

UNIVERSITÉ DE MONTRÉAL

ÉTUDE DE L'INTÉGRATION DE LA RECONNAISSANCE VOCALE
AU POSTE D'OPÉRATION DE L'INSTRUCTEUR EN SIMULATEUR
DE VOL

JOËL MIGNEAULT
DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE INDUSTRIEL)
AVRIL 2008



Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-41570-2
Our file Notre référence
ISBN: 978-0-494-41570-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

ÉTUDE DE L'INTÉGRATION DE LA RECONNAISSANCE VOCALE
AU POSTE D'OPÉRATION DE L'INSTRUCTEUR EN SIMULATEUR
DE VOL

présenté par : MIGNEAULT Joël

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. GAGNON Michel, ing., Ph.D., président

M. ROBERT Jean-Marc, Doctorat, membre et directeur de recherche

M. DESMARAIS Michel C., Ph.D., membre et codirecteur de recherche

M. KRUK Ronald V., Ph.D., membre

*Comment dédier le résultat d'autant de labeur à certaines personnes
particulières alors qu'elles ont été si nombreuses à me soutenir
tout au long de mes études.*

*Une pensée toute spéciale pour mes grand-parents
et mes parents qui ont toujours cru en moi.*

REMERCIEMENTS

Je tiens à remercier la compagnie CAE Inc. pour son support financier ainsi que pour les ressources humaines et matérielles qui m'ont permis de réaliser ce projet. Mes remerciements vont aussi au CRSNG pour son support financier via la bourse d'études supérieures à incidence industrielle sans laquelle le projet n'aurait pas pu être d'une aussi grande envergure.

Je remercie M. Sylvain Caron pour sa supervision du projet, ainsi que les concepteurs d'interfaces du poste d'opération, les ingénieