurs) : présente ou absente

L'application de ces critères permet de répartir les messages dans 9 ensembles distincts :

- Court + Information (CI)
- Court + Conseil-Explication (CC)
- Court + Conseil-Explication + Commentaire-Exemple (CCE)
- Moyen + Information (MI)
- Moyen + Conseil-Explication (MC)
- Moyen + Conseil-Explication + Commentaire-Exemple (MCE)
- Long + Information (LI)
- Long + Conseil-Explication (LC)
- Long + Information + Commentaire-Exemple (LIE)

Les messages d'un même ensemble ont été associés à un seul scénario d'animation ; par conséquent, 9 scénarios d'animation ont dû être créés pour les 46 réponses aux Questions "Quoi ?" sur les icônes de Flash.

Les Scénarios d'animation

Comportement commun à l'ensemble des animations Pour distinguer les réponses aux questions "Quoi ?" des autres catégories de messages, l'émotion globale d'Eloona est le Contentement (i.e., elle aime expliquer, enseigner, conseiller). Par conséquent, juste avant de parler, elle arbore un très léger sourire (bouche + joues) qui s'accentue légèrement après l'énoncé pour les messages de type Conseil-Explication. Elle regarde l'utilisateur avant et pendant l'énoncé (cf. les expressions Conseil et Information du lexique de C. Pélachaud), sauf pour les messages associés à une copie d'écran. Pour ces derniers, elle esquisse d'abord un sourire avant de pencher la tête vers le bas, en direction de la copie d'écran (cf. l'expression du lexique, Déictique). Après l'énoncé du message, elle relève la tête, regarde l'utilisateur et lui sourit brièvement, avant de tourner la tête vers le bord droit de l'écran.

Influence du contenu du message Pour les messages d'information, le sourire d'Eloona est très léger et disparaît pendant l'énonciation du message et après. On utilise l'expression Information du lexique proposé par C. Pélachaud. Pour les messages de type Conseil-Explication et les commentaires d'exemples, le sourire, qui est légèrement plus marqué et permanent, s'accentue à la fin de l'énoncé du message. Pour les commentaires des exemples qui accompagnent certains messages, le regard n'est pas dirigé vers l'utilisateur pendant l'énoncé du message mais vers la (ou les) copie(s) d'écran décrivant l'exemple.

Influence de la longueur Pour les messages courts, Eloona tourne rapidement la tête avant et après l'énonciation du message. Pour les moyens et les longs, ses mouvements de tête sont plus lents. Pour les longs, le mouvement de tête avant l'énoncé ralentit en milieu de course, et une expression de type Réflexion (voir le lexique présenté dans la section 1) se met en place pour un court instant. Pour les messages courts, on a choisi l'expression Suggestion, et pour les autres l'expression Conseil.

Expressions du lexique de C. Pélachaud utilisées : Conseil, Contentement, Information, Réflexion, Suggestion, Fin de message.

Scénarios détaillés

CI : Eloona tourne rapidement la tête vers l'utilisateur (i.e., vers la gauche) et le regarde avec un léger sourire qui disparaît pendant l'énoncé du message. Il reparaît brièvement après l'énoncé du message, avant qu'elle ne tourne la tête rapidement vers la droite.

CC : Eloona tourne rapidement la tête vers l'utilisateur (voir CI) en la penchant légèrement vers le bas (puisqu'elle est en haut de l'écran) et en élevant légèrement les sourcils. Elle sourit en même temps, plus largement que pour CI. Le sourire persiste pendant l'énoncé et s'accentue après, en même temps qu'elle tourne la tête vers la droite.

CCE : Eloona tourne rapidement la tête vers l'utilisateur (voir CI et CC). Elle regarde brièvement l'utilisateur en souriant comme dans CC, puis penche la tête et dirige son regard vers le bas (pour attirer l'attention de l'utilisateur vers les graphiques) en maintenant un léger sourire (voir CI) pendant l'énoncé. Elle doit tourner légèrement la tête pour regarder vers le milieu de la fenêtre d'aide en même temps qu'elle la penche vers le bas. Après l'énoncé, elle relève la tête, regarde l'utilisateur brièvement en lui souriant (sourire identique à celui de CC) avant de tourner rapidement la tête vers la droite.

MI : Identique à CI mais mouvements de tête plus lents.

MC : Eloona tourne la tête vers l'utilisateur (voir MI), le regarde en souriant comme dans MI et hoche légèrement la tête avant de parler. Elle maintient ce léger sourire pendant l'énoncé, l'accentue après l'énoncé avant de tourner la tête vers la droite à la même vitesse que pour MI.

MCE : Eloona tourne la tête vers l'utilisateur (voir MI), le regarde en souriant comme dans MI et en hochant légèrement la tête (voir MC), puis penche la tête et dirige son regard vers le bas (pour attirer l'attention de l'utilisateur vers les graphiques) en maintenant un léger sourire (voir MI) pendant l'énoncé. Elle tourne légèrement la tête pour regarder vers le milieu de la fenêtre d'aide en même temps qu'elle la penche vers le bas. Après l'énoncé, elle se comporte comme CCE, mais avec des mouvements de tête comme dans MI.

LI : Eloona tourne la tête vers l'utilisateur comme dans MI mais, à mi-parcours, elle s'arrête un court instant, paupières baissées (Réflexion), puis termine son mouvement en regardant l'utilisateur et en lui souriant légèrement un court instant. Le sourire disparaît pendant l'énoncé du message. Après l'énoncé du message son comportement est identique à MI.

LC : Eloona tourne la tête vers l'utilisateur les yeux baissés (Réflexion) à la même vitesse que dans MI, puis lève les yeux en fin de mouvement pour regarder l'utilisateur et lui sourire

(voir MC). Le sourire est maintenu pendant l'énoncé. Après l'énoncé, même comportement que MC.

LCE : Eloona tourne la tête vers l'utilisateur les yeux baissés (Réflexion), les lève pour le regarder et lui sourire avant de pencher la tête et diriger le regard vers le bas en tournant légèrement la tête pour regarder vers le milieu de la fenêtre d'aide. Le sourire est maintenu pendant l'énoncé. Après l'énoncé, elle relève la tête et la tourne vers l'utilisateur en levant les yeux pour le regarder et accentuer le sourire avant de tourner la tête vers la droite à la même vitesse que pour MI.

Prosodie des messages de réponse aux questions de type "Quoi ?"

Deux catégories de messages ont été considérées : les messages de définition des fonctions des icônes visibles de Flash, et ceux qui définissent les fonctions représentées par des items dans les menus. Dans chacune de ces catégories, deux sous-catégories de messages ont été distinguées : les définitions et les conseils d'utilisation. N'est décrite dans cette section que la prosodie des messages de définition du rôle des icônes. Nous avons affecté, comme émotion globale, aux réponses aux questions de type "Quoi ?" le *Contentement*, l'argument étant que si Eloona est coopérative, elle aime expliquer, enseigner, conseiller. Ces considérations expliquent le choix des valeurs des paramètres prosodiques ci-dessous, en particulier celui du style didactique.

Icônes
- Définitions : rythme moyen, intonation expressive, style didactique
- Conseils d'utilisation : rythme entre moyen et rapide, intonation variée, style empathique

5.3.2 Conception - Activation des messages d'aide

Comme nous avons pu le voir dans le chapitre précédent la zone d'interaction du système d'aide se compose de quatre boutons qui permettent à l'utilisateur d'interroger le système. Pour mémoire, les messages d'aide sont de la forme "Parole + Graphique", la parole étant simulée par un bouton que l'utilisateur actionne pour entendre l'énoncé et voir apparaître ou non, suivant la condition, l'ACA Eloona.

Bouton "Comment ?"
Exemple : Comment réaliser une interpolation de mouvement

Ce bouton, permet de mettre en correspondance par l'intermédiaire d'un lexique la représentation qu'a le sujet de la tâche et les fonctions offertes par le logiciel. Ainsi, quand le sujet appuie sur le bouton "Comment", la liste des actions possibles (premier niveau du lexique), exprimées dans un vocabulaire proche de celui de l'utilisateur, apparaît dans la fenêtre d'aide. En cliquant sur les items affichés, le sujet parcourt l'arborescence des actions possibles sur Flash ; les feuilles de cette arborescence correspondent aux messages d'aide fournis par le compère en réponse à des questions de type "Comment". Il peut, s'il le désire, remonter dans les niveaux de l'arborescence grâce à un bouton 'Back'. Cependant lorsqu'il clique sur une feuille, le message d'aide correspondant vient recouvrir le niveau du lexique couramment affiché.

Pour faire réapparaître le lexique, le sujet doit cliquer à nouveau sur le bouton "Comment" ; ce qui provoque l'affichage du premier niveau. A noter que l'arborescence contient trois niveaux au maximum : liste initiale, détail des éléments de la liste et titres des messages d'aide.
Certains mots, de couleur bleu, sont des liens vers les rubriques correspondantes du dictionnaire. Ils apparaissent aux différents niveaux du lexique sauf aux feuilles. En effet, sous la zone d'affichage des messages d'aide se trouve un menu déroulant, le dictionnaire, qui contient de brèves définitions des termes techniques utilisés dans le lexique ou les messages d'aide. Si l'utilisateur choisit un mot, une fenêtre de type pop-up (voir figure 5.1) vient recouvrir la fenêtre d'aide, à l'exception de son titre ; cette fenêtre contient la définition du terme sélectionné. Une fois la consultation terminée, l'utilisateur ferme la fenêtre pour revenir au lexique ou au message "Comment".

Les rubriques du dictionnaire définissent brièvement les termes apparaissant dans :
- les différents niveaux du lexique associé aux questions de type "Comment",
- les réponses au questions "Quoi" portant sur les objets visibles, les items de la barre des menus et certains items de ces sous-menus, ceux utilisés le plus souvent,
- dans les réponses aux questions "Comment".

Afin de réduire le temps de recherche/consultation, le menu déroulant du dictionnaire contient uniquement, à un instant donné, les termes techniques qui apparaissent dans le message "Quoi" ou "Comment" ou le niveau du lexique associé au "Comment" affiché à cet instant dans la fenêtre d'aide, ou dans le message (du dictionnaire) affiché dans la fenêtre pop-up du dictionnaire. Les définitions du dictionnaire peuvent en effet contenir, elles aussi, des liens vers les définitions des termes qu'elles utilisent.

En outre, quelques messages "Comment" contiennent des références à d'autres messages "Comment". Pour éviter que le sujet n'ait à repasser par le bouton "Comment" pour lancer ces messages, on place, à la fin du message courant affiché, des liens textuels qui renvoient vers ces messages. Lorsque le sujet clique sur un tel lien, une fenêtre contenant le message "Comment" correspondant au lien vient recouvrir la fenêtre d'aide, comme pour les messages du dictionnaire. Mais contrairement à ce qui se passe pour les autres messages, le menu déroulant du dictionnaire est désactivé afin d'éviter un empilement complexe des messages actifs où le sujet risquerait de se perdre. Pour obtenir la définition d'un terme technique inclus dans un tel message, il faut lancer l'affichage du message à l'aide du bouton "Comment" et du lexique associé à ce type de question. Comme pour le dictionnaire, pour revenir au message "Comment" initial, on ferme simplement cette fenêtre.

Le **contenu de la fenêtre d'aide** est, de haut en bas, le suivant (voir figure 5.1) :
- le titre du message (Texte),
- des copies d'écran (Graphique),
- au dessus de chaque copie d'écran, le(s) bouton(s) d'activation du message (ou des sous-messages) sonore(s) correspondant(s) ; les boutons sont rangés en colonne à droite de la fenêtre d'aide, dans l'ordre, de haut en bas, de leur succession dans le texte,
- pour certains messages "Comment", les liens vers d'autres messages "Comment",

– en dessous de la fenêtre d'aide, le menu déroulant des termes du dictionnaire utilisés dans le message sonore courant.

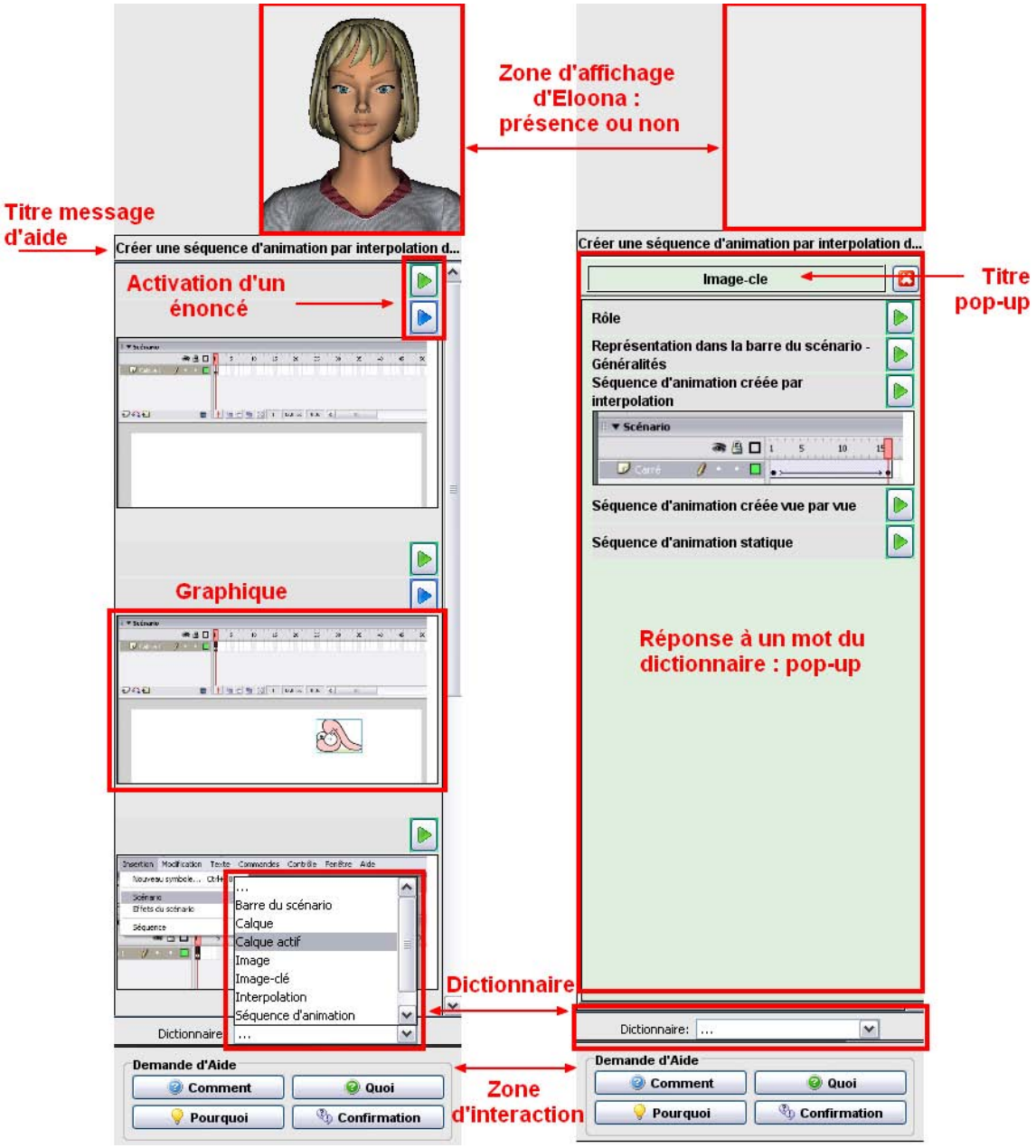


FIG. 5.1 – Interface utilisateur du système d'aide.

A noter que les copies d'écran occupent le maximum de la largeur de la fenêtre d'aide et que chaque copie peut être zoomée à l'initiative du sujet (par un clic sur celle-ci). Le graphique agrandi s'affiche au centre de l'écran dans une fenêtre dédiée. Pour le faire disparaître, le sujet ferme simplement cette fenêtre.

Pour les messages sonores, l'utilisateur a la possibilité d'arrêter l'énonciation d'un message sonore en cliquant sur le bouton qu'il a utilisé pour lancer ce message. Un bouton stop surchargerait l'interaction avec l'aide.

Messages : Certains impersonnels comme "on" ont été remplacés par des "vous" pour accentuer le caractère coopératif et humain de la relation.

Les messages comprennent trois types d'informations :
- des préconditions,
- des instructions (actions à effectuer), et
- des informations complémentaires (conseils, explications, détails procéduraux).

Chaque type d'information est exprimé dans un sous-message distinct autonome. A chaque sous-message correspond un bouton de type "Play" que l'utilisateur sélectionne lorsqu'il veut l'écouter. Les boutons d'activation des sous-messages sont de formes et de couleurs différentes selon la nature des informations qu'ils contiennent : icônes de type "Warning" pour les préconditions et icônes de type "Play" et de couleurs différentes pour les deux autres types d'informations. A noter que les boutons des instructions ont une saillance visuelle supérieure à celle des messages d'informations complémentaires.
De la sorte, on a un système d'aide adaptable, l'initiative étant donnée aux sujets de cliquer ou non, par exemple, sur les messages d'informations complémentaires. Pour ces derniers, le ton du clone est beaucoup plus coopératif/enjoué que pour les autres.

Bouton "Quoi ?" :
Quelle est la fonction, le rôle, de cette icône/menu ?

Lorsque l'utilisateur veut poser une question de type "Quoi", il clique sur le bouton "Quoi" dans le panneau d'interaction. A l'issue du clic, le message suivant apparaît dans la zone d'aide :

« A l'aide du bouton du milieu, cliquer à l'endroit voulu. »

L'utilisateur peut alors utiliser le bouton du milieu de la souris (inutilisé par Flash) pour sélectionner :
- soit une **une en-tête de menu** ; la liste des items de ce menu dont le rôle est commenté s'affiche alors dans la fenêtre d'aide. Chaque item est associé à un bouton qui permet de déclencher le message oral décrivant la fonction de l'item.

- soit **un objet visible** de l'interface de Flash ; la définition du rôle de cette icône ou de ce "Widget" s'affiche alors dans la fenêtre d'aide.

Pour réaliser cette seconde fonctionnalité, un ensemble de zones a été défini sur l'interface de Flash. Ce qui permet d'identifier l'objet graphique sur lequel porte la demande de l'utilisateur à partir de la position du clic sur l'interface de Flash, et de sélectionner le message correspondant sur le poste "Compère" ; le compère vérifie la pertinence de la sélection avant d'activer l'affichage du message sur le poste "Sujet".

Bouton "Pourquoi ?" :
Pourquoi ai je obtenu ce résultat, suite à une action ?

Lorsque le sujet effectue une requête "Pourquoi ?", le système répond positivement ou négativement. C'est le compère qui choisit entre ces deux types de réponse en fonction de la pertinence de la dernière action effectuée par le sujet.

Exemples de réponses positive et négative :
- « C'est normal. Voici un message qui vous donnera quelques explications. »
- « Vos actions ne correspondent pas à vos intentions. Je vais les défaire, puis afficher un message susceptible de vous aider à les réaliser. Vous pouvez aussi appeler le tuteur humain au 3083. »

Bouton "Confirmation ?" :
La dernière action effectuée est-elle compatible avec le but courant de l'utilisateur tel qu'il a été compris par le compère, ou avec les instructions du message "Comment ?" en cours d'exécution ?

Lorsque le sujet effectue une requête de type "Confirmation", le système répond positivement ou négativement. C'est le compère qui choisit entre ces deux types de réponse en fonction de la pertinence de la dernière action effectuée par le sujet.

Exemples de réponse positive et négative :
- « Oui, c'est ça. »
- « Non, non. Je défais les actions que vous venez d'effectuer et j'affiche un message susceptible de vous aider à les réaliser. »

5.4 Analyse des données

5.4.1 Profils des sujets - Questionnaire initial

Étant donné le caractère exploratoire de l'étude, nous avons choisi une population d'utilisateurs potentiels homogène (de façon à réduire le nombre de facteurs à prendre en compte). Compte tenu du caractère ponctuel de l'expérimentation (une seule session), nous avons sélectionné des sujets familiers des animations sur ordinateur, afin de limiter autant que possible l'effet de nouveauté susceptible d'exercer une influence prépondérante sur les comportements et attitudes des participants. En effet, l'objectif majeur est bien d'obtenir des jugements subjectifs correspondant à un usage aussi proche que possible d'un usage en situation réelle d'utilisation.

Ainsi, afin de respecter le premier critère, l'âge et le niveau culturel des participants varient peu. L'application du second critère nous a conduits à choisir des étudiants de Licence, d'un âge moyen de 20 ans. En outre, pour tenir compte d'un facteur dont l'influence sur les jugements subjectifs peut être importante, à savoir les compétences en informatique et l'intérêt pour ce domaine, les participants peuvent être divisés en deux groupes, suivant leur expérience informatique :

- Quatorze participants (**groupe GI**) sont engagés dans une formation à dominante informatique sans s'être encore vraiment spécialisés dans cette discipline (parcours informatique de la Licence Mathématiques et Informatique de l'Université Henri Poincaré, niveaux L2 et L3) ; ils possèdent donc une certaine expérience de la programmation, utilisent fréquemment un ordinateur et s'intéressent vivement à l'informatique.
- Les autres participants, au nombre de huit (**groupe GNI**) sont engagés dans d'autres cursus universitaires (avec comme dominante, sciences de la matière ou sciences du langage), ignorent tout de la programmation et ne s'intéressent pas particulièrement à l'informatique.

Le groupe GI comprend 13 étudiants et 1 étudiante et le groupe GNI comprend 3 étudiants et 5 étudiants. Ces répartitions illustrent, d'une part, la désaffection croissante des étudiantes pour certaines disciplines scientifiques dont l'informatique et, d'autre part, l'intérêt, réparti de façon plus équilibrée entre les deux sexes, pour les formations à dominante non informatique.

Donc, 22 étudiants volontaires, qui n'avaient jamais utilisé Flash auparavant ni d'outils de création d'animations (22 réponses non aux questions correspondantes du questionnaire), ont expérimenté les deux systèmes d'aide oraux à l'utilisation de Flash que nous avons réalisés, l'un doté d'un Agent Conversationnel Animé, Eloona (qui énonce les messages d'aide), l'autre non. Pour mémoire, l'ordre d'utilisation des deux systèmes a été contrebalancé ; voir la répartition correspondante des sujets en deux groupes dans le tableau 5.1). On rappelle également que les sujets ont été équipés d'un eye-tracker (de type casque) en vue d'obtenir des informations sur leur activité visuelle au cours de l'expérimentation des deux systèmes.

	Eloona	
	Présente	Absente
Scénario 1 et 2 : Décor + Animation des poissons	7 GI, 4 GNI	7 GI, 4 GNI
Scénario 3 : Site Web	7 GI, 4 GNI	7 GI, 4 GNI

TAB. 5.1 – Répartition des étudiants en deux groupes, suivant l'ordre d'utilisation des systèmes d'aide.

Les profils des sujets, tous volontaires et non rémunérés, sont décris dans le tableau 5.2. Ils ont été établis à partir des réponses des participants au questionnaire initial.

Groupes	**Age (Moyenne)**	**Scolarité**	**Nombre d'étudiantes**
GI (14 participants)	21 ans	6 L2 MI / 8 L3 MI	1
GNI (8 participants)	19 ans	3 Bac+1, 4 Bac+2, 1 Bac+3	5

TAB. 5.2 – Caractéristiques générales des participants : âge, scolarité, sexe. GI désigne le groupe des 'informaticiens', GNI celui des 'non informaticiens'.

Afin de caractériser plus précisément l'expérience informatique des participants, nous avons introduit, dans le questionnaire initial, des questions à choix multiple et des questions ouvertes

portant sur leur expérience de la programmation, des logiciels grand public et des jeux. Ces questions portent sur :

- les types d'équipements et les dispositifs d'interaction utilisés,
- les systèmes d'exploitation et les logiciels de conception graphique connus,
- les activités informatiques pratiquées (programmation, réalisation de pages Web, etc.),
- la stratégie adoptée pour se familiariser avec l'utilisation d'un nouveau logiciel (tutoriel, aide en ligne, approche par essais et erreurs etc.).

Nous avons regroupé dans le tableau 5.3 les réponses aux questions qui permettent d'établir des profils utilisateur pour les groupes GI et GNI concernant l'expérience informatique des participants et la démarche qu'ils adoptent pour se familiariser avec l'utilisation d'un nouveau logiciel.

Expérience en informatique	**Jamais**		**Parfois**		**Souvent**	
	GI	GNI	GI	GNI	GI	GNI
Activités						
Saisie ou consultation de données	1	1	7	4	5	3
Navigation sur Internet	0	0	1	1	13	7
Utilisation de logiciels grand public	0	0	3	6	11	2
Jeux	0	1	6	2	8	5
Programmation	1	7	1	1	12	0
Conception et développement d'applications	1	8	8	0	4	0
Utilisation d'interfaces de commande vocale	4	6	10	2	0	0
Prise en main d'un nouveau logiciel						
Une approche par essais-erreurs	1	1	7	3	6	4
L'aide en ligne associée au logiciel	2	2	7	5	5	1
Un manuel/notice d'utilisation	4	3	6	3	4	2
Un ouvrage de référence	3	7	10	1	1	0
Un didacticiel en ligne	2	4	9	4	3	0
Des exemples	0	2	4	3	10	3
L'assistance d'un utilisateur expérimenté	6	2	7	3	1	3

TAB. 5.3 – Expérience informatique des participants et stratégies d'apprentissage d'un nouveau logiciel.

Il ressort de l'analyse des tableaux 5.2 et 5.3 les conclusions suivantes :

- Tous les participants sont familiers de la navigation sur Internet, des jeux sur ordinateur ou console, et utilisent plus ou moins souvent des logiciels grand public.
- Les deux groupes GI et GNI se distinguent essentiellement par leur pratique d'activités de programmation, conception et développement logiciels. En effet, ces activités sont quasi étrangères aux participants du groupe GNI alors quelles sont fréquemment pratiquées par ceux du groupe GI. Les deux groupes se distinguent également l'un de l'autre par leur usage des interfaces de commande vocale, plus ou moins familières aux sujets du groupe GI, mais jamais utilisées par la plupart des sujets du groupe GNI.

– Concernant la prise en main d'un nouveau logiciel (domaine de l'aide en ligne), les approches les plus utilisées sont, pour le groupe GI, le recours à des exemples, la mise en œuvre d'une approche par essais et erreurs et, à un degré moindre, l'utilisation de l'aide en ligne associée au logiciel et/ou celle d'un didacticiel. A noter que les sujets de ce groupe font parfois appel à un utilisateur expérimenté. En ce qui concerne le groupe GNI, il ressort du tableau 5.3 que les approches les plus souvent utilisées sont l'expérimentation(i.e., démarche par essais et erreurs) et l'utilisation d'exemples.

L'analyse des jugements subjectifs des participants, qui est à l'origine des principaux résultats de notre étude, tient compte des profils individuels établis à partir des réponses au questionnaire initial.

5.4.2 Préférences des sujets : Présence ou absence d'un agent conversationnel

Jugement global - Questionnaire final

A la question « Quel système préférez-vous ? », reprise en synthèse par une question à choix multiple où les participants pouvaient choisir, pour exprimer leur préférence, entre les deux systèmes d'aide expérimentaux et les aides en ligne classiques (textuelles et éventuellement graphiques), un seul participant sur les 22 (S12, groupe GNI) a préféré les aides classiques aux deux systèmes expérimentaux. Hormis S12 qui préfère les aides classiques par conservatisme semble-t-il (« J'ai été habitué a ce mode » dit-il en commentaire), la moitié des participants préfère l'aide multimodale avec ACA à celle sans ACA et juge les deux systèmes expérimentés meilleurs que les aides classiques. Les partisans de la présence d'Eloona représentent presque la moitié du groupe GI alors qu'ils sont très largement majoritaires dans le groupe GNI (7 sur 8). En outre, ceux qui préfèrent l'aide sans ACA font tous les cinq partie du groupe GI. Les préférences globales des participants par groupe (GI, GNI) et par ordre de passation des deux conditions (aide en ligne avec ou sans ACA) sont regroupées dans le tableau 5.4.

Type d'aide		**Groupe GI**		**Gourpe GNI**	
		ACA : Sc1 et Sc2	ACA : Sc3	ACA : Sc1 et Sc2	ACA : Sc3
Multimodale (MM)	**5**	2	3		
Multimodale + ACA	**13**	2	4	4	3
MM et MM + ACA	**3**	3			
Classique	**1**				1
		7	**7**	**4**	**4**

TAB. 5.4 – Préférences des participants pour un type d'aide en ligne parmi les deux systèmes d'aide expérimentaux et les systèmes classiques. "ACA : Sc1 et Sc2" et "ACA : Sc3" font référence à l'ordre de passation ; Sc1, 2, 3 désignent les scénarios 1, 2 et 3.

L'analyse des commentaires exprimés par les sujets sur les deux systèmes expérimentaux dans les questionnaires verbaux montre néanmoins que les jugements des participants de GI sur l'aide avec ACA sont, dans une très large majorité, positifs. A noter que certains jugements,

par exemple ceux des sujets S5 et S10 du groupe GI, varient d'une question à l'autre. S5 indique dans un premier temps sa préférence pour l'aide avec ACA *« plus intéressant(e) car il y a une présence »* , puis nuance son appréciation en cochant les deux systèmes expérimentaux et en déclarant : *« Cela dépend surtout des questions et de l'utilisation que l'on veut faire du logiciel. Un lycéen préférera sûrement Eloona alors qu'un ingénieur se contentera de l'aide classique. »* . Quant à S10, après avoir de même exprimé sa préférence pour le système avec Eloona *« car plus agréable »* , il a choisi, en réponse à la question globale sur les préférences, le système sans ACA plutôt que les aides classiques ou le système avec ACA, en précisant : *« ...le clone peut nuire à la concentration lors de phases plus délicates. (Le système sans ACA) ne distrait pas comme l'autre. »*

`Conclusions sur l'expression des préférences` : Si l'on compare les deux groupes, on observe que les préférences du groupe GI se répartissent entre les deux systèmes d'aide que nous avons testé, à savoir Multimodal versus Multimodal + ACA. En revanche, pour les novices du groupe GNI, tous les participants à l'exception du sujet S12 préfèrent le système d'aide avec Eloona.

Interprétation des préférences : Évaluation des deux systèmes d'aide - Questionnaire final

Dans cette section, nous décrivons et analysons les jugements détaillés des participants sur chacun des deux systèmes expérimentaux, en vue de mieux comprendre les fondements de leurs préférences. Réponses et commentaires extraits des questionnaires verbaux sont regroupés comme suit, en fonction des préférences globales exprimées par les participants pour :

- l'aide avec ACA, groupe 'Avec' : 13 participants, **GI : S1, S16, S2, S6, S8, S22 ; GNI : S13, S15, S18, S21, S14, S17, S19** ;
- l'aide sans ACA, groupe 'Sans' : 5 participants, **GI : S3, S7, S4, S10, S20** ;
- les deux systèmes, groupe 'Les deux' : 3 participants du groupe GI **S5, S9, S11**. Nous ne tenons pas compte du sujet S9 dans l'analyse qui suit, car ce sujet s'est manifestement 'débarrassé' du questionnaire et du debriefing, par manque de motivation ou par fatigue. Concernant la réponse à la question où il doit comparer les deux systèmes, il affirme n'avoir *« pas senti de grosse différence entre les deux systèmes. »* . De plus, il ne donne aucun argument pour justifier son jugement global.

Remarque : Le groupe que l'on peut nommer 'Aucun', qui contient les participants (**S9 et S12**) qui préfèrent, soit les deux systèmes (S9), soit les systèmes classiques (S12)ne sont pas pris en compte dans l'analyse qui suit et qui porte, de ce fait, sur 20 participants.

`Efficacité et convivialité`

L'analyse des jugements subjectifs issus des comparaisons entre les deux systèmes expérimentaux, suivant les critères ergonomiques d'efficacité, de facilité d'utilisation et d'agrément, permet de mieux comprendre les motivations susceptibles de rendre compte des préférences exprimées par les participants. Les questions verbales qui fournissent les informations les plus intéressantes sont les deux suivantes « Quel système d'aide (parmi les deux expérimentés) vous a semblé le plus efficace (respectivement le plus agréable) ? Pourquoi ? ». Une question plus générale qui in-

vitait les participants à comparer les avantages et inconvénients des deux systèmes d'aide, l'un par rapport à l'autre, et chacun d'eux par rapport aux aides classiques, fournit également des informations utiles par les commentaires qu'elle a suscités.

Questions	Gr. 'Avec'		Gr. 'Sans'		Gr. 'Les deux'	
	MM+ACA	MM	MM+ACA	MM	MM+ACA	MM
Aide la plus efficace (15)	6	3	1	4	1	0
Aide la plus agréable (14)	9	0	5	0	0	0

TAB. 5.5 – Jugements des participants sur l'efficacité et la convivialité des deux aides expérimentales.

L'analyse des jugements regroupés dans le tableau 5.5 suggère plusieurs observations.

D'abord, si les participants sont divisés sur l'efficacité des systèmes expérimentaux, ils s'accordent parfaitement pour juger l'aide avec ACA plus agréable que celle sans ACA. Ainsi, 8 participants (i.e., 8 jugements sur 15 exprimés) jugent le système avec ACA plus efficace que l'autre, tandis que 7 le considèrent moins efficace. En revanche, 14 participants (14 jugements sur 14 exprimés) jugent le système avec ACA plus agréable que l'autre, et ce, quelle que soit leur préférence globale. Ce constat est important, car il montre indirectement que les choix de conception concernant l'apparence et le comportement de l'ACA sont pertinents alors même qu'ils se démarquent nettement des choix effectués à la fois pour les prototypes de recherche et les applications commercialisées.

Voici un aperçu des commentaires exprimés par les participants :

- Groupe 'Avec' :
  « Avec l'hologramme c'est plus attrayant » S14 ,
  « Avec le clone, car plus convivial » S16 ,
  « (Le système avec ACA) est plus clair » S2 ,
  « (Le système avec ACA est agréable grâce à l'animation de l'avatar » S22 ,
  « (Le système avec ACA) est le plus convivial pour quelqu'un qui ne sait pas ce qu'il faut faire » S6 .

- Groupe 'Sans' :
  « (Le système avec ACA est agréable), car il "humanifie" le système d'aide » S20 ,
  « (Le système avec ACA) m'a semblé le plus agréable grâce au clone. » S10 ,
  « (Le système avec ACA est plus agréable) à cause du clone. » S4 .

- Groupe 'Les deux' :
  « (Le système avec ACA est) plus intéressant car il y a une présence. Mais (les deux systèmes sont) aussi agréables l'un que l'autre » S5 .

On notera les avis positifs de l'ensemble des participants du groupe 'Sans' ayant exprimé leur avis sur l'agrément de la présence d'un ACA. S'ils préfèrent un système sans ACA, ce n'est pas parce qu'ils ressentent négativement la présence d'Eloona, mais parce que leur préférence globale est motivée par d'autres considérations que la convivialité, dont l'efficacité. A noter également

l'influence de la présence d'Eloona sur des jugements ayant trait à l'efficacité de l'aide ; voir, par exemple, le commentaire de S2 qui juge les messages d'aide du système avec ACA plus clairs que ceux du système sans ACA alors que les messages d'aide des deux systèmes étaient strictement identiques.

`Conclusions sur l'évaluation des deux systèmes` : Il semble que la présence d'un ACA soit perçue comme un facteur de nature à accroître la convivialité de l'aide. En revanche, les sujets sont partagés quant à l'influence d'une telle présence sur l'efficacité de l'aide. Ce dernier point mérite d'être approfondi. Il convient, en particulier, de vérifier si l'ordre de passation a exercé ou non une influence sur les jugements des participants. On peut en effet estimer surprenant les différences de jugements portés par les participants sur l'efficacité de l'aide, étant donné que les messages d'aide, l'interface de consultation et le comportement des deux systèmes étaient identiques.

`Tentative d'interprétation des divergences de jugements sur l'efficacité des systèmes`

L'effet observé dans la première ligne du tableau 5.5, à savoir une interférence éventuelle de la présence/absence d'un ACA avec les jugements subjectifs de l'efficacité de l'assistance, est beaucoup moins net pour les questions qui portent sur l'efficacité d'aspects particuliers des deux systèmes et n'imposent pas aux participants de comparer les deux systèmes entre eux sur ces aspects. C'est ce que suggère l'analyse des réponses aux questions qui se présentent sous la forme de différenciateurs sémantiques, chaque question étant répétée deux fois, une par système ; voir le tableau 5.6. Les réponses synthétisées portent, pour chaque système, sur les questions suivantes :

- Les informations (d'aide) obtenues ont-elles satisfait vos besoins ? (Pas du tout / entièrement)
- L'assistance fournie par les systèmes d'aide vous a semblé : (Inefficace / Efficace)
- Les connaissances que vous avez acquises sur la création d'animations Flash grâce à l'aide en ligne, vous semblent : (Insignifiantes / Importantes)
- Les informations fournies par l'aide en ligne vous ont-elles facilité la réalisation des scénarios ? (Pas du tout / Énormément)

Les données synthétiques présentées dans le tableau 5.6 montrent une homogénéité supérieure des jugements sur l'efficacité des deux systèmes à celle des jugements globaux comparatifs reportés sur la première ligne du tableau 5.5. En effet, si l'on cumule les réponses identiques et voisines (i.e., les réponses différant d'une seule case dans les deux échelles), on observe que plus de la moitié des participants jugent l'efficacité des deux systèmes équivalente, à l'exception de l'assistance à l'activité où les jugements d'un nombre supérieur de sujets varient d'un système à l'autre. L'analyse des réponses 'différentes' entre les deux systèmes ($>$ ou $<$ 1 case des échelles) confirme les observations faites dans le tableau 5.5, à savoir que la présence d'Eloona est susceptible d'influencer les jugements portés par certains des participants sur l'efficacité de l'aide.

Questions	Réponses (21, groupe 'Aucun' inclus		
	Identique	+ ou - 1	> ou < 1
Satisfaction des besoins	10	9	2
Efficacité de l'assistance	11	6	4
Importance des connaissance acquises	12	4	5
Assistance à la réalisation des scénarios	9	4	8

TAB. 5.6 – Réponses des participants aux questions sous forme de différenciateurs sémantiques (échelle à 7 cases) portant sur l'efficacité de chaque système.
+ ou - 1 : 1 case de différence entre les deux échelles.
> ou < 1 : différence absolue entre les deux échelles supérieure à 1 case.
Les réponses de S15 n'ont pu être prises en compte car elles ne portent que sur un seul des deux systèmes.

5.4.3 Apport potentiel d'un ACA à l'interaction avec une aide en ligne

Treize questions du questionnaire verbal (voir, en annexe B, la page 6 du questionnaire final) visent à recueillir les jugements subjectifs des participants sur l'apport éventuel, à l'efficacité et à la convivialité de l'interaction, d'une représentation par une tête parlante du système d'aide en ligne. En d'autres termes, l'objectif est d'évaluer les effets éventuels, positifs ou négatifs, de la présence d'un ACA en charge de l'assistance à la découverte et à la prise en main d'un nouveau logiciel sur :

- la motivation du novice pour la consultation de l'aide et son intérêt pour la découverte et l'apprentissage du logiciel ;
- l'attention qu'il accorde aux informations d'aide, à leur compréhension, et son investissement dans leur prise en compte dans les activités qui motivent son utilisation du logiciel ;
- son assurance, sa confiance en lui lors de l'interaction avec le logiciel, le stress lié à tout apprentissage de l'utilisation d'un nouveau logiciel ;
- l'agrément de l'interaction avec l'aide et de l'apprentissage du logiciel.

Évaluer précisément l'apport de la présence d'un ACA d'aide en ligne est un objectif difficile à atteindre, dans la mesure où les jugements subjectifs risquent d'être fortement influencés par l'apparence de l'ACA spécifique évalué, du comportement et de la « personnalité » dont celui-ci a été doté. Ces jugements ne sont pas indépendants non plus du domaine d'application dont relève le logiciel objet de l'apprentissage. Ce qui rend difficile toute généralisation des conclusions de l'analyse de ces jugements. Cette question de la portée des conclusions d'une telle étude, de leur généralité, est abordée dans la section 4, de même que celle de leur validité temporelle, qui se pose avec une acuité particulière, en raison du caractère innovant des interfaces utilisateur proposées, du fort ancrage social de la forme d'interaction homme-machine proposée, qui simule l'interaction entre humains.

Les treize questions sur les effets possibles de la présence d'Eloona sur la consultation de l'aide en ligne et l'interaction avec le logiciel utilisent des différenciateurs sémantiques (échelle à 7 cases) et peuvent être regroupées thématiquement comme suit :

1. Effets de la présence d'Eloona sur l'efficacité de l'interaction avec l'aide et sur son agrément.
   - Les effets de la présence d'Eloona sur l'efficacité des messages d'aide vous ont semblé : (Nuisibles / Bénéfiques)
   - Les effets de la présence d'Eloona sur l'agrément, la convivialité de l'interaction avec l'aide vous ont semblé : (Nuisibles / Bénéfiques)
   - Quelles ont été vos réactions à la présence d'Eloona ? (Irritation / Amusement)
2. Réactions affectives des participants à cette présence et effets de celle-ci sur leur efficacité dans l'interaction avec le logiciel dont l'utilisation est à maîtriser (Flash au cas particulier).
   - Quelles ont été vos réactions à la présence d'Eloona ? (Déconcentration / Décontraction ; Perturbation / Stimulation de l'activité)
3. Effets sur l'utilisation de l'aide : prise en compte des information d'aide, intérêt pour la consultation de l'aide.
   - La présence d'Eloona a-t-elle influencé :
   - Votre compréhension des messages oraux ? (Négativement / Positivement)
   - Votre intérêt pour l'aide en ligne et sa consultation ? (Négativement / Positivement)
4. Effets sur la motivation des participants pour la réalisation des activités interactives envisagées (au cas particulier, la réalisation des scénarios d'animation et l'apprentissage de Flash), sur l'agrément de cette activité et la qualité des résultats obtenus.
   - La présence d'Eloona a-t-elle influencé :
   - Votre investissement dans la réalisation des scénarios d'animation ? (Négativement / Positivement)
   - Votre motivation pour la découverte de Flash ? (Négativement / Positivement)
   - La qualité de la réalisation des scénarios et de l'apprentissage de Flash ? (Négativement / Positivement)
   - L'intérêt et l'agrément des tâches proposées et de l'apprentissage de Flash ? (Négativement / Positivement)
5. Effets sur l'assurance, la confiance en soi ressentie par les participants pendant la découverte du logiciel (Flash au cas particulier) et l'interaction avec lui.
   - La présence d'Eloona a-t-elle influencé : Votre confiance en vous et votre assurance lors de l'interaction avec Flash ? (Négativement / Positivement)
6. Évolution des réactions et de l'attitude des participants vis-à-vis d'Eloona au cours de l'expérimentation.
   - Votre perception d'Eloona a-t-elle évolué au cours de la session ? Si oui, comment ? (Négativement / Positivement)

Les réponses des participants aux questions ci-dessus sont analysées dans la section suivante : en premier les réponses qui expriment leur jugement subjectif global sur l'apport de la présence d'un ACA (i.e., les items 1, 2 et 5 de la liste ci-dessus), puis celles qui permettent de recueillir des jugements plus détaillés sur la contribution éventuelle de la présence d'Eloona à l'efficacité et à

la convivialité de l'interaction (items 3 et 4), et enfin, celle portant sur l'évolution des réactions et attitudes des participants au cours de la session (item 6).

Jugement global - Questionnaire final

Les jugements subjectifs concernant l'apport de la présence d'Eloona sont dans l'ensemble positifs, qu'ils soient globaux ou portent sur des points/aspects particuliers, et quelle que soit la préférence des participants pour l'une ou l'autre des aides expérimentales. Seuls les deux participants du groupe 'Aucun' (S9 et S12) expriment des jugements différents de ceux de la majorité des participants en réponse à la plupart des questions. Par conséquent, il est inutile de regrouper les réponses des participants en fonction de leurs préférences globales comme dans la section précédente.

Les réponses aux six questions destinées à recueillir les jugements globaux des participants sont présentées dans le tableau 5.7.

Questions (22 participants)	**Nombre de participants par type de réponse**						
Effets de la présence d'Eloona sur :	Nuisibles <———> Bénéfiques						
L'efficacité des messages d'aide	0	0	0	3	10	6	3
La convivialité de l'aide	0	0	0	1	3	7	11
Réactions à la présence d'Eloona :	1er différenciateur <———> 2ème différenciateur						
Irritation / Amusement	0	0	0	0	7	8	7
Déconcentration / Décontraction	0	1	0	4	7	5	5
Perturbation / Stimulation activité	0	0	0	9	4	7	2
Effets de la présence d'Eloona sur :	Négatifs <———> Positifs						
L'assurance, la confiance en soi	0	0	1	7	7	4	2

TAB. 5.7 – Jugements globaux des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction.
Les 7 colonnes de droite reproduisent les échelles, les différenciateurs correspondants figurant, soit dans la colonne de gauche, soit au dessus du groupe d'échelles auquel ils s'appliquent. Le nombre reporté dans chacune des cases représente le nombre de participants qui ont sélectionné cette case en réponse à la question correspondante.

Plusieurs constats ressortent de l'examen de ce tableau. D'abord, l'ensemble des participants jugent que la présence d'Eloona a une influence bénéfique à la fois sur la convivialité et sur l'efficacité de l'aide, mais à un degré moindre sur son efficacité (voir les deux premières lignes du tableau 5.7). L'analyse des jugements individuels montre, en outre, que même les sujets du groupe 'Sans' estiment que la présence d'Eloona a un effet positif sur l'efficacité de l'aide. Quant au jugement quasi unanime des participants sur l'apport bénéfique de la présence d'Eloona à la convivialité de l'aide, il est difficile à interpréter. En effet, les jugements seraient-ils aussi positifs au terme d'une exposition répétée au même ACA ? Quelle est la part de nouveauté dans ces avis ? Une solution pour pallier à la lassitude éventuelle des utilisateurs, pourrait être de varier l'apparence, le comportement, ainsi que la "personnalité" de l'ACA d'une session à l'autre. On pourrait également envisager de donner à l'utilisateur la possibilité de choisir l'apparence de

l'ACA comme le fait Microsoft ; l'effet de nouveauté serait ainsi prolongé. Dans cette perspective, on pourrait, à terme, proposer à l'utilisateur des outils logiciels évolués qui lui permettraient facilement de personnaliser ou définir entièrement l'apparence, le comportement et la personnalité de l'ACA, de les faire évoluer à sa guise. Cependant, des études supplémentaires dans un contexte d'utilisation prolongée sont nécessaires, avant d'orienter les recherches dans cette direction, pour s'assurer que ce phénomène de lassitude se produira effectivement pour un pourcentage significatif de la population d'utilisateurs visée. Une remarque de S3 pendant l'entretien de debriefing suscite un certain doute quant à l'apparition systématique de ce phénomène : *« C'est sympa la manière dont Eloona bouge les lèvres. Je pense que je ne me lasserai pas d'elle si Eloona n'intervient pas pour proposer de l'aide notamment comme le fait le trombone de Microsoft. Le trombone est drôle au début mais devient vite agaçant ! »* . Le succès des personnages de dessins animés qui restent inchangés d'une bande dessinée ou d'un dessin animé à l'autre confirme ce doute. Des questions du même ordre pourraient être posées à propos de la réaction générale d'amusement que suscite la présence d'Eloona. Toutefois, les résultats présentés dans le tableau 5.7 montrent la pertinence de nos choix de conception concernant l'apparence et l'animation de l'ACA (comportement, expressions faciales, timbre de la voix et prosodie).

Si la présence d'Eloona a amusé l'ensemble des participants, les effets de sa présence sur leur activité a suscité des jugements plus nuancés. 17 d'entre eux estiment qu'elle les a décontractés en réduisant le stress lié à la prise en main d'un nouveau logiciel et 13, seulement, qu'elle a stimulé leur activité et renforcé leur confiance en eux. On notera qu'aucun d'entre eux n'a jugé que cette présence les avait distraits de l'apprentissage de Flash et de la réalisation des scénarios ou avait perturbé cette activité. Un seul participant juge qu'elle a l'a distrait de son interaction avec Flash et a amoindri sa confiance en lui. Le commentaire suivant peut expliquer les jugements du nombre relativement important de sujets qui considèrent que la présence d'Eloona a été sans influence sur leur activité : *« Je pense qu'Eloona peut être distractif et peut devenir à la longue lassant, c'est pourquoi je préfère la condition sans Eloona. D'un autre côté Eloona ne m'a pas dérangé tout au long de l'expérimentation. Et, je pense qu'Eloona m'a motivé pour réaliser les animations et m'a également détendu, donc elle doit jouer sur le facteur stress. »* .

Si l'on analyse ces jugements en fonction de l'expérience en informatique des participants, c'est-à-dire en fonction de l'appartenance de leurs auteurs au groupe GI ou GNI on obtient le tableau 5.8.

Ce tableau met en évidence une différence majeure entre les jugements des groupes GI et GNI portant sur l'effet de la présence d'Eloona sur l'interaction avec Flash. La majorité des sujets du groupe GI (8 sur 14) estiment qu'elle a été sans effet sur leur activité alors que la plupart des sujets du groupe GNI (7 sur 8) considèrent qu'elle a eu un effet stimulant. A noter que les deux réactions négatives à la présence d'Eloona proviennent d'étudiants en informatique.

Pour conclure, les réactions des sujets à la présence d'Eloona, majoritairement positives, valident nos choix de conception concernant l'apparence et la « personnalité » d'Eloona. Par ailleurs, ils sont sans surprise. Ils reproduisent les stéréotypes actuels, à savoir qu'un personnage

Questions (22 participants)	**Nombre de participants par type de réponse**						
Effets de la présence d'Eloona sur :	Nuisibles <———> Bénéfiques						
L'efficacité des messages d'aide	0 / 0	0 / 0	0 / 0	1 / 2	9 / 1	3 / 3	1 / 2
La convivialité de l'aide	0 / 0	0 / 0	0 / 0	1 / 0	3 / 0	4 / 3	6 / 5
Réactions à la présence d'Eloona :	1er différenciateur <———> 2ème différenciateur						
Irritation / Amusement	0 / 0	0 / 0	0 / 0	0 / 0	6 / 1	5 / 3	3 / 4
Déconcentration / Décontraction	0 / 0	1 / 0	0 / 0	3 / 1	6 / 1	2 / 3	2 / 3
Perturbation / Stimulation activité	0 / 0	0 / 0	0 / 0	8 / 1	3 / 1	2 / 5	1 / 1
Effets de la présence d'Eloona sur :	Négatifs <———> Positifs						
L'assurance, la confiance en soi	0 / 0	0 / 0	1 / 0	4 / 3	6 / 1	2 / 2	1 / 1

TAB. 5.8 – Jugements globaux des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction en fonction de leur expérience en informatique.
A gauche du caractère '/' le cumul des réponses du groupe GI et à droite, le cumul des réponses du groupe GNI.

animé donne une dimension conviviale, ludique, à la situation dans laquelle il apparaît. Mais cette présence risque de devenir lassante dans le contexte d'un usage prolongé ; c'est du moins ce que craignent, à juste titre, les sujets qui ont une pratique informatique régulière et qu'illustre le commentaire reproduit plus haut. Un usage régulier ne rendra-t-il pas la présence d'un ACA lassante, voire irritante ? L'effet de surprise agréable passé, quelle sera l'attitude des utilisateurs vis-à-vis d'un ACA ? Une étude portant, comme la nôtre, sur une utilisation ponctuelle, ne permet pas de répondre à cette question cruciale pour l'avenir des ACAs en interaction homme-machine. Seule, l'analyse des usages sur une période de temps prolongée (e.g., plusieurs mois) permettrait d'y répondre. On peut toutefois faire l'hypothèse qu'en donnant à l'ACA une apparence et un comportement qui évoluent au cours du temps (par exemple au début de chaque session) de façon cohérente, non aléatoire, on parviendra à maintenir et même à renforcer l'intérêt des utilisateurs pour l'ACA et les effets bénéfiques de sa présence. Une autre voie de recherche pourrait être de donner à l'utilisateur les moyens de personnaliser et de faire évoluer l'apparence et le comportement de l'ACA à sa guise au cours du temps ; mais cette solution maintiendra le caractère ludique de l'ACA, avec le risque que sa présence ne distraie l'utilisateur de l'interaction avec un logiciel non ludique. A noter que cette question ne se posait pas dans le cadre du projet de recherche an partenariat avec FT R & D, car les applications envisagées sont de nature ponctuelle (e.g., aide à l'installation de la LiveBox).

`Apports de l'ACA : efficacité et convivialité :` Une très grande majorité des participants considère que la présence d'un ACA est susceptible d'améliorer l'efficacité et la convivialité d'une aide en ligne. La totalité d'entre eux ont jugé la présence d'Eloona amusante et, pour une majorité d'entre eux, elle a eu un effet décontractant. Toutefois, si la majorité des étudiants sans expérience informatique considère qu'elle a eu une influence positive sur leur activité et leur confiance en eux, la majorité des étudiants en informatique juge qu'elle a été sans effet sur leur activité.

Remarque complémentaire (SAM) : Si l'on compare les résultats des tableaux précédents aux réponses à la question globale non verbale suivante :

'Comment avez-vous ressenti, quelles ont été vos réactions par rapport à la présence d'un agent conversationnel pour « dire » les messages d'aide'

on constate que 19 participants sont satisfaits ou très satisfaits de la présence d'un ACA pour « dire » les messages d'aide (première ligne de figurines SAM). La répartition des jugements est comparable à celles obtenues pour l'évaluation verbale de l'efficacité des messages et de la convivialité de l'aide dans le tableau 5.8 (pas de différence entre les groupes GI et GNI). 13 participants ont été « excité/stimulé/intéressé », mais 5 sont indifférents ou totalement indifférents. Les réponses au questionnaire non verbal confirment donc les réponses au questionnaire verbal présentées dans le tableau 5.8. Enfin, la troisième ligne de figurines SAM fournit la même répartition que la dernière ligne du tableau 5.8. Pour 11 participants, la présence d'Eloona peut accroître la confiance en soi.

Jugements détaillés - Questionnaire final

Sont analysées dans cette section les réponses aux questions visant à déterminer l'influence, telle qu'elle a été perçue par les participants, de la présence d'Eloona sur :

- L'efficacité des messages d'aide (compréhension et prise en compte), d'une part, et l'intérêt des participants pour l'aide en ligne et leur motivation pour sa consultation, d'autre part.
- La motivation des participants pour la réalisation des scénarios d'animation et l'apprentissage de Flash, ainsi que la qualité des résultats obtenus.

Les réponses des participants aux 6 questions portant sur ces deux aspects sont présentées dans le tableau 5.9.

Questions (22 participants)	**Nombre de participants par type de réponse**						
Effets de la présence d'Eloona sur :	Négatifs <———> Positifs						
La compréhension des messages d'aide	1	0	0	6	6	7	2
L'intérêt pour l'aide en ligne	0	0	0	6	7	6	3
L'investissement dans/ la motivation pour :	Négatifs <———> Positifs						
La réalisation des scénarios	0	0	0	8	9	3	2
La découverte de Flash	0	0	0	5	4	7	5
L'intérêt/agrément des activités effectuées	0	0	0	2	8	9	2
La qualité de la réalisation de ces activités	0	0	1	3	8	7	2

TAB. 5.9 – Jugements détaillés des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction.

La majorité des participants, qu'ils préfèrent l'une ou l'autre des aides expérimentales, estiment que la présence d'Eloona est de nature à augmenter l'intérêt de l'utilisateur pour la

consultation de l'aide en ligne, sa motivation pour la réalisation des scénarios et l'apprentissage de Flash ; mais surtout, elle accroît l'agrément de ces activités pour une très large majorité d'entre eux, 19 sur 21 réponses (voir la répartition des jugements dans le tableau 5.9). Ces avis sont en accord avec les jugements globaux exprimés dans le tableau 5.7. Ils les renforcent même en élargissant leur portée à l'apprentissage de Flash et aux activités réalisées en interagissant avec lui.

Le fait qu'une très large majorité des participants ait estimé que la présence d'Eloona était propre à faciliter la compréhension des messages d'aide, accroître l'investissement dans la réalisation des scénarios, et la motivation pour la découverte de Flash, la qualité de la réalisation des scénarios d'animation et celle de l'apprentissage de Flash, soulève la même interrogation que les réponses à la question « Quel système d'aide vous a semblé le plus utile, le plus efficace... ». En effet, le contenu informationnel des messages d'aide et l'interface de consultation de l'aide étant les mêmes dans les deux systèmes exploratoires, ces réponses dénotent un effet de la présence d'Eloona sur les jugements subjectifs d'utilisateurs potentiels déjà observé dans les réponses aux questions portant sur la comparaison des deux aides exploratoires. Elles confirment et renforcent l'hypothèse de l'existence d'un tel effet. En raison de son importance, il conviendrait de consacrer à ce phénomène des études spécifiques qui permettent de mieux le comprendre, d'identifier en particulier les conditions nécessaires à son apparition, compte tenu de la diversité des réactions interindividuelles et des facteurs susceptibles d'agir sur sa durée de vie. En particulier, quel est le rôle de l'apparence et du comportement de l'ACA dans ce phénomène. Ces questions de recherche complexes, cruciales pour le succès de l'intégration des ACAs aux interfaces homme-machine, ont été abordées dans d'autres perspectives (marketing notamment) mais, à notre connaissance, elles n'ont pas fait l'objet de programmes de recherche systématique en interaction homme-machine où la problématique et donc le rôle d'un ACA sont différents.

`Jugements détaillés, efficacité et convivialité` : On observe que la présence d'un ACA influence positivement la perception de l'efficacité de l'interaction (jugement majoritaire) en accroissant l'intérêt pour l'aide et la motivation pour la réalisation des scénarios.

Elle améliore la convivialité de l'interaction en augmentant l'intérêt et l'agrément des activités.

Évolution éventuelle de la perception d'Eloona pendant la session - Questionnaire final

L'évolution éventuelle de l'attitude des participants vis-à-vis de l'ACA au fur et à mesure de l'interaction a fait l'objet d'une question verbale qui utilise comme différenciateurs sémantiques 'Négativement' et 'Positivement' sur une échelle à 7 cases.

Ainsi, les résultats à cette question (21 réponses, S10 n'a pas répondu) figurent dans le tableau 5.10.

Groupes (21 réponses)	**Nombre de participants par groupe**						
	Négativement <———> Positivement						
Groupe GI	0	0	1	3	6	2	1
Groupe GNI	0	0	0	5	1	0	2

TAB. 5.10 – Jugements détaillés des participants sur l'évolution de la perception d'Eloona au cours de la session.

8 participants sur les 21 qui ont répondu estiment que leur perception d'Eloona n'a pas évolué au cours de leur interaction avec le système d'aide qu'elle incarnait. Un seul sujet (S7) juge négativement l'évolution de sa perception d'Eloona. Pour les 12 autres, cette évolution est positive, voire très positive (pour 3 d'entre eux).
Toutefois, les jugements varient légèrement en fonction de l'expérience informatique des participants : pour le groupe GI, ils ont évolué positivement au cours de la session alors qu'ils sont restés stables pour le groupe GNI. Cette différence surprenante à première vue entre les deux groupes peut s'expliquer par les a priori généralement des informaticiens à l'égard des ACAS compte tenu du comportement souvent irritant des réalisations actuelles auxquelles ils ont exposés quotidiennement soit lorsqu'ils consultent Internet ou interagissent avec un logiciel grand public dont l'aide en ligne est incarnée par un ACA. Cette différence peut être due également à leurs a priori négatifs envers les aides en ligne classiques comme le suggère ce commentaire de l'un d'eux : *« Je ne lis jamais les aides en ligne à cause du pavé de texte qu'il faut lire, sous forme dictée ou 'en BD', c'est sûrement mieux... » S22* .

Ces résultats sont très encourageants, en particulier le fait que la perception d'un participant seulement évolue négativement au cours de la session. Ils montrent en effet qu'une exposition relativement prolongée à un ACA anthropomorphe (un quart d'heure de présentation de l'expérience par Eloona, puis entre un quart et trois quarts d'heure d'expérimentation avec Eloona incarnant l'aide en ligne) n'entraîne aucune réaction négative, lassitude par exemple, de la part des utilisateurs sauf exception (S7), et suscite une évolution positive des réactions de la moitié d'entre eux à la présence de cet ACA.

On peut avancer pour expliquer, au moins partiellement, les réactions positives de la plupart des participants à la présence d'Eloona, les choix qui ont présidé à la conception du comportement et des expressions de cette tête parlante :

– Courte présentation initiale de l'ACA en situation de monologue, l'ACA étant visualisé en grand format au centre et dans différentes zones de l'écran (voir la présentation détaillée du déroulement de l'expérience en annexe B) ; ce premier contact avec l'ACA permet à l'utilisateur de satisfaire sa curiosité vis-à-vis de l'ACA et d'évaluer le réalisme et la qualité technique de son apparence, de ses expressions faciales, et de son élocution. Cette présentation initiale est utile dans une situation d'interaction homme-machine où l'ACA est chargé d'assister l'utilisateur dans des activités cognitives complexes comme la découverte et l'apprentissage d'un nouveau logiciel ou l'installation et la configuration d'un service Web. Elle permet en effet d'éviter que l'utilisateur ne soit tenté de satisfaire sa curiosité

naturelle pour l'ACA en même temps qu'il interagit avec lui dans le cadre de l'activité pour laquelle il sollicite son assistance, ce qui risquerait de le distraire de cette activité et d'en perturber la réalisation et, par conséquent, à la longue, de le contrarier.

- Dans le souci également d'éviter ce risque, présence discrète de l'ACA pendant qu'il assiste l'utilisateur dans son activité. La mise en œuvre de cette exigence nous a conduits, au cas particulier, à :
  - Afficher l'ACA en format réduit, dans une zone fixe de l'écran proche de celle où sont visualisés les messages d'aide (voir la figure 5.1) ;
  - Lui interdire toute possibilité d'intervention « spontanée », et laisser l'initiative des échanges verbaux sous la responsabilité exclusive de l'utilisateur ;
  - Lors de ses interventions verbales, limiter son expressivité, en particulier celle de son élocution, de ses expressions faciales et mouvements de tête mais, pendant ses interventions, orienter son regard en direction de l'utilisateur ou, le cas échéant, vers les objets graphiques affichés qui sont mentionnés dans le message en cours d'énonciation ;
  - En dehors de ses interventions verbales, réduire l'animation de l'ACA à un ensemble de comportements aléatoires peu visibles (e.g., clignement des yeux, légers mouvements de la tête), et détourner son « attention » de l'utilisateur, par exemple en faisant tourner la tête à Eloona pour qu'elle se présente de trois quarts ou de profil.
- Enfin, choix d'expressions faciales, d'une élocution et d'une prosodie en rapport avec le rôle que l'ACA joue par rapport à l'utilisateur et le domaine d'application concerné. Au cas particulier, en raison de la situation d'interaction choisie (la familiarisation d'un utilisateur novice avec l'utilisation d'un nouveau logiciel), de l'apparence visuelle de l'ACA et du rôle qui lui était dévolu (assister l'utilisateur dans la prise en main du logiciel), il fallait que son comportement puisse être perçu par l'utilisateur comme coopératif, son attitude, ses expressions et sa voix comme affables ou enjouées. Eloona devait en outre paraître sérieuse, professionnelle, sans être ressentie comme maussade ni ennuyeuse. A noter que le fait de ne pas attirer l'attention de l'utilisateur sur l'ACA, de rendre la présence de celui-ci discrète, présente aussi l'avantage d'éviter que celui-ci ne se lasse du caractère répétitif et, plus globalement, du manque de « réalisme » des ACAs actuels, dus essentiellement aux limites que présentent encore aujourd'hui les techniques d'animation des ACAs anthropomorphes.

5.4.4 Facteurs à l'origine des préférences - Hypothèses

Nous analysons dans cette section les facteurs susceptibles d'avoir exercé une influence sur les préférences des participants pour l'un des deux systèmes exploratoires, avec ou sans ACA, ou les aides classiques (section 4.3) et sur leurs réactions à la présence de l'ACA Eloona (section 4.2).

Nous disposons, pour effectuer cette analyse :

- des données enregistrées et annotées (voir le logiciel d'annotation semi-automatique présenté dans le chapitre 4), à savoir la trace des interactions ainsi que les mouvements oculaires,
- du résultat des animations créées,
- des résultats du post-test destiné à évaluer l'assimilation des concepts de Flash,

– et des jugements exprimés dans les questionnaires finals et les entretiens de debriefing.
L'analyse de ces données vise également à mesurer l'influence éventuelle de la présence d'Eloona sur les performances des participants.

L'analyse porte sur une partie des données seulement, car nous nous sommes aperçus que les recours à l'aide en ligne étaient très rares au cours du troisième scénario (construction de l'introduction d'un site Web). Ceci provient du fait que l'acquisition de l'ensemble des connaissances (procédurales et conceptuelles) nécessaires à la réalisation des trois scénarios peut s'effectuer, et s'est effectuée en fait pour la plupart des participants, au cours de la réalisation des deux premiers scénarios (construction du décor de l'aquarium, scénario 1, et animation des poissons, scénario 2). Ainsi, les participants, lors de la réalisation du troisième scénario (qui représente la seconde condition), consultent l'aide uniquement pour se rappeler le détail de quelques procédures. Pour mémoire, les participants qui n'avaient pas réussi à atteindre les objectifs minimaux fixés pour la première condition (i.e., la mise en œuvre correcte des fonctions de Flash nécessaires pour réaliser les scénarios 1 et 2) bénéficiaient, avant d'entreprendre le scénario 3, d'une brève mise à niveau effectuée par l'expérimentateur qui leur expliquait les concepts non assimilés pour éviter que ces participants soient dans l'incapacité de réaliser le scénario 3. Pour ces deux raisons, nous avons limité l'analyse aux données recueillies pendant la réalisation des scénarios 1 et 2, condition plus représentative de la situation étudiée que la seconde condition (i.e., réalisation du scénario 3), à savoir la découverte ou prise en main d'un logiciel grand public avec l'assistance d'une aide en ligne multimodale (parole + graphique) matérialisée ou non par un ACA. En outre, la durée moyenne de réalisation de la première condition (40 minutes) est trois fois supérieure à celle de la seconde condition (14 minutes). Enfin, à la question : 'Quel scénario vous a semblé le plus attrayant (ou le moins ennuyeux) ? Pourquoi ?' du questionnaire final, 13 réponses sur 21 mentionnent le scénario 2 inclus dans la première condition : *« Le plus attrayant est le scénario avec l'animation de l'aquarium car il y a un effet assez agréable lors du rendu. » S6* . Cependant, c'est le scénario 2 qui est jugé très majoritairement comme le plus difficile (19 réponses sur 21). L'explication fournie par l'un des participants pour expliquer ce choix incite encore davantage à centrer l'analyse sur la première condition : *« Le plus difficile fut le scénario numéro 2 car il faut bien assimiler les notions d'animation. Mais ce n'était pas très dur quand même. » S6* .

Pour l'instant, l'analyse porte uniquement sur la première condition des 14 participants du groupe GI des étudiants engagés dans un cursus à dominante informatique, le plus intéressant à étudier de manière approfondie car le moins sensible à l'effet de nouveauté en raison de son expérience en informatique. Par conséquent, les résultats présentés dans les sections suivantes portent uniquement sur l'analyse de la trace annotée de ses interactions avec l'aide et Flash, les résultats des animations créées et du post-test, ainsi que les jugements exprimés dans les questionnaires finals et les debriefings sur les aspects innovateurs de l'aide en ligne proposée : messages multimodaux (Parole + Graphique), présence d'un ACA (apparence, comportement, « personnalité »

Activité des participants - Utilisation de l'aide

Afin de déterminer l'utilisation effective de l'aide au cours de la réalisation des scénarios 1 et 2, par participant et par condition de passation, nous nous sommes appuyés, d'une part, sur l'analyse des traces des interactions enregistrées tout au long de l'expérimentation et, d'autre part, sur les réponses aux questions portant sur l'évaluation de la qualité ergonomique de l'interface de consultation de l'aide en ligne.

`Aperçu global`

Pour mémoire, les participants ont effectué les deux conditions (avec et sans Eloona) dans un ordre contrebalancé. Ainsi, pour le groupe GI, 7 d'entre eux ont commencé par la condition 'avec ACA' pour les scénarios 1 et 2, ils forment un groupe que l'on appellera 'Groupe Avec Eloona ou Gr. AE' ; quant aux 7 autres qui ont utilisé d'abord une aide 'sans ACA' (pour réaliser les même scénarios), ils constituent un groupe nommé 'Groupe Sans Eloona ou Gr. SE'. Le tableau 5.11, présente la répartition des participants du groupe GI dans les sous-groupes AE et SE, ainsi que, pour chacun d'eux, le temps de réalisation en minutes des deux premiers scénarios.

Participants 'Sans Eloona' : Gr. SE	**S10**	**S2**	**S20**	**S22**	**S4**	**S6**	**S8**
Durée 1ère animation : Sc1 et Sc2 (min)	31	44	39	39	37	36	37
Participants 'Avec Eloona' : Gr. AE	**S1**	**S11**	**S16**	**S3**	**S5**	**S7**	**S9**
Durée 1ère animation : Sc1 et Sc2 (min)	43	34	38	38	41	36	35

TAB. 5.11 – Répartition des participants du groupe GI en fonction de l'ordre de passation et durée de réalisation des deux premiers scénarios (Sc1 et Sc2).

Il ressort de l'analyse du tableau 5.11 que la durée moyenne de passation des deux premiers scénarios est de l'ordre de 38 minutes. L'expérimentateur, au poste de compère, avait comme consigne de maintenir la durée de la passation dans des limites raisonnables. Plusieurs cas de figure se sont présentés :

- *Participant performant* : C'est le cas où le sujet a accompli rapidement les objectifs minimaux (décor + animation du poisson noir et jaune). Le compère lui laisse achever le reste de l'animation et intervient, au bout de 40 minutes environ, à la fin d'une tâche, par exemple la fin de l'animation d'un nouveau poisson. On évite de l'interrompre dans la création d'une animation pour ne pas le frustrer.
- *Participant perdu, bloqué* : Dans ce cas, le participant, n'arrive pas à créer l'animation demandée, il tente en vain diverses actions. Aussi, pour réduire son stress ou son énervement, le compère écourte la session, ce qui explique les durées inférieures à 38 minutes. C'est dans ce cas que le compère effectue une mise à niveau des connaissances conceptuelles ou procédurales du sujet avant qu'il n'aborde le scénario 3.

Afin d'avoir un aperçu global de l'activité des participants pendant la création des animations, il est intéressant d'analyser le nombre de clics (boutons gauche et droit confondus) qu'ils ont effectués et leur position à l'écran par rapport aux différentes zones définies lors de l'annotation. Pour mémoire, ces zones comprennent :

- *pour le logiciel Flash*, les zones : Menu (barre des menus de Flash), Bibliothèque (contient les objets graphiques à manipuler), Scénario (ou barre du scénario, lieu de création de l'animation) et Flash (scène + barre des outils)
- *pour le système d'aide*, les zones : Image (zoom sur un graphique du message d'aide courant), Dictionnaire (en-tête du menu déroulant d'accès au dictionnaire du système d'aide), Message (zone d'affichage des messages d'aide), Eloona (zone d'affichage de l'ACA), Dico (menu du dictionnaire affiché/déroulé) et Panneau (panneau d'interaction avec le système d'aide)
- ainsi que, pour Windows, les zones : Barre (barre des tâches où l'utilisateur clique pour obtenir l'affichage de l'animation à créer), Observation (affichage de l'animation à créer) et Sortie (interruption/terminaison de Flash.

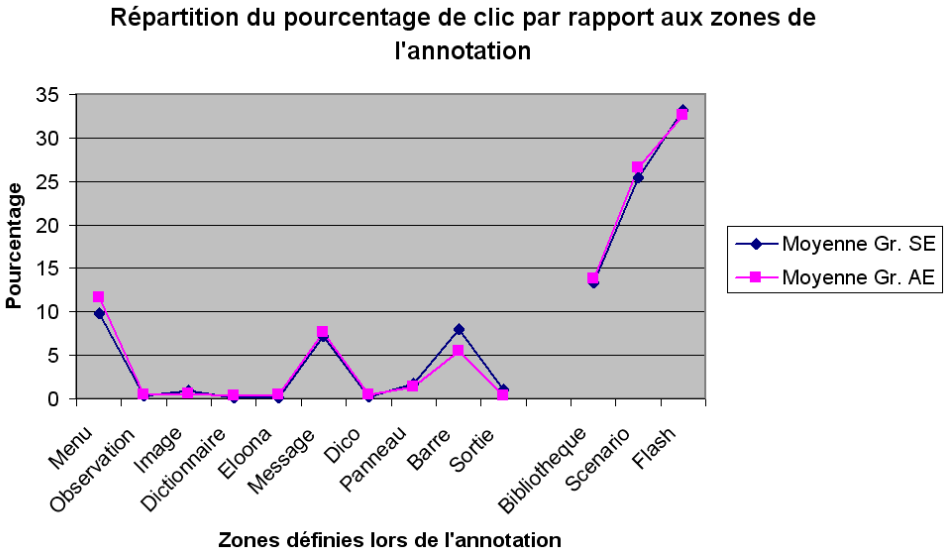


FIG. 5.2 – Distribution des clics par zone définie lors de l'annotation (pourcentages) - Sous-groupes AE et SE du groupe GI, scénarios 1 et 2.

Thèmes	**Moyenne sur le Gr. GI**
Flash	83%
Système d'aide	10%
Windows	7%

TAB. 5.12 – Distribution des clics par type d'activité (pourcentages) - Groupe GI, scénarios 1 et 2

Les résultats de l'analyse de la répartition des clics par zone, et donc par type d'activité, sont présentés, pour le groupe GI, dans la figure 5.2. On constate d'abord que la présence d'Eloona semble sans effet majeur sur l'activité des participants, puisque les courbes des groupes SE et AE se superposent quasiment ; à noter que la zone 'Eloona' ne fait l'objet d'aucune activité de la part des participants du groupe AE, ce qui est normal puisque c'est une zone non cliquable.

L'examen des zones montre également que la majorité des interactions des participants portent sur Flash et moins de 10% sur l'aide, ce qui est conforme à nos attentes compte tenu de la situation d'interaction choisie. L'étude des données oculométriques nous permettra d'approfondir ces résultats, en particulier de déterminer l'attention accordée par les participants du groupe AE à l'ACA durant leurs diverses activités.

Plus globalement, le tableau 5.12 montre que 10% et 83% de l'activité du groupe GI, sans distinction de condition et telle qu'elle apparaît à travers la répartition des clics entre Flash, l'aide et Windows, ont porté, respectivement, sur l'aide et sur Flash. Ce taux de consultation de l'aide peut sembler décevant pour un système d'aide innovateur que les participants étaient inviter à évaluer, donc à utiliser intensivement. Cependant, il n'est pas représentatif de l'utilisation effective de l'aide qu'il tend à sous-estimer. En effet, lorsque l'affichage d'un message d'aide était activé, le messge restait affiché jusqu'à l'activation d'un noveau message ; l'utilisateur pouvait donc le consulter à plusieurs reprises. L'analyse des données oculométriques nous permettra d'obtenir des indications plus fiables et plus précises sur l'utilisation effective des messages d'aide.

`Interactions avec le système d'aide`

Une partie des données issues de la trace des interactions du groupe GI pendant la réalisation des scénarios 1 et 2 sont présentées dans le tableau 5.13. Ces données décrites sujet par sujet sont regroupées par condition, AE versus SE, dans la dernière colonne du tableau.

Participants 'Sans Eloona' : Gr. SE	**S10**	**S2**	**S20**	**S22**	**S4**	**S6**	**S8**	**Moy.**
Questions de type "Comment" activés	15	4	3	4	8	7	8	7.0
Messages sonores activés	32	12	17	5	11	11	29	16.7
Questions de type "Quoi"	1	4		2		3	1	X
Questions de type "Pourquoi"								X
Questions de type "Demande de confirmation"						1	1	X
Participants 'Avec Eloona' : Gr. AE	**S1**	**S11**	**S16**	**S3**	**S5**	**S7**	**S9**	**Moy.**
Questions de type "Comment"	5	13	5	4	9	8	3	6.7
Messages sonores activés	30	31	9	8	38	32	2	21.4
Questions de type "Quoi"	2		7	1	2	1	1	X
Questions de type "Pourquoi"	1		2		1		1	X
Questions de type "Demande de confirmation"	1							X

TAB. 5.13 – Traces des interactions des participants du groupe GI, sujet par sujet, et regroupées par condition, AE vs SE, dans la dernière colonne.

Ces données, issues de l'annotation, précisent le nombre de requêtes de chaque type formulées au système d'aide, ainsi que le nombre de messages sonores activés (questions de type "Comment" et "Quoi"). Si le nombre de consultations de l'aide varie considérablement d'un participant à l'autre, le nombre moyen de questions "Comment" varie peu d'une condition à l'autre (7 contre 6.7 questions), et tous les participants ont formulé au minimum trois requêtes de type "Comment". En ce qui concerne les autres types de requêtes, les effectifs sont trop limités pour permettre des comparaisons significatives entre les deux conditions. Les nouvelles formes

de consultation de l'aide proposées ont donc été peu utilisées, à l'exception des demandes d'informations procédurales (i.e., les questions de type "Comment"), ce qui confirme les conclusions formulées par les auteurs du "minimum manual" à partir de l'observation dès les années 80.
Un résultat intéressant, en revanche, est le nombre de requêtes de type "Pourquoi". En effet, l'examen du tableau 5.13 révèle que seuls les participants qui bénéficiaient de la présence d'Eloona ont effectué ce type de requête (4 participants sur 7). Les commentaires de deux participants du sous-groupe AE suggèrent des pistes d'interprétation de cette différence : *« La présence d'Elonna rend le système plus humain/convivial. » S1 (debriefing)* et *« Eloona est vraiment un plus dans le cas présent de cette expérimentation. J'aurais été plus réticent sans Eloona et je me sens plus à l'aise avec Eloona. » S5 (debriefing)* . Ces commentaires suggèrent que les requêtes de type "Pourquoi" ne sont utilisées que si l'utilisateur pense que le système est intelligent. Ainsi, la présence de l'ACA est-elle susceptible d'influencer la perception que l'utilisateur a de l'intelligence du système ? A noter, que le recours à ce type de requête suppose, de la part de l'utilisateur, une confiance dans les capacités du système à lui fournir une explication sur le résultat de ses actions, donc la croyance en son intelligence.
Un autre résultat important, fourni par le tableau 5.13, concerne le nombre de messages sonores activés par les participants. Malgré une forte diversité inter-individuelle, le nombre moyen de messages sonores activés par les participants du sous-groupe AE est nettement supérieur à celui des messages sonores activés par le sous-groupe SE (21.4 messages contre 16.7). Il semble donc que la présence du clone stimule la consultation de l'aide qui passe par l'écoute des messages oraux. Mais le nombre de participants n'est pas suffisant pour que cette différence soit statistiquement significative. A noter que l'observation des mêmes données regroupées en fonction des préférences exprimées par les participants (i.e., les groupes 'Avec', 'Sans', 'Les deux' et 'Aucun') ne permet pas d'établir un lien entre la préférence pour l'un des systèmes d'aide et son utilisation.
Enfin, pour approfondir l'étude quantitative des messages oraux activés, nous nous sommes interrogés sur le taux des messages ré-activés, donc ré-écoutés, ce qui permet d'obtenir une idée plus précise sur l'attention et l'intérêt effectifs accordés par les participants aux informations d'aide. La méthode consiste à compter le nombre de messages déjà écoutés, puis à calculer la différence avec le nombre total de messages écoutés. Les résultats de cette analyse sont présentés dans le tableau 5.14.

Le tableau 5.14 montre à nouveau la diversité inter-individuelle des comportements des participants. Il met en outre en évidence une différence entre les sous-groupes AE et SE, liée vraisemblablement à la présence d'Eloona, à savoir que le pourcentage de ré-écoutes est moins élevé avec Eloona. La présence d'Eloona agit-elle positivement sur la concentration et l'assimilation des messages d'aide ou bien, à l'opposé, joue-t-elle sur le caractère décontractant lié à sa présence ? A l'appui de ces hypothèses, on peut se rappeler les jugements exprimés dans les tableaux 5.9 et 5.10 où la majorité des participants indiquent que la présence d'Eloona favorise la compréhension des messages d'aide, ainsi que la décontraction nécessaire à l'apprentissage d'un nouveau logiciel. Un commentaire issu d'un debriefing confirme la seconde hypothèse : *« (Il) pense qu'Eloona l'a motivé pour réaliser les animations et l'a également détendu, donc elle doit jouer sur le facteur stress. » S3 (debriefing)*

Participants 'Sans Eloona' Gr. SE	**S10**	**S2**	**S20**	**S22**	**S4**	**S6**	**S8**	**Moy.**
Messages sonores écoutés	32	12	17	5	11	11	29	16,7
Messages sonores ré-écoutés	6	1	6	0	5	1	18	5,2
Pourcentage de ré-écoute	18%	8%	35%	0%	45%	9%	62%	25,5%
Participants 'Avec Eloona' Gr. AE	**S1**	**S11**	**S16**	**S3**	**S5**	**S7**	**S9**	**Moy.**
Messages sonores écoutés	30	31	9	8	38	32	2	21,4
Messages sonores ré-écoutés	14	3	0	2	10	4	0	4,7
Pourcentage de ré-écoute	46%	9%	0%	25%	26%	12%	0%	17,1%

TAB. 5.14 – Nombres et pourcentages de ré-écoute de messages sonores, sujet par sujet, et regroupés par condition, AE vs SE, dans la dernière colonne (groupe GI).

Afin également de recueillir les réactions des participants face à l'interface de consultation de l'aide, nous avons analysé leurs réponses aux questions portant sur cet aspect dans le questionnaire final. Dans la mesure où l'interface de consultation de l'aide était identique pour les deux systèmes, il suffit de vérifier, en raison de son caractère innovant, qu'elle n'a pas suscité de réactions négatives de la part des participants, réactions qui auraient pu avoir une influence sur les jugements subjectifs qu'ils ont portés sur chaque système et sur les préférences qu'ils ont exprimées pour l'un d'eux ou pour les aides classiques.
Bien qu'ils aient très peu utilisé les possibilités d'interrogation supplémentaires, par rapport aux aides classique, qui leur étaient offertes, les participants se sont montrés satisfaits de l'interface de consultation de l'aide proposée. Les réponses au trois questions (sous forme de différenciateurs sémantiques) portant sur la difficulté/facilité d'utilisation de cette interface, son efficacité/inefficacité, le caractère fastidieux/agréable de l'interaction sont toutes positives, voire très positives pour la première question (10 réponses sur 14). Les réponses de la plupart des participants aux deux autres questions sont également positives avec toutefois 2 participants qui trouvent l'interface moyennement efficace et 2 participants qui jugent son utilisation légèrement fastidieuse.
Compte tenu du caractère positif des jugements exprimés verbalement par la majorité du groupe GI sur l'interface novatrice de consultation de l'aide commune aux deux systèmes d'aide, on peut estimer que cette composante des deux systèmes a joué un rôle négligeable dans la détermination des préférences et jugements globaux que les participants ont exprimés dans les questionnaires verbaux concernant l'apport de la représentation d'un système d'assistance par une tête parlante à l'interaction.

Analyse détaillée de l'activité des participants

Nous détaillons ici l'activité des participants du groupe GI au cours de la réalisation des scénarios 1 et 2, en distinguant comme précédemment les sous-groupes 'avec Eloona' et 'sans Eloona'. L'analyse s'est effectuée ainsi. Tout d'abord, il a fallu annoter la trace des interactions en posant des marques temporelles et en définissant des zones d'attention visuelle en fonction de l'état courant de l'affichage et des actions possibles de l'utilisateur compte tenu de cet état (voir

le tableau 4.1 du chapitre précédent). Puis, à partir des données oculométriques, on a déterminé le nombre, la position ainsi que la durée des fixations oculaires correspondant à chaque action. Enfin, on a calculé, pour chaque participant et pour chaque action, les zones pré-définies sur l'affichage qui ont mobilisé son attention visuelle, la durée du séjour du regard dans chaque zone (en pourcentage de la durée de l'activité observée), ainsi que la durée moyenne des fixations dans chaque zone. Le tableau 5.15 présente ces données pour 10 participants du groupe GI et pour les actions les plus représentatives de leur activité ; figurent dans ce tableau, les pourcentages de temps d'observation des différentes zones pré-définies, ainsi que la durée moyenne des fixations pour chaque zone (en millisecondes). Les données oculométriques des 4 autres participants du groupe GI n'ont pu être utilisées car elles se sont avérées de mauvaise qualité pour trois participants : deux ont bougé le miroir induisant la perte du regard pendant une partie de l'enregistrement et le reflet cornéen a cessé rapidement d'être détecté pour le troisième car "l'oeil s'est mis à pleurer". Quant au quatrième, un incident technique a interrompu l'enregistrement pendant un court instant, ce qui a introduit un décalage temporel impossible à déterminer avec une précision suffisante entre la trace des interactions et les mouvements oculaires.

Le tableau 5.15 fournit des informations très intéressantes sur les éléments de l'affichage qui sollicitent successivement l'attention des participants du groupe GI au cours de la création des deux premiers scénarios. Ces informations complètent l'analyse de leur activité à partir des clics souris qu'ils ont effectués. A noter que la somme des pourcentages par participant et en moyenne n'atteint pas 100% car une partie seulement des activiés effectuées par les participants au cours de la réalisation des scénarios 1 et 2 ont été prises en compte. Nous détaillons les activités que nous avons retenues car elles nous ont paru les plus représentatives par rapport à la situation choisie, l'apprentissage de l'utilisation d'un noveau logiciel avec l'assistance d'une aide en ligne multimodale représentée ou non par un ACA :

- **Menu** : On considère comme actions de type 'Menu', l'usage de la barre des menus, c'est-à-dire le clic gauche qui ouvre un menu et les raccourcis clavier de type CTRL+Z qui permettent d'annuler une action par exemple. Ainsi, on observe que l'ensemble des participants étudiés (10 participants) consacre en moyenne près de 10% du temps total d'activité à l'observation des menus, avec des variations inter-individuelles importantes puisque les pourcentages minimum et maximum sont respectivement 5% et 16%. Ces variations sont dues aux différentes stratégies des étudiants lors de la découverte d'un nouveau logiciel, certains explorant les fonctionnalités du logiciel, d'autres consultant davantage l'aide. On n'observe aucune différence notable entre les mloyennes des participants des groupes SE et AE. On peut donc conclure, au premier abord, que la présence d'Eloona est sans influence sur les accès aux menus. Nous verrons par la suite si cette conclusion tient lorsqu'on analyse en détail les menus consultés.

- **Résultat** : Cette rubrique correspond à l'affichage de l'animation que l'étudiant doit créer. Une fenêtre contenant cette animation (qui est jouée en boucle) est accessible par la barre des tâches de Windows. Les résultats montrent que les participants observent durant 3% du temps total l'animation à créer, si l'on exclut le sujet S22 qui consacre 18,8% du temps total à l'observation de l'animation des poissons ; ou bien ce participant s'est complu à regarder les poissons animés, ou bien il était très miticuleux et s'est efforcé de reproduire

à l'identique l'animation. Concernant la durée moyenne des fixations, comme il s'agit d'un décor statique et d'une scène animée, les fixations ont une durée plus courte (330ms) que pour les menus (390ms) qui font intervenir des processus de lecture et, éventuellement de réflexion/hésitation.

Participants 'Sans Eloona' SE		**S10**	**S2**	**S20**	**S22**	**S4**	**S6**	**S8**	**Moy.**
% d'attention visuelle	-> Menu	8,8	16,8	X	5,0	X	12,3	8,6	10,3
	-> Résultat	2,7	2,4	X	18,8	X	2,8	1,6	5,6
	-> Bibliothèque	8,7	7,6	X	7,1	X	7,3	8,0	7,7
	-> Scénario	16,5	20,3	X	22,4	X	36,1	17,4	22,6
	-> Flash	30,0	37,7	X	34,8	X	25,0	43,4	34,2
	-> Lexique Comment	11,8	4,8	X	1,4	X	4,4	5,7	5,6
	-> Message Comment	13,8	4,3	X	1,9	X	5,4	10,4	7,2
Durée moyenne fixations (ms)	-> Menu	442,4	450,5	X	443,3	X	501,4	310,1	429,5
	-> Résultat	328,1	294,3	X	599,5	X	337,9	341,6	380,3
	-> Bibliothèque	443,8	419,9	X	461,0	X	506,2	341,0	434,4
	-> Scénario	453,7	442,6	X	474,4	X	695,7	286,0	470,5
	-> Flash	414,8	442,5	X	449,9	X	478,1	313,9	419,8
	-> Lexique Comment	414,5	377,4	X	372,0	X	357,3	255,4	355,3
	-> Message Comment	351,0	357,3	X	320,6	X	367,7	256,4	330,6
Participants 'Avec Eloona' AE		**S1**	**S11**	**S16**	**S3**	**S5**	**S7**	**S9**	**Moy.**
% d'attention visuelle	-> Menu	11,2	9,5	6,8	X	X	12,0	14,2	10,8
	-> Résultat	3,4	2,9	5,6	X	X	3,5	1,8	3,4
	-> Bibliothèque	7,0	7,6	7,2	X	X	9,1	4,2	7,0
	-> Scénario	29,2	19,2	21,1	X	X	18,7	39,4	25,5
	-> Flash	25,2	31,7	37,6	X	X	35,5	29,4	31,9
	-> Lexique Comment	5,9	9,2	6,9	X	X	3,0	2,2	5,4
	-> Message Comment	7,7	16,0	2,9	X	X	10,0	2,4	7,8
Durée moyenne fixations (ms)	-> Menu	393,2	377,8	403,9	X	X	427,9	359,5	392,5
	-> Résultat	306,3	311,4	337,8	X	X	320,1	379,0	330,9
	-> Bibliothèque	364,2	346,3	503,1	X	X	340,7	404,9	391,8
	-> Scénario	449,7	406,8	501,7	X	X	400,1	494,6	450,6
	-> Flash	299,9	357,1	447,8	X	X	379,9	473,8	391,7
	-> Lexique Comment	286,8	325,0	300,7	X	X	297,2	275,2	297,0
	-> Message Comment	285,9	310,0	294,0	X	X	297,2	312,9	300,0

TAB. 5.15 – Répartition de l'attention visuelle des participants du groupe GI dans les zones pré-définies de l'affichage et durée moyenne des fixations dans ces zones pour certaines activités de 10 participants sur 14.

– **Bibliothèque** : Il s'agit de l'observation de la bibliothèque de Flash. Les participants, dans l'ensemble, la consultent ou la regardent durant des temps similaires. Ce qui est conforme aux attentes, dans la mesure où ils ont tous les même objets à extraire de celle-ci pour créer

l'animation. Certains cependant sont plus rapides (S9) que d'autres (S7). Pour trouver les objets graphiques pertinents et vérifier qu'ils correspondent bien aux contraintes imposées par les scénarios, ils mettent en œuvre des fixations d'une durée comparable à celles observées pour l'accès aux menus, 390ms en moyenne.

- **Scénario** : Le scénario (ou barre du scénario) est l'ensemble d'interacteurs qui permet de créer une animation. Son utilisation requiert un maximum de réflexion, ce que confirment les durées moyennes des fixations dans cette zone : 450ms en moyenne. On observe une légère différence entre les groupes AE et SE : le premier consacre davantage d'attention visuelle au scénario que le second ; cette différence est peut être à rapprocher du fait que ce groupe a consulté davantage de messages oraux que l'autre groupe.

- **Flash** : La scène où l'on pose les objets ainsi que la barre des outils constitue la zone nommée Flash. On y place, redimensionne, bouge les objets. Cela demande une moins grande concentration que la création de l'animation via la barre du scénario, d'où une durée moyenne des fixations plus faible, 390ms, que l'on peut regrouper avec celles des menus et de la bibliothèque. A la différence de la zone du scénario, le groupe AE passe un peu moins de temps dans cette zone que le groupe SE. Les participants du groupe AE accordent donc davantage (respectivement, moins) d'attention à la zone Scénario (respectivement, Flash) que les participants du groupe SE.

- **Lexique et Message "Comment"** : Nous avons fait la distinction entre l'observation du lexique (rubrique textuelle) et l'observation d'un message "Graphique+Oral". Pour l'ensemble des participants, les fixations ont une durée similaire, 326ms pour le lexique et 315ms pour les messages "Graphique+Oral" en moyenne. Ces durées sont courtes, car la recherche dans le lexique en raison de sa fréquence s'apparente davantage à une recherche de mots-clés qu'à une véritable activité de lecture. Quant à la consultation des messages multimodaux, elle relève de l'analyse d'images pour les graphiques (même lorsqu'ils contiennent des étiquettes verbales) et de la recherche visuelle (i.e., la recherche d'un bouton doté d'une couleur particulière). A noter que, très souvent, une fois le bouton de lecture d'un énoncé activé, les participants détournent leur attention de la fenêtre d'aide et reprennent immédiatement leur interaction avec Flash en parallèle avec l'écoute du message comme l'indique le commentaire suivant : *« O+G est très intéressant car l'on peut écouter et travailler en même temps. » S5 (debriefing)* . Cependant, si Eloona est présente, certains participants ont tendance à la regarder pendant toute la durée de l'énonciation des messages sonores et à ne reprendre l'interaction avec Flash que lorsque l'énoncé est terminé. L'étude détaillée de l'attention visuelle accordée à Eloona par les participants du groupe GI est présentée plus loin. Enfin, on peut remarquer le pourcentage important d'attention visuelle que les participants S10, S8, S11 et S7 ont accordé aux messages d'aide.

De manière plus générale, si l'on regroupe les activités de l'ensemble des participants selon les trois rubriques thématiques principales, Flash, Système d'aide et Windows, comme dans l'aperçu

global, on obtient le tableau 5.16 qui présente la répartition, en pourcentages, de l'attention visuelle accordée par les participants à chaque thème, associée à une estimation temporelle.

Thèmes	**Moyenne** sur 10 participants	**Estimation temporelle** sur 38 minutes d'activité (Sc1 et Sc2)
Flash	76,26%	28,9 min
Système d'aide	16,84%	6,4 min
Windows	6,88%	2,6 min

TAB. 5.16 – Activité visuelle de 10 participants du groupe GI, répartie suivant trois thèmes d'activité principaux, Flash, Aide et Windows, en pourcentages (calculés par rapport à l'activité globale) et en temps (estimation en minutes).

En moyenne, pour une session de 38 minutes environ correspondant à la première condition, les données oculométriques (Fixations + Saccades) indiquent que les participants du groupe GI ont manipulé Flash durant 29 minutes, consulté l'aide pendant 6 minutes et observé le résultat à obtenir durant 2 minutes. Elles montrent que l'aide est loin d'être délaissée avec un rapport d'environ un quart entre le temps moyen d'observation de Flash et celui d'observation de l'aide. En remarque, il faut savoir que les participants durant leur observation portent leur attention (i.e., durée des fixations) pendant 87% (moyenne sur les participants) de la durée d'observation, et que l'oeil se déplace (i.e., durée des saccades) pendant 13% du temps restant. Il s'agit de valeurs moyennes calculées sur l'activité oculaire des 10 participants du groupe GI étudiés.

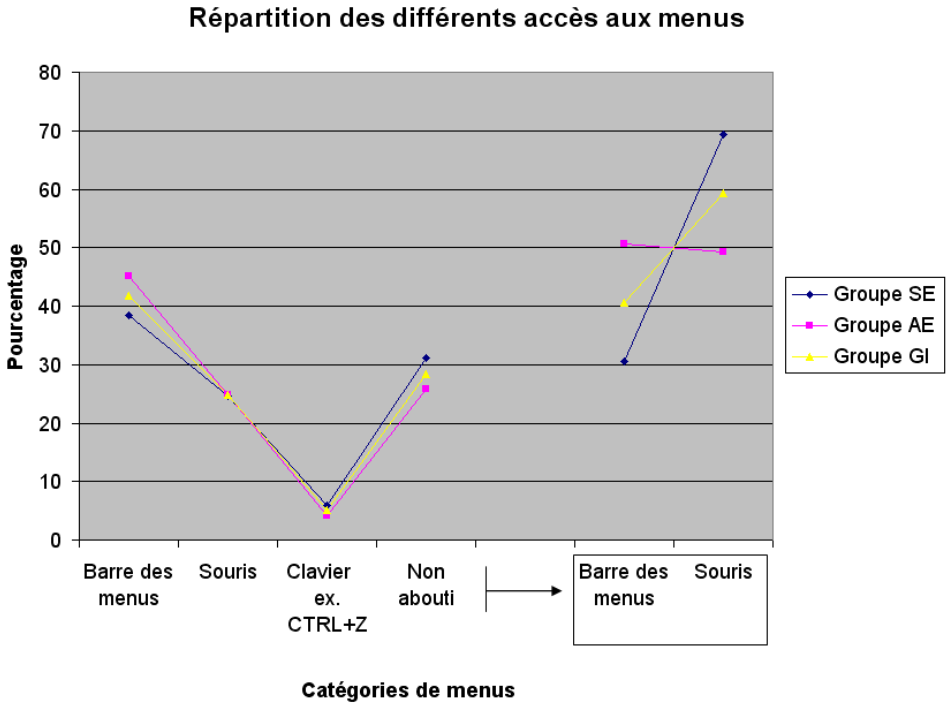


FIG. 5.3 – Accès aux menus. Groupe GI.

En complément, il est fourni le détail des accès aux menus sous forme de graphique (voir figure 5.3). Les différentes catégories de menus accédés ont été regroupées, elles correspondent à celles que nous avons citées précédemment. Le taux d'accès aux menus non aboutis a été ajouté. En effet, il est fréquent que les étudiants parcourent uniquement les menus sans sélectionner d'item, ce que nous nommons accès 'Non abouti'. Nous faisons également la distinction entre l'activation non aboutie des menus par la souris (c'est-à-dire l'activation des menus contextuels), appelée "Souris" et leur activation par la barre des menus ("Barre des menus"). Le graphique de la figure 5.3 présente ces données pour l'ensemble de GI, en distinguant les sous-groupes AE et SE.
Le premier constat qui ressort de l'analyse de ce graphique est que, pour accéder aux fonctionnalités du logiciel, les participants utilisent généralement la barre des menus plutôt que la souris, avec toutefois un pourcentage d'accès plus élevé pour le sous-groupe AE que pour le sous-groupe SE. Des questions viennent alors à l'esprit : les participants du sous-groupe AE ont-ils plus ou mieux écouté les messages oraux procéduraux que ceux du sous-groupe SE, ou bien ont-ils découvert plus de fonctionnalités de Flash ? Cette différence est à examiner avec prudence car sont totalisés dans les catégories "Barre des menus" et "Souris" les accès aux menus de type 'Édition-Annuler' ou 'Annuler-Interpolation' par exemple. Les accès à la barre des menus ne traduisent donc pas forcément le comportement des participants concernant la découverte des fonctionnalités de Flash ; ils peuvent simplement exprimer le fait que l'utilisation de la barre des menus est plus importante que celle des menus contextuels en présence d'Eloona. En revanche, un résultat qui ne peut avoir qu'une seule signification concerne les accès 'Non aboutis' qui traduisent l'exploration sans sélection d'item des menus par la souris ou par la barre des menus. Il est vrai que la différence est plutôt faible mais le groupe SE a tout de même tendance à explorer davantage le logiciel Flash. Ce résultat suscite deux pistes d'interprétation : soit les participants du sous-groupe SE sont plus souvent perdus que ceux du sous-groupe AE parce qu'ils n'utilisent pas l'aide (ils ne savent pas quel menu utiliser, donc ils explorent), soit ils sont plus performants et explorent davantage les fonctionnalités de Flash pour se familiariser avec l'utilisation de ce logiciel. Pour choisir entre ces deux interprétations, il est nécessaire de comparer les performances des deux sous-groupes telles qu'elles apparaissent à travers leurs erreurs lors de la réalisation des scénarios, la qualité des animations créées et les résultats obtenus au post-test. Dans une section ultérieure, nous analysons ces aspects en soumettant ces données à une analyse en composantes principales, afin d'identifier les stratégies d'apprentissage adoptées par les participants du groupe GI.

Pour terminer l'analyse de l'activité des participants du groupe GI, le déroulement de la session de création de la première animation est décrit brièvement à partir des accès aux différentes fonctionnalités de Flash et au système d'aide. Cette étude exploratoire a pour but d'analyser le comportement général des participants. L'idée, très simple, est d'observer leur premier accès via les menus aux fonctionnalités de Flash, à chaque étape de la réalisation des deux premières animations ; ce qui renseigne sur le déroulement de leur activité. Les observations sont présentées dans le tableau 5.17 ; les temps (en minutes) comptés à partir du début de la première condition indiquent, pour chaque type d'action, quand elle a été réalisée pour la première fois par chaque participant. On constate que les temps moyens sont très proches d'une condition à

l'autre. Ainsi, avec ce tableau, on peut retracer la construction de la première animation. Les participants commencent par déposer les objets du décor sur la scène, puis, ils cherchent à les redimensionner (à t=4,4 min Gr. GI) et à les renverser (seuls 8 participants sur les 14 ont effectué le "renversement" par l'intermédiaire du menu). Le décor créé, ils entament la phase d'animation en créant un calque (t=8,9 min Gr. GI), posent les poissons et tentent de générer l'interpolation de mouvement (t=18,9min Gr. GI).
Les petites différences entre les sous-groupes AE et SE qui semblent intéressantes se situent au niveau de la première consultation du "Comment" (i.e., l'affichage du lexique) ; le groupe 'avec Eloona' a tendance à favoriser la consultation de l'aide (t=5,2min contre t=7,5min), peut être par curiosité de voir parler l'ACA ? En effet, le sous-groupe AE écoutera un message d'aide oral avant l'autre groupe (9,0min contre 12,2min). Cela dit, ces résultats ne sont certainement pas significatifs en raison de la diversité inter-individuelle que l'on observe ; en effet, certains participants ont très peu utilisé l'aide (e.g., S9) et ont progressé de manière différente.

Temps en min.	Redimensionner un objet	Créer une interpolation	Créer un calque	Renverser un objet	1ère requête "Comment"	1ère écoute message "Comment"
Sujet 10	5,6		10,2	15,2	10,0	10,1
Sujet 2	7,5	22,1	2,9		13,8	15,3
Sujet 20	1,5	10,3	5,1		11,1	11,2
Sujet 22	2,4	25,7	3,6	6,8	1,9	2,0
Sujet 4	2,9	25,0	1,8	13,1	1,7	31,8
Sujet 6	2,7	13,3	11,7		9,5	10,8
Sujet 8	4,7	13,7	26,1	9,3	4,3	4,4
Sujet 1	7,5	12,3	12,8	35,3	10,1	10,3
Sujet 11	5,8	20,1	0,8	9,8	0,4	0,8
Sujet 16	3,7	25,5	19,2		2,3	3,5
Sujet 3	2,0	22,2	11,4		11,1	11,2
Sujet 5	7,5	29,5	6,7	4,8	1,0	2,2
Sujet 7	4,5	18,4	11,9	23,1	3,2	3,4
Sujet 9	3,0	8,0	0,6		8,5	31,9
Groupe SE	3,9	18,3	8,8	11,1	7,5	12,2
Groupe AE	4,8	19,4	9,0	18,2	5,2	9,0
Groupe GI	4,4	18,9	8,9	14,7	6,4	10,6

TAB. 5.17 – Déroulement de la création des deux premières animations (scénarios Sc1 et Sc2) - Groupe GI (14 participants).

Aide multimodale - Annotation des interactions/Questionnaires

Doter un système d'assistance de moyens d'expression oraux représente une évolution importante pour les utilisateurs des systèmes interactifs actuels encore exclusivement graphiques

et textuels. En outre, incarner cette voix dans une représentation animée du système constitue une innovation encore plus importante. Les réactions des participants à cette modalité d'expression du système, nouvelle pour certains participants et inhabituelle pour les autres, pouvaient influencer indirectement leur évaluation subjective de l'apport d'un ACA à l'interaction d'aide. Mais surtout, les choix que nous avons effectués concernant la conception des messages oraux (notamment leur longueur), l'intelligibilité de l'articulation, le timbre de la voix, le rythme d'élocution et l'expressivité de la prosodie sont susceptibles d'influer directement sur la manière dont l'ACA est perçu et donc sur les jugements portés par les participants sur Eloona. En particulier, nous avons fait l'hypothèse que la prosodie des énoncés devait être en accord avec les expressions faciales et le comportement de l'ACA pour que celui-ci soit jugé réaliste.
Une double évaluation de la qualité/pertinence des messages oraux, d'une part, en tant que nouveau mode d'expression du système et, d'autre part, en tant que composante majeure de la « personnalité » de l'ACA était donc nécessaire pour tester la validité de notre hypothèse. Trois questions dans chacun des deux questionnaires finals, verbal et non verbal, sont consacrées à cette évaluation. Nous les reproduisons ci-dessous. A noter que, les données oculométriques n'étant pas nécessaires pour effectuer cette évaluation, nous avons étudié les réactions de l'ensemble des participants (i.e., 22).

- Questions verbales :
  - « La présentation orale plutôt que textuelle des informations d'aide améliore-t-elle l'efficacité de l'assistance ? » (Pas du tout / Énormément)
  - « La présentation orale des informations d'aide améliore-t-elle la facilité d'utilisation de l'interface et son agrément ? » (Pas du tout / Énormément)
  - « Les messages d'aide vous ont semblé : » (Trop courts / Trop longs / De longueur appropriée)
- Questions non verbales (utilisation de SAM) : « Comment avez-vous ressenti, quelles ont été vos réactions, par rapport aux innovations suivantes :
  - « L'utilisation de la parole au lieu du texte dans les messages d'aide. »
  - « La qualité des énoncés de synthèse vocale. »
  - « Concernant l'agent conversationnel mis en œuvre, Eloona, son comportement global (apparence + voix). »

Les deux premières questions verbales tentent de recueillir des informations sur les réactions et l'attitude des participants vis-à-vis de l'utilisation de la parole en tant que modalité d'expression des messages système. La dernière question non verbale, ainsi que la seconde, porte sur l'évaluation de nos choix de conception des messages oraux. La première question non verbale vise à obtenir des informations sur les réactions des participants à l'utilisation de la parole en tant que modalité d'expression du système, et la dernière, entre autres, à évaluer la pertinence de nos choix de conception concernant la cohérence entre « la voix » d'Eloona et ses « expressions », ainsi que le réalisme et l'attrait de la « personnalité » que nous lui avons construite.

Nous détaillons dans la suite de cette section les indications fournies par les réponses aux questions ci-dessus, d'abord sur l'utilisation de la modalité orale pour la formulation des mes-

sages du système d'aide, puis sur nos choix de conception concernant les messages oraux et leur cohérence avec l'apparence, les expressions faciales et le comportement d'Eloona.

`Messages système oraux : réactions et attitudes des participants`

17 participants estiment que la présentation orale des informations d'aide améliore l'efficacité de l'assistance (voir la première question verbale reproduite au début de cette section) ; 5 d'entre eux jugent cette amélioration importante, tandis que 12 l'estiment mineure (5ème et 6ème différenciateurs sémantiques). Un seul des 22 participants juge son influence plutôt négative (S12).

Les jugements des participants sur la facilité d'utilisation et l'agrément de cette nouvelle modalité d'expression du système sont également positifs (voir la seconde question verbale reproduite au début de cette section). En effet, 3 participants ont un avis neutre et le sujet S12 a, comme précédemment, un avis négatif. En revanche, les 18 autres ont des avis positifs, et les commentaires dont ils accompagnent leurs réponses précisent les raisons de leurs jugements. Certains nuancent leurs avis, parfois en suggérant des évolutions susceptibles de mieux répondre à leurs attentes.

« C'est toujours mieux à mon avis d'avoir des explications orales, cela nous évite de lire tous les textes. » S6

« Cela permet de faire en temps réel. » S22

« Ceci permet de visualiser les menus et les opérations à effectuer pendant l'explication orale. » S11

« L'oral permet d'être plus attentif, (c'est) intéressant pour ceux qui n'aiment pas lire. Il faudrait peut être un bouton pour répéter. » S14

« Quand Eloona parle il faut aller à son rythme et reprendre son commentaire si on a oublié de faire quelque chose. » S19

« Cela (i.e., l'oral) rend les informations plus claires et plus précises. » S21

Les réponses au questionnaire non verbal (première question non verbale reproduite au début de cette section) montrent que les réactions et attitudes de la plupart des participants sont en accord avec leurs jugements verbaux. La parole en tant que modalité d'expression du système « plaît » à 19 participants sur 22 (première ligne de SAM). Son usage « déplaît » légèrement à 2 participants seulement, S3 du groupe 'Sans' et S7 du groupe 'Les deux' ; en revanche, l'avis de S12, qui préfère les aides classiques donc textuelles aux aides innovatrices proposées, est simplement neutre, ce qui est plutôt surprenant. Il est difficile d'expliquer les jugements de S3 et S7, dans la mesure où ces deux participants n'ont donné de réponse négative à aucune des deux questions verbales correspondantes : les réponses de S3 à ces deux questions sont neutres et celles de S7 sont, neutre pour la première et positive pour la seconde. Cependant, le commentaire de S7 reproduit ci-dessus suggère que ce participant préférerait une présentation des messages d'aide qui associe au graphique texte et parole (en alternance ou simultanément), ce que les questions verbales ne lui permettaient pas d'exprimer. De même, S3 a nuancé, au cours de l'entretien de debriefing, les jugements exprimés dans le questionnaire verbal, en précisant qu'une présentation textuelle des informations d'aide était préférable à une présentation orale qui oblige à ré-écouter le message depuis le début, même lorsqu'il s'agit simplement de retrouver une information ponc-

tuelle dans un message connu.
Les effets de la mise en œuvre de la parole (en tant que modalité d'expression du système) sur l'état émotionnel des participants, tel qu'ils peuvent le ressentir, varient d'un participant à l'autre. Les réponses non verbales concernant ce point couvrent la totalité de la plage d'expression offerte, de « Excité/Stimulé » à « Indifférent/Calme » (voir la seconde ligne de SAM). Cependant on peut observer un pic de 11 réponses de type 'Stimulé' (2ème mannequin à partir de la gauche) auxquelles il convient d'ajouter 2 réponses de type 'Très stimulé'. La majorité des 22 participants semble donc apprécier l'utilisation de la parole au lieu du texte en tant que modalité de formulation des informations d'aide.
Quant à la troisième ligne des réponses à la première question de SAM reproduite au début de la section, elle indique que les messages oraux sont ressentis par 13 participants comme susceptibles d'accroître leur assurance, tandis qu'ils sont perçus par 7 autres comme sans effet sur leur confiance en eux.

Globalement, les jugements verbaux sur l'efficacité et la convivialité de la parole en tant que modalité de présentation des messages d'aide montrent qu'une large majorité des participants (17 sur 22) juge les présentations orales plus efficaces que les présentations textuelles, et que tous sauf un les estiment plus conviviales. Le nombre des jugements positifs exprimés en réponse à la question non verbale correspondante se situe entre les deux puisqu'il est de 19 (pour la première ligne de SAM). Dans la mesure où cette question englobe efficacité et convivialité, on peut considérer qu'il y a convergence entre les jugements verbaux et non verbaux exprimés par les participants, sauf peut-être pour S3 et S7.

Évaluation de la pertinence des choix de conception des messages oraux

Nous disposons, pour évaluer comment nos choix de conception des messages oraux ont été perçus et jugés par les participants, des réponses à deux questions ; la première, verbale, porte sur la longueur des messages oraux, la seconde, non verbale, sur la qualité des énoncés d'aide. La troisième question non verbale permet de recueillir quelques indications sur la cohérence entre « la voix » d'Eloona et ses « expressions », son comportement. Ces trois questions sont reproduites au début de la section 4.4.2.
A noter qu'un premier indicateur, fourni grâce à l'annotation, est la consultation de la base des messages d'aide. Si l'on regarde la base du "Comment", fonctionnalité la plus utilisée, 74% des messages "Comment" ont été consultés par l'un ou l'autre des participants. Ce taux élevé démontre leur adéquation aux besoins des participants, d'autant plus que, pour assurer la crédibilité de la situation d'interaction proposée, la base des messages d'aide couvrait l'ensemble des fonctions de Flash accessibles par la barre des menus, les barres d'outils affichées et les widgets présents dans la fenêtre de Flash. Par conséquent, elle incluait un nombre important de messages inutiles pour réaliser les animations, ce qui explique en partie les 26% de messages non consultés par les participants.

Concernant la longueur des messages, il convient d'accorder davantage d'attention aux avis des participants sur le premier système d'aide qu'ils ont utilisé puisque, dans la seconde condition, ils ont rarement consulté l'aide. Les messages d'aide se sont avérés, pour l'ensemble des

participants (voir le tableau 5.18), de longueur et de précision appropriées. Par conséquent le découpage des informations d'aide que nous avons effectué, semble être pertinent. Reste cependant à comprendre pourquoi 5 participants ont trouvé les messages trop vagues. Une étude spécifique serait nécessaire pour expliquer ce jugement vraisemblablement dû à l'usage de la parole comme modalité d'expression des messages système qui était, pour les participants, une situation d'interaction nouvelle ou inhabituelle ?

Les messages d'aide vous ont semblé (22 réponses) :					
Trop courts	2	Trop longs	1	De longueur appropriée	19
Trop vagues	5	Trop détaillés	1	De précision appropriée	16

TAB. 5.18 – Jugements détaillés des participants sur la longueur et le contenu informationnel des messages d'aide.

L'analyse des réponses à la question non verbale portant sur les messages d'aide montre que la qualité de la synthèse vocale suscite une satisfaction unanime de la part des participants (première ligne de SAM). Concernant la seconde ligne de réponse à cette question, on observe une diversité des jugements comparable à celle observée pour la question portant sur l'appréciation de la présentation orale des informations d'aide ; à noter que la majorité des participants ont exprimé un jugement neutre. Quant à la troisième ligne, elle contient des réponses très voisines de celles données par les participants pour la question portant sur l'évaluation de la parole en tant que modalité de présentation des informations d'aide : 12 participants ont perçu la voix d'Eloona comme susceptible d'accroître leur assurance, 9 comme sans effet sur leur confiance en eux. Un seul participant (S5) estime que cette voix peut réduire sa confiance en lui.

La présence d'Eloona pour 'dire' les messages d'aide est ressentie de la même manière que l'utilisation de la parole ou lieu du texte : 19 participants la perçoivent de façon positive ou très positive (première ligne de SAM). Une différence, cependant, par rapport aux jugements sur l'usage de messages d'aide vocaux, sa présence ne 'déplaît' à personne. Concernant la seconde ligne des réponses non verbales, 13 participants (le même nombre que pour l'utilisation de la parole) ont été stimulés par la présence d'Eloona. Quant à la troisième ligne, il apparaît que, pour 14 participants, Eloona accroît davantage la confiance en soi et l'assurance ; 7 la jugent sans effet sur leur assurance et leur confiance en eux. Ce résultat est important, car il confirme notre hypothèse que la conjonction d'une représentation graphique animée du système d'assistance et d'une présentation orale des informations d'aide est susceptible d'accroître les effets bénéfiques de chacune de ces deux innovations sur l'interaction, notamment en agissant sur la satisfaction des utilisateurs et, dans le contexte de la situation choisie, sur leur confiance en eux.

Les résultats de l'analyse des questionnaires verbal et non verbal confirment donc la pertinence de nos choix de conception concernant les messages oraux (taille et intelligibilité). Ils montrent aussi indirectement l'effet bénéfique sur les jugements subjectifs des participants, de la cohérence que nous avons cherché à obtenir, lors de la conception de l'animation de l'ACA, entre ses expressions faciales, son comportement et l'expressivité des messages d'aide qu'elle a la

charge d'énoncer.

Eloona - Annotation des interactions/Questionnaires

L'analyse des jugements et réactions exprimés par les participants dans les questionnaires verbal et non verbal vise un double objectif :

- Déterminer si la « personnalité » que nous avons donnée à Eloona, les intentions qui ont présidé à sa conception ont été perçues par les participants conformément à nos attentes et en accord avec nos propres perceptions.
- Ensuite, analyser les jugements verbaux et non verbaux que ceux-ci ont portés sur Eloona. En effet, leurs réactions affectives à cet ACA particulier, attraction, indifférence ou irritation, sont susceptibles d'influencer considérablement leur jugement sur l'apport de celui-ci à l'efficacité et à la convivialité de l'interaction d'aide. La connaissance de ces réactions est donc indispensable pour être en mesure de définir la portée exacte des conclusions de l'analyse des jugements globaux exprimés par les participants sur l'apport d'une tête parlante à l'interaction avec un système d'assistance. Elle est nécessaire, en particulier, pour déterminer dans quelle mesure nos choix de conception concernant l'élocution, les expressions et le comportement d'Eloona sont pertinents, dans quelle mesure ils permettent de définir des ACAs parlants dont la présence contribue effectivement à améliorer la perception que les utilisateurs ont de l'efficacité et de la convivialité des systèmes d'assistance, et qui soient capables de séduire une vaste population d'utilisateurs qui les acceptent durablement.

Les données analysées comprennent les réponses aux questions suivantes des questionnaires verbal et non verbal :

- Questions verbales
  Elles se présentent sous la forme de différenciateurs sémantiques et visent à déterminer comment les participants ont perçu la « personnalité » d'Eloona :
  - « Eloona vous a semblé : »
  - « Son comportement vous a semblé : »

- Questions non verbales
  « Comment avez-vous ressenti, quelles ont été vos réactions par rapport aux innovations suivantes :
  - L'aspect physique d'Eloona.
  - Ses expressions faciales. »

La trace des fixations oculaires des participants (groupe GI) a également fait l'objet d'une analyse, car la direction du regard est un indice important de l'intérêt courant de l'utilisateur et de l'évolution de cet intérêt au cours de l'interaction avec une interface graphique.

Nous présentons successivement les données du questionnaire verbal, puis celles du questionnaire non verbal et ensuite les données oculométriques.

`Questions verbales`

Les réponses au questionnaire verbal sont présentées dans le tableau 5.19, regroupées en fonction de l'expérience informatique des participants (groupe GI et GNI) et des dimensions du comportement et de la 'personnalité' d'Eloona abordées dans le questionnaire verbal.
On observe que les jugements effectués par les participants sur la personnalité d'Eloona sont en accord avec les intentions qui nous ont guidés dans sa conception. En effet :

- son apparence est attractive (16 avis positifs), et son expression souriante (17 avis positifs) ; bref, son comportement est affable (19 avis positifs) ;
- dans son rôle d'assistance, elle est aussi sérieuse (9 avis)qu'enjouée (10 avis) ;
- elle est chaleureuse (19 avis) et coopérative dans son rôle d'assistance (15 avis).

Ainsi, Eloona apparaît pour la majorité des participants comme un personnage distrayant (20 avis), doté d'une personnalité attirante et agréable et faisant preuve d'une attitude très sympathique (21 avis). Il semble donc que notre objectif, qui était de créer un ACA attractif dans le rôle d'assistant à la prise en main d'un logiciel, ait été atteint. La personnalité spécifique d'Eloona ne peut donc avoir influencé négativement les jugements des participants sur la présence d'un ACA.

Questions (22 participants)	**Nombre de participants par type de réponse GI/GNI**						
Apparence physique et expressions faciales d'Eloona							
Terne-Fade / Avenante-Attractive	0/0	0/0	1/0	3/0	7/1	2/2	1/3
Maussade / Souriante	0/0	0/0	1/1	2/1	4/2	4/1	3/3
Comportement 'social' d'Eloona							
Distant-Réservé / Affable-Aimable	0/0	0/0	2/0	0/0	8/3	3/3	1/1
Comportement 'Professionnel' d'Eloona système d'aide							
Sérieux / Enjoué	0/0	0/1	7/1	3/1	3/5	0/1	1/0
Froide / Chaleureuse	0/0	0/0	0/1	2/1	9/1	2/3	1/3
Passive / Coopérative	0/0	1/0	2/2	2/2	3/0	3/3	3/3
Effet global de l'apparence et du comportement							
Antipathique / Sympathique	0/0	0/0	0/0	1/0	7/1	5/2	1/5
Ennuyeuse / Distrayante	0/0	0/0	1/0	3/0	5/2	3/2	2/3
Réalisme du comportement d'Eloona							
Artificiel / Réaliste	0/0	1/0	5/0	2/0	3/4	2/3	1/1
Monotone / Varié	0/0	3/1	1/3	4/3	1/0	3/0	2/1

TAB. 5.19 – Jugements verbaux des participants, regroupés en fonction de leur expérience informatique (GI/GNI), sur la 'personnalité' d'Eloona (différenciateurs sémantiques).

Concernant le 'réalisme d'Eloona', son comportement a tendance à être perçu comme un peu monotone. Pour éviter cette critique, il suffirait, vraisemblablement, d'accentuer ses mouvements de tête et ses clignements d'yeux lorsqu'elle est inactive, ce qui la rendrait moins monotone, puisque ces mouvements sont aléatoires. Par ailleurs, le groupe GNI dans sa majorité juge l'ani-

mation d'Eloona plutôt réaliste, alors que les informaticiens sont partagés sur ce point : 6 jugent Eloona un peu artificielle tandis que 6 autres la trouvent réaliste. Étant donné les progrès que réalise actuellement FT R & D concernant les expressions faciales d'Eloona, ce défaut pourrait être corrigé rapidement, au moins en partie. Toutefois, dans la mesure où les jugements négatifs sont minoritaires (sur l'ensemble des participants) et que les critiques formulées sont légères, on peut faire l'hypothèse que les limites du réalisme d'Eloona ont exercé une influence négligeable sur les préférences des participants pour l'un ou l'autre des deux systèmes d'aide qu'ils étaient censés évaluer.

Questions non verbales

Les réponses à la question non verbale sur 'l'aspect physique' d'Eloona, confirment les réponses aux questions verbales recensées dans le tableau précédent ; quant aux réactions non verbales suscitées par ses 'expressions faciales', elles complètent et raffinent les réponses à ces mêmes questions verbales.
Plus de la moitié des participants sont satisfaits à la fois de l'aspect physique d'Eloona (17 avis) et de ses expressions faciales (12 avis). Seuls 3 participants sont légèrement mécontents de ses expressions faciales, ce sont exclusivement des informaticiens. Au cours des debriefings, ces 3 étudiants en informatique ont fait la remarque que les mouvements des lèvres n'étaient pas toujours synchronisées avec l'énonciation pour justifier leurs jugements. Le reste des participants a réagi de façon neutre. Ces réactions présentent une certaine ressemblance avec les jugements exprimés sur l'attractivité du clone.

La seconde ligne de SAM relative aux mêmes questions montre une répartition équitable des réactions vis-à-vis de l'aspect physique d'Eloona : 8 participants sont stimulés par l'ACA alors que 9 sont indifférents à sa présence. Concernant les expressions faciales, les participants ont été encore moins sensibles à cet aspect du comportement de l'ACA : 12 participants ont exprimé leur indifférence à cet égard.

La troisième ligne, qui porte sur la confiance en soi et l'assurance, fait ressortir que l'aspect physique accroît légèrement l'assurance de 13 participants alors que les expressions faciales laissent la majorité sans avis.

Fixations oculaires sur Eloona

Les positions sur l'écran des échantillons de données oculométriques (oculomètre ASL, modèle 501), calculées à une fréquence de 60 Hz, ont été analysées afin de déterminer dans quelle mesure les participants ont observé Eloona. Comme pour l'activité des participants, nous avons analysé uniquement les mouvements oculaires d'une partie du groupe GI, soit 10 participants pour lesquels les données oculométriques enregistrées étaient de bonne qualité ; en outre, l'analyse a porté, comme précédemment et pour les mêmes raisons, sur la réalisation des deux premiers scénarios, Sc1 et Sc2 (i.e., sur la première condition). Avec un ordre de passation contrebalancé, le nombre d'étudiants ayant bénéficié de la présence d'Eloona pendant cette première période d'interaction se ramène à 5. Ce nombre peut paraître faible a priori ; il est toutefois suffisant pour permettre d'obtenir des résultats qualitatifs intéressants, en raison de l'originalité des données et

du caractère pionnier de l'étude dont la portée sera élargie ultérieurement, lorsque l'annotation de la trace des interactions des participants du groupe GNI (i.e., le groupe des étudiants 'non informaticiens') sera terminée ; 4 d'entre eux ont bénéficié de la présence d'Eloona pendant la première condition (réalisation des scénarios 1 et 2). A noter que ces 4 participants n'auront certainement pas le même regard, au sens propre et figuré, sur Eloona que les 'informaticiens' : l'effet de nouveauté, plus important, aura certainement un effet sur l'affectation de leur attention visuelle.

Participants	**Sujet 1**	**Sujet 11**	**Sujet 16**	**Sujet 7**	**Sujet 9**	**Moy.**
Eloona 'Parle'						
Nombre Fix.	40	57	58	61	16	46,4
Temps Fix. (en sec)	18,6s	18,8s	28,2s	29,4s	6s	20,2s
Durée Fix. Moy. (en ms)	467,2ms	330,3ms	486,8ms	483,1ms	375ms	428,5ms
Eloona 'Inactive'						
Nombre Fix.	46	38	34	7	20	29
Temps Fix. (en sec)	16,8s	12,0s	13,4s	2,5s	4,6s	9,9s
Durée Fix. Moy. (en ms)	365,5ms	317,4ms	395,2ms	370,2ms	230,4ms	335,8ms
Temps (en sec) et nombre	35,5s	30,8s	41,6s	32,0s	10,6s	30,1s
des Fix. sur Eloona	86	95	92	68	36	75.4
Temps total des énoncés oraux activés en sec	324,6s	345,3s	140,3s	371,7s	24,7s	241,3s
Rapport tps énoncés / tps Fix. quand Eloona 'Parle'	5,7%	5,4%	20,1%	7,9%	24,2%	12,6%

TAB. 5.20 – Nombre et durée moyenne (en millisecondes) des fixations oculaires sur Eloona, et durée (en secondes) d'observation d'Eloona, lorsqu'elle parle et lorsqu'elle est inactive. Groupe GI (5 participants).

Le tableau 5.20 récapitule les résultats de l'analyse des mouvements oculaires de 5 participants du groupe GI en fonction des deux états d'Eloona, à savoir : lorsqu'elle énonce un message d'aide oral (Eloona 'Parle') et lorsqu'elle est inactive et regarde le bord droit de l'écran, en faisant uniquement de légers mouvements de tête et clignements d'yeux aléatoires : Eloona 'Inactive'. 'Nombre Fix.' représente le nombre de fixations en direction d'Eloona pendant qu'elle parle ou lorsqu'elle est inactive ; 'Temps Fix.' est le cumul (en secondes) des durées des fixations centrées sur l'ACA., et 'Durée Fix. Moy.' la durée moyenne en millisecondes de ces fixations.

Ces données objectives montrent d'abord que la présence d'Eloona ne représente, pour les participants, qu'une source négligeable de distraction de l'interaction avec Flash et avec l'aide. En effet, sur 38 minutes de session, qui correspond à 33 minutes de fixations (rappel : le rapport temporel fixations/saccades est de 87%), Eloona est observée, tous états confondus, pendant 30 secondes, soit environ 1,5% de la durée d'observation de l'affichage pendant la première condition (réalisation des scénarios Sc1 et Sc2). Le nombre moyen de fixations que suscite l'ACA est de

75 fixations, avec un nombre moyen plus important lorsqu'Eloona 'parle' (46 fixations contre 29 en lorsqu'elle est 'inactive'). La durée moyenne des fixations varie suivant l'état d'Eloona. En effet, lorsqu'Eloona 'parle' les participants l'observent vraiment , ils regardent ses mouvements de lèvres par exemple par curiosité ou pour mieux comprendre l'énoncé, ce qui entraîne des durées moyennes des fixations plus longues de 100ms par rapport aux simples "coups d'oeil" qu'ils lui jettent lorsqu'elle est 'inactive'.
Dans la mesure où la direction du regard, l'attention visuelle de l'utilisateur, en situation d'interaction avec une interface graphique, est un indicateur fiable de son intérêt courant, on peut supposer que le personnage d'Eloona a suscité un réel intérêt de la part des participants puisqu'ils l'ont tous 'observée' (inactive ou en train de parler) entre une dizaine et une quarantaine de secondes, ce qui représente un peu moins d'une centaine de fixations oculaires. A noter que l'on ne peut pas imputer cet intérêt au seul effet de nouveauté, à la curiosité vis-à-vis de l'ACA, puisque les participants avaient déjà eu l'occasion de découvrir cet ACA et de se familiariser avec lui lors de la présentation de l'expérimentation qui était effectuée par Eloona.

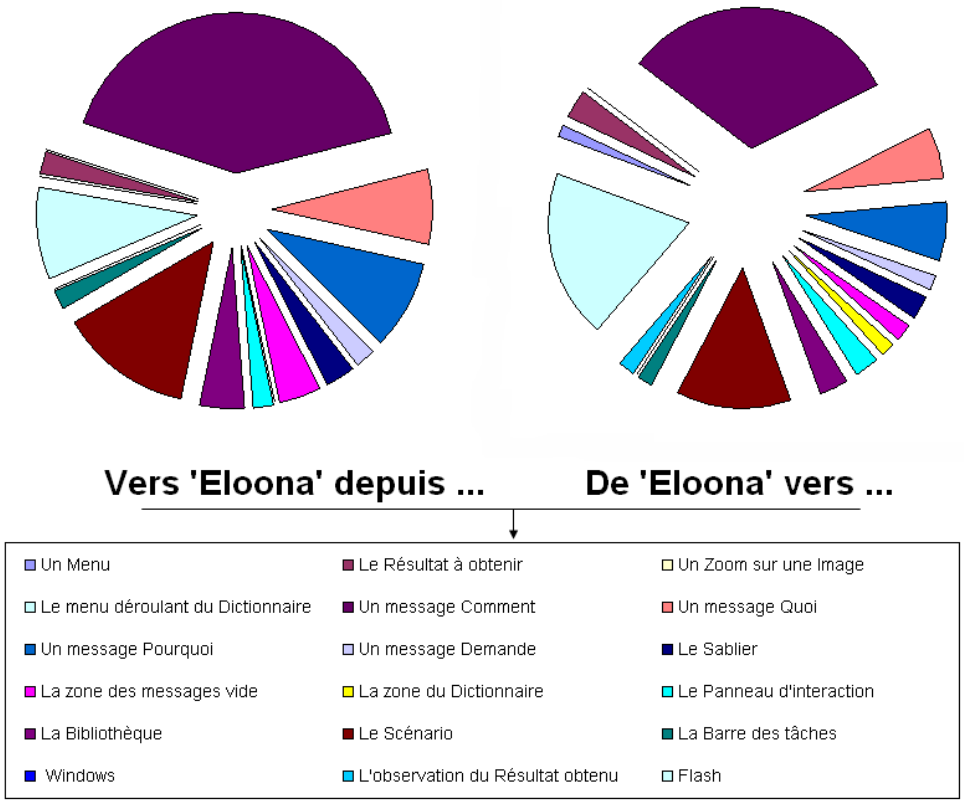


FIG. 5.4 – Origine du regard vers Eloona et direction du regard en partance d'Eloona (i.e., positions des fixations précédant et suivant celles sur le visage d'Eloona). Groupe GI (5 participants).

Ces observations valident nos choix de conception de l'animation de l'ACA pour la situation d'interaction choisie (l'assistance à la prise en main d'un logiciel), en particulier, notre parti pris de donner à la présence d'Eloona, à son animation et à son comportement, un caractère à la fois discret et attractif. Cette double caractéristique de la personnalité d'Eloona explique vraisem-

blablement pourquoi les participants n'ont pas été distraits de l'interaction avec Flash et avec l'aide par la présence de cet ACA, alors même que la plupart d'entre eux ont porté un intérêt plus ou moins vif à cette représentation animée du système.

L'analyse des parcours oculaires nous permet également de connaître le temps qu'il consacre à l'observation d'Eloona pendant qu'elle 'parle' par rapport à l'ensemble des énoncés (en pourcentage). Ainsi, les participants l'ont regardée en moyenne 12% du temps pendant lequel elle a parlé. Ce pourcentage moyen recouvre deux catégories de profil : un groupe qui l'observe entre 5 et 7% de la durée d'énonciation des messages d'aide, et un autre groupe qui l'écoute en la regardant entre 20 et 24% de ce temps.
Cette analyse est complétée par les graphiques de la figure 5.4 qui précisent ce que les participants regardaient avant et après chaque observation d'Eloona. Ces graphiques montrent que les fixations précédant l'observation de l'ACA portaient en majorité sur un message "Comment". En effet, c'est lorsqu'Eloona 'parle' qu'on trouve un maximum de fixations sur l'ACA. Ainsi, les participants lancent l'un des énoncés du message "Comment" affiché, puis vont la regarder parler. Les pourcentages de fixations précédant l'observation d'Eloona qui portent sur Flash, la barre du scénario ou des messages "Quoi" et "Pourquoi" sont voisins. Le graphique de droite montre que, en général, le regard des participants, à la suite de l'observation d'Eloona, retourne soit vers le message "Comment", pour visualiser les graphiques qui complètent le discours, soit vers Flash pour continuer à travailler. Les pourcentages des fixations suivant l'observation d'Eloona dont la barre du scénario, des messages "Quoi" et "Pourquoi" font l'objet sont également similaires.

ACP sur les participants du groupe GI

L'analyse en composantes principales (ACP) est une méthode statistique (initialement de statistique descriptive) qui a pour but de comprendre et de visualiser comment les effets de phénomènes a priori isolés se combinent. Nous l'avons appliquée sur les données recueillies au cours de l'expérimentation pour rendre compte de la diversité observée des comportements et des jugements subjectifs des participants, et tenter d'identifier des profils utilisateur caractéristiques. L'analyse a porté sur le groupe GI uniquement, en raison de la relative homogénéité des comportements des participants de ce groupe par rapport à ceux du groupe GNI et du nombre plus important de participants qu'il comprenait (14 versus 8). En outre, les effets des différences d'expérience informatique, dont cette étude a montré l'influence prépondérante sur les stratégies de découverte d'un nouveau logiciel et les réactions à la présence d'un ACA, auraient pu masquer ceux d'autres facteurs. Nous avons donc recherché les directions de l'espace qui représentaient le mieux les corrélations entre plusieurs variables aléatoires. Nous avons retenu des variables communes à l'ensemble des participants du groupe GI qui tiennent compte à la fois de leurs performances, de leur utilisation de l'aide et d'un facteur représentatif de leur stratégie d'approche d'un nouveau logiciel.
Ainsi, pour tenir compte des performances, nous avons noté, pour chaque participant, les résultats de la réalisation des deux premiers scénarios (Sc1 et Sc2) qui constituaient la première condition, celle pendant laquelle l'essentiel de l'apprentissage de l'utilisation de Flash a eu lieu, ainsi que le post-test effectué par chacun d'eux à l'issue de l'expérimentation (les barèmes d'évaluation sont

notés sur 25 points). Pour rendre compte de l'utilisation du système d'aide, nous avons retenu le nombre de messages oraux consultés. Enfin, afin de mettre en avant les différentes stratégies adoptées habituellement par les participants lors de prise en main d'un nouveau logiciel, nous avons décidé de constituer une variable qui utilise les auto-corrections effectuées par les participants, à savoir, par exemple, 'Édition-Annuler' ou tous les menus qui permettent de défaire une action comme 'Annuler-Interpolation' qui permet à l'utilisateur qui a créé une interpolation dont le résultat n'est pas celui escompté de la défaire. Le tableau 5.21 suivant regroupe les quatre variables que nous avons choisies.

Sujets	**Notes Scénarios 1 et 2**	**Notes Post-Test**	**Nombre d'auto-corrections**	**Nombre de messages**
S10	18	18	32	32
S2	15	11	16	12
S20	5	21	24	17
S22	18	20	36	5
S4	12	24	33	11
S6	12	21	6	11
S8	10	22	17	29
S1	11	18	44	31
S11	16	22	2	31
S16	16	23	12	9
S3	12	15	16	8
S5	8	17	35	38
S7	13	25	29	32
S9	12	16	16	2

TAB. 5.21 – Valeurs des 4 variables prises en compte dans l'ACP : résultats des scénarios 1 et 2 (première condition), résultats du post-test, auto-corrections, nombre de messages d'aide consultés. Groupe GI (14 participants).

En utilisant Scilab (logiciel libre développé par l'INRIA) pour calculer l'ACP sur les données du tableau 5.21, nous avons obtenu le tableau 5.22 qui permet de connaître pour chaque axe sa valeur propre ainsi que son inertie. Le cercle des corrélations permet d'identifier la (les) variable(s) associée(s) à chaque axe.

Axes	**Valeurs propres**	**Inerties**	**Inertie cumulée**	**Variable associée à l'axe**
Axe 1	1,42	0,35	0,35	Nombre de messages consultés
Axe 2	1,03	0,25	0,61	Notes Post-test
Axe 3	0,89	0,22	0,83	Notes Scénarios 1 et 2
Axe 4	0,64	0,16	1,00	—

TAB. 5.22 – Résultats de l'ACP calculée sur les données du tableau 5.21 : Valeurs propres, Inertie et Inertie cumulée.

En complément, le tableau 5.23 présente les contributions, en pourcentages, des variables à chaque axe.

Pour commenter les résultats de l'ACP, nous nous appuyons sur la figure 5.5 qui représente la répartition du groupe GI suivant l'axe 1 porté à 62% par la variable 'Nombre de messages consultés', et l'axe 2 porté à 78% par la variable 'Notes Post-test'. Sachant que l'inertie cumulée n'atteint pas 80%, il faudrait tenir compte du plan constitué par les axes 2 et 3 ; cependant, on obtient la même répartition que sur la figure 5.5 ; il est donc inutile de représenter ici cette répartition.

	Axe1	**Axe2**	**Axe3**	**Axe4**
Notes Scénarios 1 et 2	26%	6%	**67%**	1%
Notes Post-test	13%	**78%**	0%	9%
Nombre d'auto-corrections	43%	17%	19%	20%
Nombre de messages consultés	**62%**	1%	3%	33%

TAB. 5.23 – Résultats de l'ACP : Les variables portées par chaque axe.

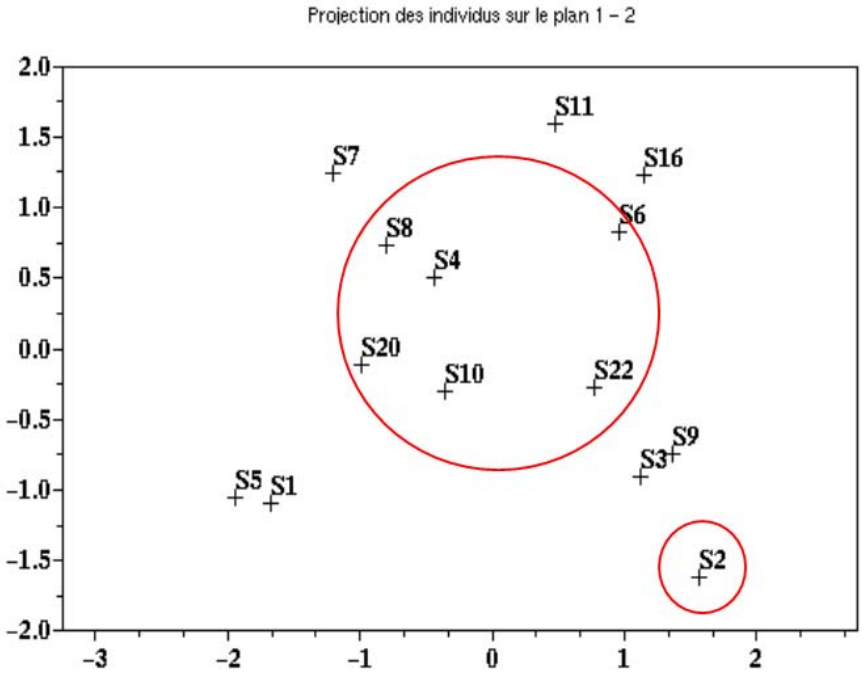


FIG. 5.5 – Projection des individus (Groupe GI, 14 participants) sur le plan représenté par les axes 1 et 2.

Il ressort de la figure 5.5 que les participants du sous-groupe SE (dont les participants ont utilisé le système d'aide sans Eloona pendant la première condition) sont regroupés dans le grand cercle rouge à l'exception du sujet 2 (dans le petit cercle rouge). Quant aux participants du sous-groupe AE (qui ont bénéficié de l'assistance d'Eloona pendant la première condition), ils

se situent sur les bords de la représentation. Cette représentation peut s'interpréter comme suit. Tout d'abord, les participants du sous-groupe SE se rassemblent autour de la moyenne, centre de la représentation (le point de coordonnées (0,0)), alors que le sous-groupe AE, beaucoup plus diversifié, se disperse aux quatre coins de la représentation. Il semble donc que la présence d'Eloona ait exercé une influence sur le sous-groupe AE dont les participants se sont soit laissés distraire par cette présence, ce qui expliquerait les mauvaises notes au post-test de certains d'entre eux, soit laissés guider par Eloona, ce qui expliquerait le nombre élevé de messages d'aide consultés. Ainsi, la présence d'Eloona aurait stimulé la consultation des messages d'aide et homogénéisé les performances des participants. Ainsi, la présence d'Eloona apparaît comme un facteur susceptible d'accroître la diversité inter-individuelle élevée que l'on observe dans les comportements cognitifs des utilisateurs mis en jeu lors d'activités telles que l'apprentissage de l'utilisation d'un nouveau logiciel, la familiarisation avec ses fonctionnalités.

Eloona au cours du temps

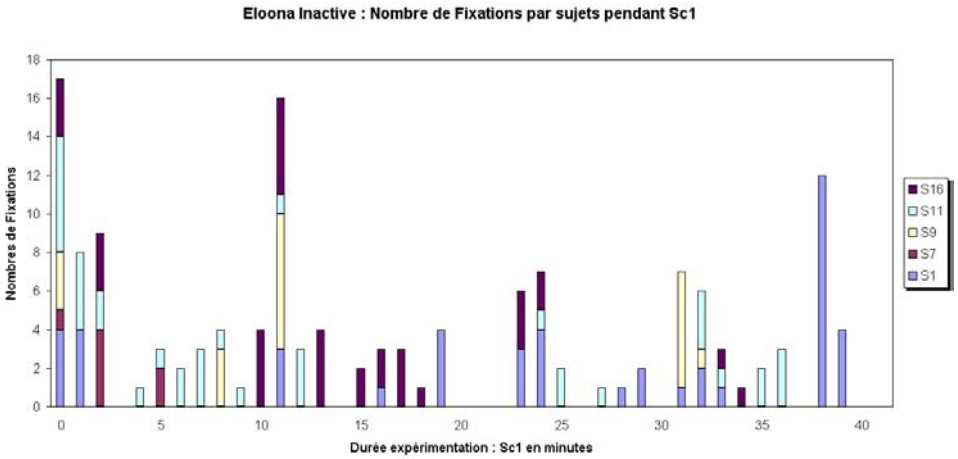


FIG. 5.6 – Nombre de fixations au cours du temps quand Eloona est inactive (Répartition par sujets).

Dans le but d'analyser l'observation d'Eloona au cours du temps (sur le premier scénario), nous présentons dans les graphiques 5.6 et 5.7 les fixations que les participants GI (au nombre de 5) équipés de l'eye tracker effectuent sur le clone, tout en faisant la distinction lorsqu'Eloona est inactive (tête fixe, clignements des yeux et mouvements de bouche) et lorsqu'Eloona énonce un message oral (mouvement de tête et animation du visage). Ainsi, l'on observe que l'ensemble du groupe observe Eloona inactive en début de session et en fin de session. Cette première observation permet de conclure que les participants ne se sont pas lassés d'elle au cours du premier scénario. Ici, la diversité interindividuelle est assez importante. En complément, les énoncés d'Eloona sont observés généralement entre 10 et 25 minutes du premier scénario, ceci correspond au moment où les participants commencent à créer l'interpolation de mouvement et tentent d'animer les poissons. Cette période est assimilée à la pointe de consultation des messages d'aide et par la même occasion, celle d'observation d'Eloona pendant les énoncés. Il serait intéressant d'étudier en détails les fixations oculaires pendant chaque énoncé, toutefois étant donné le nombre de

messages d'aide demandé assez faibles, une étude correspondant aux émotions ne serait pas représentatif.

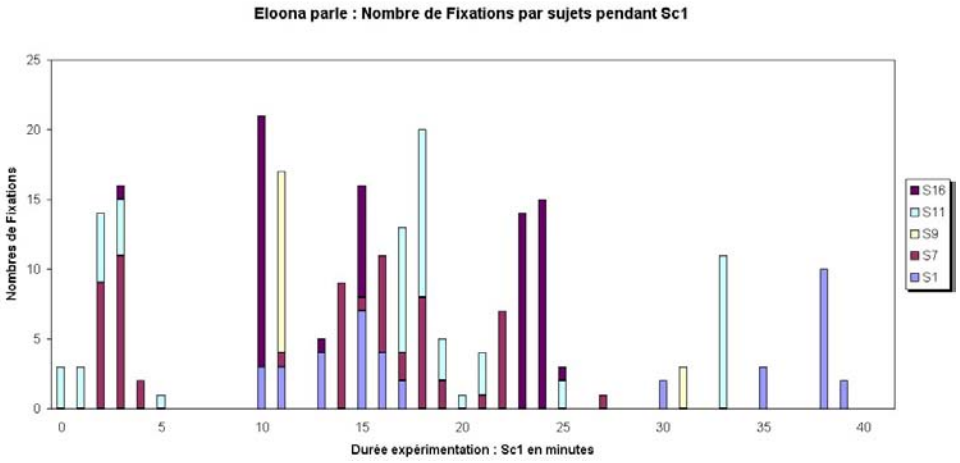


FIG. 5.7 – Nombre de fixations au cours du temps quand Eloona énonce un message d'aide (Répartition par sujets).

5.5 Conclusion et discussion

L'analyse des réponses aux questions verbales portant sur les impressions et les réactions suscitées par Eloona a montré que les participants, dans leur très grande majorité, ont exprimé des jugements positifs, confirmés par les réponses au questionnaire non verbal, sur l'apparence physique d'Eloona, son comportement ainsi que sa personnalité. En outre, le profil d'Eloona que nous avons élaboré semble avoir satisfait l'ensemble des participants, quelle que soit leur expérience en informatique.
Concernant le réalisme de l'ACA, nous avons vu que les avis étaient assez partagés. Néanmoins, il semble que le manque de réalisme de l'animation des ACAs anthropomorphes actuels ne soit pas un obstacle à leur utilisation en interaction homme-machine. On peut rendre la présence d'un ACA acceptable et attractive en donnant à celui-ci un rôle dans l'interaction qui ne soit pas simplement de divertir l'utilisateur, en le dotant d'une 'personnalité' en accord étroit avec son rôle dans l'interaction et en rendant sa présence 'physique' suffisamment discrète pour ne pas lasser l'utilisateur de sa présence et rendre acceptables les limites actuelles du réalisme de son comportement.
La présence d'Eloona a eu un effet rassurant sur une majorité de participants sur qui elle a eu des effets comparables à ceux d'une présence humaine coopérative.

Chapitre 6

Interfaces adaptatives : adaptation aux connaissances

Cette troisième étude, comme les deux études précédentes, est conçue dans le cadre d'une approche expérimentale. Elle est centrée sur l'évaluation d'un système d'aide adaptatif à l'utilisation du logiciel Flash. Ainsi, le but des participants était de se familiariser avec l'utilisation de Flash en réalisant deux animations (voir la tâche expérimentale et les scénarios Sc1, Sc2 et Sc3 décrits dans le chapitre 4), avec l'assistance d'un système d'aide adaptatif proposant des messages d'aide "Texte + Graphique". Dans cette étude, qui est exploratoire, nous simulons l'évolution du système d'aide par un expert humain (ou compère) grâce à la technique du magicien d'Oz implémentée dans la plate-forme logicielle que nous avons développée (voir le chapitre 4). 9 participants ont testé ce dispositif et nous avons recueilli, outre la trace de leurs interactions avec Flash et avec le système d'aide, leurs impressions et réactions vis-à-vis de l'adaptation dynamique des messages d'aide et de leur activation, à l'initiative de l'utilisateur mais aussi du système (c'est-à-dire du compère). Les deux formes d'adaptation dynamique mises en œuvre sont les suivantes :

- **Détection d'une intention de l'utilisateur** (par rapport à la tâche en cours) : Dès que le compère détecte une intention chez le participant, il lui envoie un message d'aide approprié à la situation. Cette stratégie évite-t-elle à l'utilisateur de perdre du temps à explorer l'aide ou à pratiquer une approche par essais-erreurs ?
- **Adaptation du contenu des messages d'aide** aux connaissances de l'utilisateur novice et à leur évolution : Chaque message d'aide est rédigé en trois versions plus ou moins détaillées ; ainsi, suivant le niveau de compétences de l'utilisateur estimé par le compère à partir de l'observation de l'interaction, le compère envoie à l'utilisateur la version du message adaptée à son niveau de connaissance de Flash et de compétence dans son utilisation. Nous détaillons dans les sections suivantes la stratégie de mise en œuvre de cette adaptation dynamique au profil de l'utilisateur courant. Une des questions auxquelles tente de répondre l'étude expérimentale présentée dans ce chapitre est la suivante : les variations du comportement d'un logiciel interactif adaptatif sont-elles perçues par l'utilisateur, plus généralement, quelles peuvent être les réactions de celui-ci à l'évolution autonome du comportement d'un tel logiciel au cours de l'interaction ?

Dans un premier temps, nous rappelons le contexte général de cette étude, puis, nous présentons la méthodologie utilisée lors de la conception du plan expérimental et du protocole. Les développements logiciels réalisés ayant été présentés au chapitre précédent, nous détaillons uniquement les composants nécessaires à la mise en œuvre du protocole de la présente étude. L'analyse des données recueillies lors de l'expérimentation est ensuite présentée et les résultats obtenus commentés.

6.1 Contexte - Généralités

6.1.1 Rappel des objectifs

L'objectif principal de l'étude présentée dans ce chapitre est d'évaluer la perception d'un système adaptatif dont le comportement évolue à l'initiative du système, dans une situation d'apprentissage interactif : l'aide en ligne à l'utilisation d'un logiciel grand public nouveau pour l'utilisateur.

Les effets, sur l'utilisateur novice de l'évolution dynamique autonome des réactions du système d'aide peuvent être d'une double nature :

1. Le fait que le système prenne l'initiative d'afficher un message d'aide susceptible de faciliter la réalisation de l'intention courante du novice, telle qu'il l'a détectée, peut accroître les performances de celui-ci. Le novice est en effet moins tenté d'entreprendre une approche par essais-erreurs qui, généralement, fait perdre du temps, ou de partir à la recherche d'informations dans une arborescence d'aide où, du fait de son ignorance des fonctions du logiciel, il risque de se perdre.
2. Le système adapte également le contenu des messages d'aide au niveau de compétence courant de l'utilisateur. Ce qui peut également accroître les performances de l'utilisateur en lui facilitant la compréhension de ces messages et, par conséquent, augmenter sa motivation pour la consultation de l'aide. On espère ainsi remédier au défaut majeur des aides classiques qui est d'être ignorées par la plupart des utilisateurs novices, comme cela a été observé dès les années 80.

L'expérience envisagée vise à recueillir les impressions, les comportements et les performances d'utilisateurs vis-à-vis d'un système d'aide adaptatif, de comprendre ce qui leur a plu ou déplu, ainsi que l'efficacité de cette forme d'aide, dans le but de proposer aux concepteurs des pistes utiles pour l'élaboration de systèmes interactifs adaptatifs qui soient efficaces, faciles et agréables à utiliser.

6.1.2 Choix des modalités d'expression des informations d'aide

Nous avons testé plusieurs modalités pour la formulation des informations d'aide : Graphique, Texte, Oral et ACA (voir les deux études précédentes). En particulier, l'étude préliminaire a montré que la majorité des participants ont préféré les messages "Texte+Graphique"). Mais

surtout, l'utilisation de messages "Oral+Graphique" aurait introduit une variable supplémentaire aux effets non négligeables (voir la seconde étude) sur les performances et les jugements subjectifs des participants ; il aurait donc été difficile de distinguer les effets de cette forme de multimodalité de ceux de l'adaptation dynamique. Nous avons donc choisi de créer des messages multimodaux Texte+Graphique. Ainsi, le seul effet de nouveauté de l'aide mise en œuvre est le caractère adaptatif de l'aide.
De même, pour éviter d'introduire un facteur d'innovation dans la présentation des informations d'aide, nous avons calqué la disposition de l'affichage sur celle des pages Web classiques.

6.2 Protocole expérimental - Définition

6.2.1 Plan expérimental - Aperçu global

Comme pour l'étude précédente, la situation d'interaction homme-machine est la consultation d'un système multimodal d'aide en ligne à la découverte et à l'utilisation de Flash, logiciel grand public de création d'animations multimédias. Nous avons fait tester un système d'aide adaptatif par 9 participants volontaires, tous issus d'un niveau Master à dominante non informatique. L'expérience en informatique de ces étudiants se réduisait à la bureautique
Le matériel utilisé (plate-forme logicielle) ainsi que les tâches expérimentales (scénarios d'animation) sont les mêmes que dans l'étude faisant intervenir Eloona, à l'exception de l'enregistrement des mouvements oculaires. Les informations objectives fournies par les mouvements du regard étaient inutiles, au cas particulier, dans la mesure où l'adaptation dynamique de l'aide était assurée par le compère. En revanche, ces informations seraient utiles pour la détection automatique des intentions de l'utilisateur car, lors de l'interaction avec une interface graphique, l'utilisateur regarde généralement l'objet graphique sur lequel il envisage de faire porter sa prochaine action avant d'entreprendre la réalisation de celle-ci. L'option d'enregistrement des mouvements oculaires a donc été désactivée lors du lancement de la plate-forme.

6.2.2 Hypothèses de travail - Mesures

Hypothèses de travail

L'étude envisagée a comme objectif de fournir des éléments de réponse aux questions de recherche suivantes :

H1. Un système d'aide qui, d'une part, adapte le contenu des messages aux compétences et connaissances de l'utilisateur courant et à leur évolution au cours de l'interaction et, d'autre part, anticipe ses besoins en informations d'aide, est susceptible de motiver l'intérêt de celui-ci pour l'aide. Comment seront perçues les initiatives du système mais aussi les variations du contenu des messages au cours de l'interaction ?

H2. Dans un contexte d'aide en ligne, un système adaptatif peut-il accroître les performances des utilisateurs, c'est-à-dire la rapidité et la qualité de leur familiarisation avec les fonctionnalités du logiciel nouveau pour eux ?

Mesures

L'évaluation de la perception du système adaptatif par les participants s'appuie essentiellement sur leurs réponses au questionnaires finals, verbal et non verbal. Ces données sont complétées par l'entretien de debriefing qui permet en effet d'approfondir certains points. En ce qui concerne les performances, comme pour l'étude précédente, nous avons utilisé l'enregistrement, grâce à la plate-forme logicielle développée, des traces des interactions des participants avec Flash et avec le système d'aide. L'analyse a porté, entre autres, sur les résultats de la réalisation des scénarios imposés (i.e., les animations créées) et les notes obtenues au post-test effectué en fin de session.

Le protocole expérimental est donc le même que pour l'expérimentation avec Eloona, à l'exception des modalités utilisées pour présenter les informations d'aide (ACA + Parole + Graphique devient Graphique + Texte) et de l'orientation des questionnaires finals (axés sur l'évaluation de l'évolution autonome du comportement du système d'aide et de ses capacités d'initiative).

6.2.3 Sujets

Nous avons orienté le choix des sujets en vue de contrôler l'âge, le niveau d'expérience informatique et le sexe des participants. Ainsi, ceux-ci sont tous des étudiants de niveau Master, engagés dans des cursus où l'informatique n'est pas une discipline dominante. La moyenne d'âge de ces participants est de 24 ans, et la fourchette d'âge est comprise en 20 et 26 ans pour 8 participants ; une participante de 32 ans se situe à l'extérieur de cette plage. Les jeunes sujets sont considérés a priori comme plus ouverts à l'innovation, plus adaptables à des situations nouvelles mais aussi plus critiques dans leurs jugements, ce qui nous intéressait particulièrement dans la mesure où l'un de nos objectifs était d'élaborer des recommandations utiles pour la conception des systèmes adaptatifs. Pour étudier l'influence de la variable sexe, 5 femmes et 4 hommes ont participé à l'expérimentation.

6.3 Protocole expérimental - Réalisation

Comme nous avons utilisé le même dispositif que pour la seconde étude (évaluation des effets de la présence d'un ACA), il n'est pas nécessaire d'expliquer ici le fonctionnement de l'interface de consultation du système d'aide, notamment en ce qui concerne les différentes requêtes que peut effectuer l'utilisateur et la manière de les exprimer. Il suffit de présenter ici uniquement la conception des messages d'aide et la disposition des informations dans la fenêtre d'aide.

Un autre point abordé ici est la stratégie mise en place par le compère lorsqu'il détecte une intention de l'utilisateur pour choisir le message d'aide (type et niveau) qu'il convient de lui

envoyer.

6.3.1 Messages d'aide - Interface et conception

Les messages d'aide, quelle que soit la requête activée par l'utilisateur ("Comment", "Quoi", "Pourquoi" et "Confirmation"), sont exprimés dans trois versions correspondant à trois niveaux d'information différents. Ces niveaux permettent au compère de simuler une aide adaptative aux connaissances courantes de l'utilisateur novice. Ainsi, on associe à chaque niveau, un profil d'expérience de l'utilisation du logiciel différent. Le compère utilise ces trois niveaux selon la stratégie suivante :

- **Niveau 1** : A utiliser la première fois que le novice pose la question (cette version est quasi identique au message oral correspondant utilisé dans le cadre de la seconde étude) ;
- **Niveau 2** (sujet expert) : A utiliser lors de la répétition d'une même question par le participant, lorsque celui-ci a su mettre à profit et utiliser correctement les informations fournies lors de la formulation initiale de la question (ou de la répétition précédente). Le message est un bref rappel des informations techniques incluses dans la formulation initiale ; en particulier, les graphiques (copies d'écrans de Flash) ne figurent pas dans cette version.
- **Niveau 3** (sujet en difficulté) : A utiliser lors de la répétition d'une même question à la suite d'une prise en compte laborieuse des informations obtenues lors de la requête précédente (formulation initiale ou répétition) ou d'une incapacité à les prendre en compte ou à les comprendre.

Le choix, pour le niveau 1, d'une formulation aussi proche que possible de celle de messages oraux utilisés dans le cadre de la seconde étude permet d'effectuer des comparaisons utiles entre cette étude et celle présentée dans ce chapitre quant à l'efficacité de l'apprentissage de Flash, puisque seule la modalité de présentation des informations d'aide varie d'une étude à l'autre.

Comme pour la version orale (voir le chapitre 5), chaque message comprend trois composantes :

- **Préconditions** : Absentes de la version 'expert' (niveau 2).
- **Instructions** (actions à effectuer) : Présentes à tous les niveaux, c'est le corps du message.
- **Informations complémentaires** (conseils, explications, détails procéduraux) : Absentes de la version 'expert' mais également du niveau 1. C'est cette composante qui permet de faire la distinction entre la version 1 et la version 3 (sujet en difficulté).

Pour ces trois composantes des messages d'aide, les choix de présentation suivants ont été adoptés : les informations procédurales et pré-conditions sont sur fond jaune clair en caractères foncés alors que les conseils et explications sont en bleu foncé sur fond bleu clair avec une police de taille réduite. On demande d'ailleurs aux participants, lors de l'entretien de debriefing si les rôles de ces différentes composantes ont bien été perçus.

Une autre différence avec la version orale existe : le 'dictionnaire' a été supprimé et remplacé par des liens dans le texte des messages. Lors de l'écoute d'un énoncé, si l'utilisateur ne connaissait pas la définition d'un mot, il pouvait consulter un menu déroulant (zone dictionnaire dans la fenêtre du système d'aide) qui contenait les définitions des principaux termes techniques relatifs à Flash. Dans la version textuelle et graphique, il était possible, et c'était plus simple pour

l'utilisateur, de remplacer le 'dictionnaire' par des 'liens' associés aux termes techniques inclus dans le texte des messages ; un clic sur l'un de ces liens permet d'obtenir une courte définition du terme correspondant, identique à celle figurant dans le dictionnaire pour ce terme. Ces liens ont la même apparence et le même fonctionnement que ceux présents dans les pages Html classiques : ce sont des mots soulignés qui peuvent apparaître dans n'importe laquelle des trois composantes. Ils figurent dans la version 1 des messages (requête initiale), n'apparaissent pas dans la version 2 (destinée aux 'experts') ; certains d'entre eux sont conservés dans la version 3 (destinée aux participants en difficulté), et les définitions correspondant aux autres termes (les principaux) sont explicitées dans le corps même des messages pour simplifier l'interaction nécessaire à leur consultation.

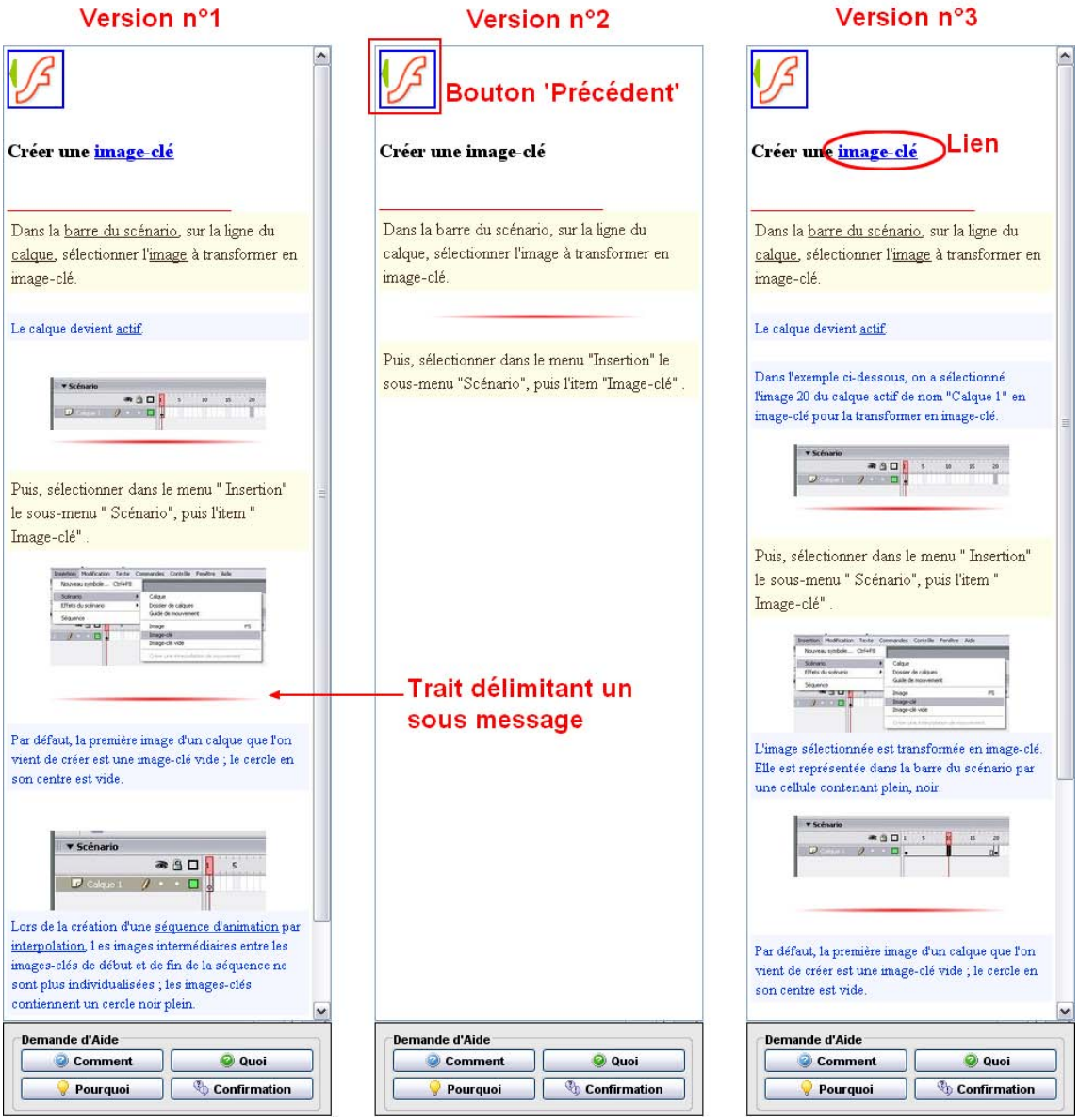


FIG. 6.1 – Présentation des trois niveaux : requête initiale, versions destinées, respectivement, aux 'experts' et aux participants en difficulté - Message d'aide : "Comment créer une image-clé".

Lorsque l'utilisateur clique sur un lien, la définition du terme correspondant recouvre le message courant, comme lors de l'activation d'un item du menu déroulant du 'dictionnaire' dans la seconde étude. Toutefois, suite aux remarques recueillies lors des debriefings des participants à la seconde étude, nous avons ajouté un bouton 'Précédent' (à côté du titre du message) qui permet de réafficher facilement le message précédent, sans avoir à reformuler leur requête. Nous n'avons pas introduit de bouton 'Suivant' (pour revenir au message courant), car cela aurait compliqué la tâche du compère inutilement.
La figure 6.1 présente, à titre d'illustration, les versions du message d'aide "Comment créer une image-clé" correspondant aux trois niveaux de complexité définis pour mettre en œuvre l'adaptation dynamique de l'aide aux compétences de l'utilisateur novice.

En résumé, la fenêtre d'aide contient :
- le titre du message (texte) ;
- un bouton 'Précédent' pour revenir au précédent message d'aide affiché ;
- les copies d'écran de Flash (ou Graphiques) ;
- au dessus de chaque copie d'écran, le texte correspondant à cette copie ;
- des traits horizontaux (en relief) pour structurer les messages complexes en séquences de sous-messages complémentaires ;
- pour certains messages "Comment", des liens vers d'autres messages "Comment" ; ces références se présentent sous forme de liens textuels ; ces liens ne sont pas différents des liens vers les définitions de termes techniques.

Pour les messages longs (i.e., nombreuses copies d'écrans), un ascenseur est prévu sur la droite de la fenêtre d'aide ; cet ascenseur que l'on voit sur la figure 6.1 est utile pour la version 3 des messages, plus détaillée que les versions 1 et 2.
Les copies d'écran peuvent être zoomées, comme dans la seconde étude, à l'initiative du sujet (par simple clic sur la copie). Le zoom fonctionne de la même manière que dans la seconde : il s'affiche dans la fenêtre Flash et, pour le faire disparaître, le sujet ferme simplement la fenêtre de zoom ou clique sur un lien.

6.3.2 Stratégie d'adaptation : Rôle du compère

Lorsque le compère identifie une intention du sujet, il envoie sur le poste 'Utilisateur', de sa propre initiative, le message "Comment" approprié, c'est-à-dire de niveau 1, 2 ou 3, correspondant aux connaissances du participant. La figure 6.2 illustre l'envoi d'un message depuis le poste de travail du compère ; on peut remarquer les trois versions (identifiées par I, II et III) de chaque message, ainsi qu'un compteur placé devant l'identifiant de chaque version qui permet au compère de connaître les versions déjà envoyées. Sur l'exemple, l'utilisateur a déjà reçu (compteur à 1 devant I) la version initiale, et le compère s'apprête à lui envoyer la version 'expert'. On rappelle que la sélection du message dans la base se fait automatiquement lors du choix de la requête par le participant et qu'il ne reste au compère que la sélection et l'envoi (clic droit sur le niveau, item envoi) de la version de niveau approprié.

On entend par identification d'une intention, une anticipation de la prochaine action ou tâche que le participant va entreprendre. Un critère parmi d'autres, mais important pour déclencher

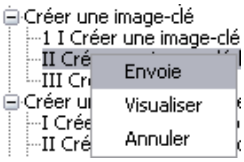


FIG. 6.2 – Le compère s'apprête à envoyer la version 2 (niveau 'expert') du message "Comment créer une image-clé" ; le niveau 1 a déjà été envoyé 1 fois (compteur présent devant l'identifiant de la version de niveau 1).

l'envoi d'un message, est le fait que le sujet ne fait rien pendant un temps de l'ordre de 5 secondes ou qu'il se lance dans une démarche de type essais-erreurs ou encore qu'il adopte un comportement incohérent.
Le message "Comment" qu'affiche le compère correspond :

- Lors de l'identification d'une intention du participant correspondant à une tâche, au message contenant les informations procédurales nécessaires pour réaliser la première étape de cette tâche ; le compère peut s'aider du scénario en cours pour identifier la tâche envisagée par le participant, mais il doit attendre que celui-ci ait entamé la réalisation de la tâche pour envoyer le message.
- Lors de l'identification d'une intention correspondant à une procédure, au message décrivant la procédure entière, et non au sous-message décrivant la première action atomique dont elle implique la réalisation.
- Lors de l'identification d'une action atomique, au message (sous-message) décrivant cette action.

Dans tous les cas, quel que soit le niveau de granularité de l'anticipation réalisée par le compère, celui-ci doit attendre que le sujet ait esquissé un début de réalisation pour envoyer le message ou qu'il soit inactif ou hésitant, ou encore qu'il s'engage dans une démarche de type essais-erreurs. Le message reste affiché jusqu'à la prochaine anticipation du compère ou jusqu'à la prochaine question du sujet à partir des boutons du panneau d'interaction avec l'aide.
Lorsqu'un message est affiché, le compère peut lui superposer l'affichage du contenu des liens (définitions des termes techniques ou messages "Comment") inclus dans ce message, de sa propre initiative. C'est à la charge du sujet de revenir au message initial contenant le lien en utilisant le bouton 'Précédent'.

Pour simplifier sa tâche, le compère dispose sur papier d'une classification regroupant, pour chaque étape de chaque scénario, la liste des différents messages "Comment" susceptibles d'être activés. Tous les niveaux de granularité sont représentés dans ces listes, de l'action atomique (e.g., 'Créer un calque', 'Afficher la bibliothèque') aux actions complexes et aux procédures (e.g., 'Créer une image-clé', 'Créer une interpolation de mouvement').

Les participants ne sont pas prévenus que le système d'aide est capable d'afficher des messages de sa propre initiative, ce qui constitue un choix de conception du protocole dont l'incidence sur leurs jugements subjectifs peut s'avérer importante.

6.4 Analyse des données

6.4.1 Profils des sujets - Questionnaire initial

9 étudiants volontaires, qui n'avaient jamais utilisé Flash auparavant (ni aucun autre outil de création d'animations puisqu'on relève 9 réponses négatives dans le questionnaire initial), ont expérimenté le système d'aide adaptatif décrit précédemment pour la création des deux animations, Aquarium et site Web ; ils ont donc réalisé les mêmes scénarios (Sc1, Sc2 et Sc3) que les participants à la seconde étude. Ils étaient informés qu'ils allaient évaluer un système d'aide innovant, mais la nature de l'innovation (i.e., l'adaptation dynamique du système au profil de l'utilisateur courant) ne leur était pas précisée. Les profils des participants, issus des réponses au questionnaire initial, sont décrits dans le tableau 6.1.

Groupes	Age (Moyenne)	Scolarité
F (5 participantes)	24 ans	3 Master, 2 écoles d'ingénieur
H (4 participants)	24 ans et demi	3 Master, 1 kinésithérapie

TAB. 6.1 – Caractéristiques générales des participants : âge, scolarité, sexe. F désigne le groupe des 'femmes' et H le groupe des 'hommes'.

Afin de caractériser davantage l'expérience informatique des participants, le questionnaire initial, identique à celui utilisé dans le cadre de la seconde étude, comprenait des questions à choix multiple sur, d'une part, la nature précise et la régularité de leurs activités informatiques et, d'autre part, leur approche et méthode pour prendre en main un nouveau logiciel. Le tableau 6.2 synthétise les réponses à ces questions pour l'ensemble des participants, ainsi que pour les groupes 'F' et 'H'.

Il ressort du tableau 6.2 les caractéristiques suivantes :

- Tous les participants sont familiers de la navigation sur Internet et utilisent souvent les logiciels grand public. Concernant les autres activités mentionnées dans le questionnaire initial, on observe une diversité inter-individuelle forte, sans aucune relation avec le sexe des participants.
- Lors de la prise en main d'un nouveau logiciel, les approches les plus souvent utilisées sont : par essais-erreurs et à l'aide d'exemples. Quant aux réponses aux questions portant sur l'utilisation d'autres approches, elles fluctuent entre "Jamais" et "Parfois", à l'exception de la lecture d'un manuel/notice qui est une approche que la plupart des participants n'utilisent "Jamais".

A souligner que l'analyse de ces jugements subjectifs ne fait apparaître aucune différence notable entre les groupes 'H' et 'F'.

6.4.2 Expression des préférences : aide adaptative versus aides classiques

Afin d'étudier les préférences des participants, nous avons procédé en trois étapes. Tout d'abord, les jugements globaux exprimés dans les questionnaires finals sont analysés, ce qui permet d'obtenir des informations sur la manière dont les participants jugent le système d'aide

Expérience en informatique	Jamais		Parfois		Souvent	
	F	H	F	H	F	H
Activités						
Saisie ou consultation de données	1	1	3	1	1	2
Navigation sur Internet	0	0	0	0	5	4
Utilisation de logiciels grand public	0	0	1	0	4	4
Jeux	1	1	2	2	2	1
Programmation	1	2	2	2	1	0
Conception et développement d'applications	3	2	1	1	0	1
Prise en main d'un nouveau logiciel						
Une approche par essais-erreurs	0	0	2	0	3	4
L'aide en ligne associée au logiciel	0	0	3	3	2	1
Un manuel/notice d'utilisation	0	0	3	4	1	0
Un ouvrage de référence	3	1	1	2	1	0
Un didacticiel en ligne	2	0	2	3	1	1
Des exemples	1	0	1	3	3	1
L'assistance d'un utilisateur expérimenté	0	1	3	1	2	2

TAB. 6.2 – Expérience informatique des participants : activités et stratégies de prise en main d'un nouveau logiciel.

adaptatif en le comparant aux aides classiques. Ensuite, nous rendons compte de leur évaluation des aspects suivants de l'aide adaptative : les messages, l'assistance et l'interface d'interaction. Enfin, pour compléter cette analyse, nous présentons leurs impressions et réactions, telles qu'ils les ont exprimées au cours des entretiens de debriefing.

Jugements globaux

A la question « Quel système d'aide en ligne préférez-vous ? », reprise en synthèse par une question où l'utilisateur pouvait choisir entre le système d'aide expérimental et les aides en ligne classiques (celles de Word ou Excel par exemple), 7 participants ont préféré le dispositif expérimental et les 2 autres ont privilégié les aides classiques. Toutefois, les justifications avancées par les partisans des aides classiques restent assez vagues : *« (Il choisit les aides classiques) par habitude. » S27* et *« (Les aides classiques sont) plus rapides (on rame moins), mais pas forcément plus pratiques sur certain points. » S30* .

Pour les partisans du système d'aide adaptatif, voici, tout d'abord, les différences principales qu'ils ont observées par rapport aux aides classiques, ainsi que les avantages que présente pour eux le système d'aide adaptatif :

- *« Beaucoup de systèmes d'aide en ligne proposent des recherches par mots-clés et ne renvoient pas toujours les informations adéquates. » S25* ,
- *« Le système qui devine les intentions est très intéressant, on trouve de plus ici la méthode complète pour réaliser une ou plusieurs actions. » S26* ,

- *« Le système est adapté à une personne qui maîtrise peu l'informatique (le langage est simple.) » S28* ,
- *« Anticipation des actions et conseils en rapport avec mes actions. » S29* ,
- *« Celui testé est plus précis. » S24* ,
- *« Le système d'aide m'a paru plus ciblé et se disperse moins que les aides en ligne classiques. » S31* ,
- *« Le fait que le système s'adapte aux actions est un point fort. Avoir la possibilité de demander une confirmation doit être vraiment très pratique pour des personnes non coutumières de l'utilisation des logiciels. Le "comment" et "quoi" sont des fonctions plus habituelles mais restent toujours très pratiques. » S32* .

Ensuite, nous présentons les arguments qu'ils avancent pour justifier leur choix en faveur du système adaptatif :

- *« (Il) cible directement les actions dont l'utilisateur a besoin. » S25* ,
- *« Réponses identiques que pour la comparaison avec les aides classiques. » S26* ,
- *« Simple et intuitif. » S28* ,
- *« J'aime l'accompagnement de l'aide, qui est efficace tout en restant discrète. Je ne supporte pas les aides par pop-up intempestives et j'ai tendance à les désactiver et à me priver d'une aide parfois bienvenue. » S29* ,
- *« Il est précis et apporte des réponses pratiques. » S24* ,
- *« Mieux conçu et plus ciblé. » S31* ,
- *« Parce qu'il réagit aux actions, le système de confirmation est vraiment pertinent pour des personnes inexpérimentées. » S32* .

Les conclusions qui émergent de ces commentaires sont que le système d'aide qui a été proposé aux participants, leur a plu en raison principalement des caractéristiques suivantes : sa capacité d'anticipation, la simplicité de son mode de fonctionnement, ainsi que la qualité des informations qu'il propose. En outre, il n'y a eu aucun rejet de la part des participants car même S27, qui préfère les aides classiques, juge positivement la capacité d'anticipation de l'aide proposée, *« L'aide vient quand on en a besoin »* , et évoque comme seul défaut : *« On ne peut pas "poser de question", y mettre un mot clé. »* , qui ne porte pas sur les capacités d'adaptation dynamique à l'utilisateur courant du système d'aide proposé.

Évaluation du système d'aide proposé

Cette section reprend certaines questions du questionnaire final afin d'orienter l'analyse vers les aspects que les participants ont évoqués dans leur jugement global.

Concernant l'adaptation dynamique, c'est-à-dire l'anticipation des besoins de l'utilisateur novice et l'évolution du contenu des messages en fonction de celle des compétences de celui-ci, deux questions permettent de connaître les avis des participants sur cet aspect. Les questions sont les suivantes :

– « Le système d'aide vous a-t-il envoyé des messages différents dans des situations identiques ? »
– « Lorsque le système devinait vos intentions, il affichait un message susceptible de vous aider à les réaliser. Par rapport aux systèmes d'aide courants, jugez-vous cette stratégie d'assistance : ... »

	Gr. F	**Gr. H**
Observation de l'adaptation du contenu ?		
A observé	0	1
N'a pas remarqué	5	3

TAB. 6.3 – Adaptation du contenu des messages - Répartition par sexe.

	Répartition suivant les groupes F / H : 9 participants						
Stratégie d'initiative du système :							
Inefficace <———> Efficace	0/0	0/0	0/0	0/1	1/1	2/1	2/1
Irritante <———> Satisfaisante	0/0	0/0	0/1	0/0	1/0	3/3	1/0

TAB. 6.4 – Évaluation de la stratégie d'assistance du système d'aide : l'anticipation des besoins en information de l'utilisateur.

Il ressort du tableau 6.3 que la plupart des participants n'a pas perçu l'évolution du contenu des messages au cours de l'interaction. Un seul l'a remarquée, notamment pour le message d'aide relatif à l'interpolation de mouvement. Toutefois, nous verrons dans l'analyse des debriefings que lorsqu'on commente l'expression "messages différents dans des situations identiques", on constate à travers les explications des participants qu'ils se sont aperçus du changement de comportement du système. La question était vraisemblablement mal posée étant donné la divergence que l'on constate entre les réponses au questionnaire et les commentaires formulés au cours des entretiens. Concernant l'anticipation des besoins des utilisateurs, qui constitue la caractéristique majeure de la stratégie d'assistance du système, 8 participants jugent cette stratégie satisfaisante (6 réponses au 6ème différenciateur) ; un seul la trouve irritante (S30). Ce jugement positif sur la stratégie d'assistance du système se traduit également par des avis positifs sur l'efficacité du système : 8 participants le trouvent efficace, voire très efficace, un seul porte un jugement neutre sur ce point.

Pour résumer l'opinion générale que les participants ont eu à l'égard du système d'aide et de son mode de fonctionnement, nous avons regroupé les réponses aux questions pertinentes du questionnaire final dans le tableau 6.5.

Les données présentées dans le tableau 6.5 permettent d'observer que 7 participants jugent utile et efficace l'assistance fournie par le système d'aide. Toutefois, les 2 autres participants la considèrent légèrement inefficace et ont un avis neutre quant à son utilité. Les participants ont donc, en majorité (7 participants sur 9), une opinion générale plutôt positive sur le fonctionnement du système d'aide. Pour ce qui est de l'interface d'interaction, on observe des jugements comparables : 7 et 8 participants (suivant la question) trouvent l'interface de consultation de

	Répartition suivant les groupes F / H : 9 participants						
Assistance fournie par le système :							
Inefficace <———–> Efficace	0/0	0/0	0/2	0/0	2/2	3/0	0/0
Inutile <———–> Utile	0/0	0/0	0/0	1/1	0/2	3/1	1/0
Interface de consultation de l'aide :							
Difficile <—> Facile (à utiliser)	0/0	0/0	0/0	0/0	0/3	2/1	3/0
Inefficace <———–> Efficace	0/0	0/0	0/1	0/0	1/1	3/2	1/0
Fastidieuse <—> Agréable (à utiliser)	0/0	0/0	0/0	0/1	1/1	2/2	2/0

TAB. 6.5 – Jugements sur le fonctionnement du système d'aide : Aperçu global.

l'aide facile, efficace et agréable à utiliser.
On notera que le groupe 'F' a toujours tendance à être plus positif dans ses jugements que le groupe 'H'.

Il convient pour terminer l'évaluation du système d'aide, de recueillir les jugements des participants sur les messages d'aide. Ces jugements nous renseignent sur la qualité de l'aide et la validité de nos choix de conception des messages. Les réponses aux questions sont présentées dans les tableaux 6.7 et 6.6.

Les données présentées dans ces tableaux montrent, comme pour l'expérimentation avec Eloona, que la conception des messages d'aide a été réussie : la majorité des participants en effet jugent les messages de longueur et de précision appropriée dans le contexte de la situation d'apprentissage étudiée.

Les messages d'aide vous ont semblé :					
Trop courts	0	Trop longs	1	De longueur appropriée	8
Trop vagues	1	Trop détaillés	1	De précision appropriée	5

TAB. 6.6 – Évaluation des messages d'aide.

En outre, le contenu des messages d'aide, en d'autres termes, les informations fournies, semblent claires et utiles à la majorité d'entre eux. Cependant, on observe un avis général neutre concernant le caractère suffisant des informations : davantage d'informations étaient peut être attendues. Malgré les avis positifs formulés sur le contenu des messages, on peut constater que 6 participants seulement sont satisfaits des informations obtenues, mais pas entièrement ; quant aux trois autres, ils ont l'impression que les informations fournies par le système d'aide n'ont pas répondu à leur besoin. Il ressort des commentaires associés à ces réponses des remarques comme les suivantes : *« Des fois on n'arrive pas à appliquer ce qui est indiqué. » S30* , *« Une fois la barrière de la découverte des bases de Flash franchie, on est un peu livré à nous-mêmes. » S26* , ou encore *« Les textes en bleu sur fond bleu sont moins lisibles, les liens hypertexte sont pratiques. » S32* . Ces remarques sont plutôt axées sur la manipulation du logiciel que sur les messages d'aide, hormis la remarque qui signale que le texte sur fond bleu n'est pas lisible.

	Répartition suivant les groupes F / H : 9 participants						
Les informations d'aide sont :							
Peu <—> Parfaitement (claires)	0/0	0/0	0/1	0/1	3/1	1/1	1/0
Inutiles <—> Indispensables	0/0	0/0	0/0	0/0	1/1	2/2	2/1
Insuffisantes <—> Suffisantes	0/0	0/0	1/0	2/2	2/2	0/0	0/0
Les informations obtenues ont-elles satisfait vos besoins ?							
Pas du tout <—> Entièrement	0/0	0/0	1/2	0/0	3/2	1/0	0/0

TAB. 6.7 – Évaluation de la qualité des messages d'aide.

Informations complémentaires - Les debriefings

Pour compléter l'analyse des jugements subjectifs des participants, nous présentons ici brièvement quelques commentaires extraits des debriefings qui peuvent, comme nous l'avons signalé précédemment, apporter des éléments d'information supplémentaires sur la façon dont les participants ont perçu les deux formes d'adaptation dynamique (au profil de l'utilisateur courant et à son évolution au cours de l'interaction) qu'ils ont eu l'occasion d'expérimenter.

Tout d'abord, l'ensemble des participants ont réagi à la stratégie d'anticipation et de prise d'initiatives du système d'aide. Voici quelques unes de leurs réactions :

- *« L'idée est bonne. » S30* ,
- *« Le mieux, c'est quand il anticipe. » S29* ,
- *« C'est un jeu avec la machine. » S28* ,
- *« J'ai bien aimé, cela dépanne bien que la machine trouve ce que je voulais. » S26* ,
- *« On avance, sans s'énerver, la machine voit ce qu'on doit faire et on est quitte de chercher. » S32* ,
- *« Quand on est perdu c'est bien, et cela devient drôle. » S25* .

D'abord, les participants ne se sont pas rendu compte que l'assistance fonctionnait à l'aide d'un compère et se sont pris au jeu. Certains attendaient même que le système envoie les messages pour éviter chercher eux-mêmes les informations (S32). La forme d'assistance évoluée proposée présente donc le risque de rendre l'utilisateur passif.

Ayant observé que la question portant sur l'adaptation du contenu n'avait pas été comprise par les participants, j'ai reformulé cette question lors des entretiens de debriefing. Ce qui m'a permis de constater que 6 participants sur les 9 se sont aperçus de l'évolution du contenu des messages au cours de l'interaction alors qu'un seul d'entre eux l'avait mentionné dans le questionnaire final. Leurs commentaires portent principalement sur la taille des messages ; voici, à titre d'illustration, trois d'entre eux : *« C'était plus court. » S30* , *« Le message devient plus petit. » S27* et *« L'on obtient un résumé. » S32* . En effet, lors du passage de la version 1 à la version 2 ('expert'), la taille du message diminue considérablement, étant donné que les graphiques, les pré-conditions et les explications complémentaires disparaissent. Cependant, la présence et les effets de cette forme d'adaptation ne sont peut-être pas suffisamment sensibles, en raison de la brièveté de

l'interaction avec l'aide qui limite le nombre de messages affichés à l'initiative du système ou sollicités par les participants.

Concernant la composition des messages (pré-conditions, instructions et conseils/explications), seuls trois participants ont remarqué et compris la signification des couleurs qui n'était pas expliquée lors de la présentation de l'expérimentation. Voici, à titre d'exemples, les interprétations de deux participants : *« En bleu c'est complémentaire et en jaune c'est l'essentiel. » S29* et *« En bleu c'est du détail et en jaune, c'est l'endroit où cliquer. » S25* .

En résumé, l'avis général sur le système d'aide est que sa discrétion et sa prise d'initiatives le rendent très agréable. Cependant quelques critiques formulées par les participants sont à prendre en compte concernant, notamment, les questions suivantes : quand le système doit-il faire évoluer le contenu des messages et dans quel contexte doit-il afficher de sa propre initiative un message (e.g., l'utilisateur est engagé dans la réalisation d'une action), et comment l'évolution autonome de son comportement peut-elle être utilisée pour faciliter la découverte de nouvelles fonctionnalités du logiciel et favoriser l'approfondissement de l'apprentissage de son utilisation.

6.4.3 Avantages et limites du système d'aide expérimenté

Nous rappelons que les analyses qui vont suivre ont été réalisées sur la première animation 'Aquarium' (Sc1 : construction du décor et Sc2 : animation des poissons).

Activité des participants / Utilisation de l'aide

Dans cette expérimentation, l'ordre de passation a été le même pour tous les participants. On distingue uniquement deux groupes ('F' et 'H') en fonction du sexe des participants pour déterminer si ce facteur peut exercer une influence sur les jugements subjectifs et les performances. Le groupe 'F' comporte 5 femmes et le groupe 'H' 4 hommes. Nous commençons par présenter dans le tableau 6.8 la répartition des participants, accompagnée des durées de réalisation de la première animation (scénarios Sc1 et Sc2).

Participants du Gr. F	**S25**	**S26**	**S27**	**S28**	**S29**
Durée 1ère animation : Sc1 et Sc2 (min)	41	33	44	42	42
Participants du Gr. H	**S24**	**S30**	**S31**	**S32**	
Durée 1ère animation : Sc1 et Sc2 (min)	35	39	46	50	

TAB. 6.8 – Répartition des participants dans les groupes F et H et durées de réalisation de la première animation.

Il ressort des données présentées dans le tableau 6.8 que la durée moyenne de réalisation de la première animation est de l'ordre de 40 minutes, soit 2 minutes de plus que pour le groupe GI dans la seconde étude (voir le chapitre 5). Les règles d'intervention du compère pour terminer la réalisation de la première animation sont identiques à celles adoptées dans le cadre de la seconde étude.

Pour analyser l'utilisation de l'aide par les participants, nous nous sommes appuyés sur la trace de leurs interactions avec Flash et l'aide, enregistrées tout au long de l'expérimentation grâce à la plate-forme logicielle décrite dans le chapitre 4. comme pour la seconde étude, nous commençons par calculer les pourcentages de clics (clics gauche et droit confondus) effectués dans chaque zone de l'interface : Flash, Système d'aide et Windows.

Thèmes	**Moyenne**
Flash	80 %
Système d'aide	11 %
Windows	9 %

TAB. 6.9 – Pourcentages de clics effectués au cours de la réalisation de la première animation, regroupés par thèmes.

Comme l'indique le tableau 6.9, on observe des pourcentages similaires à ceux obtenus dans le cadre de la précédente expérimentation (voir le chapitre 5). La navigation en termes 'de clics', quelles que soient les modalités utilisées, "Graphique + Texte" ou "Graphique + Oral", est d'une complexité semblable. En effet, comme l'occupation d'un message utilisant l'oral est plus faible, l'utilisateur n'a pas besoin de cliquer sur l'ascenseur pour consulter l'intégralité du message, mais il doit en revanche cliquer sur les boutons d'activation des énoncés ; il semble donc que les interactions nécessaires pour consulter les messages d'aide présentés sous forme "Graphique + Texte" ou "Graphique + Oral" soient d'une complexité équivalente. Concernant la répartition des clics dans les zones de l'affichage définies lors de l'annotation, nous avons rassemblé dans la figure 6.3 les données fournies par la présente étude et la précédente pour faciliter leur comparaison.

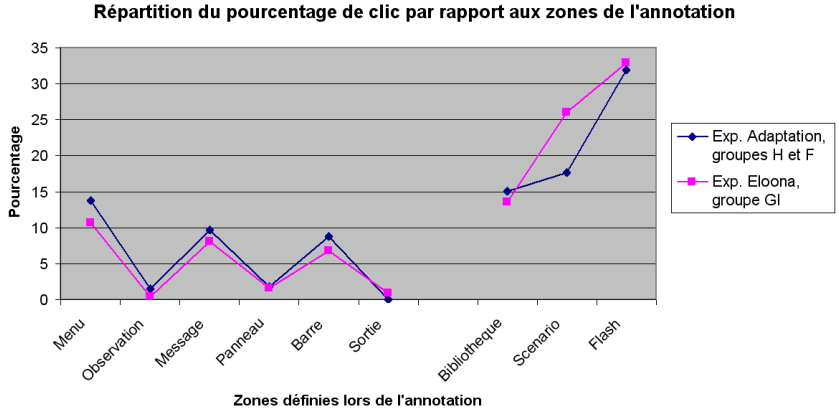


FIG. 6.3 – Pourcentage de clics par rapport aux zones définies lors de l'annotation. Comparaison avec l'expérimentation précédente.

L'analyse du graphique présenté dans la figure 6.3 montre que les courbes d'activité sont très proches pour les deux systèmes d'aide utilisés, adaptatif ou avec ACA. Toutefois, une différence majeure apparaît, concernant les pourcentages de clics dans la zone de la barre du scénario. En effet, le pourcentage de clics effectués dans cette zone par les participants à la seconde étude est supérieur de 10 % environ à celui observé pour les participants à l'étude décrite dans ce chapitre.

Cette différence a une signification forte, dans la mesure où c'est dans cette zone que se crée l'animation. Les participants qui ont utilisé le système adaptatif ont eu moins d'interactions avec la barre du scénario lors de l'animation des poissons que ceux qui disposaient des messages oraux d'Eloona.
Cette différence importante dans l'activité des participants aux deux études mérite d'être expliquée. Mais pour être en mesure de proposer des interprétations pertinentes, il est nécessaire d'étudier la répartition des accès aux menus pour déterminer si l'adaptation du contenu des messages aux compétences courantes de l'utilisateur et l'affichage des menus à l'initiative du système évitent ou réduisent les recours à une approche par essais-erreurs.

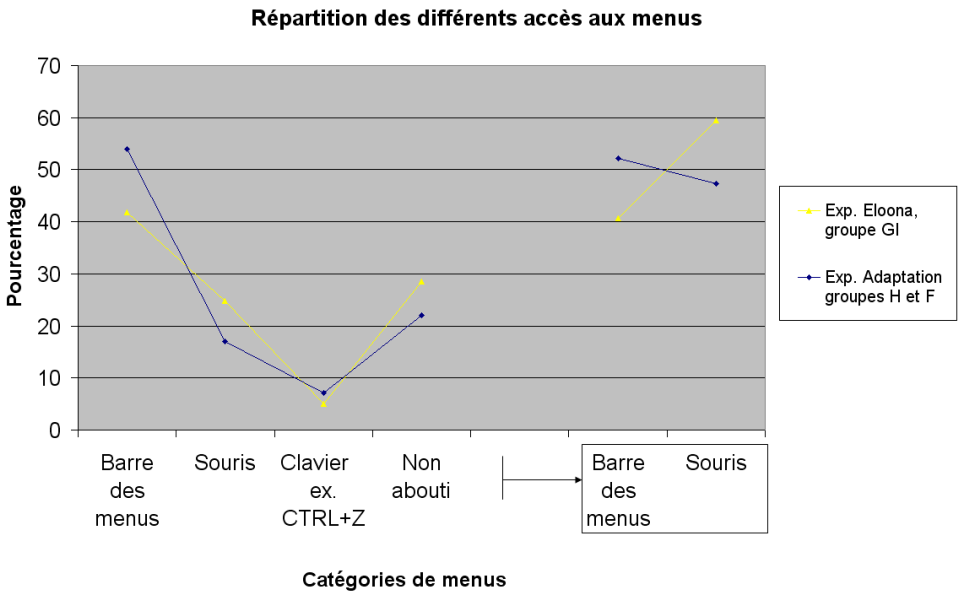


FIG. 6.4 – Répartition des différents accès aux menus. Comparaison avec l'expérimentation précédente.

Le détail des accès aux menus présenté sous forme de graphique (figure 6.4) permet d'observer que les participants qui ont bénéficié de l'assistance du système adaptatif recourent moins souvent à une approche par essais-erreurs que ceux de la seconde étude où les deux systèmes d'aide expérimentaux appliquaient une stratégie classique (i.e., statique) d'intervention. En effet, les pourcentages d'accès à la barre des menus montrent que les participants à l'étude décrite dans ce chapitre vont directement dans les menus pertinents pour la réalisation de leurs intentions et parcourent rarement des menus sans sélectionner d'item (pourcentage d'accès 'non abouti' plus faible) ; leur taux d'accès aux menus via la souris est aussi inférieur (pourcentage 'Souris' plus bas). Le système d'aide adaptatif semble donc fournir une meilleure assistance, un guidage des novices plus efficace dans la découverte des fonctionnalités d'un nouveau logiciel que les systèmes d'aide classiques dont le comportement est passif et n'évolue pas au cours de l'interaction de façon autonome.

Observons à présent les interactions réalisées avec le système d'aide.

Participants du Gr. F	**S25**	**S26**	**S27**	**S28**	**S29**	**Moy.**
Nombre total de "Comment" affichés	30	22	22	17	18	21.8
Messages "Comment" demandés par l'utilisateur	7	5	6	4	4	5.2
Messages "Quoi" demandés	2	2	9		1	X
Messages "Pourquoi" demandés	1	1	2			X
Messages "Demandes confirmations" demandées		1			1	X
Participants du Gr. H	**S24**	**S30**	**S31**	**S32**		**Moy.**
Nombre total de "Comment" affichés	16	28	32	15		22.7
Messages "Comment" demandés par l'utilisateur	4	6	15	11		9
Messages "Quoi" demandés	5	7	2			X
Messages "Pourquoi" demandés		1				X
Messages "Demandes confirmations" demandées			1			X

TAB. 6.10 – Traces des interactions des participants regroupées en fonction de la distinction homme-femme (H/F).

Dans le tableau 6.10, le nombre total de messages affichés dans chaque catégorie ("Comment", "Quoi", ...) correspond à la somme du nombre de messages de la catégorie demandée par l'utilisateur et du nombre de messages de la même catégorie affichée à l'initiative du système (i.e., du compère). Concernant les messages "Comment", le système a envoyé pour 7 participants trois fois plus de messages que ceux-ci n'en ont sollicités ; les participants S31 et S32 ont un ratio différent car ils ont sollicité eux-mêmes un nombre important de messages (respectivement 15 et 11), nettement supérieur aux nombres de requêtes de type "Comment" formulées par les autres participants. Toutefois, quel que soit leur sexe, les participants ont reçu en moyenne 20 messages "Comment" durant la réalisation de la première animation.
En ce qui concerne les messages "Quoi", on remarque une utilisation légèrement plus forte qu'avec Eloona ; il semble que les liens favorisent davantage la consultation de définition de termes que le dictionnaire mis en place lors de la précédente expérimentation. Ceci provient peut être de l'emplacement que nous avions attribué au dictionnaire, situé à distance des boutons d'activation des messages sonores et en bas de l'écran à droite.
Pour les messages "Pourquoi", il convient de remarquer que 4 participants seulement ont utilisé cette fonctionnalité. C'est un résultat intéressant au sens où les participants à l'étude précédente qui ne disposaient pas de la présence d'Eloona (groupe SE) n'employaient pas ce type de requête alors que la présence d'Eloona favorisait l'emploi du "Pourquoi". Ici, semble-t-il, le fait que le système prenne l'initiative le rend en quelque sorte intelligent, de manière semblable à la présence d'Eloona, représentation du système qui par son apparence et son comportement quasi humains incite à croire que le système possède des capacités cognitives proches des nôtres.
Dans cette expérimentation comme dans la précédente, les requêtes "Demande de confirmation" ont été très peu utilisées : 3 demandes pour les 9 participants. L'utilité de ce nouveau type de fonctionnalité de l'aide ne semble pas avoir été comprise par les participants aux deux études qui la mettaient en œuvre.

Performances des participants

Afin d'évaluer les performances des participants, nous avons comparé les résultats de la création de l'animation 'Aquarium' et les notes obtenues au post-test par les participants à l'étude décrite ici avec les résultats et les notes obtenus par les participants à la seconde étude décrite dans le chapitre 5. A noter que la comparaison fait intervenir, pour la seconde étude, les données des groupes GI et GNI.

	Note Aquarium (moyenne)	**Note Post-Test** (moyenne)
Exp. ACA, groupe GI (14 participants)	12,7	19,5
Exp. ACA, groupe GNI (8 participants)	12,6	17,6
Exp. Adaptation, groupes 'F' et 'H' (9 participants)	11,1	15,4

TAB. 6.11 – Performances des participants (résultats de la création de la première animation et notes obtenues au post-test). Comparaison avec les performances des groupes GI et GNI (expérience informatique) des participants à la seconde étude.

Les notes regroupées dans le tableau 6.11 montrent que les résultats de la création de la première animation sont semblables (variation des moyennes de 11,1 à 12,7), quel que soit le niveau d'expérience en informatique, avec tout de même une moyenne légèrement inférieure pour les participants qui ont utilisé le système adaptatif. Un autre aspect important des résultats concerne l'assimilation des concepts et des savoir-faire que nous avons évaluée via le post-test. Ainsi, on observe que les notes varient sensiblement en fonction du niveau de compétence en informatique mais aussi en fonction du système utilisé.
On observe une diversité inter-individuelle importante dans les notes de création de la première animation pour les 'informaticiens' (Gr. GI). Mais celle-ci disparaît au niveau du post-test, ce qui permet à ce groupe d'obtenir la meilleure moyenne. Le groupe GI fait donc preuve ainsi d'une certaine aisance pour assimiler les fonctionnalités d'un nouveau logiciel et à exprimer les concepts et savoir-faire assimilés dans les réponses aux questions du post-test.
Les non-informaticiens du groupe GNI réalisent des performances inférieures, avec deux points de moins par rapport au groupe GI. Cette différence semble due à la difficulté que les 'non informaticiens' ont à se souvenir des procédures à suivre (suite d'items à activer), à assimiler et définir explicitement les concepts nouveaux (animation statique, par exemple) évoqués dans le post-test.

A noter que si l'on ne fait plus la distinction entre les niveaux d'expérience informatique des participants à la seconde étude, et que l'on prend comme critère de groupement la présence de l'ACA, on obtient le tableau 6.12 où toutes les notes sont semblables avec cependant un léger avantage pour le groupe qui a bénéficié de la présence de l'ACA.

En ce qui concerne les participants qui bénéficiaient du système adaptatif, on constate que ce groupe recueille les plus mauvaises notes. Pour ce qui est des résultats de la création de la

	Note Aquarium (moyenne)	**Note Post-Test** (moyenne)
Exp. ACA, groupe 'sans Elonna' (11 participants)	12,5	18,6
Exp. ACA, groupe 'avec Eloona' (11 participants)	12,8	19,0

TAB. 6.12 – Performances des participants à la seconde étude (résultats de la création de la première animation et notes obtenues au post-test) regroupés en fonction de la présence (groupe AE) ou de l'absence (groupe SE) d'Eloona.

première animation, les performances inférieures obtenues sont en accord avec le faible pourcentage de clics dans la zone de la barre du scénario signalé plus haut, et s'expliquent par un état de l'animation moins avancé, d'où une note moins élevée.
Toutefois, peut-on réellement comparer les conditions "Texte+Graphique" et "Oral+Graphique" ? Dans le chapitre 3, nous avions fait l'hypothèse que l'utilisation de l'"Oral+Graphique" pouvait s'avérer légèrement plus performante que celle du "Texte+Graphique", dans la mesure où les participants pouvaient manipuler en même temps qu'ils écoutaient un énonce. Mais cette hypothèse est-elle applicable à l'expérimentation décrite dans ce chapitre ? Pour tenter de répondre à cette question, nous avons mesuré la durée cumulée des moments d'inactivité des participants aux deux études (assistance assurée par un système adaptatif ou par un ACA). Nous avons considéré qu'un participant était inactif lorsque la souris était immobile, ce qui est le cas, entre autres, lorsque celui-ci lit le texte d'un message. Ainsi, les durées des moments où, pendant la réalisation de la première animation, la souris était immobile ont été cumulés ; les pourcentages présentés dans le tableau 6.13 ont été calculés sur la durée totale de la réalisation de la première animation.

	Taux d'inactivité (moyenne en %)
Exp. ACA, groupe GI (14 participants)	52,40
Exp. ACA, groupe GNI (8 participants)	53,11
Exp. Adaptation, groupes 'F' et 'H' (9 participants)	61,58

TAB. 6.13 – Pourcentages d'inactivité des participants (souris immobile) durant la première animation (groupes H et F). Comparaison avec la seconde étude (groupes GI et GNI).

Les résultats présentés dans le tableau 6.13 montrent que les deux groupes GI et GNI (seconde étude) ont un pourcentage d'inactivité inférieur de 10 % à celui des participants à l'étude décrite ici. Comme seule la présentation des messages d'aide varie d'une expérience à l'autre, on peut conclure que la lecture des messages induit un retard dans la création de l'animation. Ainsi, même si le système adaptatif prend l'initiative d'envoyer les messages appropriés au bon

moment, il conduit à une réalisation de la première animation plus lente que celle obtenue avec l'assistance de messages oraux ; la lecture des messages textuels ralentit sensiblement la réalisation de la première animation, ce qui explique vraisemblablement la légère différence que l'on observe en faveur des participants à la seconde étude concernant les résultats de la création de la première animation.

Une autre différence, plus importante (pouvant atteindre 4 points), toujours en faveur des participants à la seconde étude, apparaît dans les résultats au post-test. Une explication pourrait être que, comme le système envoie les "bons" messages aux "bons" moments, les participants appliquent "bêtement", sans vraiment réfléchir, les instructions reçues ; l'assimilation des concepts et la mémorisation des termes techniques sont donc inférieures à celles obtenues avec les systèmes classiques dont la passivité impose un apprentissage actif aux utilisateurs novices. Ce résultat constitue une piste de recherche intéressante : la prise en charge et le guidage des participants, qui distingue l'aide adaptative des systèmes d'assistance classiques, conduisent-ils à une assimilation des concepts et des procédures à acquérir pour maîtriser l'utilisation d'un nouveau logiciel, plus médiocre que les systèmes classiques ? Cette question a une portée très générale qui englobe non seulement l'aide en ligne mais également tous les systèmes interactifs adaptatifs développés dans le cadre des environnements intelligents d'assistance à l'apprentissage, notamment ceux destinés à l'enseignement. Les données empiriques et expérimentales sur l'efficacité de l'assistance à l'apprentissage fournie par les systèmes adaptatifs sont encore trop peu nombreuses pour que l'on soit en mesure de répondre à cette question de façon définitive.

Comparaison des résultats de l'ACP pour le groupe GI et les groupes 'F' et 'H'

L'idée ici est d'appliquer l'ACP aux données recueillies sur l'utilisation du système d'aide adaptatif en utilisant les mêmes variables qu'au chapitre précédent, à savoir : les résultats de la réalisation de la première animation, les notes du post-test, le nombre d'auto-corrections effectuées par chaque participant et le nombre de messages d'aide qu'il a consultés de sa propre initiative (ne sont pas comptabilisés les messages envoyés à l'initiative du système).
En effet, pour que la variable relative aux messages soit commune aux deux études expérimentales, celle où les utilisateurs novices utilisent le système d'aide adaptatif et celle où l'assistance est fournie oralement par un ACA, nous avons choisi de prendre en compte le nombre de messages d'aide sollicités par chaque participant.
Précisons également que, pour l'étude présentée dans le chapitre précédent, nous avons retenu uniquement les données du sous-groupe AE du groupe GI, constitué des participants 'informaticiens' qui ont bénéficié de l'assistance d'Eloona pour réaliser la première animation. Ce choix est motivé par la volonté de comparer les performances d'utilisateurs potentiels des deux systèmes d'aide innovants que nous proposons.

Les valeurs des quatre variables sélectionnées ont été soumises à Scilab. Les résultats de l'exploitation de ces données sont présentés dans le tableau 6.14 qui précise la répartition des participants suivant l'axe 1 porté à 56 % par la variable 'Post-test', et l'axe 2 porté à 66 % par la variable 'Nombre d'erreurs'. Sachant que l'inertie cumulée n'atteint pas 80 % sur les axes 1 et 2, il faudrait, en toute rigueur, tenir compte du plan défini par les axes 2 et 3. Cependant, les

calculs permettent de constater que la prise en compte de ce plan n'apporte aucune information complémentaire ; les données corresspondantes ne seront pas représentées ici.

	Axe1	**Axe2**	**Axe3**	**Axe4**
Notes 1ère animation	47%	7%	**43%**	2%
Post-test	**56%**	8%	17%	19%
Nombre d'erreurs	16%	**66%**	2%	16%
Nombre de messages d'aide demandés	40%	31%	10%	19%

TAB. 6.14 – Résultats de l'ACP pour les participants qui ont bénéficié de l'assistance du système d'aide adaptatif. Variables portées par chacun des principaux axes.

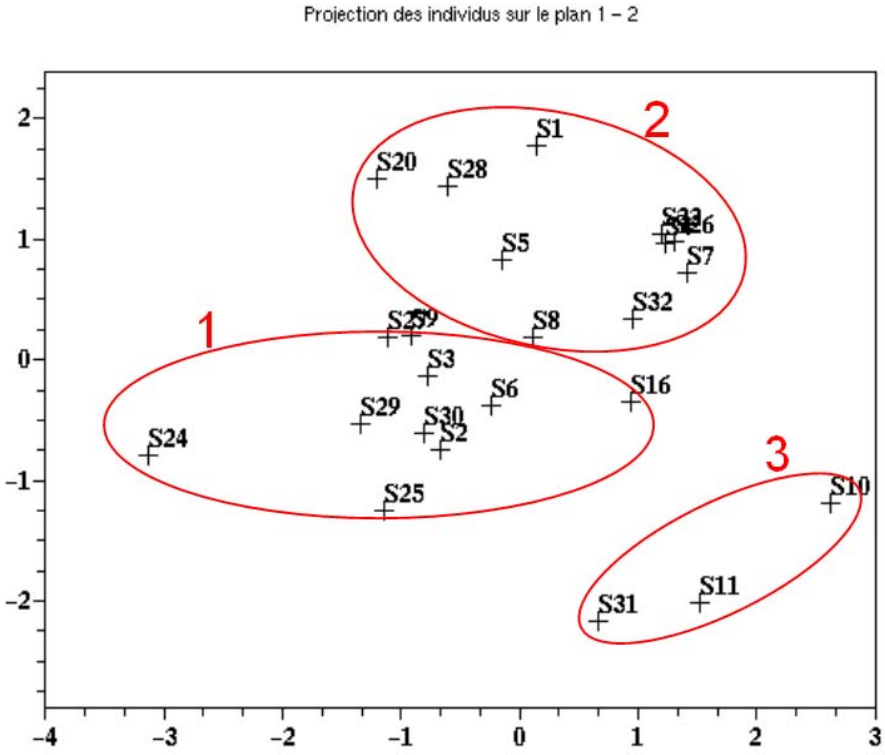


FIG. 6.5 – Projection des individus (Groupe GI et Groupes 'H' et 'F' confondus) sur le plan défini par les axes 1 et 2.

Appliquées aux données représentées dans la figure 6.5, les méthodes de clustering permettent d'identifier trois groupes entourés par des cercles rouges et numérotés de 1 à 3.
Voici, en résumé, les caractéristiques communes à chacun de ces groupes :

- **Groupe 1** : Proche de la moyenne, ce premier groupe consulte relativement peu l'aide, ne commet pas d'erreur, mais obtient une note médiocre au post-test. Il suit les instructions sans vraiment assimiler ni mémoriser les concepts. Il est composé en majorité de non informaticiens.

- **Groupe 2** : Ce groupe consulte peu de messages, commet énormément d'erreurs, mais obtient de bonnes notes (réalisation de l'animation et post-test). Il adopte une stratégie par essais-erreurs qui lui permet d'obtenir la meilleure note au post-test. Il est constitué d'une majorité d'informaticiens.

- **Groupe 3** : Comme le groupe précédent, celui-ci obtient de bonnes notes (réalisation de l'animation et post-test) mais, contrairement au précédent, il n'hésite pas à consulter l'aide fréquemment et commet peu d'erreurs. Vu sa petite taille, il constitue un profil vraiment à part, d'ailleurs, il est composé d'un 'non informaticien' et de deux 'informaticiens'.

6.5 Conclusion

L'analyse des réponses aux questions verbales sur les impressions et réactions suscitées par le système d'aide adaptatif a montré que les participants, dans leur grande majorité, ont exprimé des jugements positifs quant à la prise d'initiative du système et à la qualité des informations qu'il dispense. En outre, la conception de la stratégie d'adaptation, ainsi que la conception des messages d'aide, est en accord avec les besoins et les attentes des participants.
Ainsi, le système est bien perçu dans sa globalité. Les nouvelles fonctionnalités d'aide (voir les questions de type "Pourquoi" et "Demande de confirmation") sont encore peu utilisées ; cependant, les requêtes de type "Pourquoi" ont suscité des réactions positives de la part des participants. Toutefois, il semble que ces nouvelles fonctions de l'aide ne puissent être acceptées que si les participants ont l'impression qu'ils interagissent avec un système intelligent. Pour mémoire, les requêtes de type "Pourquoi" n'ont été utilisées, dans le cadre de la seconde étude, qu'en présence de l'ACA.
Concernant les performances, il est difficile de se prononcer étant donné la difficulté que présente l'évaluation d'un système adaptatif. Pour réellement évaluer l'influence d'un tel système sur les performances, il faudrait comparer son utilisation, par les mêmes participants, avec une version statique "Texte+Graphique" du système de façon à éliminer les effets des différences cognitives inter-individuelles, ce qui est impossible à réaliser ou comparer l'utilisation, par deux groupes de taille importante (plusieurs dizaines, voire centaines de participants), de chaque système, ce qui est irréaliste compte tenu du coût de l'annotation manuelle de la trace des interactions (voir l'évaluation des résultats de la création de la première animation). De même, la comparaison des performances des participants qui ont utilisé le système adaptatif avec celles des participants qui ont bénéficié de l'assistance de l'ACA ou du système "Oral+Graphique" est biaisée en partie, car l'inactivité (en termes d'interaction) due à la lecture des messages textuels (voir le tableau 6.13) peut avoir une influence négative sur les notes des participants qui ont utilisé le système d'aide adaptatif.
Enfin, les commentaires des participants utiles à prendre en compte dans la conception de logiciels d'assistance adaptatifs est que le système doit éviter de prendre l'initiative et de fournir des informations d'aide lorsque l'utilisateur est engagé dans une activité même si ces informations tiennent compte de façon pertinente des besoins courants de celui-ci dans la mesure où elles se fondent sur l'anticipation de ses intentions, des actions qu'il envisage de réaliser avec l'assistance ou la coopération du système.

Chapitre 7

Conclusions et perspectives

D'après l'état de l'art présenté dans le chapitre 2, on observe que les ACAs et les interfaces adaptatives sont des thèmes de recherche en plein essor. Les problématiques encore peu abordées, suscitées principalement par les avancées techniques récentes dont les possibilités sont à explorer, visent à comprendre et à évaluer l'influence de la présence des ACAs sur l'interaction, et à répondre aux nombreuses questions d'ordre ergonomique que soulèvent les interfaces adaptatives quant à leur apport effectif à l'interaction.

D'une part, les études portant sur l'évaluation de l'incidence effective des ACAs sur l'interaction n'utilisent généralement que des techniques de recueil des jugements subjectifs d'utilisateurs potentiels, à savoir, questionnaires, interviews, 'focus groups'. En effet, le recours, à l'observation des comportements et à l'analyse des performances et de l'usage (trace des interactions) est rare. Nous nous sommes donc intéressés à l'observation et à l'analyse de l'apport effectif de la présence d'un ACA, en tant que variable libre, à l'efficacité et à l'utilisabilité de l'interaction : présence/absence d'un ACA dans une situation réaliste, l'assistance à l'utilisation d'un logiciel grand public.

D'autre part, malgré les nombreuses difficultés méthodologiques que soulève l'étude ergonomique des systèmes d'interaction adaptatifs, nous avons tenté de répondre aux questions suivantes : Quel contrôle donner à l'utilisateur sur l'évolution du comportement du système ? La transparence de l'adaptation est-elle souhaitable ? Ne vaut-il pas mieux notifier l'utilisateur de l'évolution spontanée du comportement du système au cours de l'interaction et, si oui, sous quelle forme ? Ces questions, ainsi que les observations des participants à l'étude préliminaire font que nous avons choisi d'étudier les réactions d'utilisateurs face une interface qui s'adapte de manière autonome aux connaissances et compétences de l'utilisateur novice et à leur évolution au cours de l'interaction, sans l'informer de cette évolution.

L'objectif final était d'élaborer des recommandations utiles pour la conception de systèmes d'aide intégrant des techniques d'adaptation et des ACAs, qui soient à même de répondre aux besoins et attentes effectifs des utilisateurs.
Nos travaux contribuent à offrir aux concepteurs des informations utiles pour la conception d'interfaces utilisateur capables d'accompagner le grand public dans la prise en main d'un logiciel courant grâce à la présence d'un ACA ou d'un mécanisme d'adaptation dynamique à l'utilisa-

teur courant. Ils visent, d'une part, à accroître les capacités d'assistance et de coopération des systèmes interactifs et, d'autre part, à les rendre plus attractifs et plus agréables à utiliser, voire à humaniser l'interaction et lui donner une dimension affective.

Nous avons adopté une démarche expérimentale en vue d'atteindre cet objectif. L'étude préliminaire que nous avons menée à titre exploratoire avait pour objectif d'évaluer l'efficacité et la convivialité de messages d'aide en ligne intégrant la parole ou le texte, et le graphique. L'observation des résultats et réactions des utilisateurs lors de cette étude préliminaire nous a permis d'entreprendre la conception, la réalisation (technique du magicien d'Oz) et l'évaluation ergonomique d'une première interface intégrant un ACA, Eloona (chapitre 5), puis d'une seconde interface mettant en œuvre une technique d'adaptation dynamique du comportement du système au profil de l'utilisateur courant et de son évolution (chapitre 6). Ces deux expérimentations ont été réalisées dans une situation d'assistance à l'utilisation d'un logiciel grand public Flash MX, inconnu des participants.
La réalisation et à la mise en œuvre de ces études expérimentales ont nécessité le développement d'une plate-forme logicielle qui fournit une assistance évoluée au compère dans le cadre de la technique du magicien d'Oz et permet d'enregistrer une trace des interactions utilisateur-système relativement riche : les actions (souris, clavier) de l'utilisateur avec Flash et les deux systèmes d'aide expérimentaux et les réactions du système, éventuellement ses mouvements oculaires et ses énoncés, sont datés et sauvegardés. Pour faciliter l'analyse de ces données comportementales, un logiciel intégré à la plate-forme permet de rejouer ces données de façon synchronisée et de les annoter semi-automatiquement.
En ce qui concerne les participants, 24 utilisateurs ont testé le système avec Eloona, 12 informaticiens et 12 autres ayant un parcours où l'informatique n'est pas dominante. Pour le système adaptatif, 9 étudiants non informaticiens (5 femmes et 4 hommes) ont expérimenté ce système capable d'adapter le contenu des messages aux compétences de l'utilisateur et d'anticiper les besoins de celui-ci en affichant, de sa propre initiative, des informations d'aide pertinentes dans le contexte d'interaction courant.

Évaluation ergonomique d'un ACA : Eloona en situation d'assistance

Tout d'abord, la situation expérimentale créée est réaliste, au sens où les participants se sont réellement investis dans l'activité proposée, l'apprentissage de Flash, et y ont pris un réel intérêt ; on peut faire l'hypothèse qu'ils ont "oublié" le caractère expérimental de la situation proposée. Toutefois, les jugements recueillis au terme de cette session ponctuelle sont à prendre avec prudence, étant donné le caractère innovant de l'environnement logiciel proposé ; en particulier, l'effet de nouveauté lié à la présence d'Eloona a pu influencer positivement les jugements subjectifs des participants, mais cet effet risque de disparaître rapidement avec une pratique prolongée. On ne peut affirmer avec certitude que leurs jugements seraient aussi positifs au terme d'un usage régulier de l'interface. Nous nous sommes efforcés de limiter les effets de ce facteur. L'expérimentation a été conçue pour qu'il influence le moins possible les jugements des participants. En effet, nous avons familiarisé ceux-ci avec la présence d'Eloona avant qu'ils ne s'engagent dans la

réalisation des scénarios, puisque c'est Eloona (en plein écran) qui a assuré la présentation de l'expérimentation à l'ensemble des participants.

L'analyse des réponses aux questions verbales portant sur les impressions et les réactions suscitées par la présence d'Eloona a montré que les participants, dans leur très grande majorité, ont exprimé des jugements positifs, confirmés par les réponses au questionnaire non verbal, sur l'apparence physique d'Eloona, son comportement ainsi que sa personnalité. En outre, le profil d'Eloona, à savoir, son apparence, ses expressions faciales, sa personnalité et la prosodie de ses énoncés semblent avoir donné satisfaction à l'ensemble des participants, quelle que soit leur expérience en informatique. Les principes qui ont guidé la conception de l'animation d'Eloona et que nous avons présentés dans le chapitre 5 semblent donc pertinents.

Le premier principe préconise de donner à un ACA dont le rôle est d'assister l'utilisateur dans son activité une présence discrète, de façon à ne pas distraire l'utilisateur de cette activité. Cette discrétion offre l'avantage supplémentaire de rendre moins visible le manque de réalisme actuel de l'animation, opinion partagée par une majorité participants. Les comparaisons faites spontanément par les participants avec le compagnon Office de Microsoft montrent que ceux-ci ont apprécié la discrétion de la présence d'Eloona, ce qui valide ce choix de conception. En effet, Eloona est là quant on en a besoin mais ne dérange pas !

Le second principe recommande de donner à l'ACA un comportement et une "personnalité" en accord avec le rôle qu'il joue dans l'interaction. L'application de ce principe à la conception d'Eloona impliquait, compte tenu de ses fonctions d'assistant/tuteur, que son comportement soit à la fois sérieux et professionnel pour être crédible, affable, chaleureux et empathique pour paraître coopératif. Les participants ont effectivement perçu ces dimensions que nous avions voulu donner à la "personnalité" d'Eloona. Nombreux sont, par exemple, ceux qui ont apprécié son comportement coopératif, souriant, voire joyeux et qui l'ont jugée sympathique. Ces réactions des participants montrent le bien-fondé de nos choix de conception de la personnalité d'Eloona.

Cependant, l'analyse des réactions des participants met en évidence la diversité inter-individuelle de leurs attentes et de leurs intérêts. Même si la majorité des participants se sont montrés satisfaits, la "personnalité" d'Eloona n'a pas fait l'unanimité. Ainsi, pour satisfaire au mieux les exigences de l'ensemble de la population d'utilisateurs visée, il semble nécessaire d'offrir la possibilité à ceux-ci, soit de choisir l'apparence et le comportement de "leur" ACA, soit de personnaliser l'ACA proposé à l'aide d'outils simples à utiliser. La personnalisation, sous la forme d'une adaptation statique ou dynamique des ACAs aux préférences de l'utilisateur courant apparaît comme un domaine de recherche vaste, difficile et peu exploré, mais dont l'intérêt pour le succès de la diffusion des ACAs en interaction homme-machine et dans la société paraît incontestable.
Les résultats de l'ACP offrent quelques pistes pour identifier les facteurs à l'origine de la diversité inter-individuelle observée. En effet, cette analyse permet dans un premier temps de distinguer les utilisateurs confrontés à l'ACA en premier (utilisateurs en dehors des cercles rouges) mais aussi d'observer des réactions non homogènes face à la présence d'Eloona (les quatre coins de la représentation). La présence d'Eloona n'agit donc pas de la même manière sur l'ensemble des individus. De plus, tous les résultats (hors questionnaires) ont tendance à montrer que la présence d'un ACA apporte un plus dans l'interaction, mais étant donné le faible nombre d'utilisateurs

et la diversité inter-individuelle, il est impossible d'entreprendre une analyse statistique pour déterminer si cette observation constitue un résultat significatif ou non.

Cela dit, la présence d'Eloona a eu un effet rassurant sur une majorité de participants sur qui elle a eu des effets comparables à ceux d'une présence humaine coopérative. Ainsi, les résultats de cette étude suggèrent qu'Eloona pourrait être intégrée avec profit à l'interface utilisateur des logiciels, aides en ligne ou didacticiels, destinés à faciliter la prise en main d'un nouveau logiciel. Les participants ont jugé sa présence utile pour faciliter la découverte d'un logiciel ; en revanche, comme il est impossible pour l'instant d'anticiper les réactions des utilisateurs à long terme, on pourrait envisager de restreindre la présence de l'ACA à la prise en main d'un nouveau logiciel ou à l'interaction avec des logiciels d'utilisation ponctuelle et/ou occasionnelle (e.g., l'installation de la LiveBox proposée par France Télécom ou l'installation et la configuration d'un système d'exploitation, Windows ou Linux par exemple).

Évaluation du système adaptatif

Comme pour la seconde étude, la tâche expérimentale a plu à l'ensemble des participants de la troisième étude qui leur a permis de bénéficier d'une aide en ligne adaptative. Aucune remarque négative n'a été formulée sur l'intérêt des activités proposées. Malgré une durée d'expérimentation assez longue, les participants affirment n'avoir pas vu le temps passer, tant ils étaient absorbés par la découverte de Flash.

L'objectif était de recueillir et d'analyser les réactions d'utilisateurs non avertis des prises d'initiative du système concernant l'affichage et le contenu informationnel des messages d'aide. Le constat majeur est une bonne perception de l'adaptation dans sa globalité. Cependant, ici encore il faut prendre garde à l'effet de nouveauté qu'était susceptible de créer ce système innovant. Une étude complémentaire sur une durée beaucoup plus longue serait nécessaire pour valider le caractère positif de ces jugements. Mais la réalisation d'une telle étude imposerait le développement d'un prototype opérationnel capable de reproduire la stratégie d'assistance du compère, en particulier son anticipation des besoins en informations de l'utilisateur courant.

Les participants ont remarqué et apprécié la capacité du système d'aide à envoyer des messages "appropriés" au "bon moment". La pertinence des apports d'information du système et celle du "timing" de ses interventions tiennent essentiellement à la simulation de l'adaptation dynamique de l'interaction par un opérateur humain. Il serait extrêmement difficile actuellement, compte tenu de l'état de l'art dans le domaine des interfaces adaptatives, de créer un système intelligent opérationnel capable de détecter avec autant de pertinence l'intention courante de l'utilisateur et d'adapter le contenu des messages d'aide à ses compétences et à leur évolution au cours de l'interaction. Sans la simulation de l'intelligence du système par le compère, les jugements des participants auraient certainement été moins positifs. Les entretiens de debriefing ont montré que certains participants avaient critiqué l'envoi de messages par le système pendant qu'ils travaillaient. Cette discussion montre les limites de la technique du magicien d'Oz. La principale difficulté est moins d'identifier l'intention courante de l'utilisateur que d'envoyer le message d'aide approprié lorsque celui-ci en a besoin afin de ne pas perturber son activité, ce qui est en dehors des capacités des systèmes adaptatifs actuels et risque de l'être encore pen-

dant longtemps, puisque même les humains commettent parfois des erreurs. Cependant, si cette difficulté n'est pas surmontée, les systèmes d'assistance adaptatifs dotés d'initiative risquent de susciter un rejet de la part des utilisateurs.

Concernant l'adaptation du contenu des messages aux compétences courantes de l'utilisateur, il est difficile d'observer ses effets ou, plus précisément, les effets de la navigation effectuée par le compère dans les versions de niveaux différents d'un même message, dans le cadre de cette expérimentation qui ne permet la consultation que d'un nombre relativement faible de messages. Pour les 9 participants, les niveaux le plus souvent utilisés ont été le niveau 1 (première consultation) et le niveau 2 ('expert'). Nombreux ont été les sujets qui ont perçu l'évolution du contenu des messages et qui ont apprécié d'obtenir un résumé de la version initiale d'un message "Comment" à la suite d'une première exécution réussie des instructions détaillées dans la version initiale. Cependant, là aussi l'intelligence humaine, au cas particulier celle du compère, est nécessaire pour répondre de façon satisfaisante aux attentes des utilisateurs ; en effet, l'adaptation du contenu des messages peut s'avérer complexe dans certains contextes, par exemple lorsque le résumé du message ne suffit pas et que le système doit être capable de s'en rendre compte et d'envoyer une version plus détaillée du message, à choisir entre la version initiale et celle destinée aux utilisateurs 'en difficulté'. Or les systèmes actuels sont incapables de traiter ces situations de façon pertinente

Du point de vue des performances, il est également difficile de se prononcer, étant donné la difficulté que présente l'évaluation d'un système adaptatif. Au cas particulier, pour réellement tester les performances d'utilisateurs potentiels, il faudrait les comparer avec celles qu'ils auraient en utilisant une version "Texte+Graphique" de l'aide qui soit non adaptative pour l'apprentissage du même logiciel, ce qui est de toute évidence impossible. Les différences cognitives inter-individuelles sont telles que, pour obtenir des résultats significatifs à partir de la comparaison de deux groupes de sujets différents, il faudrait que chaque groupe inclue plusieurs dizaines, voire davantage, de participants, ce qui représente un coût prohibitif lorsqu'on annote et analyse les interactions de chaque participant comme nous le faisons. La comparaison avec le système "Oral+Graphique", malgré ces réserves, montre que les participants du groupe GNI, de profil semblable à ceux des participants à la troisième étude ont des performances meilleures que ces derniers. Concernant les résultats de la création de la première animation, cette différence peut s'expliquer par l'utilisation de modalités différentes pour présenter les informations d'aide. En effet, pendant la lecture d'un message textuel, l'utilisateur ne peut interagir avec le logiciel, ce qui est possible lorsque les messages sont oraux. En revanche, la différence entre les notes obtenues au post-test, au bénéfice des participants qui ont bénéficié d'une assistance orale suggère des pistes de recherche intéressantes car nouvelles. On peut interpréter cette différence de deux manières : ou bien l'oral a permis une assimilation meilleure des concepts de Flash et une meilleure mémorisation des procédures d'activation de ses fonctionnalités, ou bien les initiatives de l'aide adaptative, son anticipation des besoins en informations des utilisateurs novices a favorisé, chez les participants qui ont utilisé cette forme d'aide en ligne le développement d'une certaine passivité susceptible d'avoir nui à l'assimilation et à la mémorisation des concepts et procédures de Flash. Des études empiriques et expérimentales complexes sont nécessaires pour déterminer la pertinence de ces deux hypothèses.

Enfin, certains participants ont suggéré, dans leurs commentaires sur le système d'aide adaptatif que celui-ci devrait favoriser la découverte/exploration des fonctionnalités avancées du logiciel par l'utilisateur, ne pas se limiter à celle de ses fonctions de base. Ce qui nécessiterait d'accroître sensiblement les capacités d'adaptation du système, bien au delà des capacités des prototypes actuels, ce qui est peut-être irréaliste. En effet, pour qu'un tel système recueille une large adhésion, il faudrait non seulement qu'il dispose d'un modèle évolué de l'utilisateur et de stratégies d'adaptation sophistiquées, mais encore qu'il soit robuste et transparent, c'est-à-dire que l'utilisateur puisse sans difficulté "comprendre/interpréter" son comportement, et éventuellement le contrôler. Ces deux derniers aspects qui interviennent largement dans la qualité ergonomique d'un logiciel, sont loin d'être acquis et nécessitent un effort de recherche important en intelligence artificielle et en ergonomie des logiciels.

Aide en ligne, nouvelles fonctionnalités

L'exploitation des données a montré que les fonctionnalités de l'aide proposées n'ont pas toutes été largement utilisées. En effet, le "Comment" est d'après les participants la fonction de l'aide la plus utile et la plus efficace. Cependant, même si l'objectif qui leur était assigné était d'évaluer le système d'aide, les participants se sont tellement absorbés dans la création des animations qu'ils n'ont pas à tester les fonctionnalités originales des systèmes d'aide expérimentaux à leur disposition, vérifiant une nouvelle fois le paradoxe de motivation mis en évidence dès les années 80 dans Carroll [Carroll et Rosson, 1987]. Un point important, c'est que la fonction "Pourquoi" n'a été utilisée que par les groupes assisté par Eloona ou par le système adaptatif. En dotant le système d'aide de comportements "intelligents", à savoir, anticipation et initiative, ou représentation par une tête parlante, a-t-on incité les participants à le considérer différemment des systèmes classiques et donc à lui poser des questions analogues à celles qu'ils poseraient à un assistant/tuteur humain ? Ce résultat offre des pistes de recherche intéressantes pour donner à l'utilisateur une image réaliste du système avec lequel il interagit, de ses capacités effectives et de ses limites.

Trace des interactions : les données oculométriques

L'utilisation de l'oculomètre a été relativement bien perçue par les participants, à l'exception du poids du casque qui s'est fait sentir en fin de session pour certains d'entre eux. En effet, comme l'expérimentation dure environ 1h, le poids de 400 grammes du dispositif devient progressivement intrusif. En outre, étant donné la durée de l'expérimentation, il faudrait tous les quarts d'heure par exemple effectuer un contrôle pour vérifier que le miroir n'a pas bougé et que les données sont toujours de bonne qualité. Toutefois, ces interventions risqueraient de perturber l'activité des participants et d'introduire un biais dans l'expérimentation, ne serait-ce qu'en rappelant la présence de l'oculomètre que de nombreux participants ont réussi à oublier.
On observera qu'il n'a pas été fait d'analyses approfondies des mouvements du regard, mais que les premières observations faites sur les durées moyennes des fixations oculaires suffisent pour distinguer les différentes activités des utilisateurs. Par exemple, la durée des fixations sur Eloona

varie selon qu'elle est inactive ou qu'elle parle. De même, les activités demandant davantage de réflexion (construction de l'animation) engendrent des fixations plus longues.

Perspectives

Les perspectives de ce travail sont nombreuses. D'abord, avant d'intégrer un ACA doué de capacités d'expression orale à un logiciel d'assistance commercialisé ou diffusé à grande échelle, il serait utile de mener des recherche complémentaires dans les deux directions suivantes :

- Pour résoudre les problèmes d'utilisation inhérents aux présentations orales d'informations par rapport aux messages textuels, à savoir l'obligation de réécouter un message familier depuis le début lors de la recherche d'une information ponctuelle, il serait utile, comme l'ont suggéré plusieurs sujets, de proposer des solutions qui combinent texte et parole et d'évaluer la qualité ergonomique de ces propositions.
  Une autre approche, plus ambitieuse, consisterait à donner la possibilité à l'utilisateur d'accéder par mots-clés à des fragments du message oral de façon à lui fournir un accès au contenu des messages oraux comparable, en termes de facilité et de rapidité, à celui dont il dispose pour l'information visuelle, notamment textuelle.
- Il serait également prudent de tester la robustesse des résultats de l'étude ergonomique des effets de la présence d'Eloona sur l'interaction, en effectuant une étude d'usage qui impliquerait un nombre de participants beaucoup plus important, avec des profils plus variés, dans un contexte d'utilisation réelle, c'est-à-dire à domicile ou sur le lieu de travail, et éventuellement sur une période de quelques mois si un usage régulier, plutôt qu'occasionnel ou unique, est visé.

Pour le système adaptatif, les messages d'aide étant sous forme HTML, une étude envisagée consisterait à rendre disponible ce système via le Web, ce qui permettrait une évaluation à grande échelle. Toutefois, il faudrait implémenter la stratégie du compère, par exemple sous forme de règles d'anticipation des besoins courants en information de l'utilisateur, pour que le système d'aide active un message pertinent dans la version appropriée et au moment où l'utilisateur novice a besoin de cette information, ce qui est complexe si le système doit être robuste et son comportement facilement compréhensible par l'utilisateur. Dans ce contexte, la définition et la mise à jour par le système du profil de l'utilisateur (i.e., de ses compétences et activité courantes) représente un véritable défi scientifique.

Concernant les données oculométriques recueillies, l'idée, à court terme, est d'étudier précisément les aller-retours du regard entre le système d'aide oral et le logiciel, en présence et en l'absence d'Eloona. Cette analyse est en cours.

Annexe A

Étude préliminaire - compléments

L'annexe A détaille certains points abordés lors de la présentation de l'étude préliminaire.

Description de l'eye-tracker

Sur le schéma ci-dessous, vous pouvez visualiser le dispositif de l'eyetracker qui permet de capturer les mouvements oculaires.

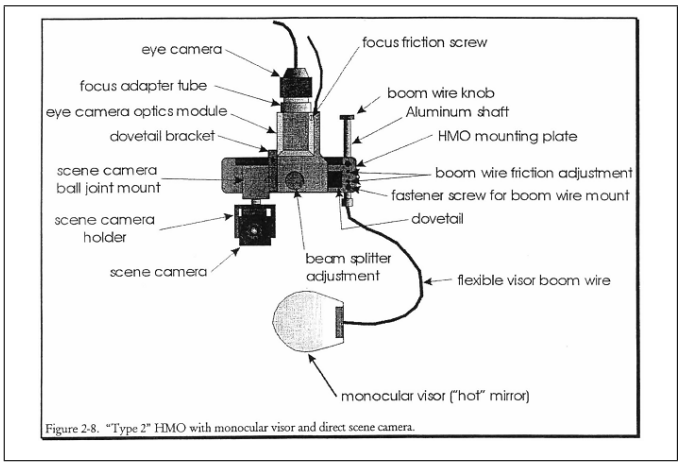


FIG. A.1 – Description de l'eye-tracker.

Architecture de l'expérimentation (Différents lieux)

Lieu n°1 : Salle sans ordinateur, on y effectue la présentation de l'expérimentation, la saisie des questionnaires et l'entretien de débriefing.

Lieu n°2 : Salle d'expérimentation (deux ordinateurs) où l'on effectue la présentation de Flash, l'expérimentation (calibrage, manipulation sous Flash) et le post test. Sa description se

trouve ci-dessous :

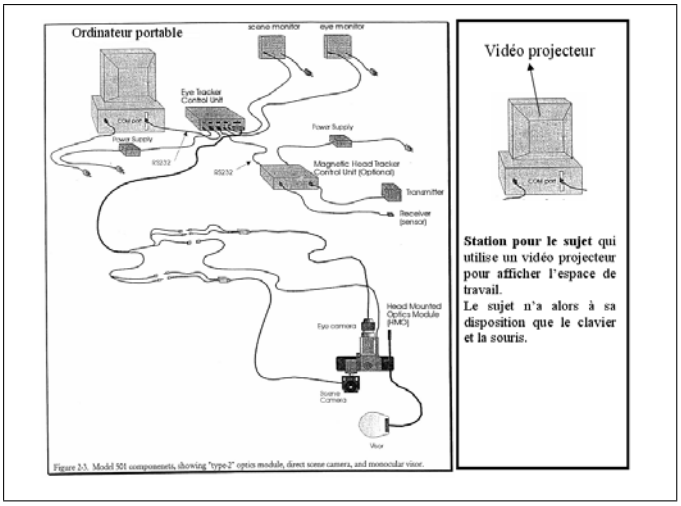


FIG. A.2 – Architecture de l'expérimentation.

Messages d'aide

Nous montrons ici deux messages d'aide, un pour la condition "G+T" et un pour la condition "O+G", la totalité des messages est disponible sur CD.

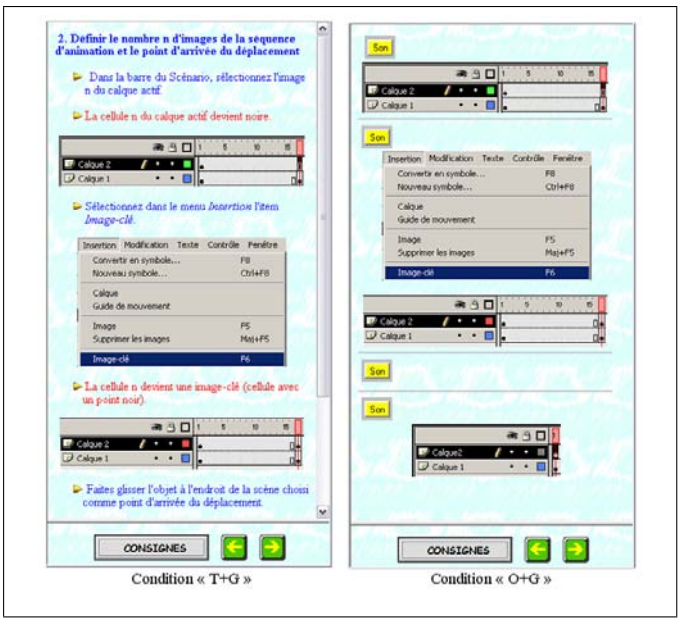


FIG. A.3 – Messages d'aide.

Post-test scénarios 1 et 2

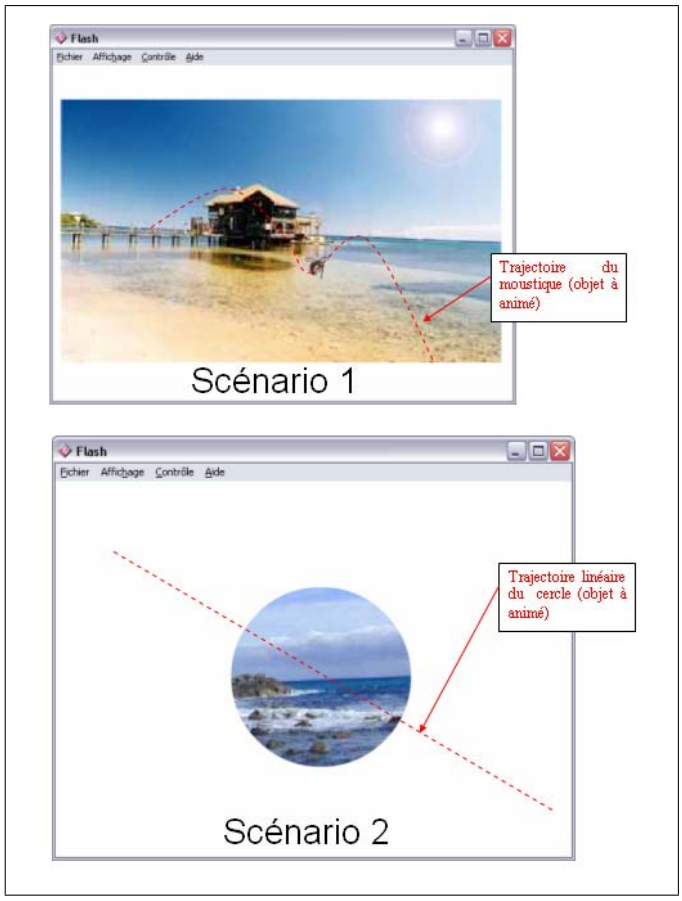


FIG. A.4 – Post-test scénarios 1 et 2.

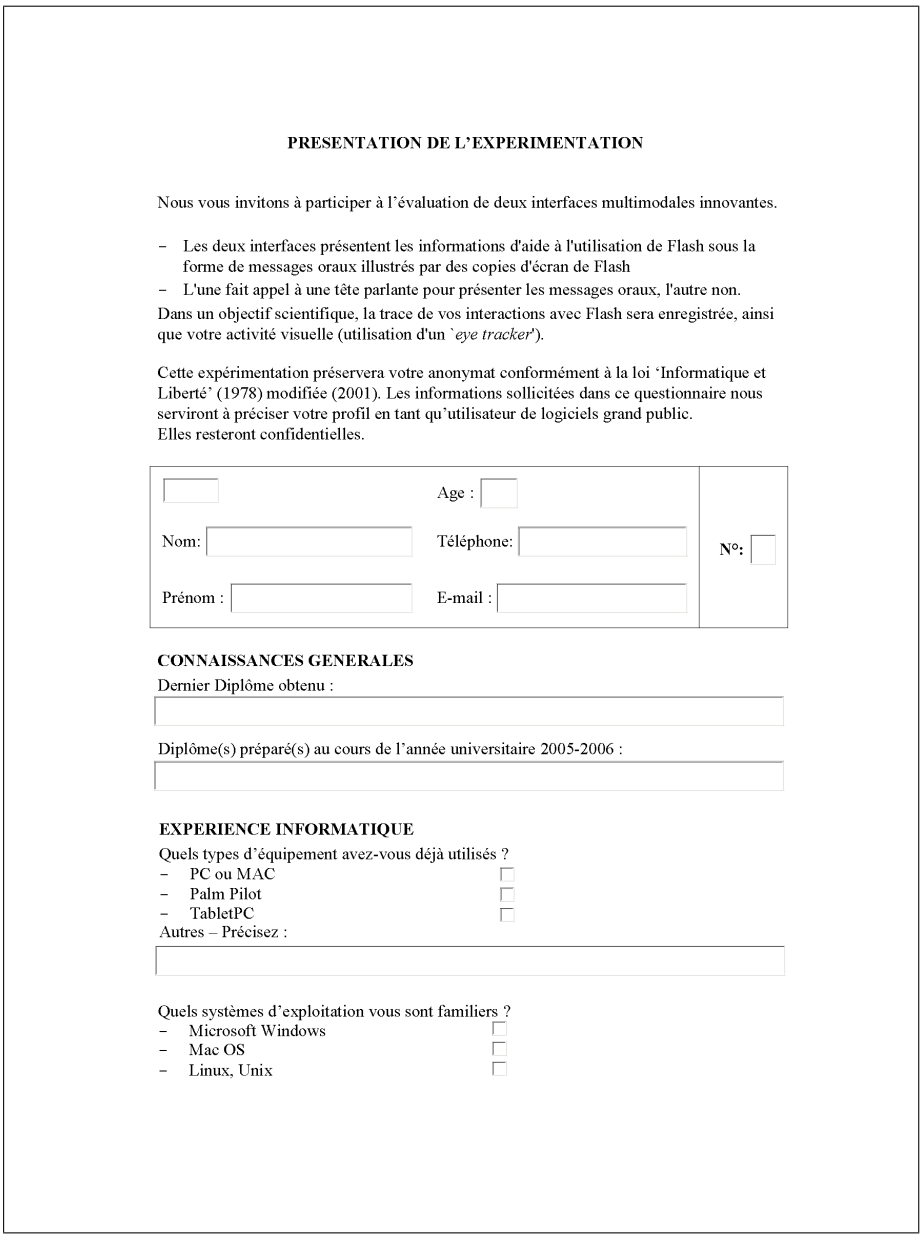
Annexe B

Études 2 et 3 : Détails du déroulement d'une passation

L'annexe B contient le détail de certaines parties du déroulement d'une passation entreprise lors des études 2 et 3. Les passations étaient individuelles, dans une salle d'expérimentation dédiée, et comprenaient plusieurs étapes. Ainsi, l'annexe permet de visualiser ce que les participants ont expérimenté mais permet également de distinguer les variances effectuées entre les deux études.

Questionnaire d'identification sous Adobe (i.e., caractérisation du profil de l'utilisateur)

Ce questionnaire contient deux pages et son contenu est similaire d'une étude à l'autre. Les informations qui sont demandées, et toutes les informations relatives à l'expérience resteront bien entendu confidentielles.



PRESENTATION DE L'EXPERIMENTATION

Nous vous invitons à participer à l'évaluation de deux interfaces multimodales innovantes.

- Les deux interfaces présentent les informations d'aide à l'utilisation de Flash sous la forme de messages oraux illustrés par des copies d'écran de Flash
- L'une fait appel à une tête parlante pour présenter les messages oraux, l'autre non.

Dans un objectif scientifique, la trace de vos interactions avec Flash sera enregistrée, ainsi que votre activité visuelle (utilisation d'un `*eye tracker*').

Cette expérimentation préservera votre anonymat conformément à la loi 'Informatique et Liberté' (1978) modifiée (2001). Les informations sollicitées dans ce questionnaire nous serviront à préciser votre profil en tant qu'utilisateur de logiciels grand public.
Elles resteront confidentielles.

Age :

Nom: Téléphone: N°:

Prénom : E-mail :

CONNAISSANCES GENERALES

Dernier Diplôme obtenu :

Diplôme(s) préparé(s) au cours de l'année universitaire 2005-2006 :

EXPERIENCE INFORMATIQUE

Quels types d'équipement avez-vous déjà utilisés ?

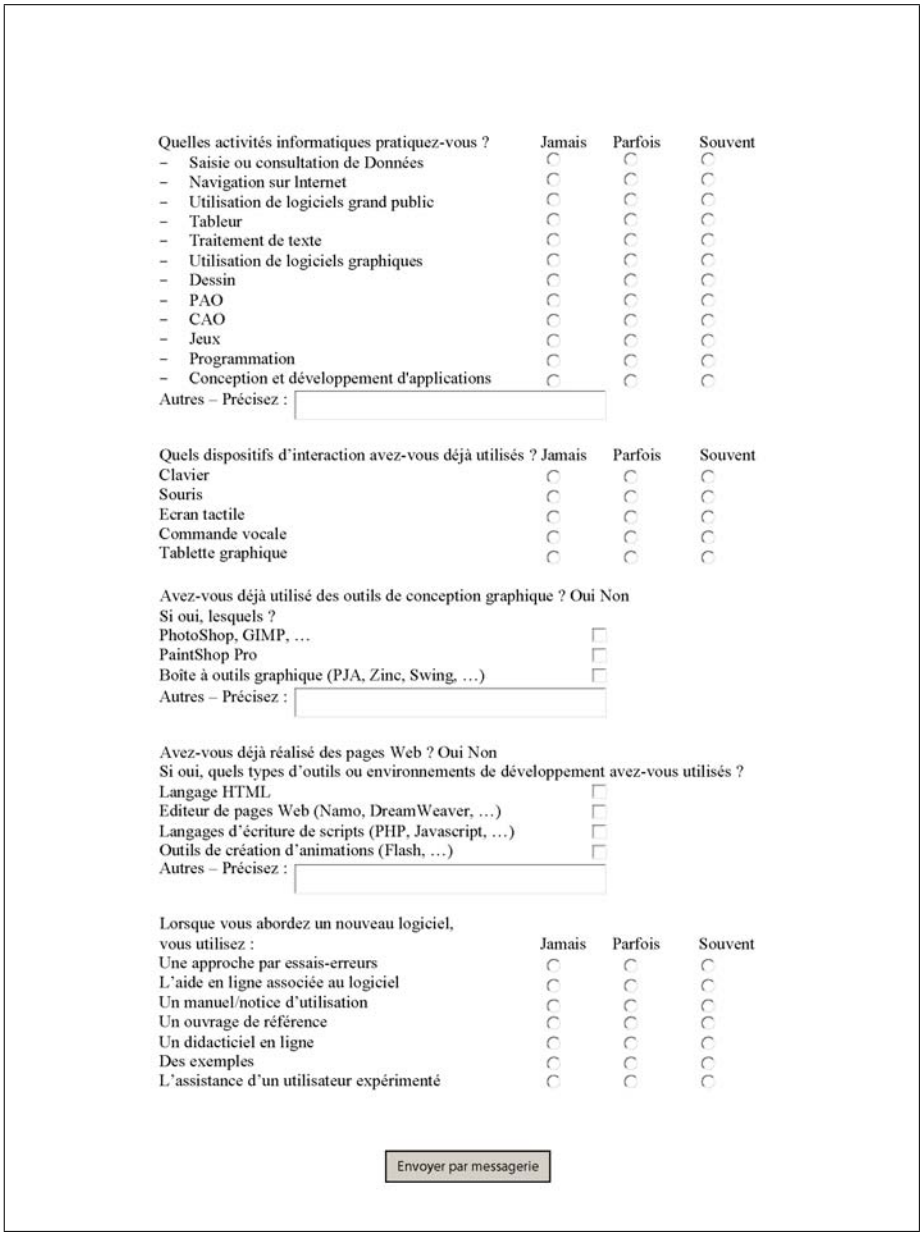
- PC ou MAC
- Palm Pilot
- TabletPC

Autres – Précisez :

Quels systèmes d'exploitation vous sont familiers ?

- Microsoft Windows
- Mac OS
- Linux, Unix

FIG. B.1 – Premier questionnaire : Statut du participant - Page 1/2.



Quelles activités informatiques pratiquez-vous ?	Jamais	Parfois	Souvent
– Saisie ou consultation de Données	○	○	○
– Navigation sur Internet	○	○	○
– Utilisation de logiciels grand public	○	○	○
– Tableur	○	○	○
– Traitement de texte	○	○	○
– Utilisation de logiciels graphiques	○	○	○
– Dessin	○	○	○
– PAO	○	○	○
– CAO	○	○	○
– Jeux	○	○	○
– Programmation	○	○	○
– Conception et développement d'applications	○	○	○

Autres – Précisez :

Quels dispositifs d'interaction avez-vous déjà utilisés ?	Jamais	Parfois	Souvent
Clavier	○	○	○
Souris	○	○	○
Ecran tactile	○	○	○
Commande vocale	○	○	○
Tablette graphique	○	○	○

Avez-vous déjà utilisé des outils de conception graphique ? Oui Non
Si oui, lesquels ?
PhotoShop, GIMP, … ☐
PaintShop Pro ☐
Boîte à outils graphique (PJA, Zinc, Swing, …) ☐
Autres – Précisez :

Avez-vous déjà réalisé des pages Web ? Oui Non
Si oui, quels types d'outils ou environnements de développement avez-vous utilisés ?
Langage HTML ☐
Editeur de pages Web (Namo, DreamWeaver, …) ☐
Langages d'écriture de scripts (PHP, Javascript, …) ☐
Outils de création d'animations (Flash, …) ☐
Autres – Précisez :

Lorsque vous abordez un nouveau logiciel, vous utilisez :	Jamais	Parfois	Souvent
Une approche par essais-erreurs	○	○	○
L'aide en ligne associée au logiciel	○	○	○
Un manuel/notice d'utilisation	○	○	○
Un ouvrage de référence	○	○	○
Un didacticiel en ligne	○	○	○
Des exemples	○	○	○
L'assistance d'un utilisateur expérimenté	○	○	○

Envoyer par messagerie

FIG. B.2 – Premier questionnaire : Statut du participant - Page 2/2.

Présentation de l'expérimentation

Les études 2 et 3 ont des présentations qui différent :

- "Étude ACA" : Présentation de l'expérimentation et des tâches à réaliser par l'ACA nommé "Eloona", puis lecture par le sujet du tutoriel multimédia en dehors de la présence de l'expérimentateur (15 minutes environ).
- "Étude Adaptation" : Présentation de l'expérimentation et des tâches à réaliser par l'expérimentateur, puis lecture du tutoriel multimédia par le sujet en dehors de la présence de l'expérimentateur (15 minutes environ).

Étude avec Eloona

Une présentation de Flash est proposée, sous la forme d'un diaporama PowerPoint qui permet au sujet de se familiariser avec les concepts de base de ce logiciel. Il peut naviguer dans la présentation à sa guise. Le temps de consultation est libre. « Vous disposez, pour interagir avec cette présentation, et ultérieurement avec Flash, des dispositifs classiques d'entrée, clavier et souris. »

La présentation se déroule ainsi :

- Eloona se présente, explique le fonctionnement du tutoriel et informe le sujet qu'elle va lancer le tutoriel.
- Ensuite, le tutoriel se lance ; l'utilisateur doit cliquer sur « Echap » dès qu'il a fini le tutoriel pour poursuivre la présentation.
- Le tutoriel achevé, Eloona présente l'environnement logiciel que le sujet va utiliser.
- Une fois la présentation de l'environnement terminée, Eloona montre les animations à réaliser.

1. Eloona se présente...

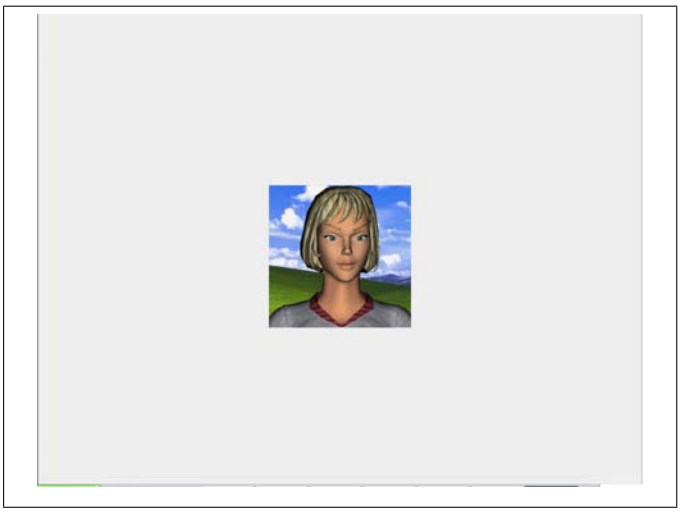


FIG. B.3 – Eloona se présente.

« Bonjour, je m'appelle Eloona et je vais vous accompagner pendant votre découverte du logiciel Flash. Vous aller vous familiariser avec les concepts de base de ce logiciel à l'aide d'un court tutoriel multimédia, puis l'utiliser pour créer trois animations très simples. Le tutoriel va se lancer automatiquement. Utilisez les flèches « droite » et « gauche » du clavier pour parcourir les diapositives. Lorsque vous aurez terminé, appuyer sur la touche « Echap ». A tout de suite... »

2. Le tutoriel est lancé... L'utilisateur navigue à l'aide du clavier ou de la souris.

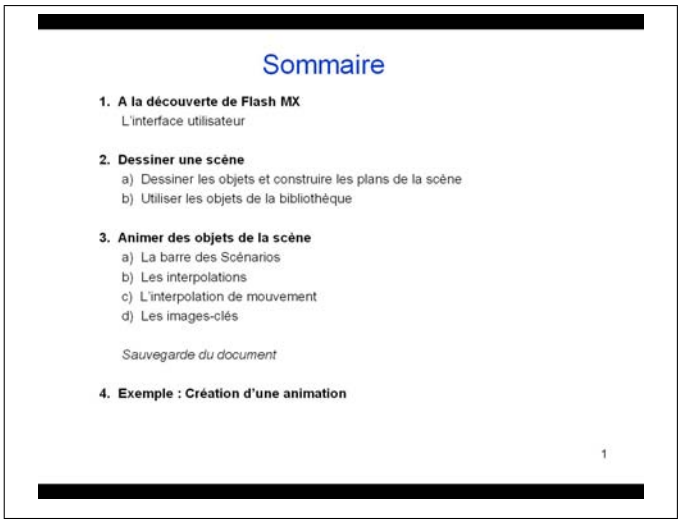


FIG. B.4 – Première diapositive du tutoriel (PowerPoint).

3. L'utilisateur a terminé la lecture du tutoriel. Eloona commence la présentation de l'environnement logiciel que le sujet va utiliser.



FIG. B.5 – Eloona débute la présentation de l'environnement logiciel expérimental.

A noter, la présence d'un distracteur visuel au centre de l'écran évite que le sujet focalise son attention visuelle uniquement sur l'ACA.

3.1. Environnement logiciel : Emplacement d'Eloona pendant l'expérimentation.

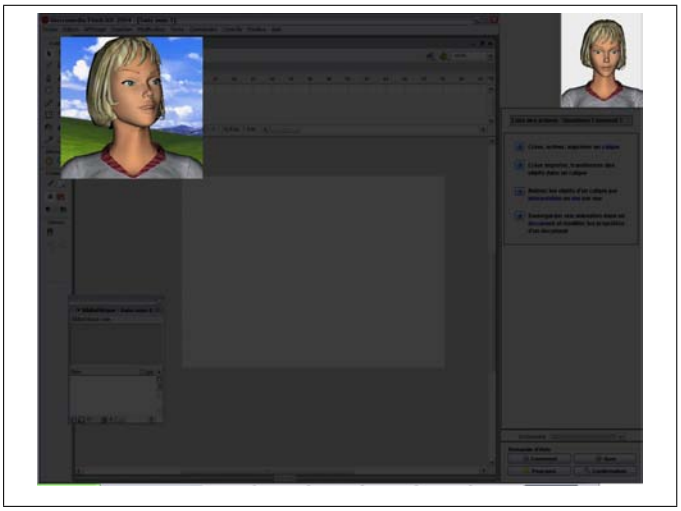


FIG. B.6 – Emplacement d'Eloona pendant l'expérimentation.

« C'est dans cette fenêtre que je serai pendant une partie de l'expérience. »

3.2. Environnement logiciel : Zone d'affichage des messages d'aide.

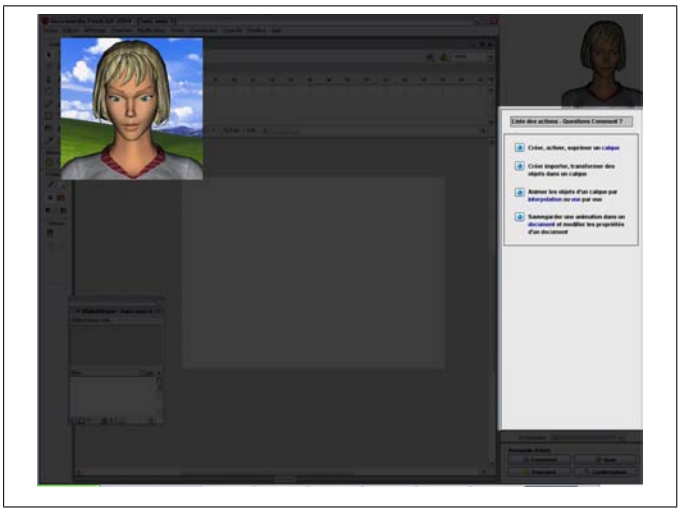


FIG. B.7 – Fenêtre d'affichage des messages d'aide.

« Ici, la fenêtre d'affichage des messages d'aide. Elle comprend :le lexique que vous utiliserez pour préciser le sujet de vos questions de type 'Comment ?' ; et le dictionnaire, dans lequel vous trouverez

de courtes définitions des termes techniques utilisés dans les messages d'aide. »

3.3 Environnement logiciel : Zone d'affichage du dictionnaire.

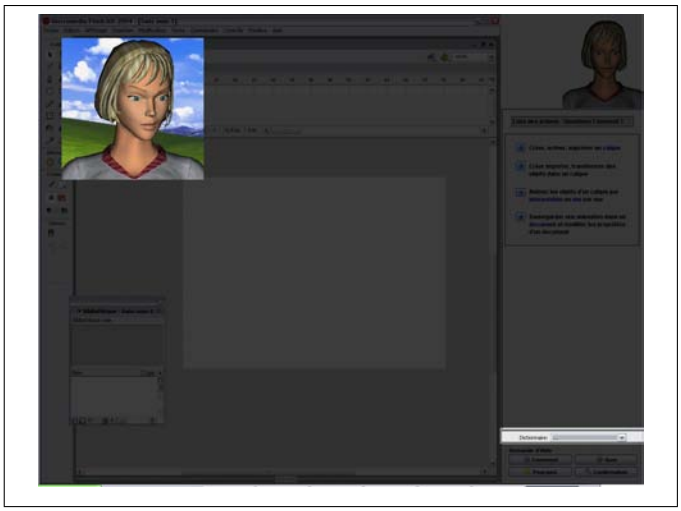


FIG. B.8 – Zone d'affichage du dictionnaire.

« Là, le dictionnaire, sous la forme d'un menu déroulant. Si vous ne comprenez pas l'un des termes techniques utilisés dans le message d'aide courant, vous obtiendrez une courte définition de ce terme en cliquant sur l'item correspondant dans le menu déroulant. »

3.4. Environnement logiciel : Zone d'interaction avec le système d'aide.

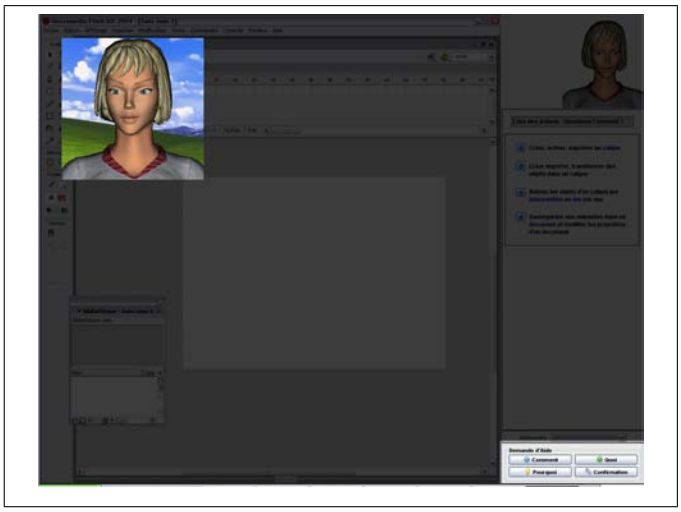


FIG. B.9 – Zone d'interaction avec l'aide (4 boutons).

« C'est à l'aide de ces quatre boutons que vous allez formuler vos demandes d'aide. Le bouton « Comment ? » permet d'afficher un lexique hiérarchique dans lequel vous pourrez naviguer jusqu'à

trouver le message d'aide qui vous permettra de réaliser l'action que vous souhaitez effectuer sur l'animation Flash en cours de création.
Le bouton « Quoi ? » vous permet d'obtenir des informations sur la fonction d'une icône ou le rôle d'un menu de Flash en cliquant sur l'objet graphique ou le menu avec le bouton du milieu de la souris. Si le résultat de la dernière action que vous avez effectuée sur l'animation en cours de création ne correspond pas à vos attentes, vous pouvez cliquer sur le bouton « Pourquoi ? ». Le système d'aide « défera » les actions erronées et éventuellement affichera dans la fenêtre d'aide un message approprié. Si vous n'êtes pas sûr que l'action ou la procédure que vous envisagez d'effectuer sur l'animation en cours de création est correcte, vous pouvez cliquer sur le bouton « Demande de confirmation ». Le système d'aide vous répondra par l'affirmative ou la négative. En cas de réponse négative, il pourra, éventuellement, afficher un message d'aide approprié ou défaire les actions antérieures erronées. »

3.5. Environnement logiciel : Fenêtre principale de Flash

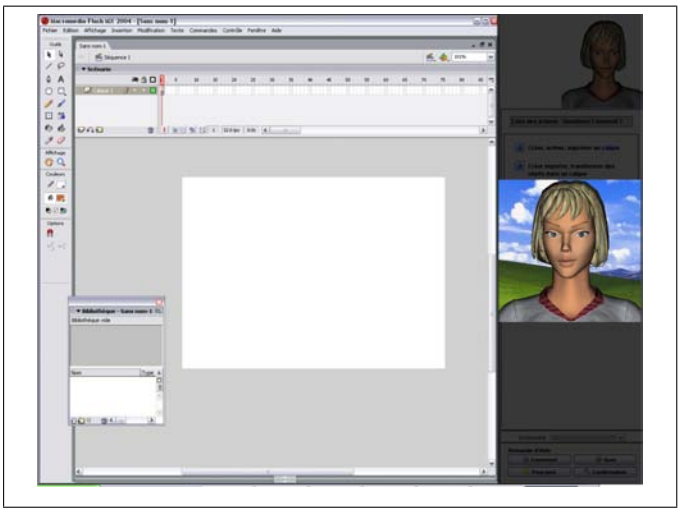


FIG. B.10 – La fenêtre principale de Flash.

« Ici, enfin, la fenêtre principale de Flash où vous allez créer vos animations. »

4. Ensuite, Eloona présente les scénarios à réaliser (même dis tracteur que précédemment).



FIG. B.11 – Eloona débute la présentation des scénarios.

« Parlons maintenant des animations à réaliser. »

Présentation des animations à réaliser. Puis Eloona salue le sujet qui appelle ensuite, par téléphone, l'expérimentateur.

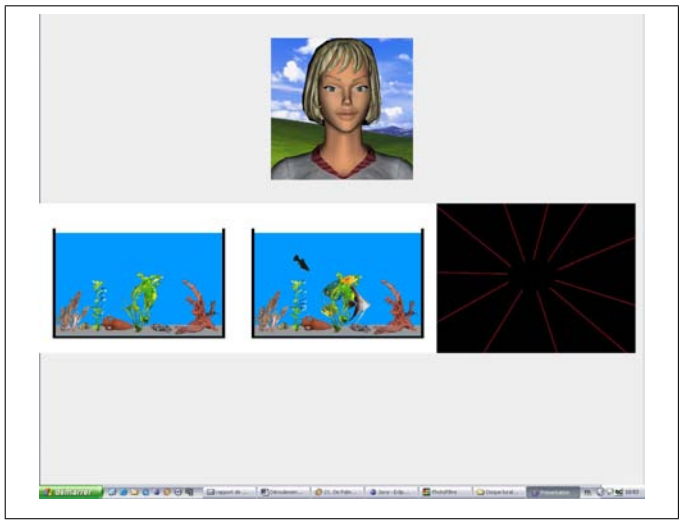


FIG. B.12 – Présentation des scénarios d'animation à réaliser.

« A gauche la première partie de l'animation 'Site Web', au centre la second partie de cette animation, et à droite l'animation 'Aquarium'. Si vous avez des questions, l'expérimentateur est à votre disposition pour y répondre. A bientôt ! »

Étude avec Adaptation - Schéma de présentation de l'expérimentation

Vous aller vous familiariser avec les concepts de base de Flash à l'aide d'un court tutorial multimédia. Ensuite, vous utiliserez ce logiciel pour créer trois animations très simples.
Le tutoriel va se lancer automatiquement. Utiliser les flèches " droite " et " gauche " du clavier pour parcourir les diapositives.
Lorsque vous aurez terminé, appuyer sur la touche " Echap " et appelez le 3083.
Vous serez assisté dans la réalisation des animations par un système d'aide en ligne innovant dont nous vous demanderons ensuite d'évaluer l'efficacité et la convivialité. Vous répondrez d'abord à un questionnaire d'évaluation. Vous préciserez ensuite vos réactions et votre jugement au cours d'un entretien avec l'organisateur de cette évaluation ergonomique.

L'expérimentateur quitte la salle.

Retour de l'expérimentateur dans la salle :
Souhaitez-vous des explications complémentaires sur le contenu du tutoriel ? Sinon, je vais vous présenter l'environnement logiciel que vous allez utiliser.

[L'expérimentateur affiche une fenêtre Flash et une fenêtre d'aide et explique le fonctionnement du système d'aide.]
L'écran est divisé en deux parties :
* A gauche, la fenêtre de Flash dans laquelle vous allez réaliser les scénarios d'animation.
* A droite, la fenêtre du système d'aide que vous pourrez interroger en cas de besoin.

Quatre boutons, en bas de la fenêtre où s'affichent les messages d'aide, vous aideront à formuler vos demandes d'informations au système d'aide.
Le bouton " Comment " permet d'afficher un lexique hiérarchique dans lequel vous pourrez naviguer jusqu'à trouver le message d'aide qui vous assistera dans la réalisation de l'action que vous souhaitez effectuer sur l'animation Flash en cours de création. En cliquant sur ce bouton, vous afficherez dans la fenêtre d'aide un lexique des actions qu'il est possible d'effectuer sur l'animation en cours de création. Ce lexique est organisé hiérarchiquement. Vous trouverez les messages d'aide de type " Comment " au niveau hiérarchique le plus bas. Pour descendre dans la hiérarchie, utiliser les boutons à gauche de chaque item. Pour remonter d'un niveau, utiliser le bouton rouge situé en haut à gauche de la fenêtre du lexique. [Insister oralement sur le bouton qui permet de remonter dans l'arborescence du lexique.]
Le bouton " Quoi " vous permet d'obtenir des informations sur la fonction d'une icône ou le rôle d'un menu de Flash. Il vous suffit de cliquer sur l'objet graphique ou le menu avec le bouton du milieu de la souris pour obtenir les informations associées à celui-ci.
Si le résultat de la dernière action que vous avez effectuée sur l'animation en cours de création ne correspond pas à vos attentes, vous pouvez cliquer sur le bouton " Pourquoi ". Le système d'aide " défera " les actions erronées et, éventuellement, affichera dans la fenêtre d'aide un message approprié.
Si vous n'êtes pas sûr que l'action ou la procédure que vous envisagez d'effectuer sur l'animation

en cours de création est correcte, vous pouvez cliquer sur le bouton " Demande de confirmation ". Le système d'aide vous répondra par l'affirmative ou la négative. En cas de réponse négative, il pourra, éventuellement, afficher un message d'aide approprié ou défaire les actions antérieures erronées.

Les mots soulignés dans les messages d'aide représentent des liens vers un " dictionnaire " des termes techniques utilisés dans les messages. En cliquant sur un de ces liens vous obtiendrez une courte définition, éventuellement illustrée, du terme correspondant. Cette définition s'affichera dans la fenêtre d'aide en recouvrant le message contenant le lien. Pour faire réapparaître ce message, il suffit de fermer le message du dictionnaire en cliquant sur le bouton rouge à droite du titre du message du dictionnaire.
Vous pouvez agrandir les copies d'écran incluses dans les messages en cliquant sur elles. La copie d'écran apparaît, agrandie, dans une fenêtre dédiée qui recouvre partiellement la fenêtre Flash.

Parlons maintenant des animations à réaliser.
A gauche, le résultat du premier scénario qui correspond à la première partie de l'animation " Aquarium ", au centre, le résultat du second scénario, qui consiste à terminer l'animation " Aquarium ", à droite, le résultat de l'animation " Site Web ".

Je suis à votre disposition pour répondre à vos questions ou vous aider en cas de problème. Appelez le 3083.

Après avoir répondu aux questions éventuelles du sujet, l'expérimentateur affiche le résultat des animations à effectuer sur l'ordinateur portable présent à côté de l'écran, puis lance la plateforme avant de quitter la salle

Installation et calibrage de l'eye-tracker par l'expérimentateur

Rappel :
"Étude avec ACA" : Installation et calibrage de l'eye-tracker par l'expérimentateur (10 minutes environ).
"Étude avec Adaptation" : Eye tracker, non utilisé.

Installation de l'eye-tracker et calibrage
L'eye-tracker dont nous disposons est le modèle 501 de chez ASL (fréquence d'échantillonnage utilisée : 60 Hz). Le dispositif est fixé sur un casque que l'utilisateur doit porter sur la tête. Une source infrarouge illumine l'oeil dont l'image est réfléchie sur un miroir translucide, le reflet est alors enregistré par une caméra fixée sur le casque. C'est donc à partir de la pupille et du reflet cornéen que le logiciel calcule la position du regard sur l'écran. Une autre caméra, appelée caméra de scène, peut enregistrer la scène regardée par le sujet. Au cas particulier, l'utilisation de cette caméra est inutile puisque la plate-forme logicielle que nous avons développée sauvegarde une trace des interactions qui permet un « rejeu » analogue de l'évolution de l'affichage sur l'écran du sujet. Un capteur head-tracker magnétique complète l'équipement du casque. Grâce à une source magnétique fixe, disposée dans la même salle que l'eye-tracker, on connaît la position instantanée de la tête du sujet par rapport à cette source. Le calibrage de ce dispositif consiste à déterminer à mesurer par laser la distance entre la source magnétique et l'écran du sujet.

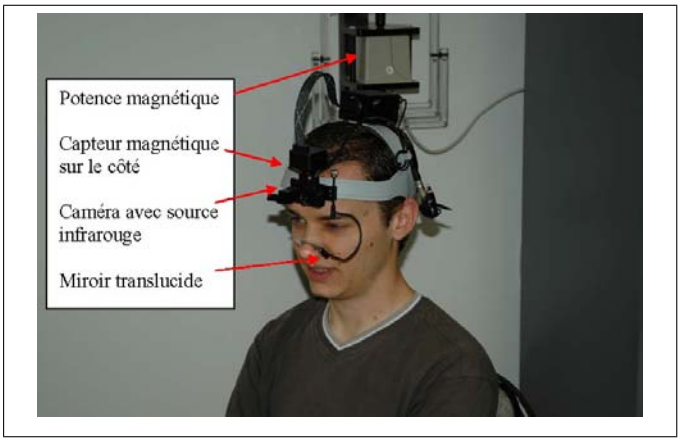


FIG. B.13 – Installation du casque, réglage du miroir de façon à pouvoir détecter la pupille quel que soit le point de l'écran regardé.

A noter que l'on récupère sur le port série du poste de travail du sujet les trames en provenance de l'eye-tracker ; une trame contient les coordonnées x, y (dans le plan de l'écran) du point d'intersection entre la direction du regard et l'écran à un instant donné ou « point de regard » (gaze point), ainsi que la distance entre le capteur magnétique et l'écran du sujet. Les coordonnées des points de regard sont synchronisées avec les événements Windows (souris, système et clavier) générés par l'interaction du sujet avec Flash et l'aide. Ces informations sont sauvegardées sur le poste du sujet.

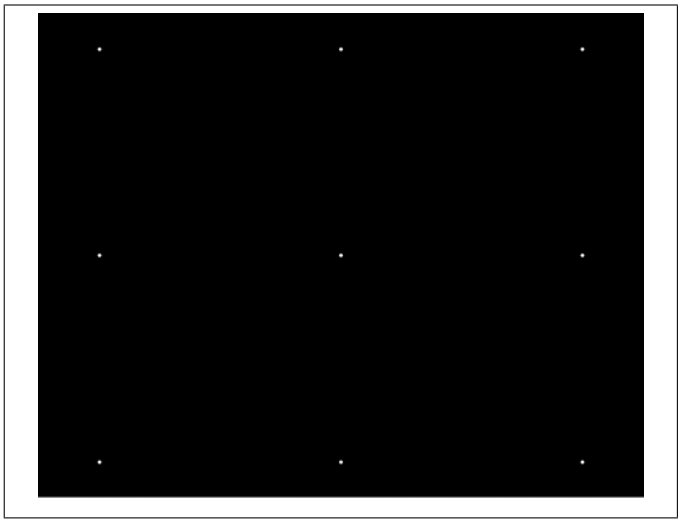


FIG. B.14 – Affichage de la mire à 9 points utilisée pour calibrer l'eye-tracker.

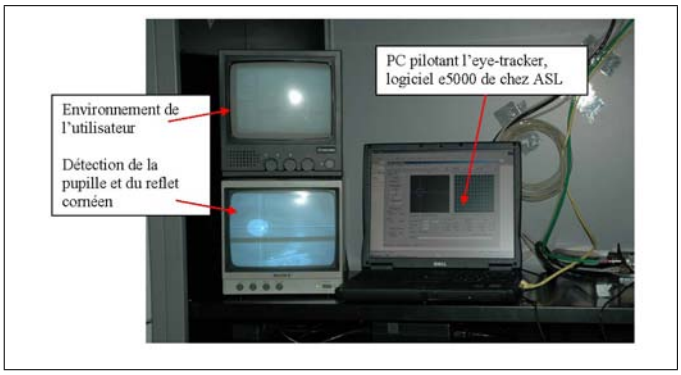


FIG. B.15 – Équipement de pilotage de l'eye-tracker.

Questionnaires d'évaluation et post-test

Afin de compléter les données obtenues au cours de l'étape d'expérimentation, nous avons demandé aux sujets de donner leurs impressions et jugements sur les deux types de présentation des messages d'aide, par écrit et de manière anonyme, en remplissant :

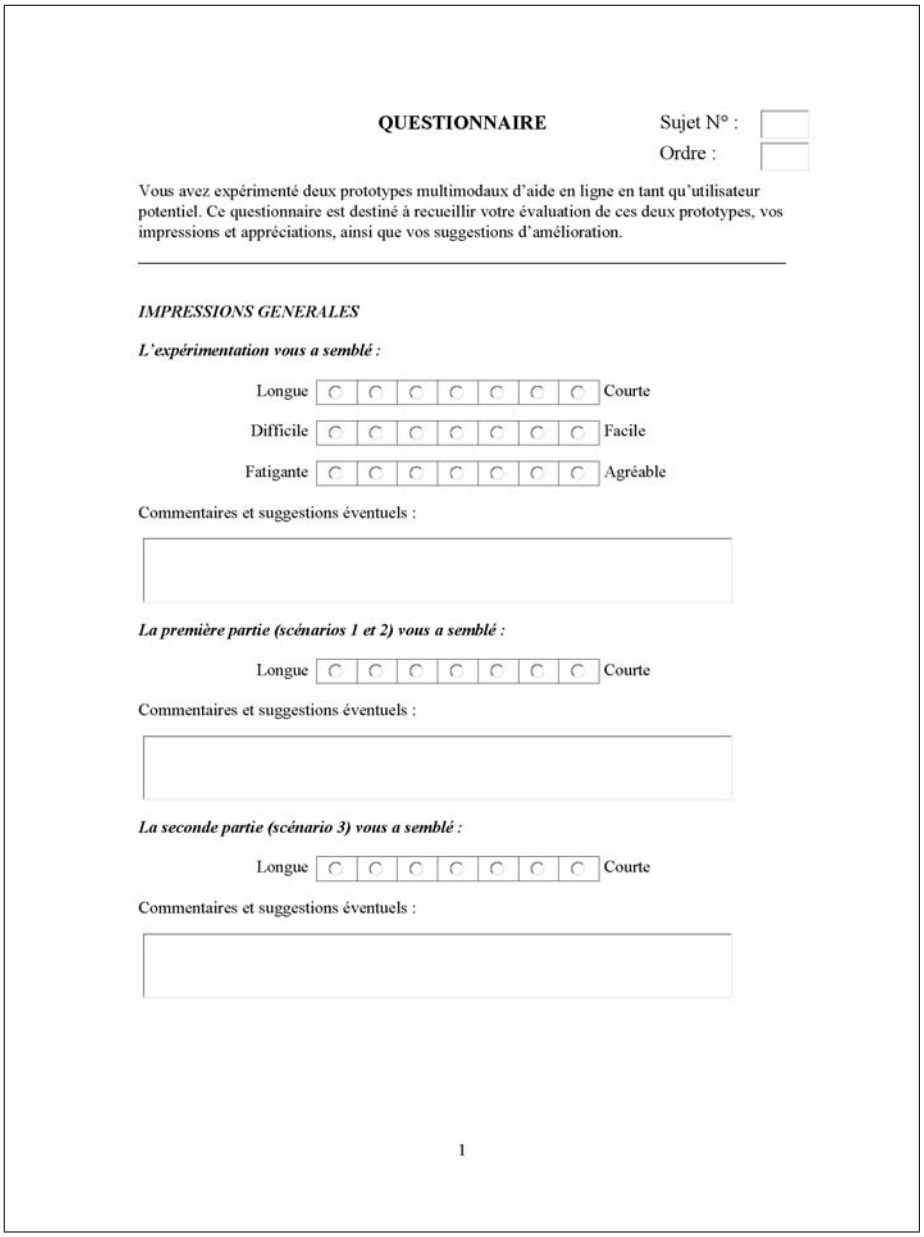
- un questionnaire verbal d'évaluation qui diffère d'une étude à l'autre,
- un questionnaire non verbal (SAM) qui diffère également d'une étude à l'autre, et
- un questionnaire destiné à évaluer les concepts et savoir-faire acquis au cours de l'expérimentation (post-test), commun aux deux études.

Complément : Le questionnaire non verbal SAM est un outil basé sur la présentation d'un personnage doté de différentes émotions et réactions affectives dans le but d'évaluer les émotions d'un participant. Ainsi, trois émotions se déclinent en cinq niveaux d'intensité (voir les figures de la page suivante), nous décrivons ici textuellement les extrêmes (niveau 1 et niveau 5) de chaque émotion :

1. Satisfait / Content / Heureux (niveau 1) à Insatisfait / Mécontent / Irrité (niveau 5) ;
2. Excité / Stimulé / Intéressé (niveau 1) à Calme Amorphe / Indifférent (niveau 5) ;
3. Doutant de soi / Timide / Dominé / Inférieur (niveau 1) à Sûr de soi / Assuré / Dominant / Supérieur (niveau 5) .

Afin de recueillir le ressenti et les réactions des participants face aux innovations auxquelles ils ont été confronté, nous leur posons diverses questions et pour chaque question, le participant doit choisir une vignette associée à un niveau de chaque émotion.

Questionnaire verbal pour l'étude avec Eloona



QUESTIONNAIRE Sujet N° : Ordre :

Vous avez expérimenté deux prototypes multimodaux d'aide en ligne en tant qu'utilisateur potentiel. Ce questionnaire est destiné à recueillir votre évaluation de ces deux prototypes, vos impressions et appréciations, ainsi que vos suggestions d'amélioration.

IMPRESSIONS GENERALES

L'expérimentation vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Fatigante ○ ○ ○ ○ ○ ○ ○ Agréable

Commentaires et suggestions éventuels :

La première partie (scénarios 1 et 2) vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

Commentaires et suggestions éventuels :

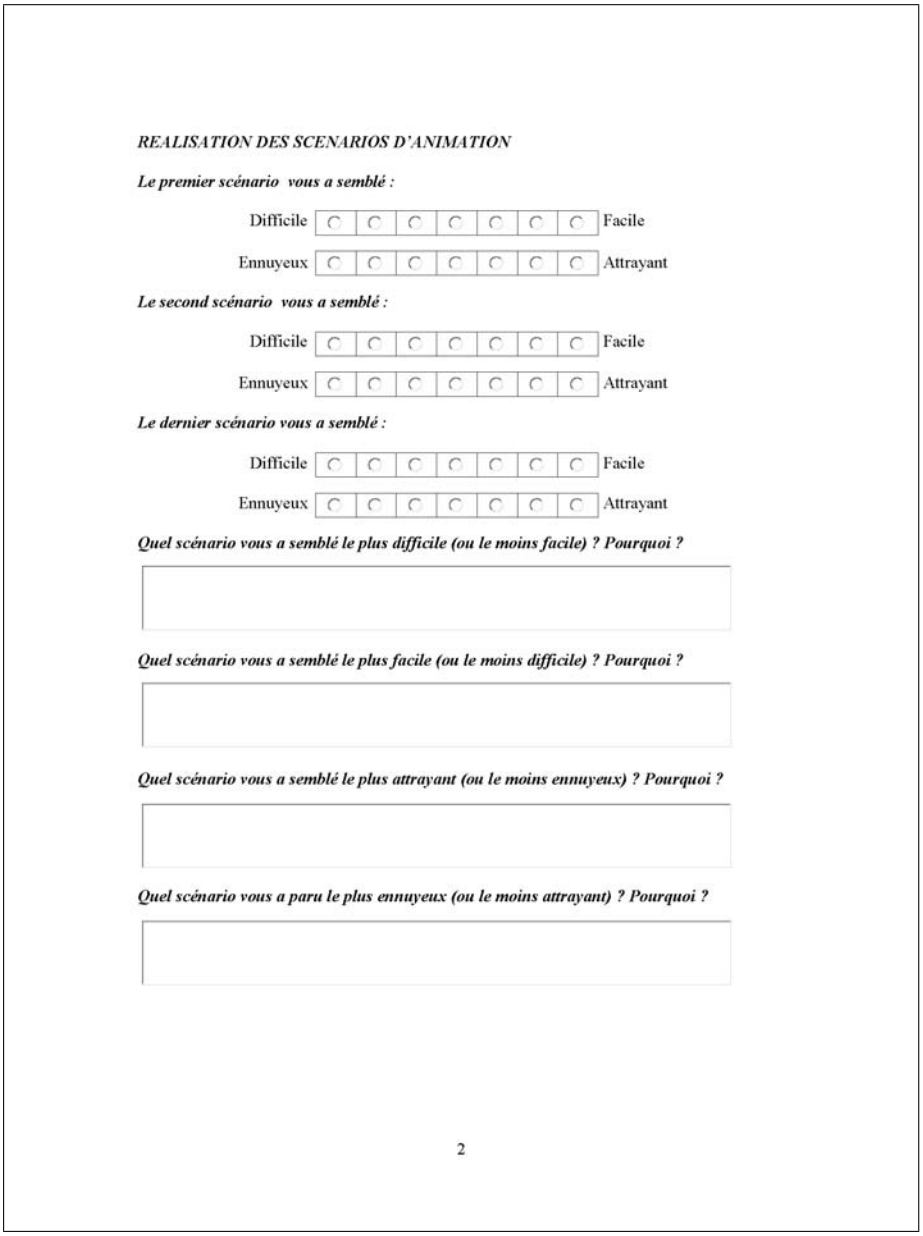
La seconde partie (scénario 3) vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

Commentaires et suggestions éventuels :

1

FIG. B.16 – Questionnaire verbal : Étude avec ACA - Page 1/10.



REALISATION DES SCENARIOS D'ANIMATION

Le premier scénario vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Le second scénario vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Le dernier scénario vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Quel scénario vous a semblé le plus difficile (ou le moins facile) ? Pourquoi ?

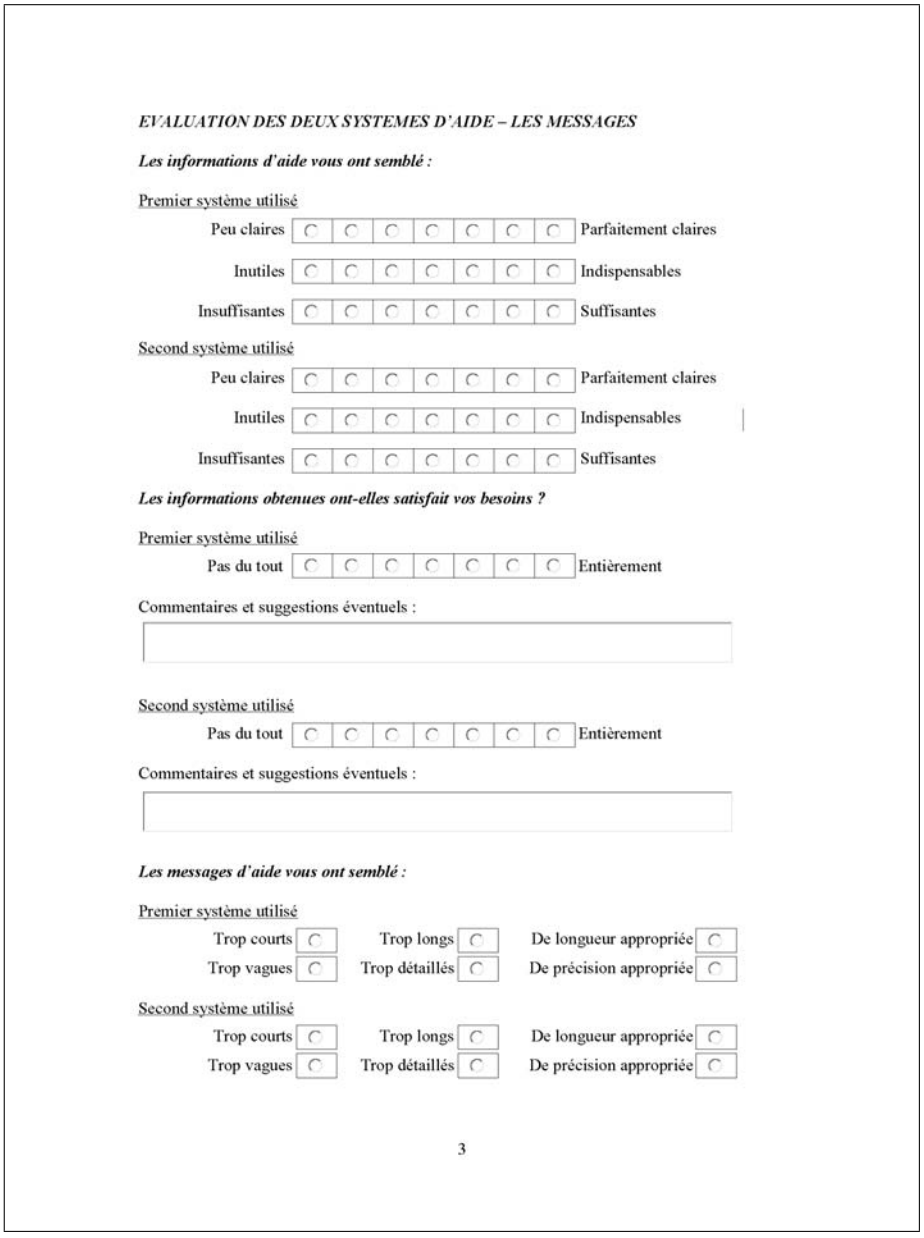
Quel scénario vous a semblé le plus facile (ou le moins difficile) ? Pourquoi ?

Quel scénario vous a semblé le plus attrayant (ou le moins ennuyeux) ? Pourquoi ?

Quel scénario vous a paru le plus ennuyeux (ou le moins attrayant) ? Pourquoi ?

2

FIG. B.17 – Questionnaire verbal : Étude avec ACA - Page 2/10.



EVALUATION DES DEUX SYSTEMES D'AIDE – LES MESSAGES

Les informations d'aide vous ont semblé :

Premier système utilisé

Peu claires ○ ○ ○ ○ ○ ○ ○ Parfaitement claires

Inutiles ○ ○ ○ ○ ○ ○ ○ Indispensables

Insuffisantes ○ ○ ○ ○ ○ ○ ○ Suffisantes

Second système utilisé

Peu claires ○ ○ ○ ○ ○ ○ ○ Parfaitement claires

Inutiles ○ ○ ○ ○ ○ ○ ○ Indispensables

Insuffisantes ○ ○ ○ ○ ○ ○ ○ Suffisantes

Les informations obtenues ont-elles satisfait vos besoins ?

Premier système utilisé

Pas du tout ○ ○ ○ ○ ○ ○ ○ Entièrement

Commentaires et suggestions éventuels :

Second système utilisé

Pas du tout ○ ○ ○ ○ ○ ○ ○ Entièrement

Commentaires et suggestions éventuels :

Les messages d'aide vous ont semblé :

Premier système utilisé

Trop courts ○ Trop longs ○ De longueur appropriée ○

Trop vagues ○ Trop détaillés ○ De précision appropriée ○

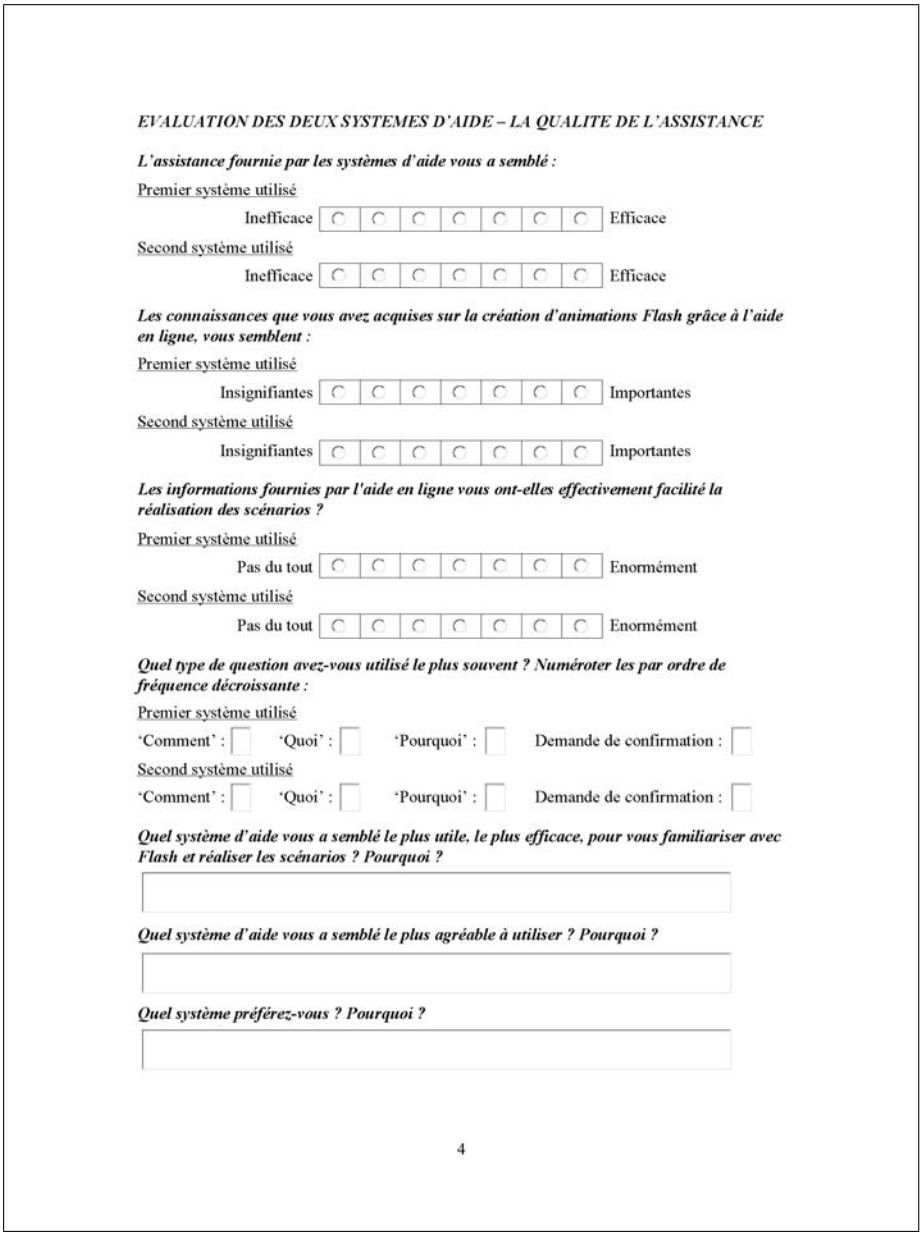
Second système utilisé

Trop courts ○ Trop longs ○ De longueur appropriée ○

Trop vagues ○ Trop détaillés ○ De précision appropriée ○

3

FIG. B.18 – Questionnaire verbal : Étude avec ACA - Page 3/10.



EVALUATION DES DEUX SYSTEMES D'AIDE – LA QUALITE DE L'ASSISTANCE

L'assistance fournie par les systèmes d'aide vous a semblé :

Premier système utilisé

Inefficace ○ ○ ○ ○ ○ ○ ○ Efficace

Second système utilisé

Inefficace ○ ○ ○ ○ ○ ○ ○ Efficace

Les connaissances que vous avez acquises sur la création d'animations Flash grâce à l'aide en ligne, vous semblent :

Premier système utilisé

Insignifiantes ○ ○ ○ ○ ○ ○ ○ Importantes

Second système utilisé

Insignifiantes ○ ○ ○ ○ ○ ○ ○ Importantes

Les informations fournies par l'aide en ligne vous ont-elles effectivement facilité la réalisation des scénarios ?

Premier système utilisé

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Second système utilisé

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Quel type de question avez-vous utilisé le plus souvent ? Numéroter les par ordre de fréquence décroissante :

Premier système utilisé

'Comment' : ☐ 'Quoi' : ☐ 'Pourquoi' : ☐ Demande de confirmation : ☐

Second système utilisé

'Comment' : ☐ 'Quoi' : ☐ 'Pourquoi' : ☐ Demande de confirmation : ☐

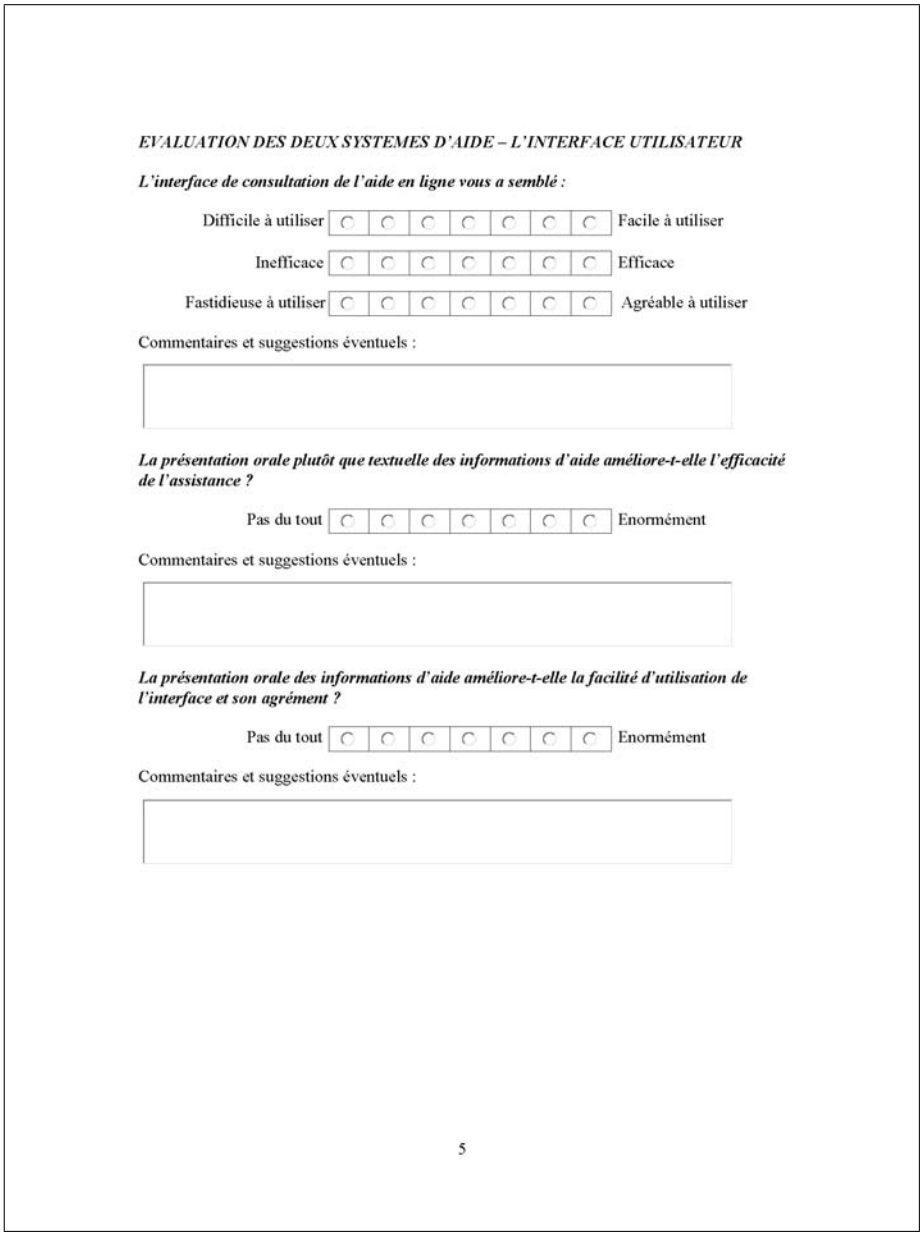
Quel système d'aide vous a semblé le plus utile, le plus efficace, pour vous familiariser avec Flash et réaliser les scénarios ? Pourquoi ?

Quel système d'aide vous a semblé le plus agréable à utiliser ? Pourquoi ?

Quel système préférez-vous ? Pourquoi ?

4

FIG. B.19 – Questionnaire verbal : Étude avec ACA - Page 4/10.



EVALUATION DES DEUX SYSTEMES D'AIDE – L'INTERFACE UTILISATEUR

L'interface de consultation de l'aide en ligne vous a semblé :

Difficile à utiliser ○ ○ ○ ○ ○ ○ ○ Facile à utiliser

Inefficace ○ ○ ○ ○ ○ ○ ○ Efficace

Fastidieuse à utiliser ○ ○ ○ ○ ○ ○ ○ Agréable à utiliser

Commentaires et suggestions éventuels :

La présentation orale plutôt que textuelle des informations d'aide améliore-t-elle l'efficacité de l'assistance ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Commentaires et suggestions éventuels :

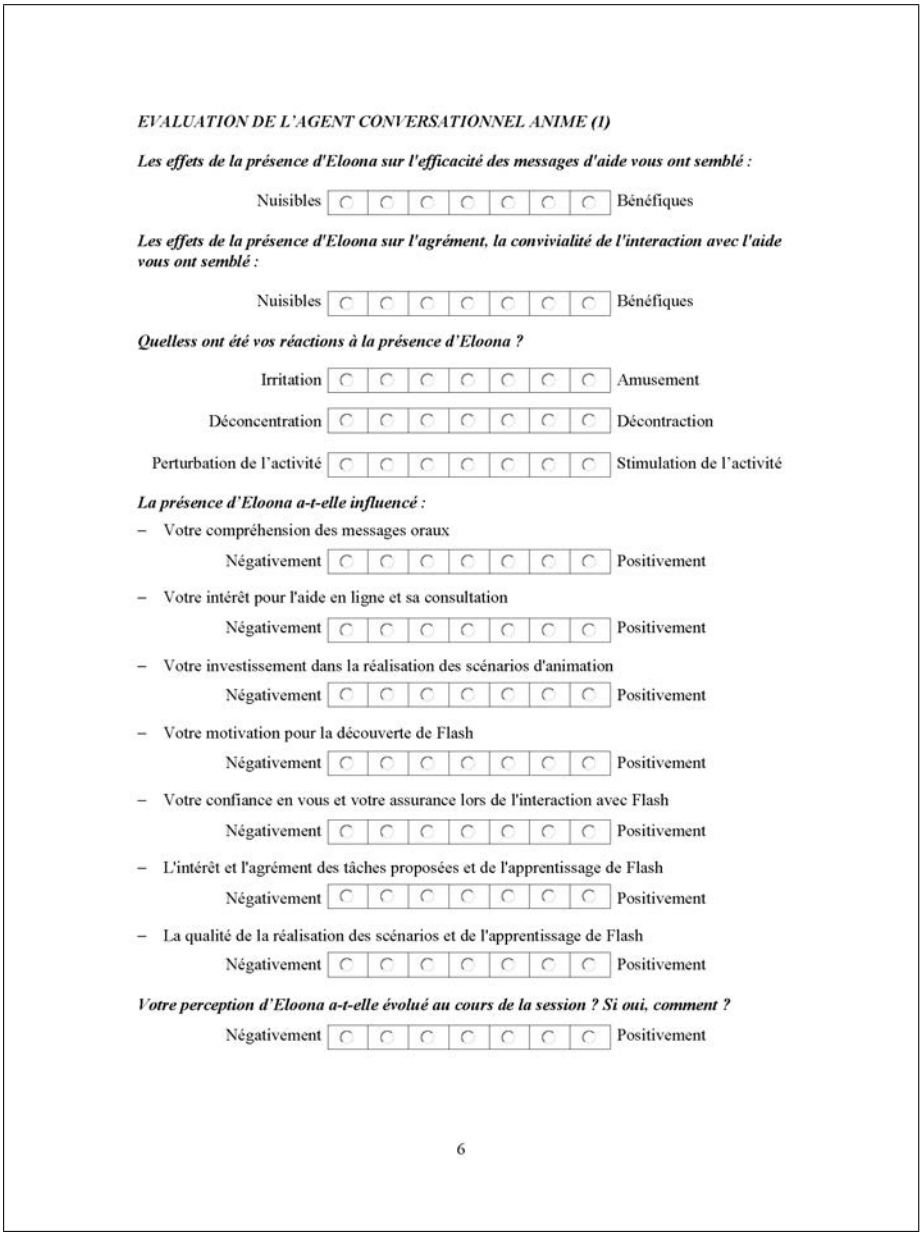
La présentation orale des informations d'aide améliore-t-elle la facilité d'utilisation de l'interface et son agrément ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Commentaires et suggestions éventuels :

5

FIG. B.20 – Questionnaire verbal : Étude avec ACA - Page 5/10.



EVALUATION DE L'AGENT CONVERSATIONNEL ANIME (1)

Les effets de la présence d'Eloona sur l'efficacité des messages d'aide vous ont semblé :

Nuisibles ○ ○ ○ ○ ○ ○ ○ Bénéfiques

Les effets de la présence d'Eloona sur l'agrément, la convivialité de l'interaction avec l'aide vous ont semblé :

Nuisibles ○ ○ ○ ○ ○ ○ ○ Bénéfiques

Quelless ont été vos réactions à la présence d'Eloona ?

Irritation ○ ○ ○ ○ ○ ○ ○ Amusement

Déconcentration ○ ○ ○ ○ ○ ○ ○ Décontraction

Perturbation de l'activité ○ ○ ○ ○ ○ ○ ○ Stimulation de l'activité

La présence d'Eloona a-t-elle influencé :

- Votre compréhension des messages oraux

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- Votre intérêt pour l'aide en ligne et sa consultation

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- Votre investissement dans la réalisation des scénarios d'animation

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- Votre motivation pour la découverte de Flash

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- Votre confiance en vous et votre assurance lors de l'interaction avec Flash

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- L'intérêt et l'agrément des tâches proposées et de l'apprentissage de Flash

  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

- La qualité de la réalisation des scénarios et de l'apprentissage de Flash

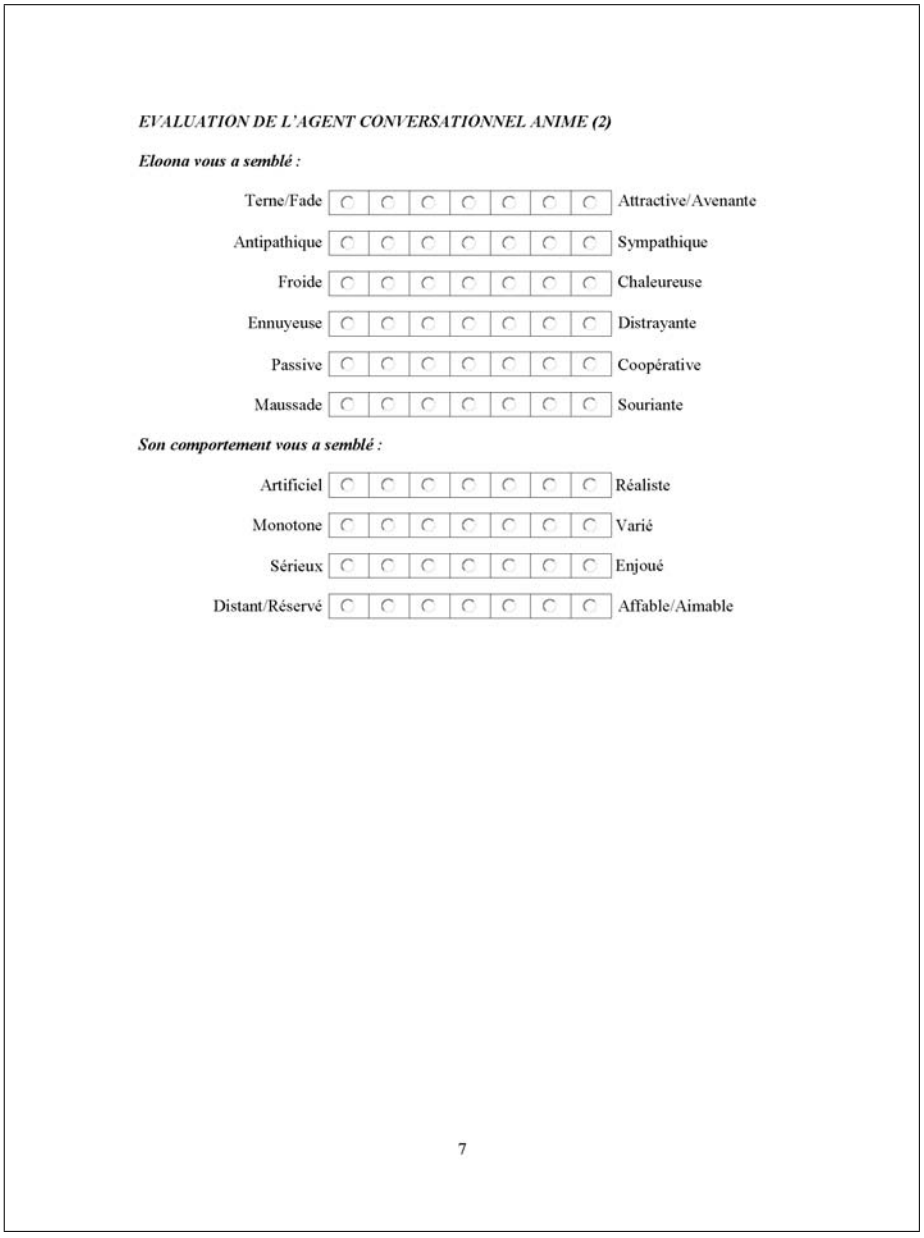
  Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

Votre perception d'Eloona a-t-elle évolué au cours de la session ? Si oui, comment ?

Négativement ○ ○ ○ ○ ○ ○ ○ Positivement

6

FIG. B.21 – Questionnaire verbal : Étude avec ACA - Page 6/10.



EVALUATION DE L'AGENT CONVERSATIONNEL ANIME (2)

Eloona vous a semblé :

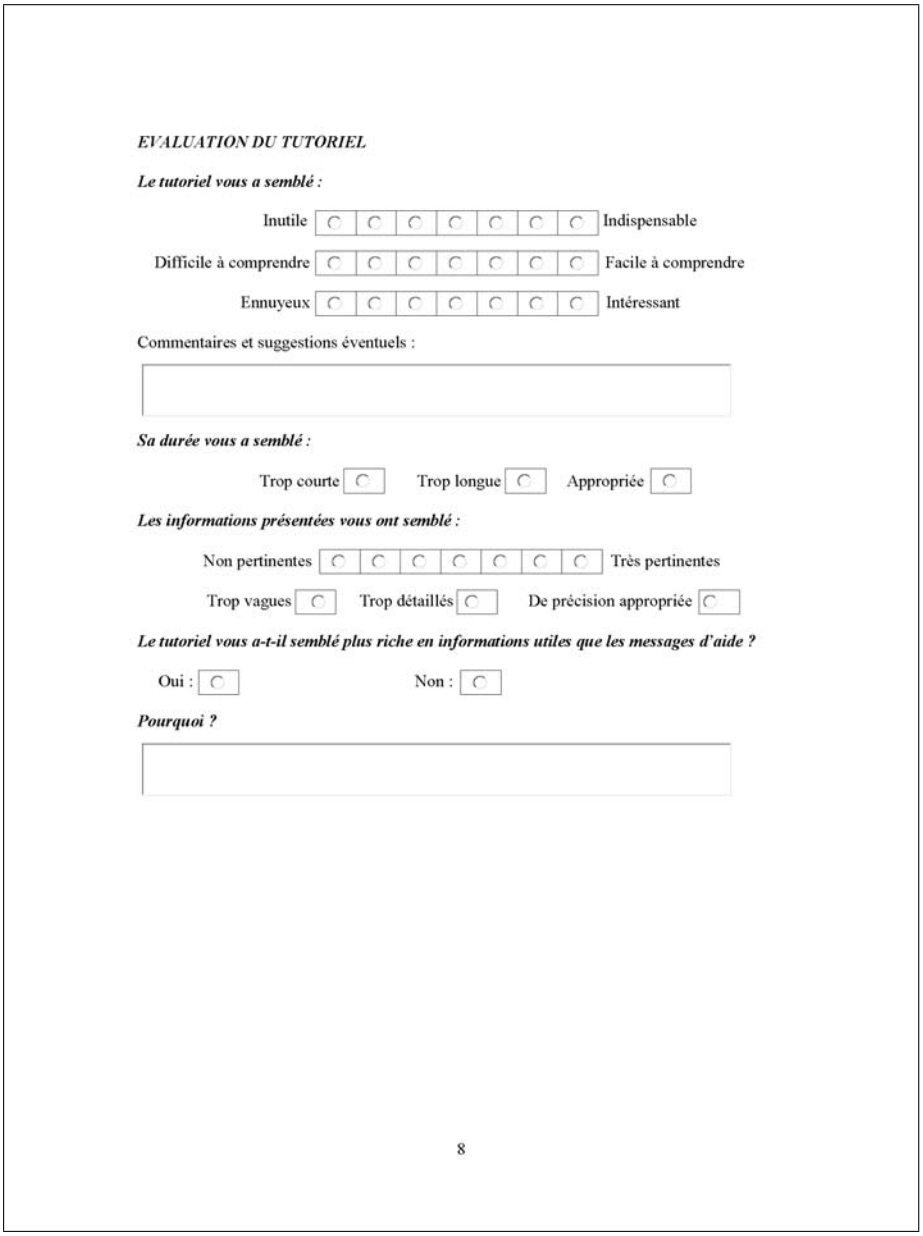
Terne/Fade	○ ○ ○ ○ ○ ○ ○	Attractive/Avenante
Antipathique	○ ○ ○ ○ ○ ○ ○	Sympathique
Froide	○ ○ ○ ○ ○ ○ ○	Chaleureuse
Ennuyeuse	○ ○ ○ ○ ○ ○ ○	Distrayante
Passive	○ ○ ○ ○ ○ ○ ○	Coopérative
Maussade	○ ○ ○ ○ ○ ○ ○	Souriante

Son comportement vous a semblé :

Artificiel	○ ○ ○ ○ ○ ○ ○	Réaliste
Monotone	○ ○ ○ ○ ○ ○ ○	Varié
Sérieux	○ ○ ○ ○ ○ ○ ○	Enjoué
Distant/Réservé	○ ○ ○ ○ ○ ○ ○	Affable/Aimable

7

FIG. B.22 – Questionnaire verbal : Étude avec ACA - Page 7/10.



EVALUATION DU TUTORIEL

Le tutoriel vous a semblé :

Inutile ○ ○ ○ ○ ○ ○ ○ Indispensable

Difficile à comprendre ○ ○ ○ ○ ○ ○ ○ Facile à comprendre

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Intéressant

Commentaires et suggestions éventuels :

Sa durée vous a semblé :

Trop courte ○ Trop longue ○ Appropriée ○

Les informations présentées vous ont semblé :

Non pertinentes ○ ○ ○ ○ ○ ○ ○ Très pertinentes

Trop vagues ○ Trop détaillés ○ De précision appropriée ○

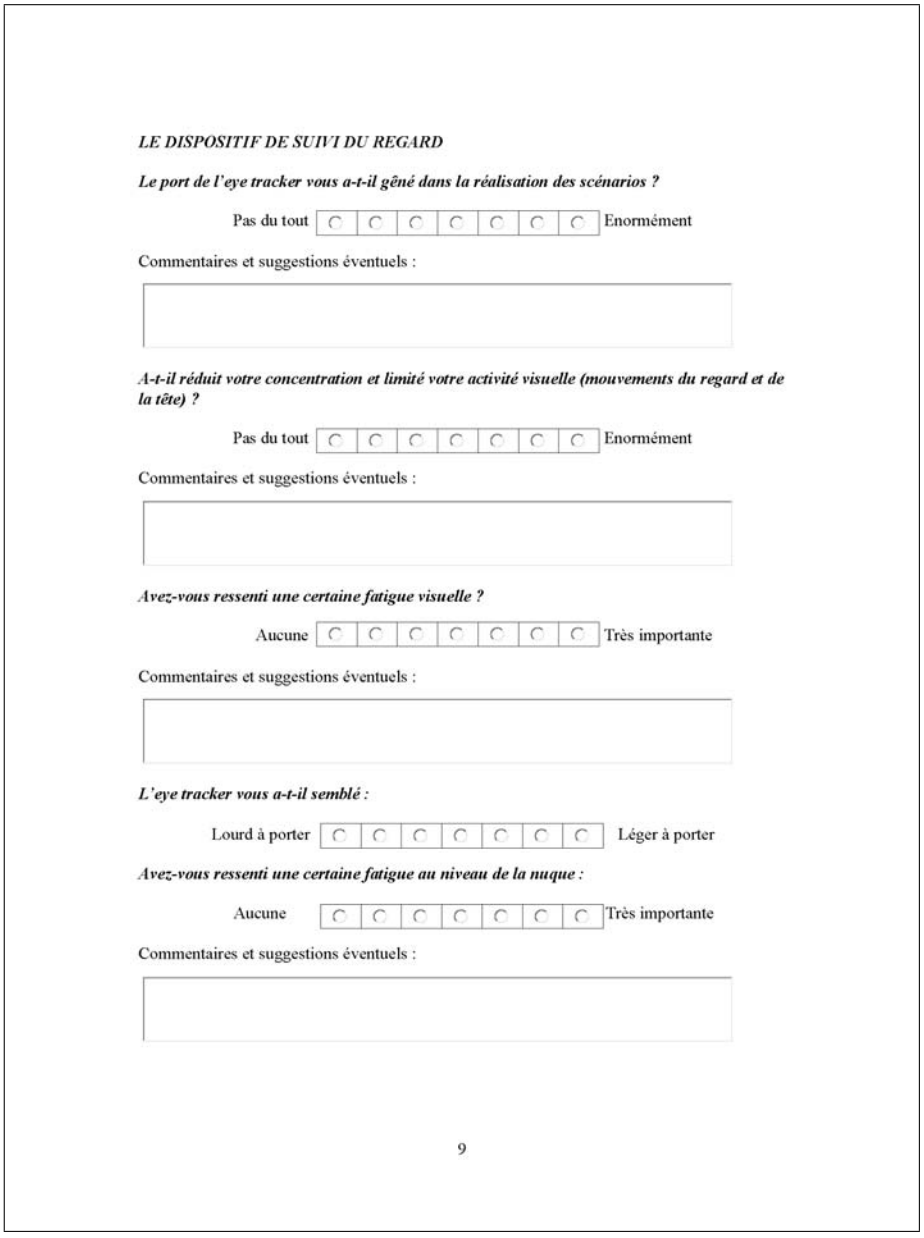
Le tutoriel vous a-t-il semblé plus riche en informations utiles que les messages d'aide ?

Oui : ○ Non : ○

Pourquoi ?

8

FIG. B.23 – Questionnaire verbal : Étude avec ACA - Page 8/10.



LE DISPOSITIF DE SUIVI DU REGARD

Le port de l'eye tracker vous a-t-il gêné dans la réalisation des scénarios ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Commentaires et suggestions éventuels :

A-t-il réduit votre concentration et limité votre activité visuelle (mouvements du regard et de la tête) ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Commentaires et suggestions éventuels :

Avez-vous ressenti une certaine fatigue visuelle ?

Aucune ○ ○ ○ ○ ○ ○ ○ Très importante

Commentaires et suggestions éventuels :

L'eye tracker vous a-t-il semblé :

Lourd à porter ○ ○ ○ ○ ○ ○ ○ Léger à porter

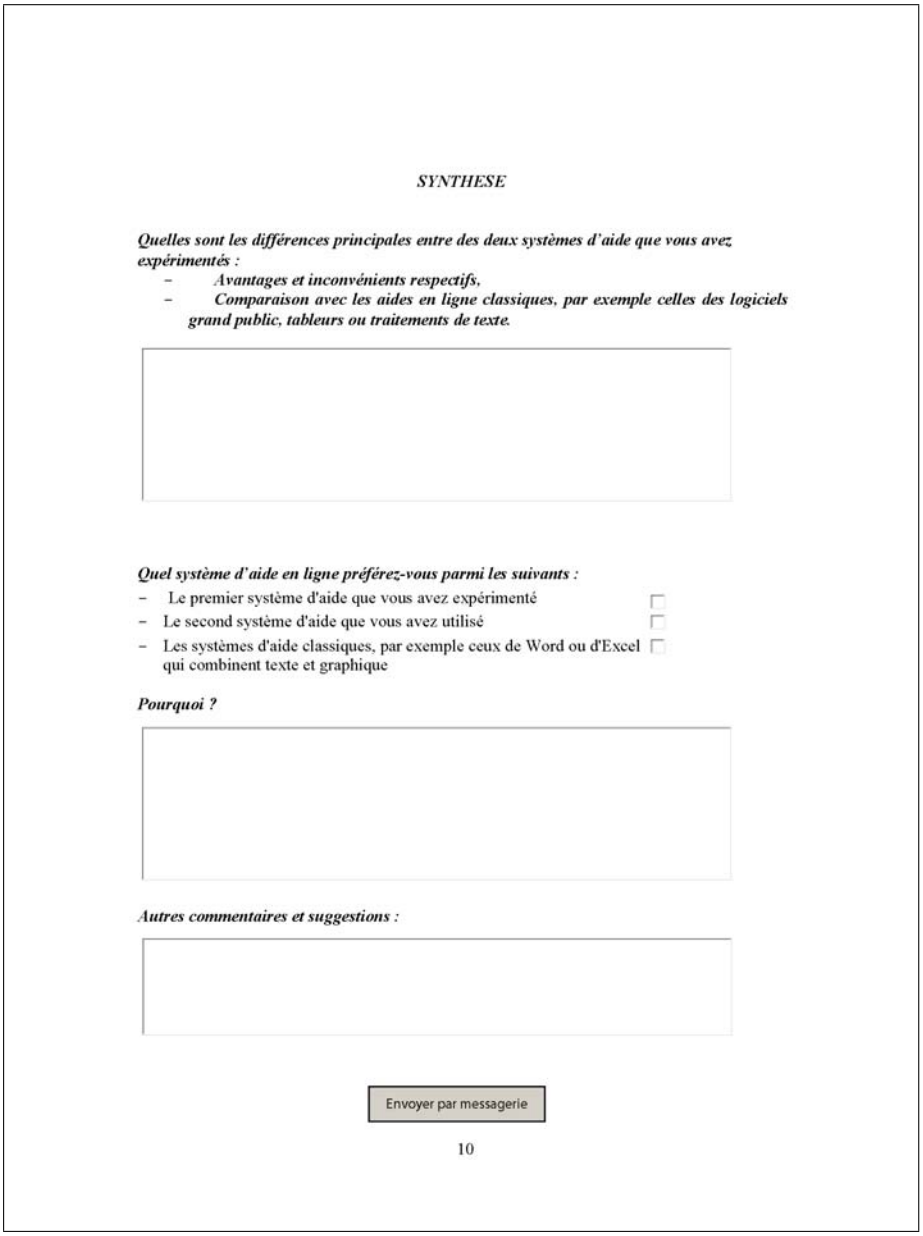
Avez-vous ressenti une certaine fatigue au niveau de la nuque :

Aucune ○ ○ ○ ○ ○ ○ ○ Très importante

Commentaires et suggestions éventuels :

9

FIG. B.24 – Questionnaire verbal : Étude avec ACA - Page 9/10.



SYNTHESE

Quelles sont les différences principales entre des deux systèmes d'aide que vous avez expérimentés :

- ***Avantages et inconvénients respectifs,***
- ***Comparaison avec les aides en ligne classiques, par exemple celles des logiciels grand public, tableurs ou traitements de texte.***

Quel système d'aide en ligne préférez-vous parmi les suivants :

- Le premier système d'aide que vous avez expérimenté ☐
- Le second système d'aide que vous avez utilisé ☐
- Les systèmes d'aide classiques, par exemple ceux de Word ou d'Excel qui combinent texte et graphique ☐

Pourquoi ?

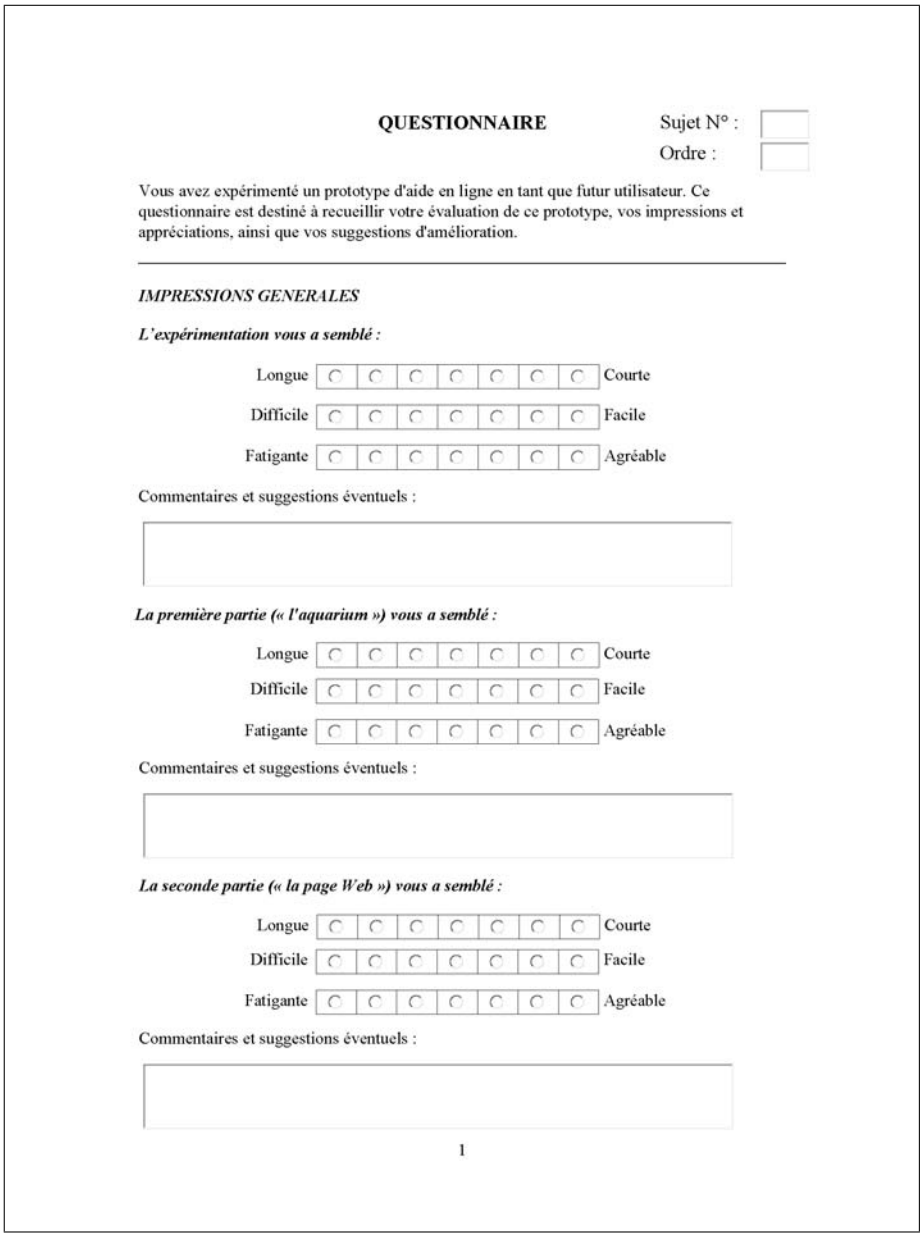
Autres commentaires et suggestions :

Envoyer par messagerie

10

Fig. B.25 – Questionnaire verbal : Étude avec ACA - Page 10/10.

Questionnaire verbal pour l'étude avec l'adaptation



QUESTIONNAIRE

Sujet N° :

Ordre :

Vous avez expérimenté un prototype d'aide en ligne en tant que futur utilisateur. Ce questionnaire est destiné à recueillir votre évaluation de ce prototype, vos impressions et appréciations, ainsi que vos suggestions d'amélioration.

IMPRESSIONS GENERALES

L'expérimentation vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Fatigante ○ ○ ○ ○ ○ ○ ○ Agréable

Commentaires et suggestions éventuels :

La première partie (« l'aquarium ») vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Fatigante ○ ○ ○ ○ ○ ○ ○ Agréable

Commentaires et suggestions éventuels :

La seconde partie (« la page Web ») vous a semblé :

Longue ○ ○ ○ ○ ○ ○ ○ Courte

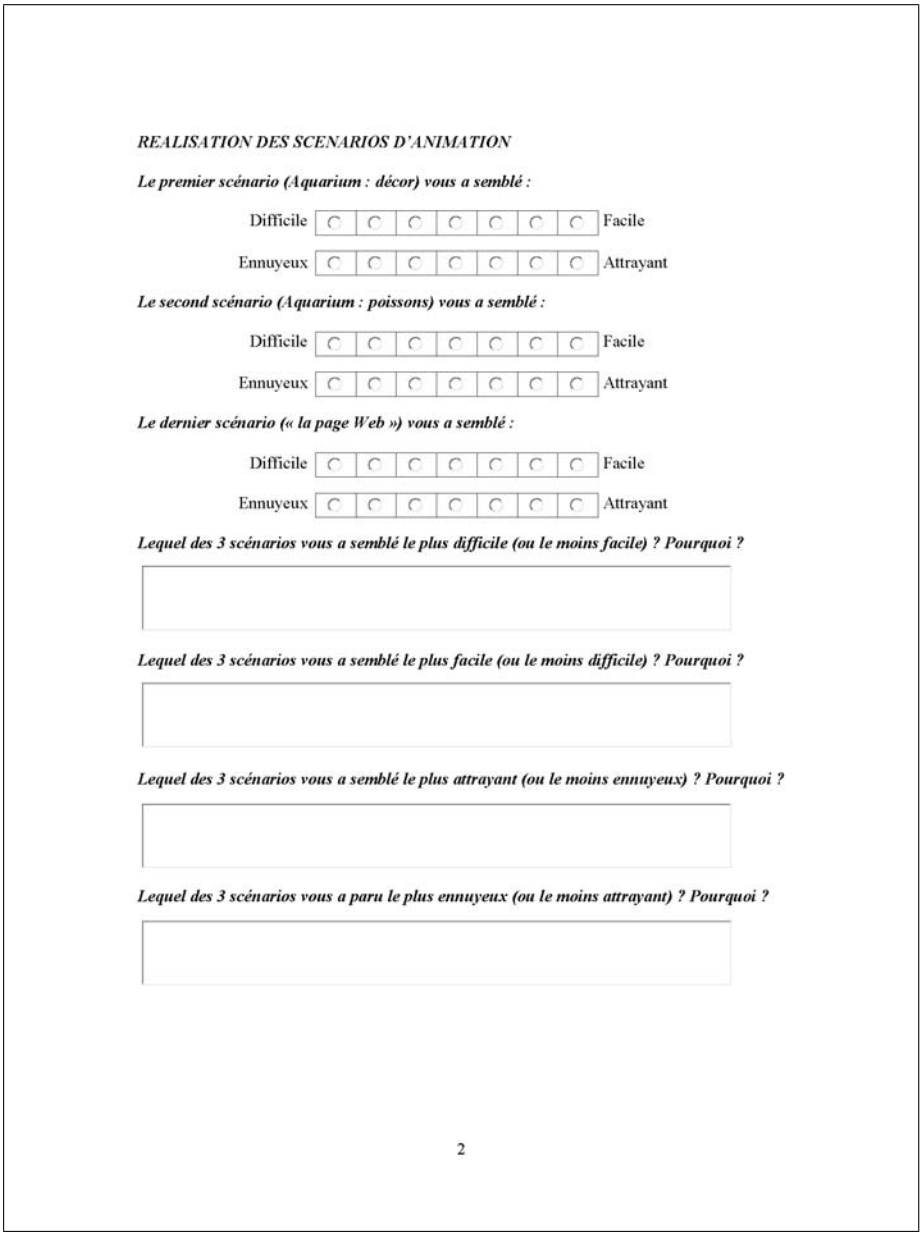
Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Fatigante ○ ○ ○ ○ ○ ○ ○ Agréable

Commentaires et suggestions éventuels :

1

FIG. B.26 – Questionnaire verbal : Étude avec Adaptation - Page 1/7.



REALISATION DES SCENARIOS D'ANIMATION

Le premier scénario (Aquarium : décor) vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Le second scénario (Aquarium : poissons) vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Le dernier scénario (« la page Web ») vous a semblé :

Difficile ○ ○ ○ ○ ○ ○ ○ Facile

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Attrayant

Lequel des 3 scénarios vous a semblé le plus difficile (ou le moins facile) ? Pourquoi ?

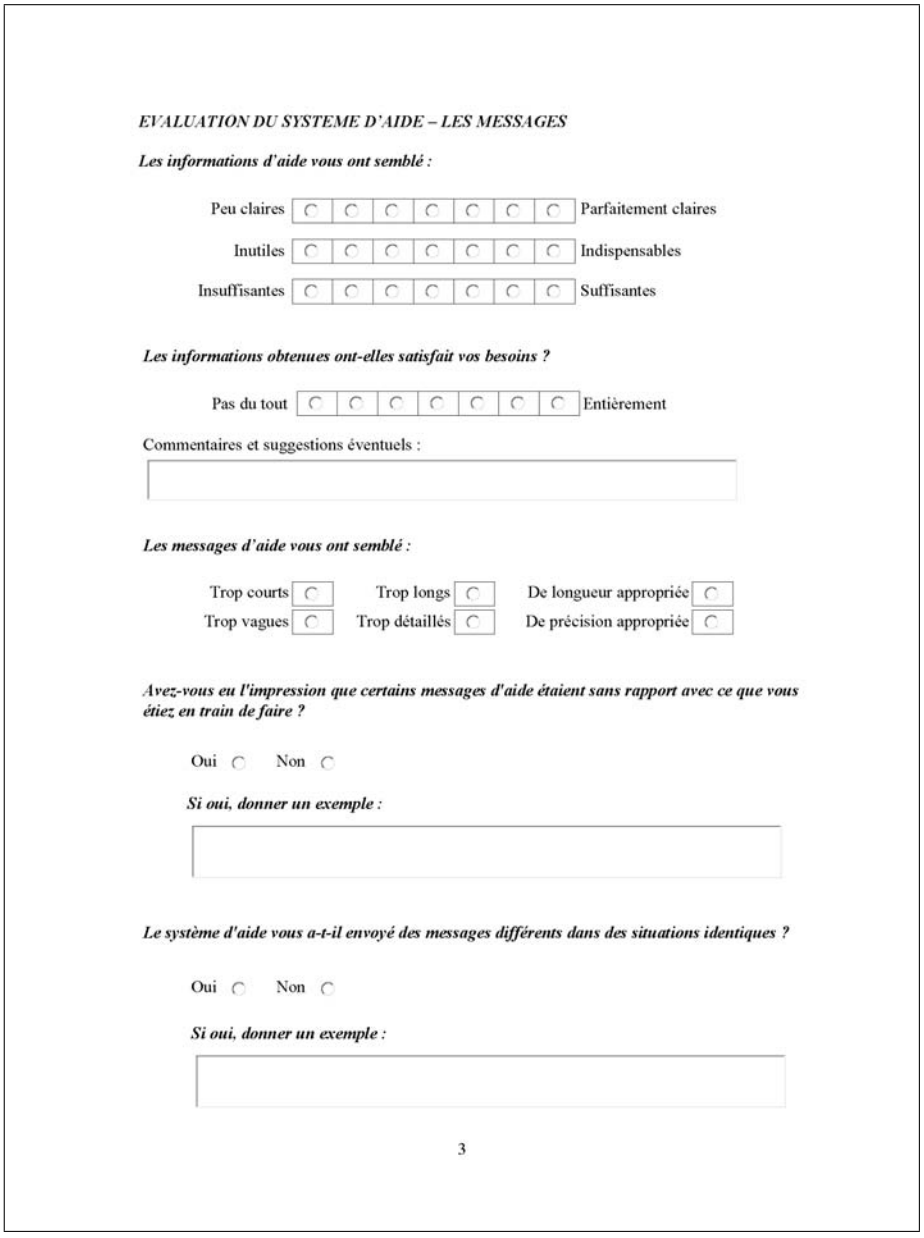
Lequel des 3 scénarios vous a semblé le plus facile (ou le moins difficile) ? Pourquoi ?

Lequel des 3 scénarios vous a semblé le plus attrayant (ou le moins ennuyeux) ? Pourquoi ?

Lequel des 3 scénarios vous a paru le plus ennuyeux (ou le moins attrayant) ? Pourquoi ?

2

FIG. B.27 – Questionnaire verbal : Étude avec Adaptation - Page 2/7.



EVALUATION DU SYSTEME D'AIDE – LES MESSAGES

Les informations d'aide vous ont semblé :

Peu claires ○ ○ ○ ○ ○ ○ ○ Parfaitement claires

Inutiles ○ ○ ○ ○ ○ ○ ○ Indispensables

Insuffisantes ○ ○ ○ ○ ○ ○ ○ Suffisantes

Les informations obtenues ont-elles satisfait vos besoins ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Entièrement

Commentaires et suggestions éventuels :

Les messages d'aide vous ont semblé :

Trop courts ○ Trop longs ○ De longueur appropriée ○

Trop vagues ○ Trop détaillés ○ De précision appropriée ○

Avez-vous eu l'impression que certains messages d'aide étaient sans rapport avec ce que vous étiez en train de faire ?

Oui ○ Non ○

Si oui, donner un exemple :

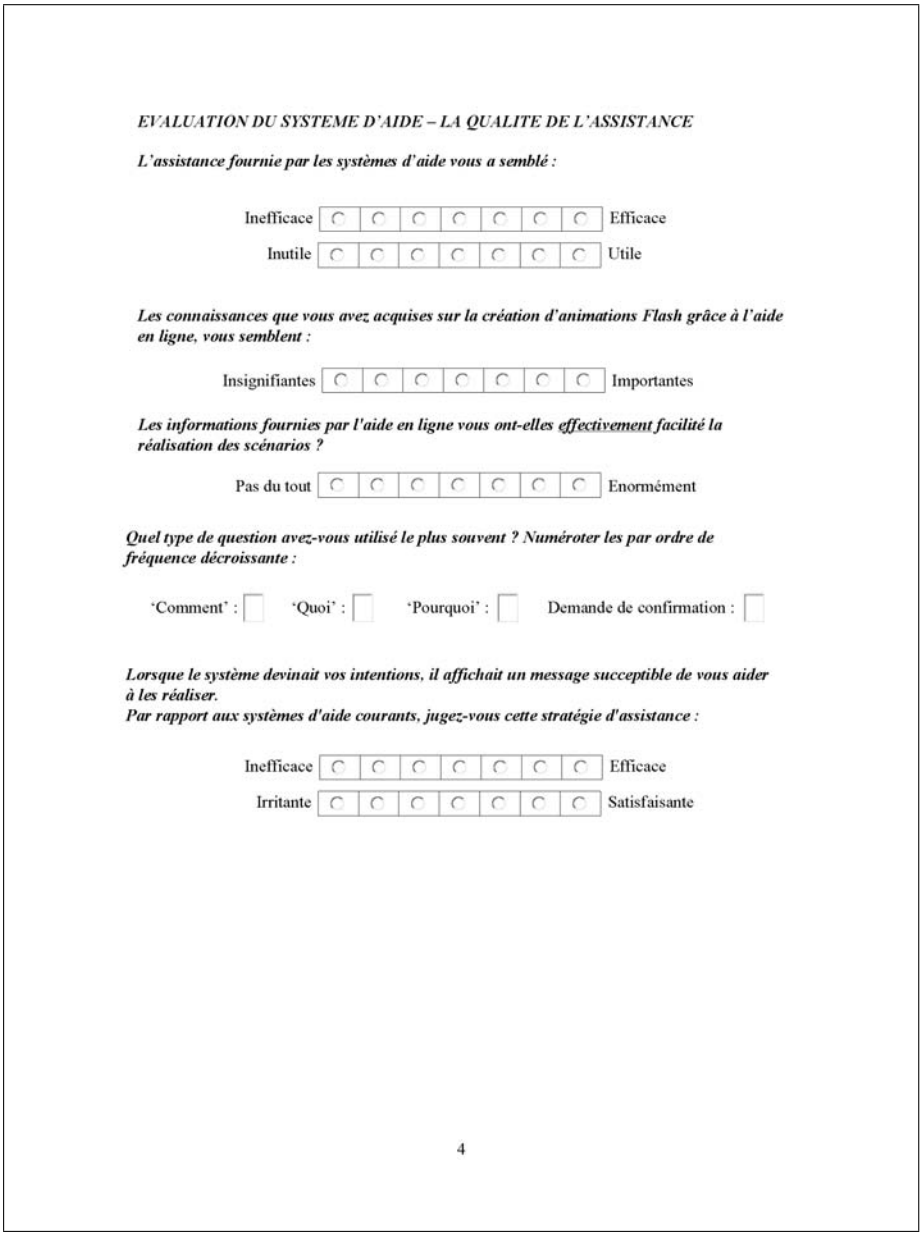
Le système d'aide vous a-t-il envoyé des messages différents dans des situations identiques ?

Oui ○ Non ○

Si oui, donner un exemple :

3

FIG. B.28 – Questionnaire verbal : Étude avec Adaptation - Page 3/7.



EVALUATION DU SYSTEME D'AIDE – LA QUALITE DE L'ASSISTANCE

L'assistance fournie par les systèmes d'aide vous a semblé :

Inefficace ○ ○ ○ ○ ○ ○ ○ Efficace

Inutile ○ ○ ○ ○ ○ ○ ○ Utile

Les connaissances que vous avez acquises sur la création d'animations Flash grâce à l'aide en ligne, vous semblent :

Insignifiantes ○ ○ ○ ○ ○ ○ ○ Importantes

Les informations fournies par l'aide en ligne vous ont-elles effectivement facilité la réalisation des scénarios ?

Pas du tout ○ ○ ○ ○ ○ ○ ○ Enormément

Quel type de question avez-vous utilisé le plus souvent ? Numéroter les par ordre de fréquence décroissante :

'Comment' : ☐ 'Quoi' : ☐ 'Pourquoi' : ☐ Demande de confirmation : ☐

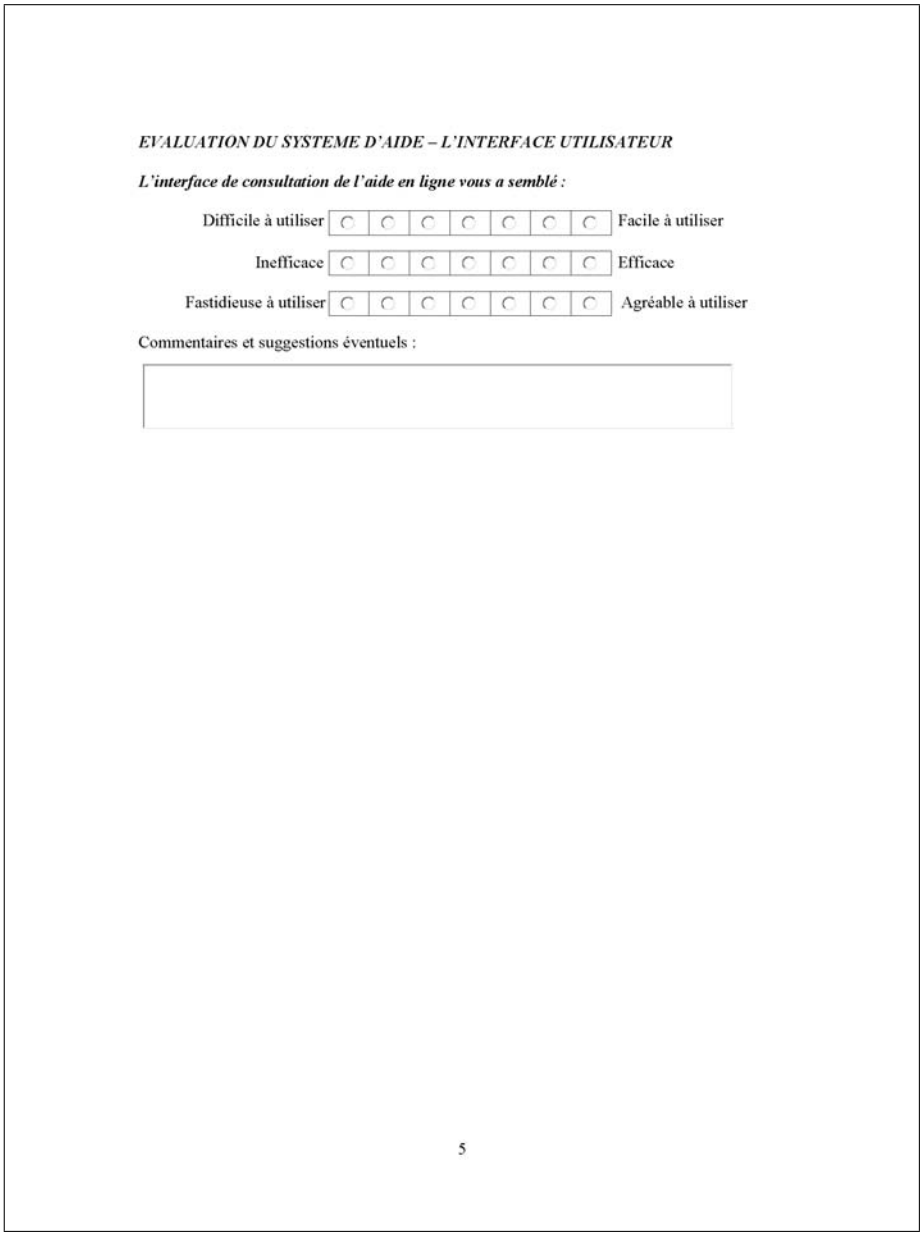
Lorsque le système devinait vos intentions, il affichait un message succeptible de vous aider à les réaliser.
Par rapport aux systèmes d'aide courants, jugez-vous cette stratégie d'assistance :

Inefficace ○ ○ ○ ○ ○ ○ ○ Efficace

Irritante ○ ○ ○ ○ ○ ○ ○ Satisfaisante

4

FIG. B.29 – Questionnaire verbal : Étude avec Adaptation - Page 4/7.



EVALUATION DU SYSTEME D'AIDE – L'INTERFACE UTILISATEUR

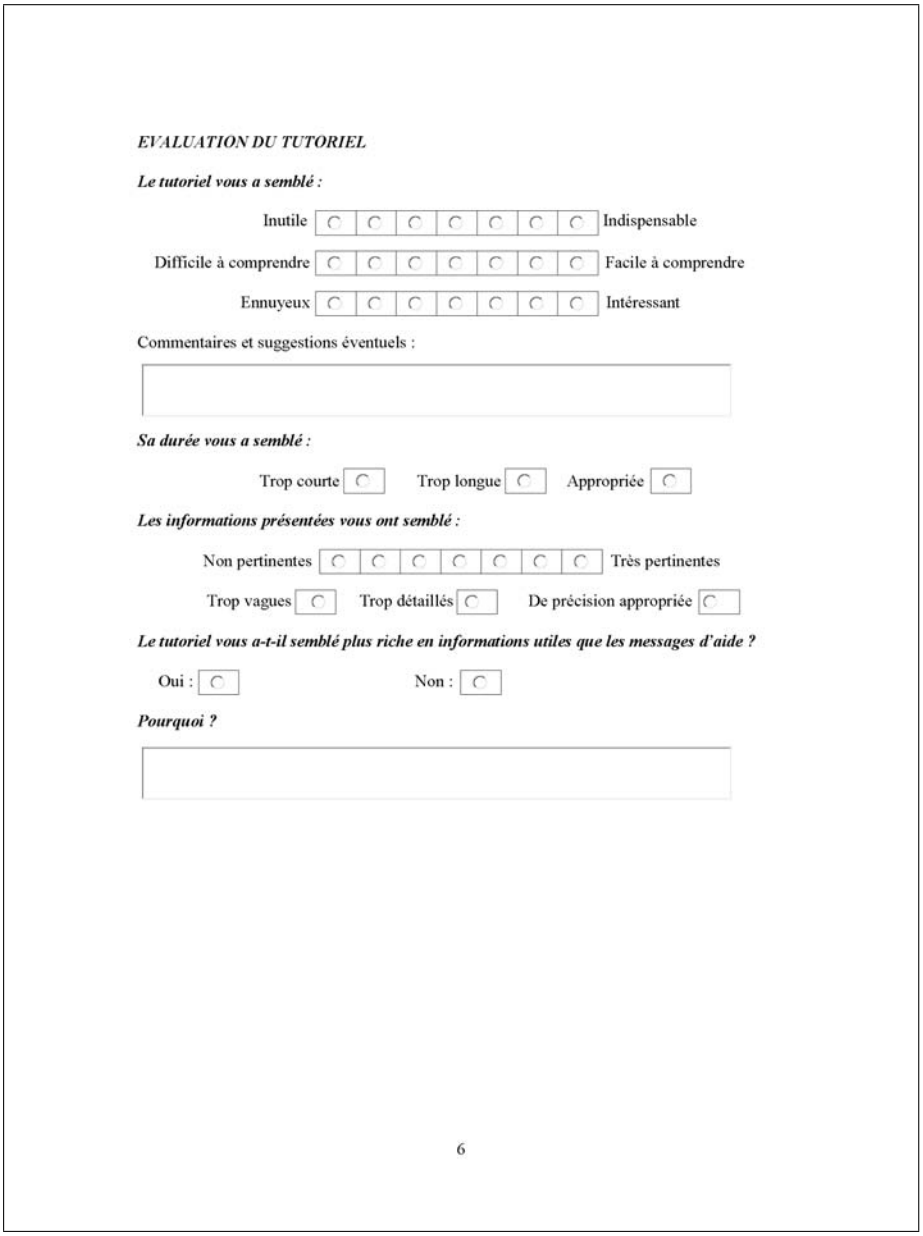
L'interface de consultation de l'aide en ligne vous a semblé :

Difficile à utiliser	○	○	○	○	○	○	○	Facile à utiliser
Inefficace	○	○	○	○	○	○	○	Efficace
Fastidieuse à utiliser	○	○	○	○	○	○	○	Agréable à utiliser

Commentaires et suggestions éventuels :

5

FIG. B.30 – Questionnaire verbal : Étude avec Adaptation - Page 5/7.



EVALUATION DU TUTORIEL

Le tutoriel vous a semblé :

Inutile ○ ○ ○ ○ ○ ○ ○ Indispensable

Difficile à comprendre ○ ○ ○ ○ ○ ○ ○ Facile à comprendre

Ennuyeux ○ ○ ○ ○ ○ ○ ○ Intéressant

Commentaires et suggestions éventuels :

Sa durée vous a semblé :

Trop courte ○ Trop longue ○ Appropriée ○

Les informations présentées vous ont semblé :

Non pertinentes ○ ○ ○ ○ ○ ○ ○ Très pertinentes

Trop vagues ○ Trop détaillés ○ De précision appropriée ○

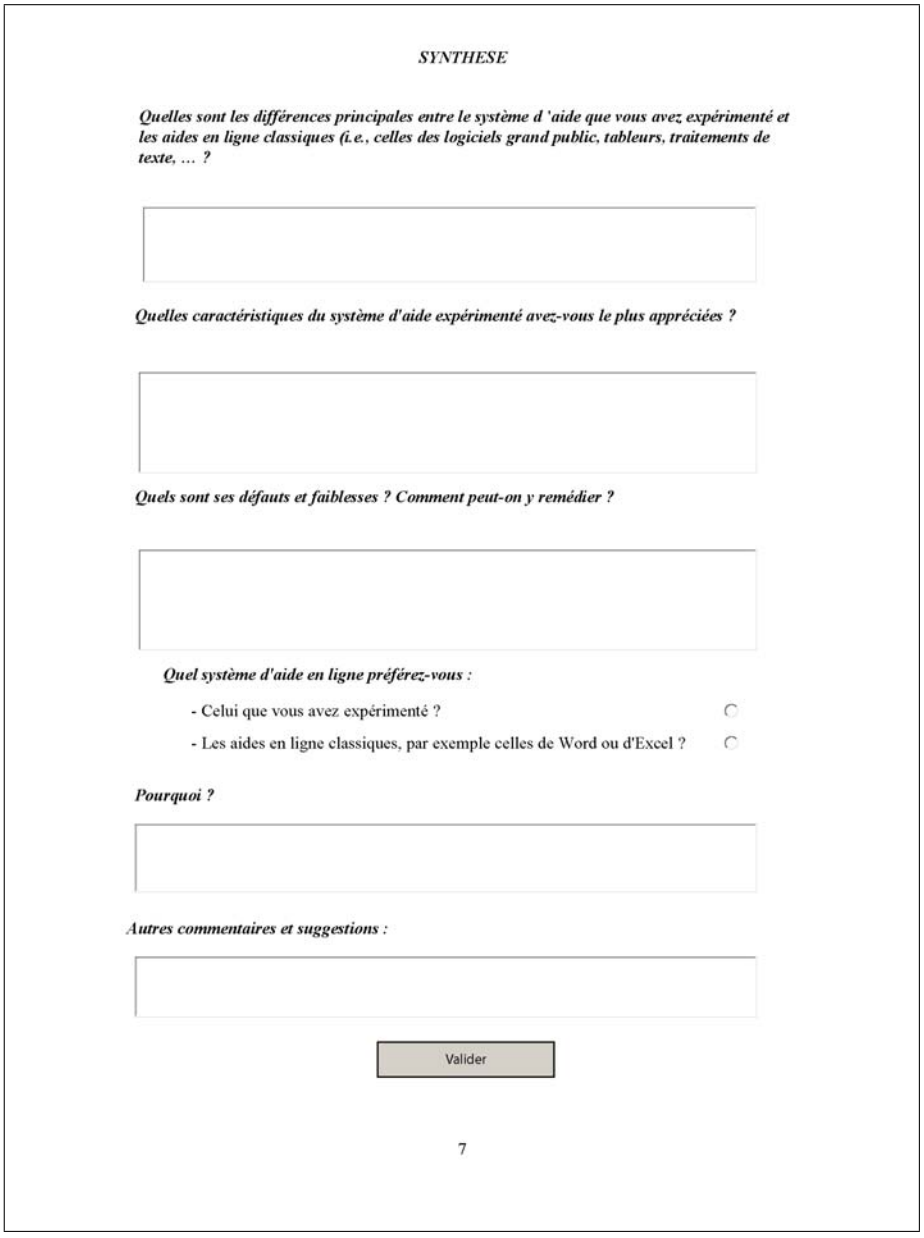
Le tutoriel vous a-t-il semblé plus riche en informations utiles que les messages d'aide ?

Oui : ○ Non : ○

Pourquoi ?

6

FIG. B.31 – Questionnaire verbal : Étude avec Adaptation - Page 6/7.



SYNTHESE

Quelles sont les différences principales entre le système d 'aide que vous avez expérimenté et les aides en ligne classiques (i.e., celles des logiciels grand public, tableurs, traitements de texte, ... ?

Quelles caractéristiques du système d'aide expérimenté avez-vous le plus appréciées ?

Quels sont ses défauts et faiblesses ? Comment peut-on y remédier ?

Quel système d'aide en ligne préférez-vous :

- Celui que vous avez expérimenté ?
- Les aides en ligne classiques, par exemple celles de Word ou d'Excel ?

Pourquoi ?

Autres commentaires et suggestions :

Valider

7

FIG. B.32 – Questionnaire verbal : Étude avec Adaptation - Page 7/7.

`Questionnaire non verbal pour l'étude avec Eloona`

QUESTIONNAIRE NON VERBAL - SAM

Pour exprimer vos impressions, vous disposez de caricatures représentant différentes émotions et réactions affectives d'un personnage appelé SAM. Le tableau ci-dessous précise l'interprétation (de la gauche vers la droite) des trois lignes de caricatures décrivant les états émotionnels de SAM.

Ligne	*Dimensions des variations des émotions de SAM*	
Première	Satisfait / Content / Heureux	Insatisfait / Mécontent / Irrité
Seconde	Excité / Stimulé / Intéressé	Calme / Amorphe / Indifférent
Troisième	Doutant de soi / Timide / Dominé / Inférieur	Sûr de soi / Assuré / Dominant / Supérieur

Comment avez-vous ressenti, quelles ont été vos réactions par rapport aux innovations suivantes :

1. L'utilisation de la parole au lieu du texte dans les messages d'aide

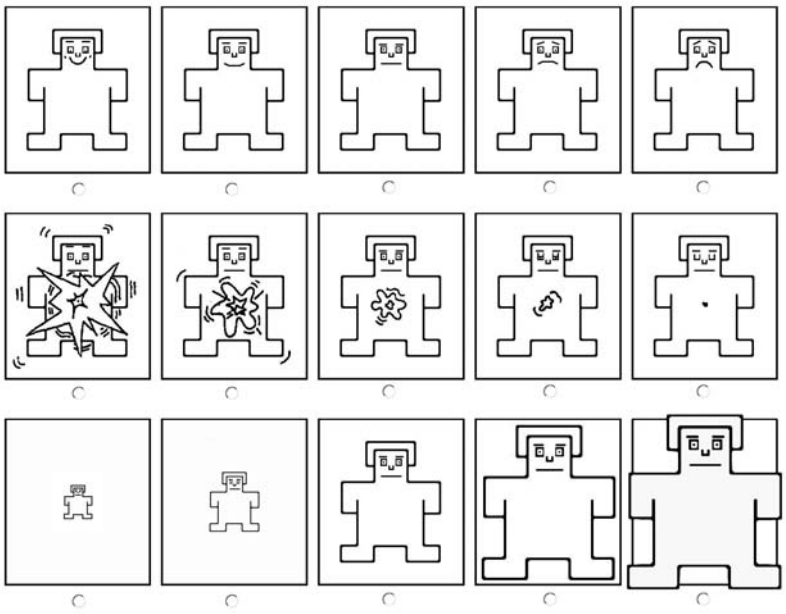


FIG. B.33 – Questionnaire non verbal : Étude avec Eloona - Page 1/6.

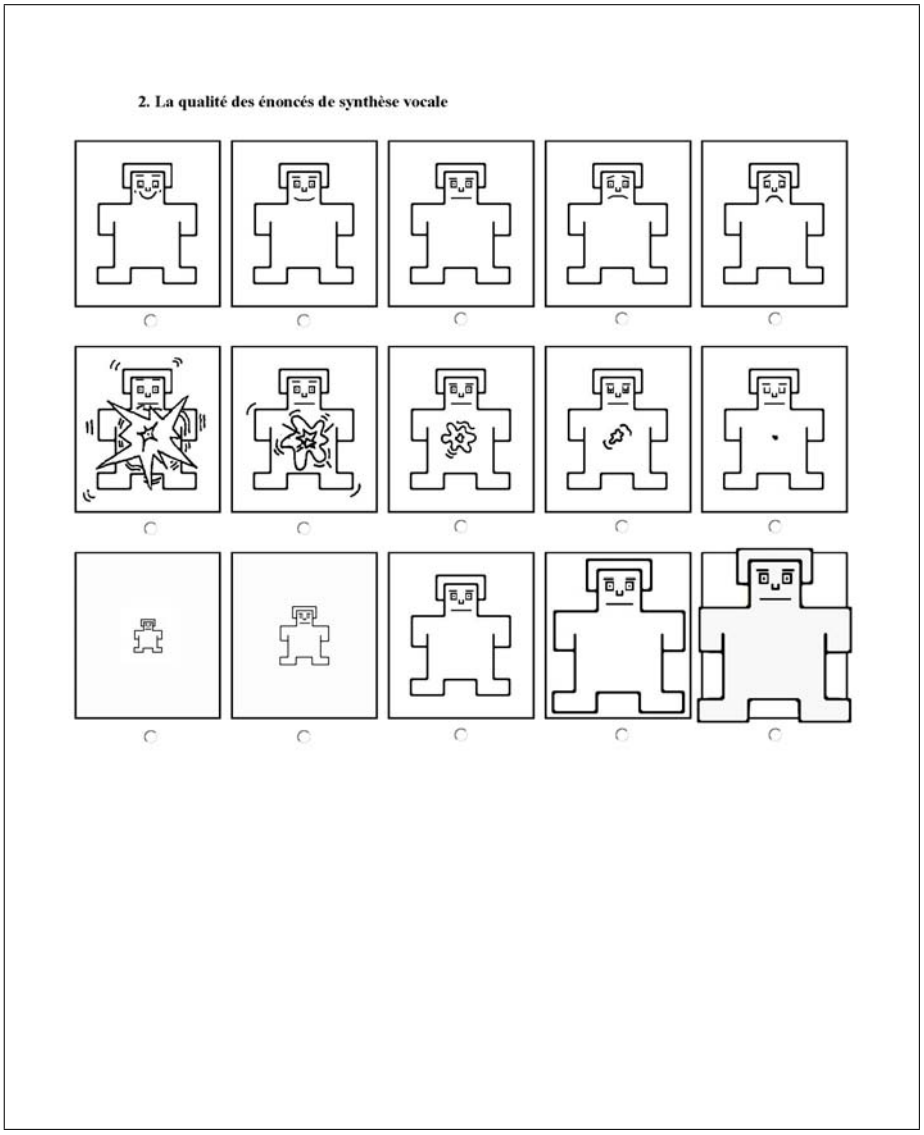


FIG. B.34 – Questionnaire non verbal : Étude avec Eloona - Page 2/6.

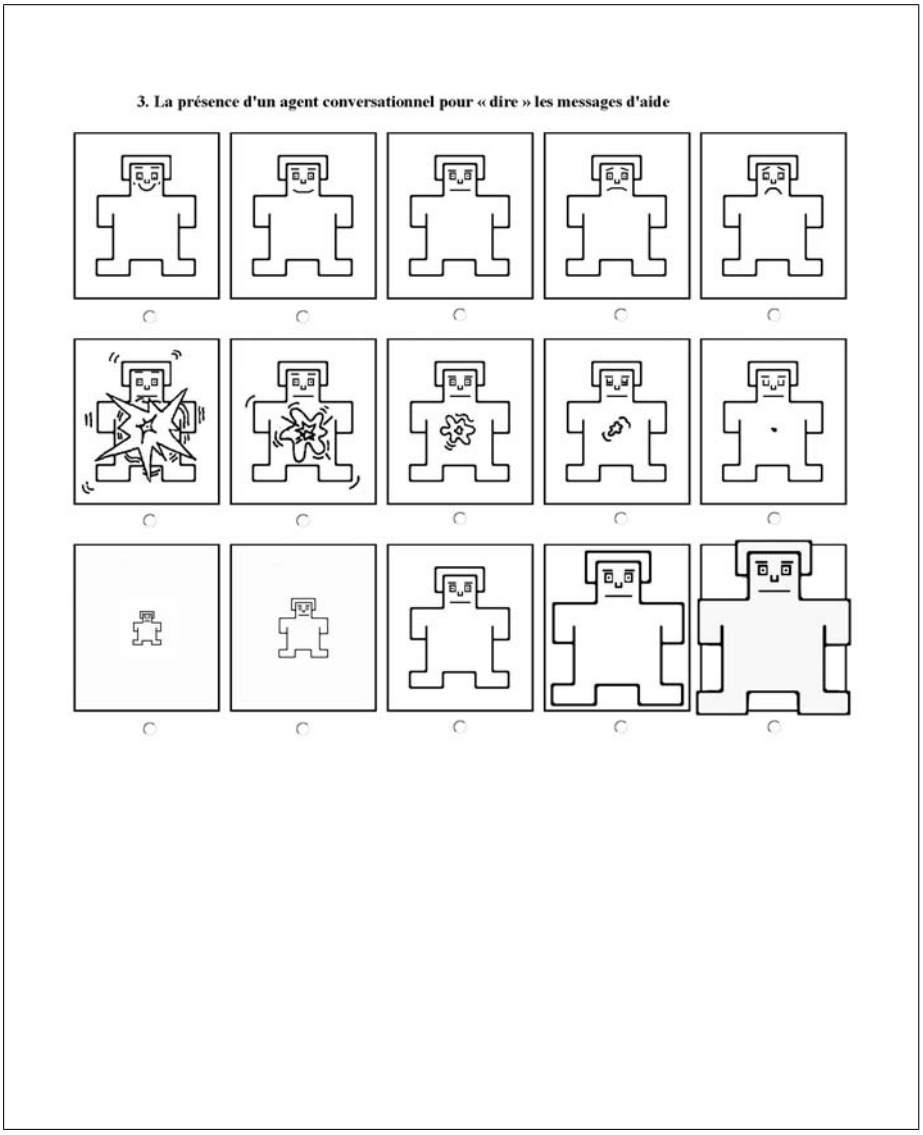


FIG. B.35 – Questionnaire non verbal : Étude avec Eloona - Page 3/6.

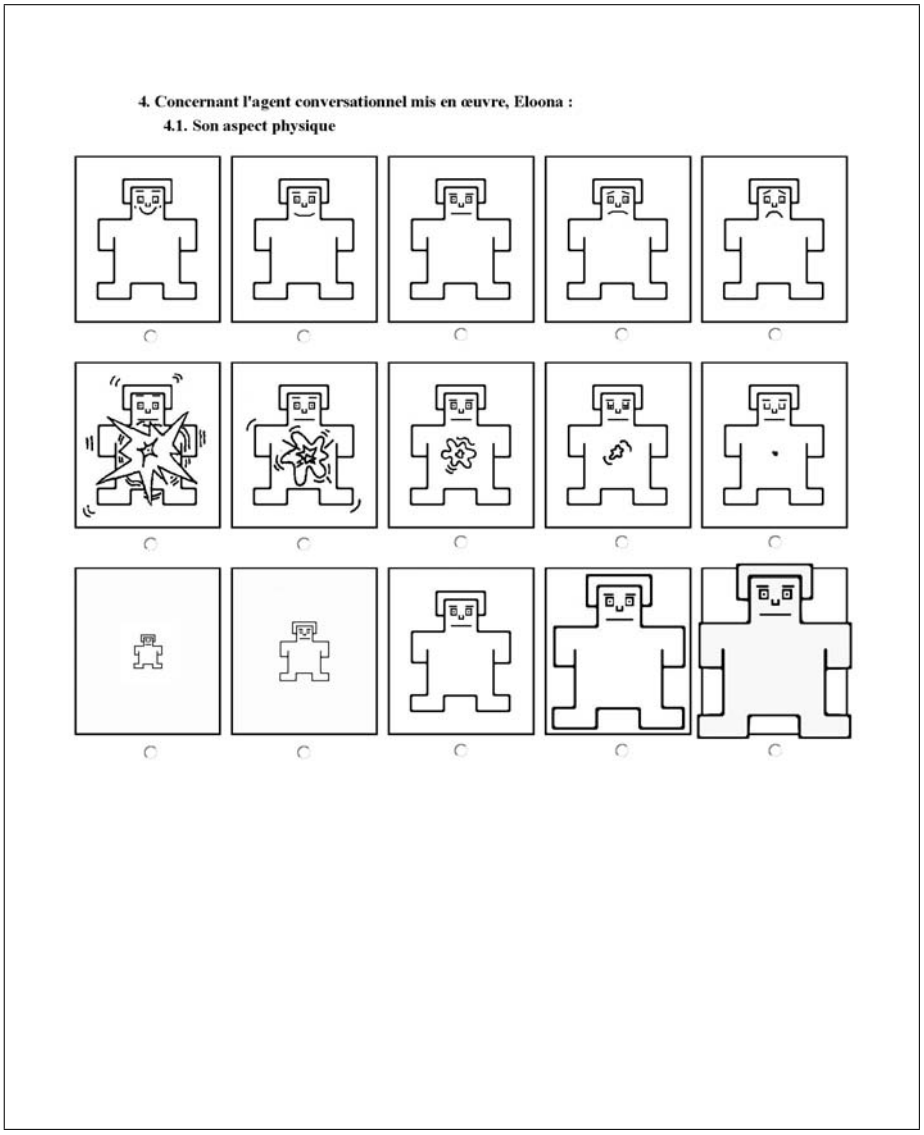


FIG. B.36 – Questionnaire non verbal : Étude avec Eloona - Page 4/6.

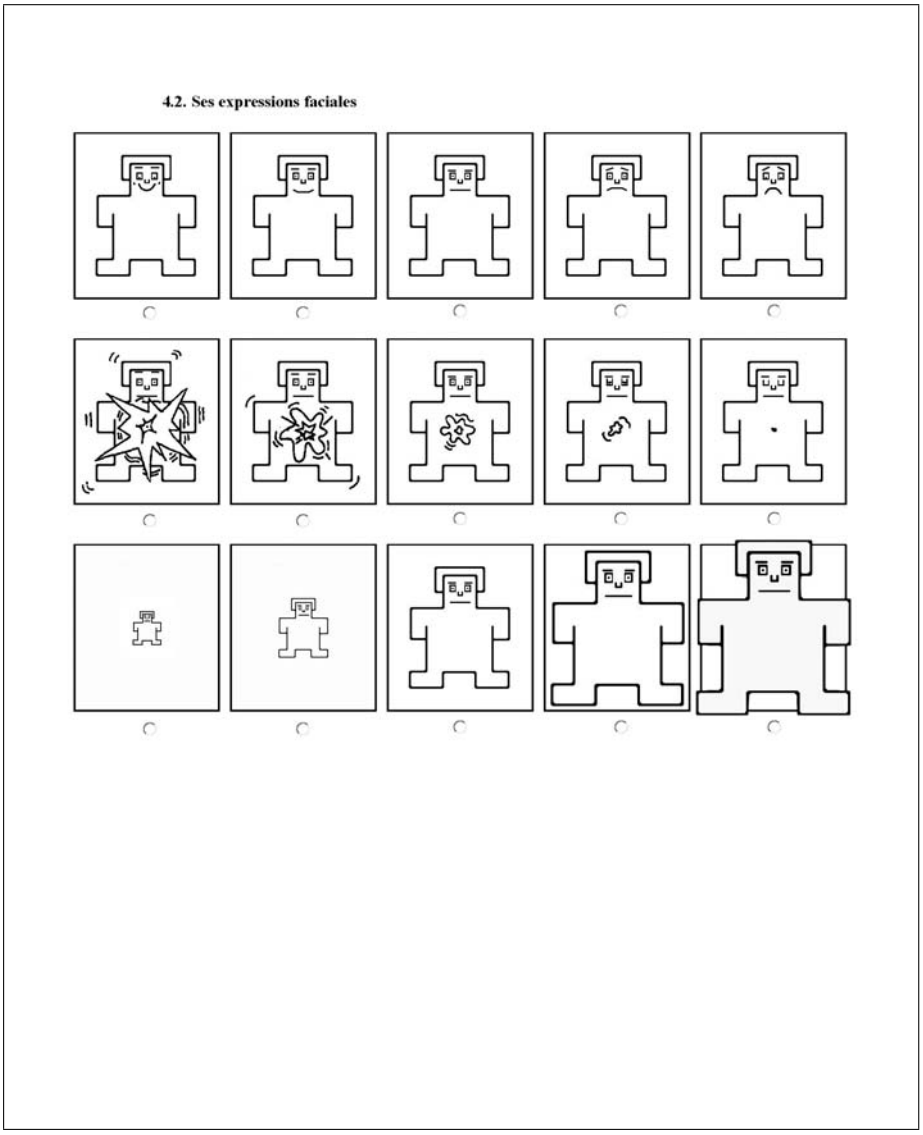


FIG. B.37 – Questionnaire non verbal : Étude avec Eloona - Page 5/6.

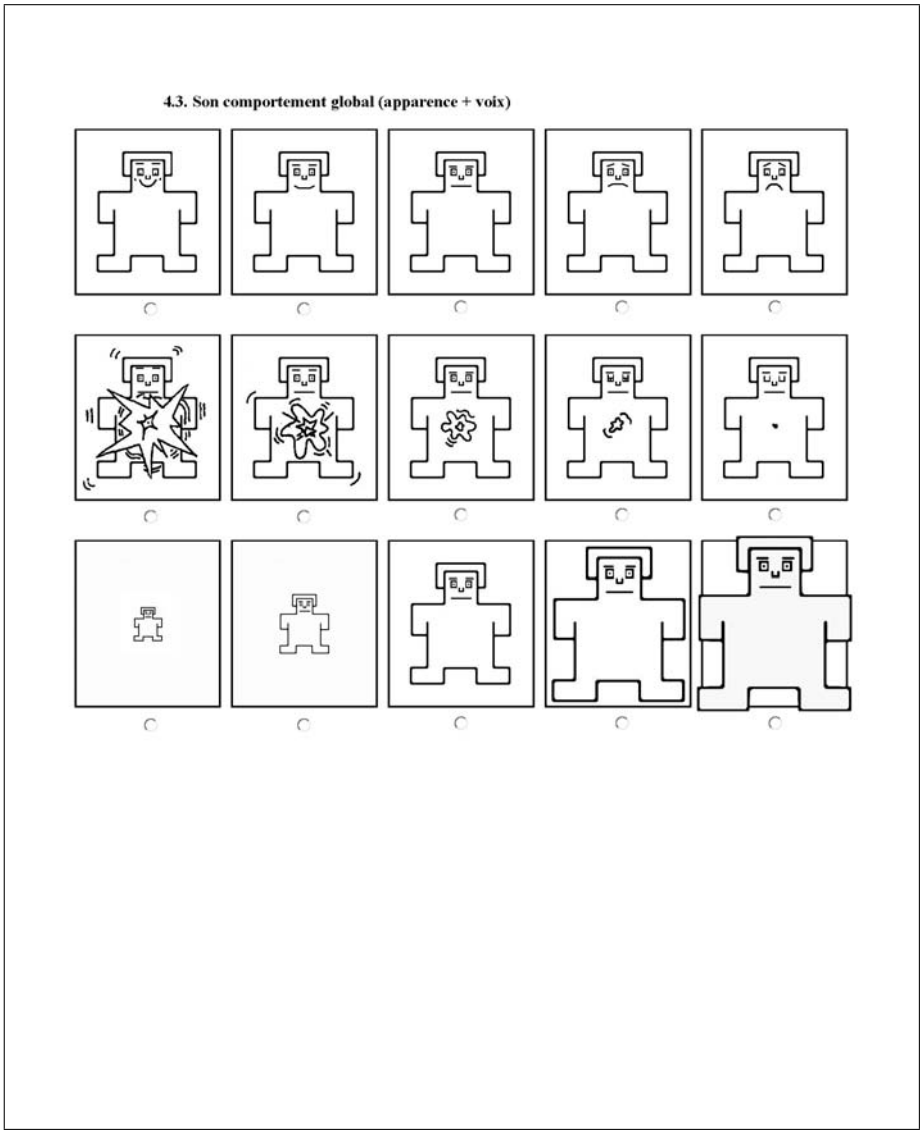
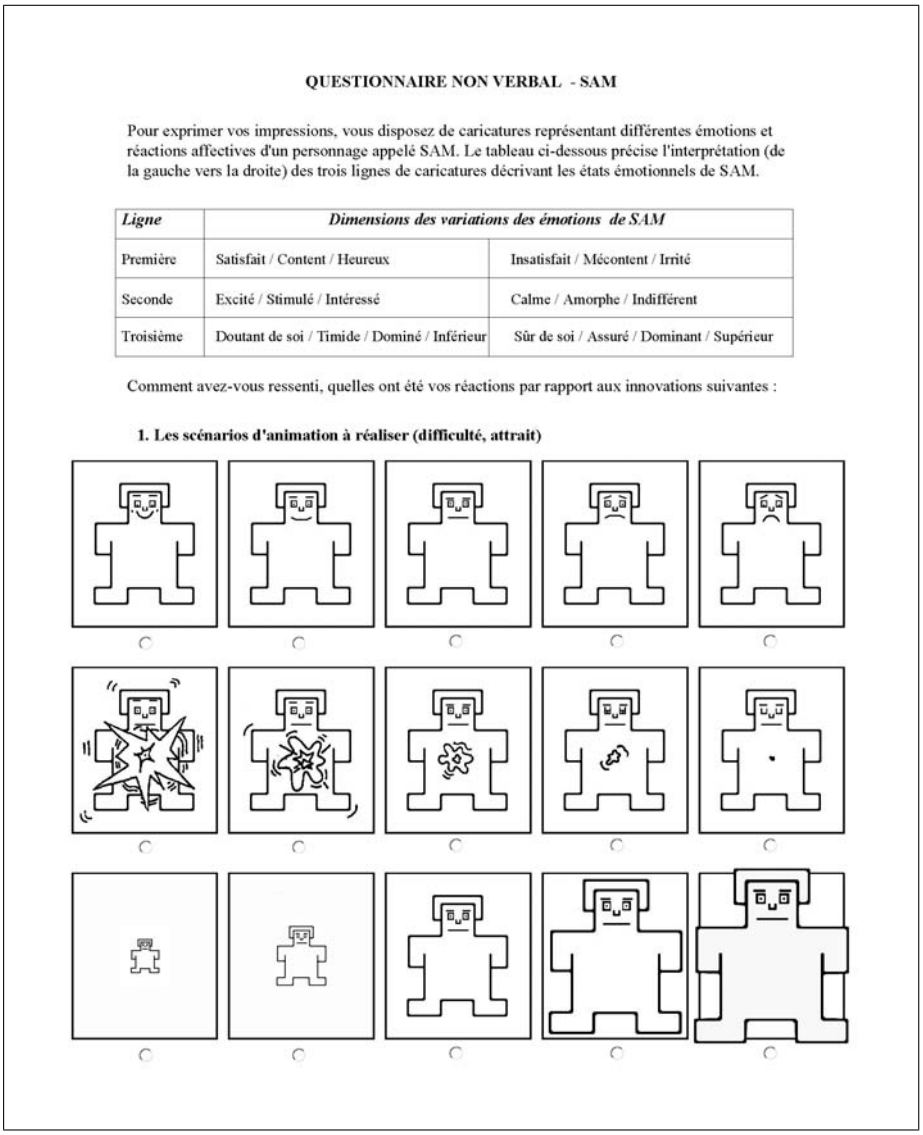


FIG. B.38 – Questionnaire non verbal : Étude avec Eloona - Page 6/6.

`Questionnaire non verbal pour l'étude avec Adaptation`



QUESTIONNAIRE NON VERBAL - SAM

Pour exprimer vos impressions, vous disposez de caricatures représentant différentes émotions et réactions affectives d'un personnage appelé SAM. Le tableau ci-dessous précise l'interprétation (de la gauche vers la droite) des trois lignes de caricatures décrivant les états émotionnels de SAM.

Ligne	*Dimensions des variations des émotions de SAM*	
Première	Satisfait / Content / Heureux	Insatisfait / Mécontent / Irrité
Seconde	Excité / Stimulé / Intéressé	Calme / Amorphe / Indifférent
Troisième	Doutant de soi / Timide / Dominé / Inférieur	Sûr de soi / Assuré / Dominant / Supérieur

Comment avez-vous ressenti, quelles ont été vos réactions par rapport aux innovations suivantes :

1. Les scénarios d'animation à réaliser (difficulté, attrait)

FIG. B.39 – Questionnaire non verbal : Étude avec Eloona - Page 1/6.

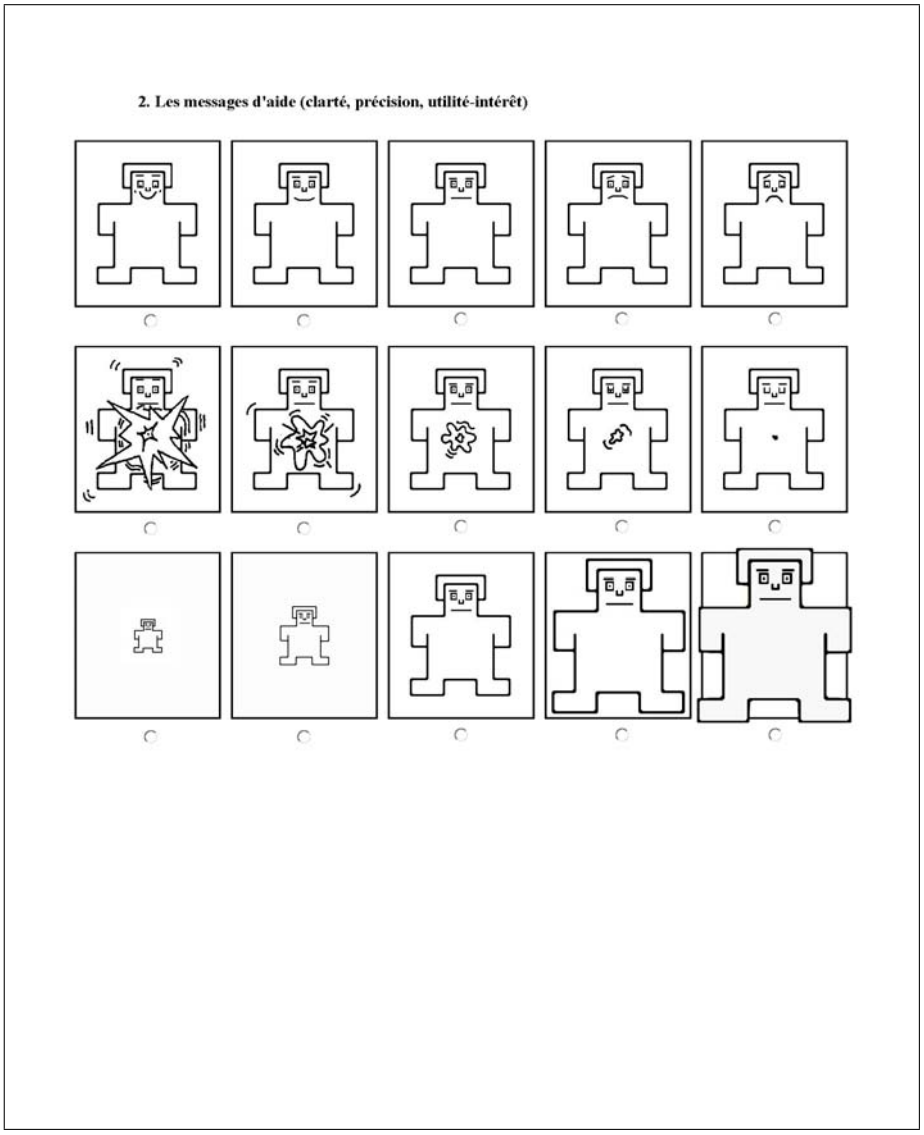


FIG. B.40 – Questionnaire non verbal : Étude avec Eloona - Page 2/6.

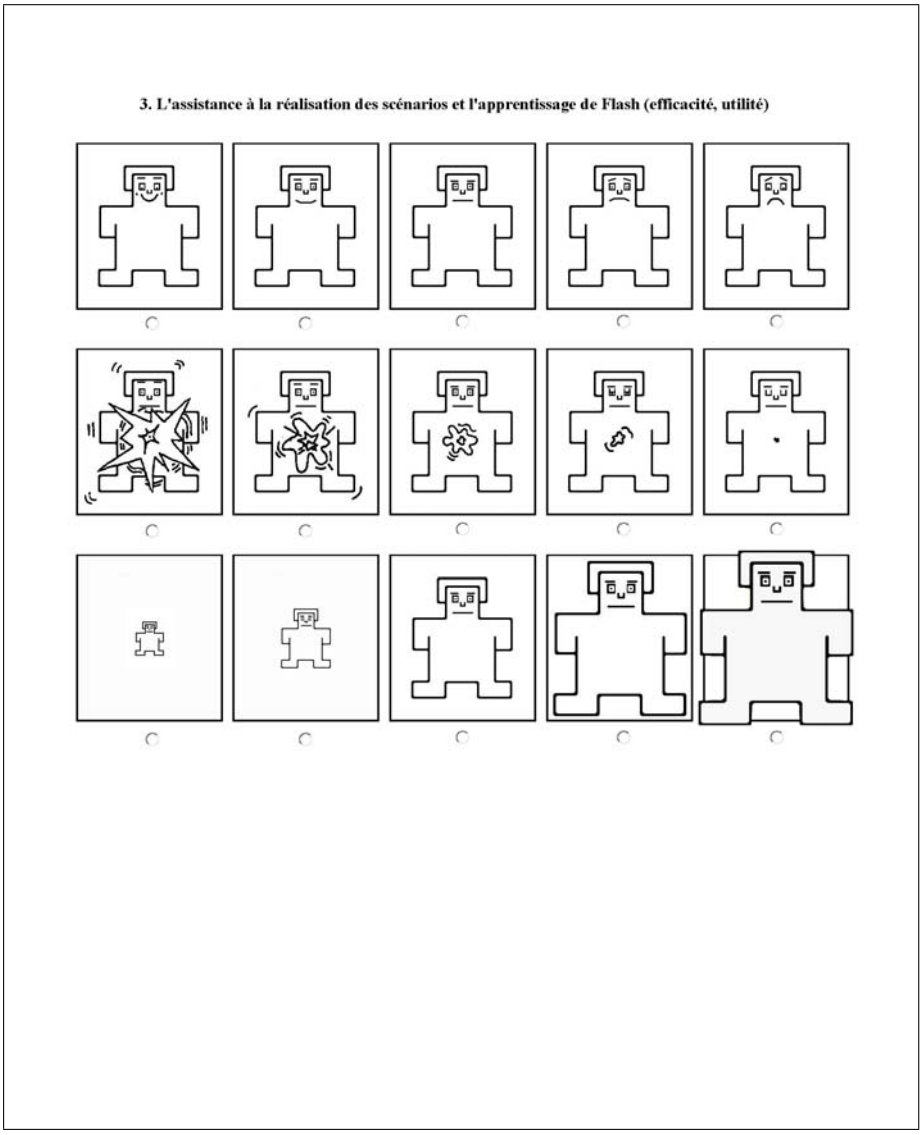


FIG. B.41 – Questionnaire non verbal : Étude avec Eloona - Page 3/6.

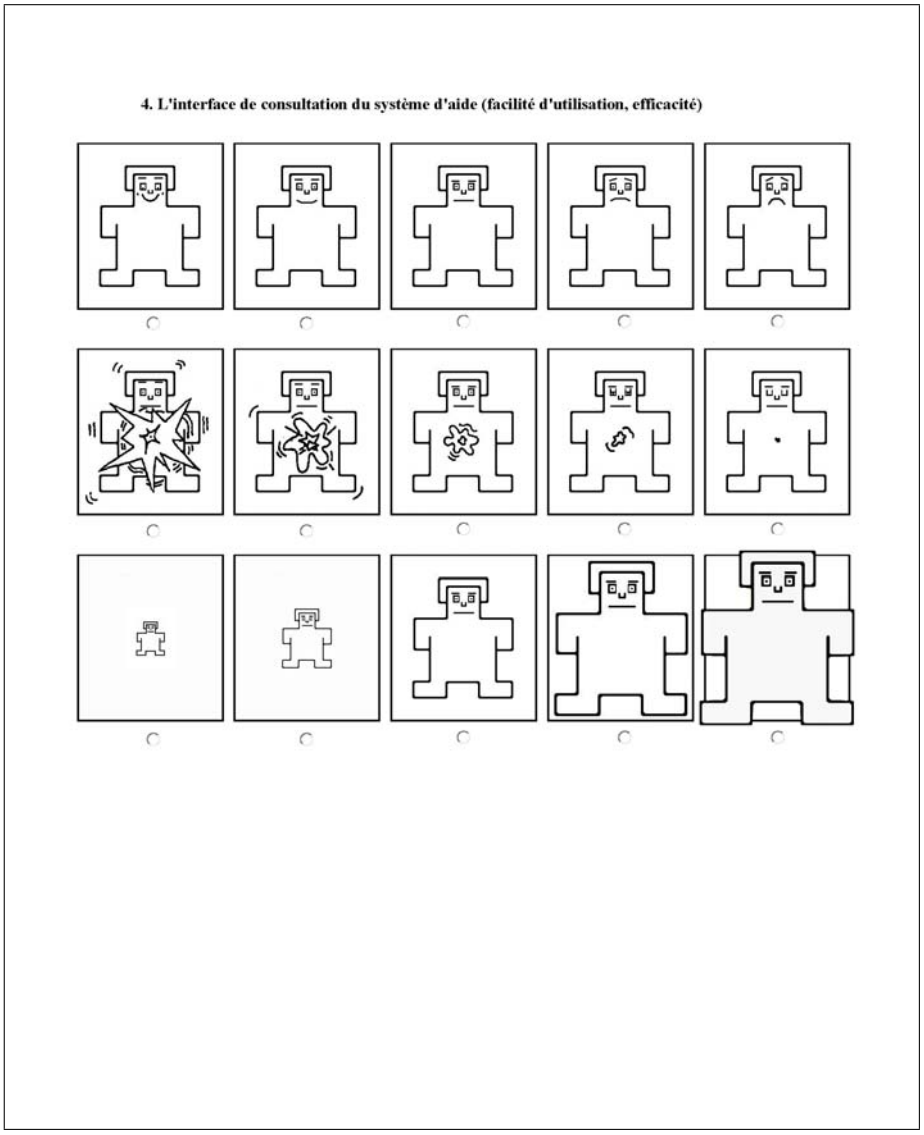


FIG. B.42 – Questionnaire non verbal : Étude avec Eloona - Page 4/6.

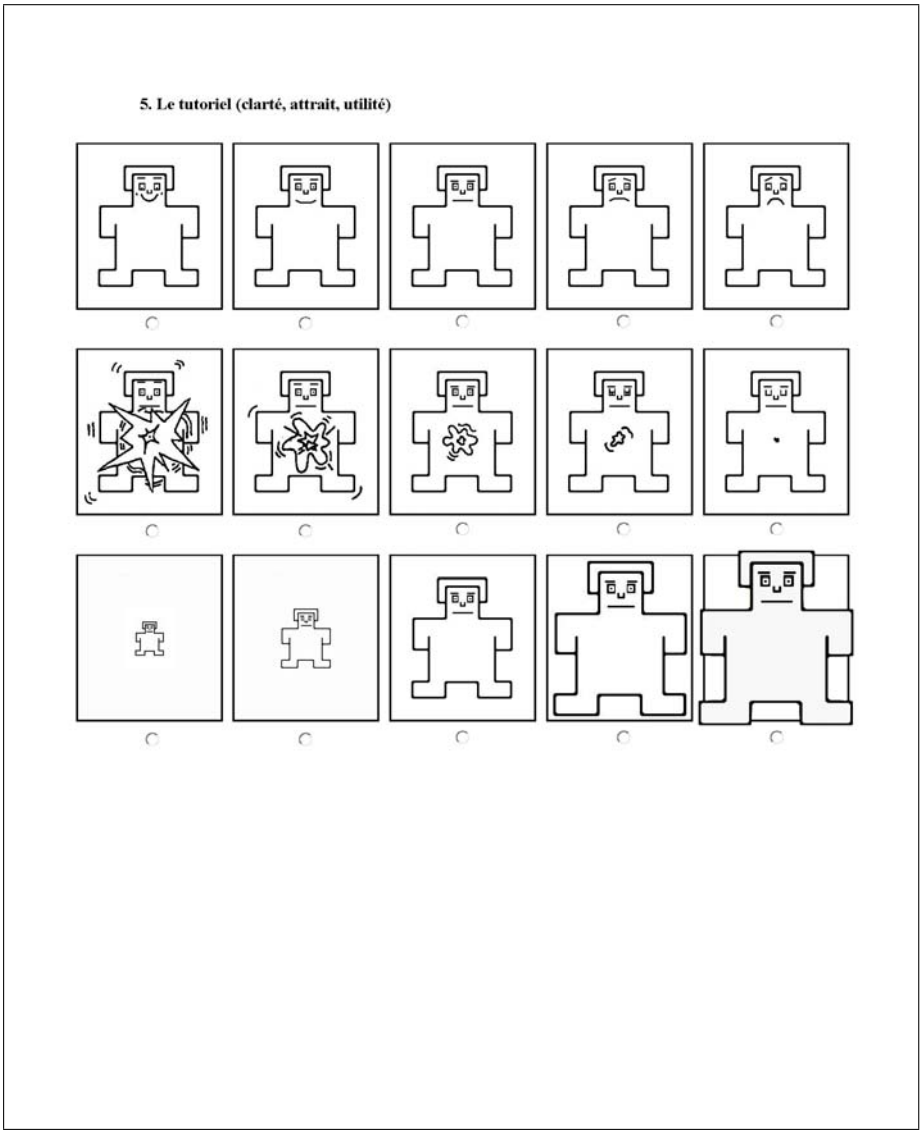


FIG. B.43 – Questionnaire non verbal : Étude avec Eloona - Page 5/6.

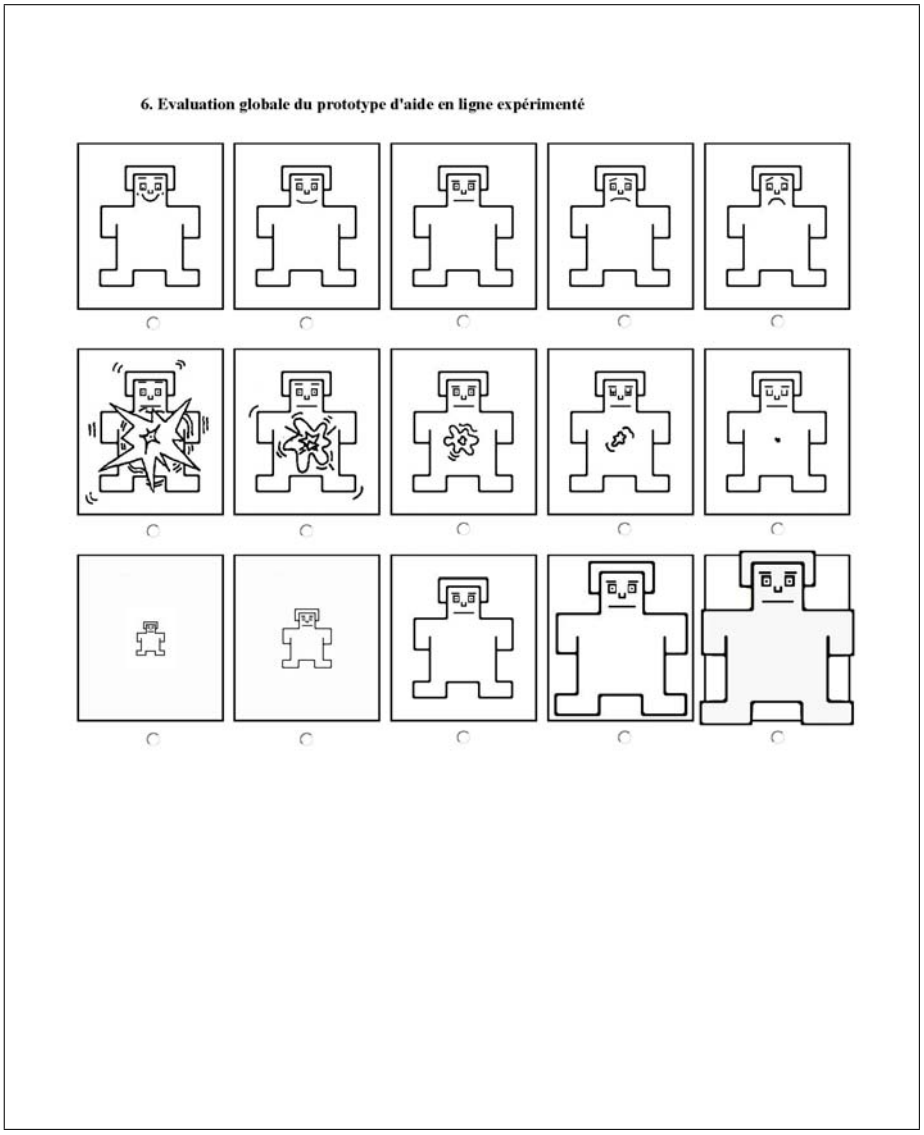
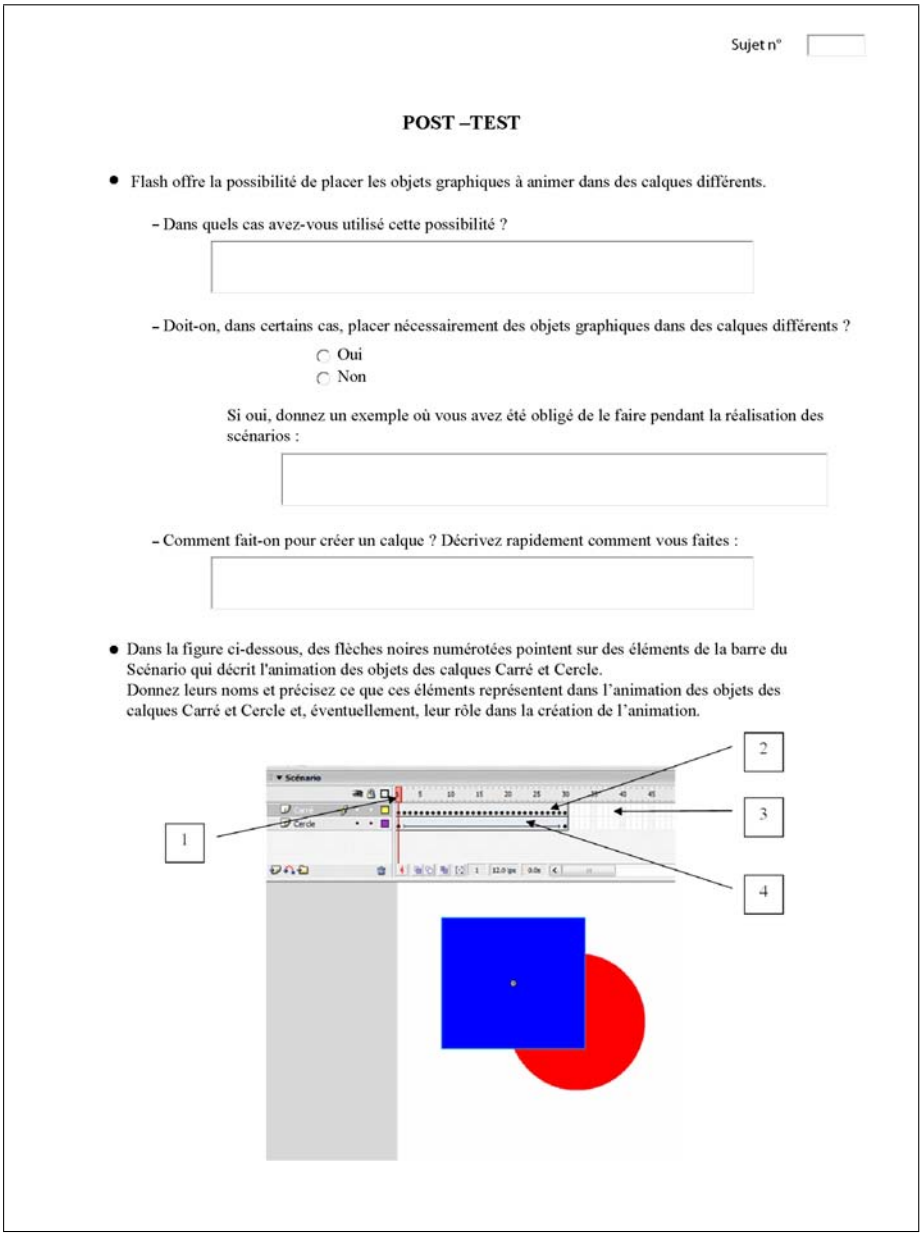


FIG. B.44 – Questionnaire non verbal : Étude avec Eloona - Page 6/6.

`Post-test commun aux deux études`



Sujet n°

POST –TEST

- Flash offre la possibilité de placer les objets graphiques à animer dans des calques différents.

  – Dans quels cas avez-vous utilisé cette possibilité ?

  – Doit-on, dans certains cas, placer nécessairement des objets graphiques dans des calques différents ?

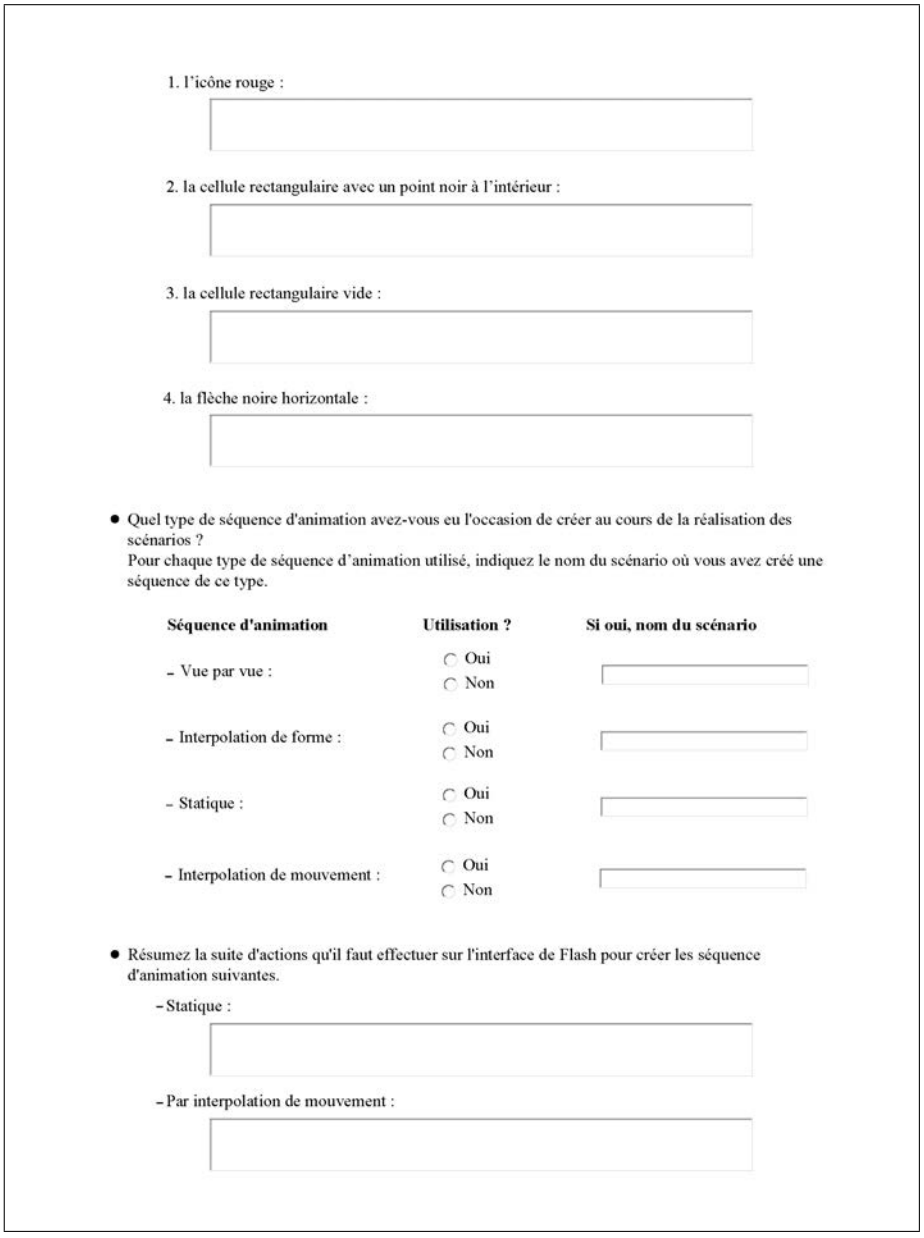
  ○ Oui
  ○ Non

  Si oui, donnez un exemple où vous avez été obligé de le faire pendant la réalisation des scénarios :

  – Comment fait-on pour créer un calque ? Décrivez rapidement comment vous faites :

- Dans la figure ci-dessous, des flèches noires numérotées pointent sur des éléments de la barre du Scénario qui décrit l'animation des objets des calques Carré et Cercle.
  Donnez leurs noms et précisez ce que ces éléments représentent dans l'animation des objets des calques Carré et Cercle et, éventuellement, leur rôle dans la création de l'animation.

FIG. B.45 – Post-test- Page 1/3.



1. l'icône rouge :

2. la cellule rectangulaire avec un point noir à l'intérieur :

3. la cellule rectangulaire vide :

4. la flèche noire horizontale :

- Quel type de séquence d'animation avez-vous eu l'occasion de créer au cours de la réalisation des scénarios ?
  Pour chaque type de séquence d'animation utilisé, indiquez le nom du scénario où vous avez créé une séquence de ce type.

Séquence d'animation	**Utilisation ?**	**Si oui, nom du scénario**
- Vue par vue :	○ Oui ○ Non	
- Interpolation de forme :	○ Oui ○ Non	
- Statique :	○ Oui ○ Non	
- Interpolation de mouvement :	○ Oui ○ Non	

- Résumez la suite d'actions qu'il faut effectuer sur l'interface de Flash pour créer les séquence d'animation suivantes.
  - Statique :
  - Par interpolation de mouvement :

FIG. B.46 – Post-test- Page 2/3.

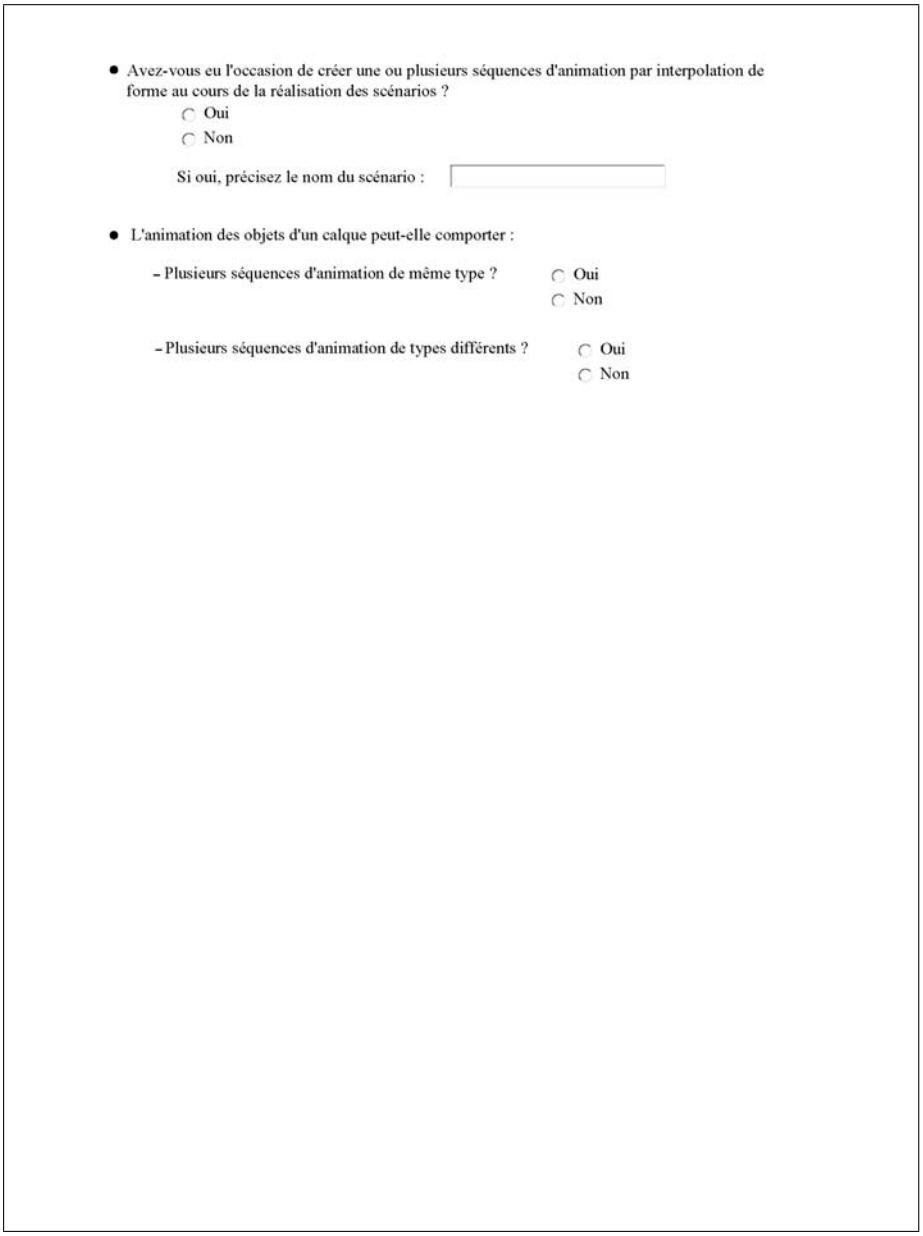
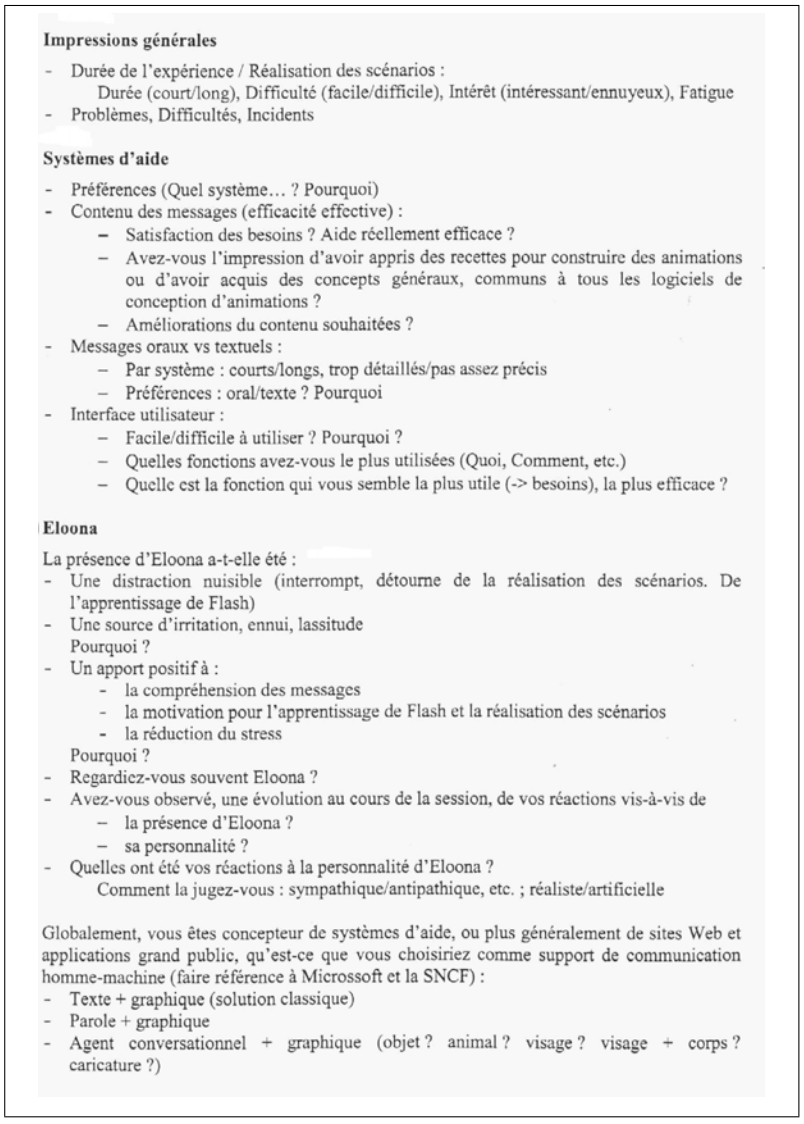


FIG. B.47 – Post-test- Page 3/3.

Entretien de debriefing

L'expérimentation se termine par un entretien avec l'un des expérimentateurs dont l'objectif est de recueillir les observations, réactions, etc., du sujet. Le canevas/scénario de cet entretien semi-directif de débriefing est donné ci-dessous. Là encore il existe une distinction entre les deux études.

`Entretien lors de l'étude avec ACA`



Impressions générales

- Durée de l'expérience / Réalisation des scénarios :
  Durée (court/long), Difficulté (facile/difficile), Intérêt (intéressant/ennuyeux), Fatigue
- Problèmes, Difficultés, Incidents

Systèmes d'aide

- Préférences (Quel système… ? Pourquoi)
- Contenu des messages (efficacité effective) :
  - Satisfaction des besoins ? Aide réellement efficace ?
  - Avez-vous l'impression d'avoir appris des recettes pour construire des animations ou d'avoir acquis des concepts généraux, communs à tous les logiciels de conception d'animations ?
  - Améliorations du contenu souhaitées ?
- Messages oraux vs textuels :
  - Par système : courts/longs, trop détaillés/pas assez précis
  - Préférences : oral/texte ? Pourquoi
- Interface utilisateur :
  - Facile/difficile à utiliser ? Pourquoi ?
  - Quelles fonctions avez-vous le plus utilisées (Quoi, Comment, etc.)
  - Quelle est la fonction qui vous semble la plus utile (-> besoins), la plus efficace ?

Eloona

La présence d'Eloona a-t-elle été :
- Une distraction nuisible (interrompt, détourne de la réalisation des scénarios. De l'apprentissage de Flash)
- Une source d'irritation, ennui, lassitude
  Pourquoi ?
- Un apport positif à :
  - la compréhension des messages
  - la motivation pour l'apprentissage de Flash et la réalisation des scénarios
  - la réduction du stress
  Pourquoi ?
- Regardiez-vous souvent Eloona ?
- Avez-vous observé, une évolution au cours de la session, de vos réactions vis-à-vis de
  - la présence d'Eloona ?
  - sa personnalité ?
- Quelles ont été vos réactions à la personnalité d'Eloona ?
  Comment la jugez-vous : sympathique/antipathique, etc. ; réaliste/artificielle

Globalement, vous êtes concepteur de systèmes d'aide, ou plus généralement de sites Web et applications grand public, qu'est-ce que vous choisiriez comme support de communication homme-machine (faire référence à Microsoft et la SNCF) :
- Texte + graphique (solution classique)
- Parole + graphique
- Agent conversationnel + graphique (objet ? animal ? visage ? visage + corps ? caricature ?)

FIG. B.48 – Schéma de l'entretien de debriefing lors de l'étude avec ACA.

Entretien lors de l'étude avec Adaptation

Impressions générales

- Durée de l'expérience / Réalisation des scénarios :
  Durée (court/long), Difficulté (facile/difficile), Intérêt (intéressant/ennuyeux), Fatigue
- Problèmes, Difficultés, Incidents

Système d'aide

- Différences par rapport aux systèmes d'aide en ligne classiques des applications destinées au grand public ? Appréciées ou non ?
- Perception du caractère évolutif des messages d'aide ? Jugement subjectif ?
- Perception des initiatives du système d'aide ? Adéquation avec ce que voulait (ou était en train de) faire le sujet, ses intentions ? Jugement subjectif ?
- Contenu des messages (efficacité effective) :
  - Satisfaction des besoins ? Aide réellement efficace ? Clarté ? Précision ?
  - Avez-vous l'impression d'avoir appris des recettes pour construire des animations ou d'avoir acquis des concepts généraux, communs à tous les logiciels de conception d'animations ?
  - Améliorations du contenu souhaitées ?
- Messages texte + graphique :
  - Trop courts/longs, trop détaillés/pas assez précis
  - Utilité du graphique
  - Comparaison avec les messages des systèmes d'aide des logiciels grand public
- Interface utilisateur :
  - Facile/difficile à utiliser ? Pourquoi ?
  - Quelles fonctions avez-vous le plus utilisées (Quoi, Comment, etc.)
  - Quelle est la fonction qui vous semble la plus utile (-> besoins), la plus efficace ?
  - Concernant le mode de consultation par boutons, qu'en pensez-vous ?

Globalement, vous êtes concepteur de systèmes d'aide, ou plus généralement de sites Web et applications grand public, qu'est-ce que vous choisiriez comme support de communication homme-machine (faire référence à Microssoft et la SNCF) :

- Texte + graphique (solution classique) ?
- Consultation par mots-clés, table des matières hiérarchiques, question en pseudo LN avec/sans agent conversationnel ?
- Initiatives de l'aide pour conseiller, anticiper les actions de l'utilisateur (voir Microsoft)
- Donner, outre des informations procédurales, des conseils d'utilisation dans une couleur différente ? Avez-vous regardé ces conseils (en bleu) ? Si oui, les jugez-vous utiles ?

FIG. B.49 – Schéma de l'entretien de debriefing lors de l'étude avec adaptation.

Annexe C

ACA de FT R & D : Eloona

Eloona (voir figure C.2), l'ACA fournit par France Télécom R & D est un visage de femme et de technologie ActiveX. FT R & D a choisi cette technologie dans le but d'utiliser l'ACA sur le web via des pages HTML. Afin, de comprendre le fonctionnement du moteur d'animation et du rendu sous OpenGL qui a été developpé par FT et qui tient compte d'un énoncé, il est conseillé de voir l'article suivant [Breton *et al.*, 2001]. Toutefois, pour notre étude intégrant l'ACA, nous avons intégré ce module ActiveX dans une application type Java. Les problèmes liés à cette intégration figurent dans la section "Intégration de l'ActiveX".



FIG. C.1 – ACA : Nestor.



FIG. C.2 – ACA : Eloona.

Il faut savoir que FT disposait également d'un agent gestuel nommé Nestor, voir figure C.1 (pour des raisons de contrat seul ces deux ACAs étaient à notre disposition), mais nous ne l'avons pas retenu pour plusieurs raisons :

- Tout d'abord, l'interface que nous envisagions : Flash + Système d'aide + ACA n'offrait pas assez de place pour un agent de pied en cap. Ce dernier aurait occulté la zone des messages d'aide comprenant les boutons d'activation des énoncés vocaux ainsi que les copies d'écran.
- Puis, l'agent Nestor était quant à lui beaucoup moins développé qu'Eloona en terme de rendu : très artificiel. Eloona étant plutôt réaliste favorise les émotions ainsi que les réactions affectives dans le rapport avec le participant. Nestor était également assez limité dans

ces fonctions de pointage avec le geste, nous avons ici choisi pour compenser ce manque, de diriger le regard d'Eloona vers les objets dont elle faisait référence lors de ces énoncés.
– Enfin, les énoncés oraux ont été enregistrés par une locutrice, donc le choix d'Eloona s'imposait.

Eloona, Emotions

Lorsque l'utilisateur décide d'activer un énoncé verbal, Eloona, suit un scénario général d'animation (voir figure C.3). Ce schéma varie suivant le type du message qu'Eloona énonce, les différentes catégories d'émotions sont répertoriées dans le chapitre 5, accompagnées d'une description détaillée de l'animation du clone.

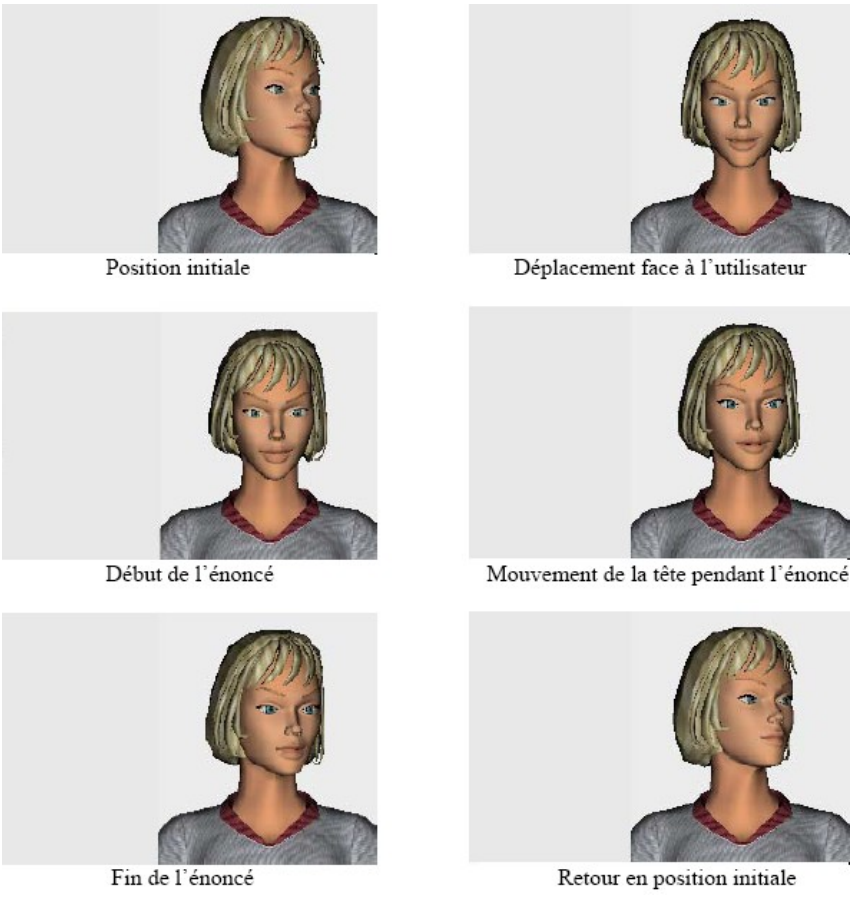


FIG. C.3 – Scénario général d'animation.

En dehors de ses interventions, Eloona ne regarde pas l'utilisateur de façon à éviter de le distraire de son intervention avec Flash ou de la consultation des graphiques affichés dans la zone des messages de la fenêtre d'aide.

Eloona, ActiveX

Nous donnons ci-après une liste des problèmes rencontrés lors de l'utilisation de l'ActiveX fourni par FT R & D (Rennes), ainsi que quelques suggestions d'amélioration.

- Mauvaise qualité de la synthèse par diphones, par conséquent, nécessité d'utiliser des fichiers WAV (environ 320 fichiers WAV et 320 fichiers smil).
- Nécessité de spécifier des chemins complets à l'ActiveX pour que celui-ci puisse jouer l'énoncé et/ou les émotions. Une erreur dans le chemin fait planter l'avatar, ce qui constitue un inconvénient majeur. Pour améliorer la portabilité et faciliter l'intégration d'un clone, il serait souhaitable de pouvoir utiliser des chemins relatifs.
- Utilisation de fichiers *.WAV : le clone plante si le format WAV est de 32 Bits PCM, il faut travailler avec le format 16 Bit PCM, ce qui peut s'avérer une limitation gênante pour certaines applications.
- Utilisation de fichiers *.smil : la balise <par> n'est pas reconnue alors que son utilisation serait un plus en permettant de jouer plus facilement les animations ; cette balise appartient pourtant au langage smil.
- Utilisation de la WebCam pour orienter le clone vers l'utilisateur : ne fonctionne pas correctement et utilise 99% du processeur.
- Observation : l'événement de fin d'énoncé n'est pas toujours réalisé lorsque le processeur est surchargé.

Remarques : Pour notre expérimentation, nous avons utilisé la balise <seq> qui marque un flux vidéo, c'est à dire l'animation qu'Eloona doit jouer, et un flux sonore, c'est-à-dire le fichier WAV qu'elle doit énoncer. L'inconvénient est qu'il faut que l'animation possède une durée similaire (un peu inférieure, sous peine de plantage) au fichier WAV. Une amélioration importante serait de pouvoir définir un "behavior" qui se cale automatiquement sur le fichier WAV de façon à éviter d'avoir à se soucier de la durée de ce behavior. Le fichier WAV est au format 16 Bits PCM ; des fichiers d'une minute (> 500 ko) ne font pas planter l'ActiveX.
L'animation quant à elle fait appel uniquement à des behaviors qui ne contiennent que des émotions ; ces émotions n'utilisent pas les commandes portant sur les mouvements de tête ou d'yeux, par exemple (head x x x). Si on utilise ces commandes, le behavior ne réalise pas toujours ce que l'on désire. On peut le vérifier en incluant des commandes Head x x x dans l'émotion Happy : lors de l'appel, celle-ci ne s'exécute pas toujours, au moins avec notre version de l'ActiveX.
Le clone, sans énoncé, consomme 15-20% de la charge du processeur. Le clone avec un énoncé et une animation 40-50% de la charge du processeur.

Table des figures

2.1 Classification générale des formes de mise en œuvre du concept d'adaptation en IHM. Voir [Simonin, Carbonell, 2004]. 21

3.1 Eye-tracker modèle 501 (cf. Annexe A). 38
3.2 Eye-head modèle 501. 38
3.3 Interface utilisateur. 39
3.4 Exemple de message dans la condition "O+G". 40
3.5 Résultat du scénario 1. 41
3.6 Résultat du scénario 2. 41
3.7 Temps passé par les sujets à la réalisation de chaque scénario (Temps en min). . . 44
3.8 Temps moyen (en min.) passé dans chaque condition (T+G et O+G) pour les deux groupes confondus. 44
3.9 Données oculométriques : Temps moyen (en min.) passé dans chaque zone, aide et Flash, tous sujets confondus. 45
3.10 Temps moyen (en min.) passé dans chacune des tâches tous sujets confondus. . . 46
3.11 Pourcentage global de réussite par condition. 47
3.12 Nombres moyens de fixations "perceptives" et "cognitives" pendant chaque condition. 47

4.1 Décor Aquarium. 54
4.2 Animation Poissons. 54
4.3 Site Web. 54
4.4 Plate-forme logicielle : Architecture. 59
4.5 Interface de lancement de l'application, côté "Utilisateur". 60
4.6 Interface du système d'aide. A gauche une configuration avec ACA (ACA + Parole + Graphique), à droite la configuration simulant l'aide en ligne adaptative (Texte + Graphique). 61

4.7 Poste "Compère". 63

4.8 Application de "Rejeu" et d'annotation. Sur la gauche, affichage extrait du rejeu de la trace d'interactions enregistrées au cours d'une passation. A droite, les menus d'assistance à l'analyse utilisés pour l'annotation. En bas de l'écran, sur la gauche, des informations temporelles sur l'activité de l'utilisateur. 65

5.1 Interface utilisateur du système d'aide. 84

5.2 Distribution des clics par zone définie lors de l'annotation (pourcentages) - Sous-groupes AE et SE du groupe GI, scénarios 1 et 2. 104

5.3 Accès aux menus. Groupe GI. 111

5.4 Origine du regard vers Eloona et direction du regard en partance d'Eloona (i.e., positions des fixations précédant et suivant celles sur le visage d'Eloona). Groupe GI (5 participants). 122

5.5 Projection des individus (Groupe GI, 14 participants) sur le plan représenté par les axes 1 et 2. 125

5.6 Nombre de fixations au cours du temps quand Eloona est inactive (Répartition par sujets). 126

5.7 Nombre de fixations au cours du temps quand Eloona énonce un message d'aide (Répartition par sujets). 127

6.1 Présentation des trois niveaux : requête initiale, versions destinées, respectivement, aux 'experts' et aux participants en difficulté - Message d'aide : "Comment créer une image-clé". 134

6.2 Le compère s'apprête à envoyer la version 2 (niveau 'expert') du message "Comment créer une image-clé" ; le niveau 1 a déjà été envoyé 1 fois (compteur présent devant l'identifiant de la version de niveau 1). 136

6.3 Pourcentage de clics par rapport aux zones définies lors de l'annotation. Comparaison avec l'expérimentation précédente. 144

6.4 Répartition des différents accès aux menus. Comparaison avec l'expérimentation précédente. 145

6.5 Projection des individus (Groupe GI et Groupes 'H' et 'F' confondus) sur le plan défini par les axes 1 et 2. 150

A.1 Description de l'eye-tracker. 161

A.2 Architecture de l'expérimentation. 162

A.3 Messages d'aide. 162

A.4 Post-test scénarios 1 et 2. 163

B.1 Premier questionnaire : Statut du participant - Page 1/2. 166

B.2 Premier questionnaire : Statut du participant - Page 2/2. 167
B.3 Eloona se présente. 168
B.4 Première diapositive du tutoriel (PowerPoint). 169
B.5 Eloona débute la présentation de l'environnement logiciel expérimental. 169
B.6 Emplacement d'Eloona pendant l'expérimentation. 170
B.7 Fenêtre d'affichage des messages d'aide. 170
B.8 Zone d'affichage du dictionnaire. 171
B.9 Zone d'interaction avec l'aide (4 boutons). 171
B.10 La fenêtre principale de Flash. 172
B.11 Eloona débute la présentation des scénarios. 173
B.12 Présentation des scénarios d'animation à réaliser. 173
B.13 Installation du casque, réglage du miroir de façon à pouvoir détecter la pupille quel que soit le point de l'écran regardé. 176
B.14 Affichage de la mire à 9 points utilisée pour calibrer l'eye-tracker. 177
B.15 Équipement de pilotage de l'eye-tracker. 177
B.16 Questionnaire verbal : Étude avec ACA - Page 1/10. 179
B.17 Questionnaire verbal : Étude avec ACA - Page 2/10. 180
B.18 Questionnaire verbal : Étude avec ACA - Page 3/10. 181
B.19 Questionnaire verbal : Étude avec ACA - Page 4/10. 182
B.20 Questionnaire verbal : Étude avec ACA - Page 5/10. 183
B.21 Questionnaire verbal : Étude avec ACA - Page 6/10. 184
B.22 Questionnaire verbal : Étude avec ACA - Page 7/10. 185
B.23 Questionnaire verbal : Étude avec ACA - Page 8/10. 186
B.24 Questionnaire verbal : Étude avec ACA - Page 9/10. 187
B.25 Questionnaire verbal : Étude avec ACA - Page 10/10. 188
B.26 Questionnaire verbal : Étude avec Adaptation - Page 1/7. 189
B.27 Questionnaire verbal : Étude avec Adaptation - Page 2/7. 190
B.28 Questionnaire verbal : Étude avec Adaptation - Page 3/7. 191
B.29 Questionnaire verbal : Étude avec Adaptation - Page 4/7. 192
B.30 Questionnaire verbal : Étude avec Adaptation - Page 5/7. 193
B.31 Questionnaire verbal : Étude avec Adaptation - Page 6/7. 194
B.32 Questionnaire verbal : Étude avec Adaptation - Page 7/7. 195
B.33 Questionnaire non verbal : Étude avec Eloona - Page 1/6. 196

B.34 Questionnaire non verbal : Étude avec Eloona - Page 2/6. 197
B.35 Questionnaire non verbal : Étude avec Eloona - Page 3/6. 198
B.36 Questionnaire non verbal : Étude avec Eloona - Page 4/6. 199
B.37 Questionnaire non verbal : Étude avec Eloona - Page 5/6. 200
B.38 Questionnaire non verbal : Étude avec Eloona - Page 6/6. 201
B.39 Questionnaire non verbal : Étude avec Eloona - Page 1/6. 202
B.40 Questionnaire non verbal : Étude avec Eloona - Page 2/6. 203
B.41 Questionnaire non verbal : Étude avec Eloona - Page 3/6. 204
B.42 Questionnaire non verbal : Étude avec Eloona - Page 4/6. 205
B.43 Questionnaire non verbal : Étude avec Eloona - Page 5/6. 206
B.44 Questionnaire non verbal : Étude avec Eloona - Page 6/6. 207
B.45 Post-test- Page 1/3. 208
B.46 Post-test- Page 2/3. 209
B.47 Post-test- Page 3/3. 210
B.48 Schéma de l'entretien de debriefing lors de l'étude avec ACA. 211
B.49 Schéma de l'entretien de debriefing lors de l'étude avec adaptation. 212

C.1 ACA : Nestor. 213
C.2 ACA : Eloona. 213
C.3 Scénario général d'animation. 214

Liste des tableaux

3.1 Principes de Mayer. 40

3.2 Répartition des sujets en fonction des situations expérimentales. 43

3.3 Pourcentages de réussite par scénario et par sujet. 46

3.4 Taux de réussite du scénario 2 réalisé pendant l'expérimentation. 48

3.5 Taux de réussite du scénario 1 réalisé pendant le post-test. 48

3.6 Taux de réussite du scénario 2 réalisé pendant le post test, par sujet. 48

3.7 Taux de réussite du scénario 2 réalisé pendant le post-test. 49

4.1 Données issues de l'annotation. 67

5.1 Répartition des étudiants en deux groupes, suivant l'ordre d'utilisation des systèmes d'aide. 87

5.2 Caractéristiques générales des participants : âge, scolarité, sexe. GI désigne le groupe des 'informaticiens', GNI celui des 'non informaticiens'. 87

5.3 Expérience informatique des participants et stratégies d'apprentissage d'un nouveau logiciel. 88

5.4 Préférences des participants pour un type d'aide en ligne parmi les deux systèmes d'aide expérimentaux et les systèmes classiques. "ACA : Sc1 et Sc2" et "ACA : Sc3" font référence à l'ordre de passation ; Sc1, 2, 3 désignent les scénarios 1, 2 et 3. 89

5.5 Jugements des participants sur l'efficacité et la convivialité des deux aides expérimentales. 91

5.6 Réponses des participants aux questions sous forme de différenciateurs sémantiques (échelle à 7 cases) portant sur l'efficacité de chaque système.
+ ou - 1 : 1 case de différence entre les deux échelles.
$>$ ou $<$ 1 : différence absolue entre les deux échelles supérieure à 1 case.
Les réponses de S15 n'ont pu être prises en compte car elles ne portent que sur un seul des deux systèmes. 93

5.7 Jugements globaux des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction.
Les 7 colonnes de droite reproduisent les échelles, les différenciateurs correspondants figurant, soit dans la colonne de gauche, soit au dessus du groupe d'échelles auquel ils s'appliquent. Le nombre reporté dans chacune des cases représente le nombre de participants qui ont sélectionné cette case en réponse à la question correspondante. 95

5.8 Jugements globaux des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction en fonction de leur expérience en informatique.
A gauche du caractère '/' le cumul des réponses du groupe GI et à droite, le cumul des réponses du groupe GNI. 97

5.9 Jugements détaillés des participants sur les effets, positifs ou négatifs, de la présence d'un ACA sur l'efficacité et la convivialité de l'interaction. 98

5.10 Jugements détaillés des participants sur l'évolution de la perception d'Eloona au cours de la session. 100

5.11 Répartition des participants du groupe GI en fonction de l'ordre de passation et durée de réalisation des deux premiers scénarios (Sc1 et Sc2). 103

5.12 Distribution des clics par type d'activité (pourcentages) - Groupe GI, scénarios 1 et 2 104

5.13 Traces des interactions des participants du groupe GI, sujet par sujet, et regroupées par condition, AE vs SE, dans la dernière colonne. 105

5.14 Nombres et pourcentages de ré-écoute de messages sonores, sujet par sujet, et regroupés par condition, AE vs SE, dans la dernière colonne (groupe GI). 107

5.15 Répartition de l'attention visuelle des participants du groupe GI dans les zones pré-définies de l'affichage et durée moyenne des fixations dans ces zones pour certaines activités de 10 participants sur 14. 109

5.16 Activité visuelle de 10 participants du groupe GI, répartie suivant trois thèmes d'activité principaux, Flash, Aide et Windows, en pourcentages (calculés par rapport à l'activité globale) et en temps (estimation en minutes). 111

5.17 Déroulement de la création des deux premières animations (scénarios Sc1 et Sc2) - Groupe GI (14 participants). 113

5.18 Jugements détaillés des participants sur la longueur et le contenu informationnel des messages d'aide. 117

5.19 Jugements verbaux des participants, regroupés en fonction de leur expérience informatique (GI/GNI), sur la 'personnalité' d'Eloona (différenciateurs sémantiques). 119

5.20 Nombre et durée moyenne (en millisecondes) des fixations oculaires sur Eloona, et durée (en secondes) d'observation d'Eloona, lorsqu'elle parle et lorsqu'elle est inactive. Groupe GI (5 participants). 121

5.21 Valeurs des 4 variables prises en compte dans l'ACP : résultats des scénarios 1 et 2 (première condition), résultats du post-test, auto-corrections, nombre de messages d'aide consultés. Groupe GI (14 participants). 124

5.22 Résultats de l'ACP calculée sur les données du tableau 5.21 : Valeurs propres, Inertie et Inertie cumulée. 124

5.23 Résultats de l'ACP : Les variables portées par chaque axe. 125

6.1 Caractéristiques générales des participants : âge, scolarité, sexe. F désigne le groupe des 'femmes' et H le groupe des 'hommes'. 137

6.2 Expérience informatique des participants : activités et stratégies de prise en main d'un nouveau logiciel. 138

6.3 Adaptation du contenu des messages - Répartition par sexe. 140

6.4 Évaluation de la stratégie d'assistance du système d'aide : l'anticipation des besoins en information de l'utilisateur. 140

6.5 Jugements sur le fonctionnement du système d'aide : Aperçu global. 141

6.6 Évaluation des messages d'aide. 141

6.7 Évaluation de la qualité des messages d'aide. 142

6.8 Répartition des participants dans les groupes F et H et durées de réalisation de la première animation. 143

6.9 Pourcentages de clics effectués au cours de la réalisation de la première animation, regroupés par thèmes. 144

6.10 Traces des interactions des participants regroupées en fonction de la distinction homme-femme (H/F). 146

6.11 Performances des participants (résultats de la création de la première animation et notes obtenues au post-test). Comparaison avec les performances des groupes GI et GNI (expérience informatique) des participants à la seconde étude. 147

6.12 Performances des participants à la seconde étude (résultats de la création de la première animation et notes obtenues au post-test) regroupés en fonction de la présence (groupe AE) ou de l'absence (groupe SE) d'Eloona. 148

6.13 Pourcentages d'inactivité des participants (souris immobile) durant la première animation (groupes H et F). Comparaison avec la seconde étude (groupes GI et GNI). 148

6.14 Résultats de l'ACP pour les participants qui ont bénéficié de l'assistance du système d'aide adaptatif. Variables portées par chacun des principaux axes. 150

Bibliographie

[André *et al.*, 1998] E. André, T. Rist, et J. Müller. Webpersona: A life-like presentation agent for the world-wide web. In *Knowledge-based Systems, 11(1)*, pages 25–36, 1998.

[André *et al.*, 1999] E. André, T. Rist, et J. Müller. Employing ai methods to control the behavior of animated interface agents. In *Applied Artificial Intelligence*, pages 415–448, 1999.

[André *et al.*, 2000] E. André, T. Rist, S. Van Mulken, M. Klesen, et S. Baldes. The automated design of believable dialogues for animated presentation teams. In *Embodied Conversational Agents*, pages 220–255, Cambridge, 2000. J. Cassell, J. Sullivan, S. Prevost,E. Chruchill (Eds.).

[Ardissono *et al.*, 2001] L. Ardissono, I. Console, et I. Torre. An adaptive system for the personalized access to news. In *AI Communications*, volume 14, pages 129–147, 2001.

[Ardissono et Goy, 2000] L. Ardissono et A. Goy. Tailoring the interaction with users in web stores. In *User Modeling and User-Adapted Interaction*, volume 10, pages 251–303, 2000.

[Baylor, 1999] A.L. Baylor. Intelligent agents as cognitive tools for education. In *Educational Technology*, pages 36–40, 1999.

[Berner et Rieger, 2003] U. Berner et T. Rieger. A scalable avatar for conversational user interfaces. In *Universal Access: Theoretical Perspectives, Practice, and Experience, LNCS 2615*, pages 350–359, Berlin, Heidelberg, New-York: Springer-Verlag, 2003. N. Carbonell, C. Stephanidis, (Eds.).

[Berthomé-Montoy et Fouet, 1995] A. Berthomé-Montoy et J-M. Fouet. Utilisation de métaconnaissances pour adapter l'aide à l'utilisateur. In *INFORSID XIII Congres*, pages 329–342, Grenoble, 1995.

[Beskow *et al.*, 2004] J. Beskow, I. Karlsson, J. Kewley, et G. Salvi. Synface: A talking head telphone for the hearing-impaired. In *Proc. 9th International Conference on Computers helping people with special needs (ICCHP 2004)*, pages 1178–1186, Irsee, Germany, 2004. Springer.

[Bevacqua et Pelachaud, 2004] E. Bevacqua et C. Pelachaud. Expressive audio-visual speech. In *Journal of Visualization and Computer Animation*, volume 15, pages 297–304, 2004.

[Billsus et Pazzani, 2000] D. Billsus et M. Pazzani. User modeling for adaptive news access. In *User Modeling and User-Adapted Interaction*, volume 10, pages 251–303, 2000.

[Bobillier-Chaumon *et al.*, 2005] M.-F. Bobillier-Chaumon, S. Carvallo, F. Tarpion-Bernard, et J. Vacherand-Revel. Adapter ou uniformiser les interactions personnes-systèmes ? In *Revue d'Interaction Homme-Machine*, volume 6, pages 91–129, 2005.

[Brajnik et Tasso, 1994] G. Brajnik et C. Tasso. A flexible tool for developing user modelling applications with non monotonic reasoning capabilities. In *International Journal of Human-Computer Studies*, volume 40, pages 31–62, 1994.

[Breton *et al.*, 2001] Gaspard Breton, Christian Bouville, et Danielle Pelé. Faceengine a 3d facial animation engine for real time applications. In *Web3D '01: Proceedings of the sixth international conference on 3D Web technology*, pages 15–22, New York, NY, USA, 2001. ACM Press.

[Buisine *et al.*, 2004] S. Buisine, S. Abrilian, et J.-C. Martin. Evaluation of multimodal behaviour of embodied agents. In *From brows to trust: Evaluating embodied conversational agents*, pages 217–238. C. Pélachaud, Z. Ruttkay (Eds.), 2004.

[Burke *et al.*, 1997] R.D. Burke, K.J. Hammond, et B.C. Young. The findme approach to assisted browsing. In *IEEE Expert*, volume 12, pages 32–40, 1997.

[Calvary *et al.*, 2005] G. Calvary, O. Daassi, J. Coutaz, et A. Demeure. Des widgets aux comets pour la plasticité des systèmes interactifs. In *Revue d'Interaction Homme-Machine*, volume 5, pages 33–53, 2005.

[Capobianco et Carbonell, 2003] A. Capobianco et N. Carbonell. Online help for the general public: specific design issues and recommendations. In *Universal Access in the Information Society Journal (UAIS)*, volume 2/3, pages 265–279. Special issue 'Countering design exclusion', 2003.

[Capobianco, 2002a] A. Capobianco. Demandes d'aide procédurale: spécification des besoins d'utilisateurs novices. In *IHM '02: Proceedings of the 14th French-speaking conference on Human-computer interaction (Conférence Francophone sur l'Interaction Homme-Machine)*, pages 57–64, New York, NY, USA, 2002. ACM Press.

[Capobianco, 2002b] A. Capobianco. *Stratégies d'aide en ligne contextuelles : acquisition d'expertises, modélisation et évaluation expérimentale.* PhD thesis, Université Henry Poincaré Nancy 1, Nancy, France, 2002.

[Carroll *et al.*, 1987] J.M. Carroll, P.L. Smith-Kerker, J.R. Ford, et S.A. Mazur-Rimetz. The minimal manual. In *Human-Computer Interaction*, volume 3(2), pages 123–153, 1987.

[Carroll et Mack, 1984] John M. Carroll et Robert L. Mack. Learning to use a word processor: by doing, by thinking, and by knowing. pages 13–51, 1984.

[Carroll et Rosson, 1987] J.-M. Carroll et M.B. Rosson. Paradox of the active user. In *Interfacing thought: cognitive aspects of human-computer interaction*, pages 81–111, Cambridge, 1987. J.M. Carroll.

[Cassell *et al.*, 2000] J. Cassell, T. Bickmore, L. Camphell, H. Vilhjalmsson, et H. Yan. The human conversation as a system framework: Designing embodied conversational agents. In *Embodied Conversational Agents*, pages 29–63, Cambridge, 2000. J. Cassell, J. Sullivan, S. Prevost, E. Chruchill (Eds.).

[Cerrato, 2002] L. Cerrato. A comparison between feedback strategies in human-to-human and human-machine communication. In *Proc. ICSLP'02*, pages 557–560, 2002.

[Chou *et al.*, 2003] C.-Y. Chou, T.-W. Chan, et C.-J. Lin. Redefining the learning companion: the past, present, and future of educational agents. In *Computers & Education*, pages 255–269, 2003.

[Conati *et al.*, 1997] C. Conati, A.S. Gertner, K. VanLehn, et M.J. Druzdzel. On-line student modeling for coached problem solving using bayesian networks. In *Proc. 6th Int. Conference on User Modeling (UM'97)*, pages 231–242, Chia Laguna, Italy, Wien - New York: Springer, 1997. A. Jameson, C. Paris, C. Tasso (Eds.).

[Dam, 1997] A. Van Dam. Post-WIMP user interfaces. *Communications of the ACM*, 40(2):63–67, 1997.

[Darves *et al.*, 2002] C. Darves, S. Oviatt, et R. Coulston. The impact of embodiment on animated character design. In *First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'02)*, page 5 pp, Bologna, 2002. New York: ACM Press.

[Darves et Oviatt, 2004] C. Darves et S. Oviatt. Talking to digital fish: Designing effective conversational interfaces for educational software. In *From brows to trust:Evaluating embodied conversational agents*, page chapter 10. Pélachaud, Z. Ruttkay (Eds.), 2004.

[Dieterich *et al.*, 1993] H. Dieterich, U. Malinowski, T. Kühme, et M. Schneider-Hufschmidt. State of the art in adaptive user interfaces. In *Adaptive User Interfaces, Principles and Practice*, pages 13–48, 1993.

[Duffy *et al.*, 1989] T. M. Duffy, B. Mehlenbacher, et J. Palmer. The evaluation of online help systems: a conceptual model. pages 362–387, 1989.

[Duffy *et al.*, 1992] T. M. Duffy, B. Mehlenbacher, et J. E. Palmer. Online help: Design and evaluation. 1992.

[Dufresne, 1992] A. Dufresne. Ergonomie cognitive, hypermédias et apprentissage. In *"Hypermédias et apprentissage", Actes des premières journées scientifiques Hypermédias et Apprentissage*, pages 121–132, Chatenay-Malabry, Paris, 1992. B. De Lapassardière et G.L. Baron.

[Edmonds, 1987] E.A. Edmonds. Adaptability response and knowledge. In *Knowledge Based Systems*, volume 1, pages 3–10, 1987.

[Eliëns *et al.*, 2002] A. Eliëns, Z. Huang, J. Hoorn, et C. Visser. A platform for embodied conversational agents based on distributed logic programming. In *AAMAS'02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'*, Bologna,Italy, 2002.

[Fink et Kobsa, 2000] J. Fink et A. Kobsa. A review and analysis of commercial user modeling servers for personalization on the world wide web. In *User Modeling and User-Adapted Interaction*, volume 10, pages 209–249, 2000.

[Gentner et Nielsen, 1996] D. Gentner et J. Nielsen. The anti-mac interface. In *Communications of the ACM*, volume 39, pages 70–82, 1996.

[Grudin, 1993] J. Grudin. *Obstacles to participatory design in large product development organizations*, pages 99–119. Lawrence Erlbaum Associates, 1993.

[Guinn et Hubal, 2004] C. Guinn et R. Hubal. An evaluation of virtual human technology in informational kiosks. In *Proc. 6th International Conference on Multimodal Interfaces (ICMI'04)*, pages 297–301, State College, PA, 2004.

[Gunes *et al.*, 2004] H. Gunes, M. Piccardi, et T. Jan. Bimodal modelling of facial and upper-body gesture for affective hci. In *Proc. Annual Conference of the Human Factors and Ergonomics Society (OZCHI 2004)*, Wollongong, Australia, 2004.

[Höök *et al.*, 1999] K. Höök, P. Persson, J. Holm, K. Tullgren, M. Sjölinder, et J. Karlgren. Spatial or narrative: A study of the agneta & frida system. In *Workshop on Affect in Interactions: Towards a New Generation of Interfaces*, conference in Siena, 1999.

[Hämäläinen et Höysniemi, 2003] P. Hämäläinen et J. Höysniemi. A computer vision and hearing based user interface for a computer game for children. In *AAMAS'02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'*, pages 299–318, Berlin, Heidelberg, New-York: Springer-Verlag, 2003. N. Carbonell, C. Stephanidis, (Eds.), Universal Access: Theoretical Perspectives, Practice, and Experience, LNCS 2615.

[Horvitz *et al.*, 1998] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, et K. Rommelse. The lumiere project: Bayesian user modeling for inferring the goals and needs of software users. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, Madison, WI*, pages 256–265, San Francisco, CA: Morgan Kaufmann, 1998. G.F. Cooper, S. Moral (Eds.).

[Jameson, 2003] A. Jameson. *Human-Computer Interaction Handbook, Adaptive Interfaces and Agents*, chapitre 15, pages 305–330. J. Jacko, A. Sears, Mahwah, NJ: Erlbaum, 2003.

[Jameson, 2006] A. Jameson. Adaptive interfaces and agents. In *Human-computer interaction handbook (2nd ed.)*, Mahwah, NJ: Erlbaum, 2006. J. Jacko & A. Sears (Eds.).

[Jamet et Heurley, 2002] E. Jamet et L. Heurley. L'apport des nouvelles technologies de l'information et de la communication dans la conception des documents techniques. In *Psychologie française*, volume 47, pages 33–40, INIST-CNRS, 2002. Elsevier, Paris, FRANCE (1956) (Revue).

[Kearsley *et al.*, 1988] G. Kearsley, R. L. Campbell, J. Elkerton, W. Judd, et J. Walker. Online help systems: design and implementation issues (panel). In *CHI '88: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 287–288, New York, NY, USA, 1988. ACM Press.

[Kleinberg et Sandler, 2004] J. Kleinberg et M. Sandler. Using mixture models for collaborative filtering. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 569–578, Chicago, IL, New York: ACM Press, 2004.

[Kopp *et al.*, 2004] S. Kopp, P. Tepper, et J. Cassell. Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proc. 6th International Conference on Multimodal Interfaces (ICMI'04)*, pages 97–104, State College, PA, 2004.

[Langley et Fehling, 1998] P. Langley et M. Fehling. The experimental study of adaptive user interfaces (technical report 98-3). In *Institute for the Study of Learning and Expertise*, Palo Alto, CA, 1998.

[Langley, 1992] P. Langley. on user modeling . banff, alberta: Springer. 18 an adaptive conversational interface, 1992.

[Lesh *et al.*, 2004] N. Lesh, J. Marks, C. Rich, et C.L. Sidner. Man-computer symbiosis revisited: Achieving natural communication and collaboration with computers. *IEICE Transactions on Electronics*, 2004.

[Lester *et al.*, 1997] J. Lester, S. Converse, S. Kahler, T. Barlow, B. Stone, et R. Bhogal. The persona effect: Affective impact of animated pedagogical agents. In *Proc. ACM International Conference on Human Factors in Computing Systems (CHI'97)*, pages 359–366, Atlanta, 1997.

[Lester *et al.*, 1999] J. Lester, J.L. Voerman, S.G. Towns, et C.B. Callaway. Deictic believability: Coordinated gesture, locomotion and speech in lifelike pedagogical agents. In *Applied Artificial Intelligence*, pages 383–414, 1999.

[Lieberman et Maulsby, 1996] H. Lieberman et D. Maulsby. Instructible agents: software that just keeps getting better. In *IBM Systems Journal*, volume 35, pages 539–556, 1996.

[MacDorman *et al.*, 2005] K.F. MacDorman, T. Minato, M. Shimada, S. Itakura, S. Cowley, et H. Ishiguro. Assessing human likeness by eye contact in an android testbed. In *Proc. 27th Annual Conference of the Cognitive Science Society (COGSCI 2005)*, Stresa, 2005.

[Maes, 1994] P. Maes. Agents that reduce work and information overload. In *Communications of the ACM*, volume 37, pages 30–40, 1994.

[Malinoski *et al.*, 1992] U. Malinoski, T. Kühme, Hartmut Dieterich, et M. Schneider-Hufschmidt. A taxonomy of adaptive user interfaces. In *People and Computers VII*, pages 391–414, The British Computer Society, London, 1992. Monk, Diaper and Harrison.

[Massaro *et al.*, 1998] D.W. Massaro, M.M. Cohen, J. Beskow, S. Daniel, et R.A. Cole. Developing and evaluating conversational agents. In *Proc. First Workshop on Embodied Conversational Characters (WECC)*, pages 137–147, Granlibakken Lake Tahoe, CA, 1998.

[Massaro et Stork, 1998] D.W. Massaro et D.G. Stork. Speech recognition and sensory integration. In *American Scientist, 86*, pages 236–244, 1998.

[Mayer, 1989] Richard E. Mayer. Systematic thinking fostered by illustrations in scientific text. In *Journal of Educational Psychology*, volume 81, pages 240–246, 1989.

[Medina-Medina *et al.*, 2002] N. Medina-Medina, L. Garcia-Cabrera, J. Rodriguez-Fortiz, et J. Parets-Llorca. Adaptation in an evolutionary hypermedia system: Using semantic and petri nets. In *AH '02: Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 284–295, London, UK, 2002. Springer-Verlag.

[Mitchell *et al.*, 1994] T. Mitchell, R. Caruana, D. Freitag, J. McDermott, et D. Zabowski. Experience with a learning personal assistant. In *Communications of the ACM*, volume 37, pages 81–91, 1994.

[Mori, 1970] M. Mori. Bukimi no tani [the uncanny valley]. In *Energy 7*, pages 33–35, 1970.

[Nanard, 1990] J. Nanard. *La Manipulation Directe en Interface Homme-Machine.* PhD thesis, Université Montpellier II, Montpellier, France, 1990.

[Nielsen, 1993] J. Nielsen. *Usability engineering.* Academic Press, Boston, 1993.

[Oppermann, 2002] R. Oppermann. Introduction. In *Adaptive User Support*, pages 1–13, Hillsdale, New Jersey, 2002. Oppermann R.

[Pagel *et al.*, 1996] V. Pagel, N. Carbonell, et Y. Laprie. "a new method for speech delexicalization, and its application to the perception of french prosody". In *Proc. 4th International Conference on Spoken Language Processing (ICSLP'96), Vol. 2*, pages 821–824, Philadelphia, PA, 1996.

[Paiva *et al.*, 1994] A. Paiva, J. Self, et J.R. Hartley. On the dynamics of learner models. In *Proceedings of European Conference on Artificial Intelligence (ECAI'94)*, pages 178–182, Amsterdam, 1994.

[Paiva *et al.*, 2003] A. Paiva, M. Costa, R. Chaves, M. Piedade, D. Mourao, D. Sobral, K. Höök, G. Andersson, et A. Bullock. Sentoy: an affective sympathetic interface. In *International Journal of Human-Computer Studies, 59(1)*, pages 227–235. N. Carbonell, C. Stephanidis, (Eds.), Universal Access: Theoretical Perspectives, Practice, and Experience, LNCS 2615, 2003.

[Paiva et Self, 1994] A. Paiva et J. Self. A learner model reason maintenance system. In *Proceedings of European Conference on Artificial Intelligence (ECAI'94)*, pages 193–196, Amsterdam, 1994.

[Payr, 2003] S. Payr. The virtual university's faculty: an overview of educational agents. In *Applied Artificial Intelligence*, pages 1–19, 2003.

[Pélachaud *et al.*, 1996] C. Pélachaud, N.I. Badler, et M. Steedman. *Generating facial expressions for speech*, pages 1–46. Cognitive Science, 20, 1996.

[Pélachaud *et al.*, 2002] C. Pélachaud, V. Carfiglio, B. De Carolis, et F. De Rosis. Embodied contextual agent in information delivering application. In *First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'02)*, pages 758–765, Bologna, 2002. New York: ACM Press.

[Prendinger *et al.*, 2005] H. Prendinger, J. Mori, et M. Ishizuka. Recognizing, modeling, and responding to users' affective states. In *Proceedings of 10th International Conference on User Modeling (UM 2005)*, pages 60–69, Edinburgh, UK, Vienna, New York: Springer, 2005. L. Ardissono, P. Bra, A. Mitrovic (Eds.).

[Rickel et Johnson, 1999] J. Rickel et W.L. Johnson. Animated agents for procedural training in virtual reality: perception, cognition and motor control. In *Applied Artificial Intelligence*, pages 343–382, 1999.

[Ringle et Halstead-Nussloch, 1989] M.D. Ringle et R. Halstead-Nussloch. Shaping user input: A strategy for natural language dialogue design. *Interacting with Computers*, 1(3):227–244, 1989.

[Rist *et al.*, 2003] T. Rist, E. André, et S. Baldes. A flexible platform for building applications qith life-like characters. In *Proc. ACM International Conference on Intelligent User Interfaces (IUI'03)*, pages 158–165, Miami, 2003.

[Rossi, 1993] M. Rossi. A model for predicting the prosody of spontaneous speech (pps model). In *Speech Communication, 13*, pages 87–107, 1993.

[Ruttkay *et al.*, 2002] Zs. Ruttkay, C. Dormann, et H. Noot. Evaluating ecas - what and how? In *AAMAS'02 Workshop on 'Embodied conversational agents - let's specify and evaluate them!'*, Bologna, Italy, 2002.

[Salvendy *et al.*, 2002] G. Salvendy, L. Rothrock, R. Koubek, F. Fuchs, et M. Haas. Review and reappraisal of adaptive interfaces: Toward biologically-inspired paradigms. In *Theoretical Issues in Ergonomic Science*, 2002.

[Schäfer et Weyrath, 1997] R. Schäfer et T. Weyrath. Assessing temporally variable user properties with dynamic bayesian networks. In *Proceedings of the Sixth International Conference on User Modeling (UM'97)*, pages 377–388, Chia Laguna, Sardinia, Italy, Vienna, New York: Springer, 1997. A. Jameson, C. Paris, C. Tasso.

[Schneiderman, 2000] B. Schneiderman. *Designing the User Interface - Strategies for Effective Human-Computer Interaction.* Addison-Wesley, 2000.

[Sears et Shneiderman, 1994] A. Sears et B. Shneiderman. Split menus: Effectively using selection frequency to organize menus. In *ACM Transactions on Computer-Human Interaction*, volume 1, pages 27–51, 1994.

[Segal et Kephart, 1999] R. Segal et J. Kephart. An intelligent assistant for organizing e-mail. In *Proceedings of 3rd International Conference on Autonomous Agents (Agents'99)*, Seattle, WA, 1999.

[Semeraro *et al.*, 2003] G. Semeraro, H.H.K. Andersen, V. Andersen, P. Lops, et F. Abbattista. Evaluation and validation of a conversational agent embodied in a bookstore. In *Universal Access: Theoretical Perspectives, Practice, and Experience, LNCS 2615*, pages 360–371, Berlin, Heidelberg, New-York: Springer-Verlag, 2003. N. Carbonell, C. Stephanidis, (Eds.).

[Shaw *et al.*, 1999] E. Shaw, W.L. Johnson, et R. Ganeshan. Pedagogical agents on the web. In *Third International Conference on Autonomous Agents*, pages 283–290, 1999.

[Shneiderman, 1983] B. Shneiderman. Direct manipulation: A step beyond programming languages. In *IEEE Computer*, volume 16, pages 57–69, 1983.

[Shneiderman, 1987] B. Shneiderman. *Designing the user interface: strategies for effective human-computer interaction.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1987.

[Shneiderman, 1993] B. Shneiderman. *Sparks of Innovation in Human-Computer Interaction.* Greenwood Publishing Group Inc., Westport, CT, USA, 1993.

[Shneiderman, 1999] B. Shneiderman. Universal usability: Pushing human-computer interaction research to empower every citizen. Rapport Technique CS-TR-4043, 1999.

[Sidner *et al.*, 2004] C.L. Sidner, C. Lee, C. Kidd, et N. Lesh. Explorations in engagement for humans and robots. In *Proc. IEEE RAS/RSJ International Conference on Humanoid Robots*, Los Angeles, 2004.

[Sidner et Lee, 2005] C.L. Sidner et C. Lee. *Robots as Laboratory Hosts*, pages 24–25. Interactions, 12(2), 2005.

[Smith et Gordon, 1997] R.W. Smith et S.A. Gordon. Effects of variable initiative on linguistic behavior in human-computer spoken natural language dialog. In *Computational Linguistics*, pages 141–168, 1997.

[Stephanidis, 2001] C. Stephanidis. Adaptive techniques for universal access. In *User Modeling and User-Adapted Interaction*, volume 11, pages 159–179, 2001.

[Thévenin *et al.*, 2003] D. Thévenin, J. Coutaz, et G. Calvary. A reference framework for the development of plastic user interfaces. In *Multi-Device and Multi-Context User Interfaces: Engineering and Applications Frameworks*, pages 29–51, Chichester, UK: John Wiley & Sons, Ltd, 2003. A. Seffah, H. Javahery.

[Totterdell *et al.*, 1987] P. A. Totterdell, Michael A. Norman, et D. P. Browne. Levels of adaptivity in interface design. In *INTERACT 87 - 2nd IFIP International Conference on Human-Computer Interaction September 1-4*, pages 715–722, Stuttgart, Germany, 1987. Bullinger, Hans-Jorg and Shackel, Brian.

[Vanderdonckt *et al.*, 2005] J. Vanderdonckt, D. Grolaux, P. Van Roy, Q. Limbourg, B. Macq, et B. Michel. A design space for context-sensitive user interfaces. In *In Proceedings of 14th International Conference on Intelligent and Adaptive Systems and Software Engineering (IASSE 2005)*, pages 207–214, Toronto, Canada, 2005.

[Vaudry, 2002] C. Vaudry. *Composition Dynamique d'Informations dans le Cadre de la Communication Homme Machine - La problématique de la Pertinence dans la CHM.* PhD thesis, UNIVERSITE MONTPELLIER II (sciences et techniques du Languedoc), 12 2002. tel-00003610.

[Weber et Specht, 1997] G. Weber et M. Specht. User modeling and adaptive navigation support in www-based tutoring systems. In *User Modeling: Proceedings of the Sixth International Conference, Modeling (UM'97)*, pages 289–300, Chia Laguna, Sardinia, Italy, Vienna, New York: Springer, 1997. A. Jameson, C. Paris, C. Tasso (Eds.).

Résumé

La thèse vise à évaluer de nouvelles formes d'interaction Homme-Machine. Plus précisément, la thématique générale de ce travail, divisée en deux axes, vise à mettre en œuvre et à évaluer l'apport d'*Agents Conversationnels Animés* (ACAs) et d'*Interfaces dites "Adaptatives"* à l'utilisation d'un logiciel. Ainsi, comportements et réactions d'utilisateurs sont recueillis à l'aide de méthodes ergonomiques et techniques d'eye-tracking.

Une approche expérimentale a été adopté afin d'évaluer l'apport de chaque axes. Pour cela, des participants (étudiants de niveau Licence) ont manipulé un logiciel de création d'animation qui leur était inconnu (Flash) afin de réaliser trois scénarios. Tout au long de leur découverte du logiciel, les participants étaient accompagnés d'un dispositif d'aide intégrant, suivant l'expérimentation, un ACA (fourni par FT R & D) ou une technique d'adaptation (détection d'intention et évolution suivant les connaissances).

Les différentes études réalisées montrent que les deux sources d'innovation employées ont été perçues positivement par la majorité des participants. Elles ont montré d'autre part qu'un ACA a un effet rassurant et qu'il peut vraisemblablement être utilisé lors de la prise en main d'un logiciel. Pour le système adaptatif, le fait que le système évolue de manière autonome n'a pas perturbé les participants, mais n'améliore guère les performances.

Mots-clés: Multimodalité, Aide en ligne, Agent conversationnel animé, Dispositif de suivi du regard, Interfaces adaptatives, Interaction Homme-Machine

Abstract

The goal of the thesis is to evaluate new forms of Human-Computer Interaction. More precisely, the general set of themes of this work, divided into two parts at aims implementing and evaluating the contribution of Embodied Conversational Agent (ECAs) and of adaptive interfaces for the use of a software. Thus, behaviors and reactions of users are collected using ergonomic methods and eye-tracking technologies.

An experimental approach was adopted in order to evaluate the contribution of each parts. For that, participants (student of bachelor's degree) handled a software of animation's creation which was unknown for them (Flash) in order to carry out three scenarios. Throughout their exploration of the software, participants were accompanied by a help system. In the first experiment an ECA (provided by FT R & D) enunciate help messages ; in the second one a adaptive system (detection of intention and evolution according to knowledge) was used.

The various studies carried out show that the two innovations employed were perceived positively by the majority of the participants. They showed in addition that a ECA has a reassuring effect and that it can probably be used in first experiences with a software. For the adaptive system, the fact that the system evolves in an autonomous way did not disturb the participants, but hardly improves the performances.

Keywords: Multimodality, Help online, Embodied Conversational Agent, Eye tracker, Adaptive interfaces, Human-Computer Interaction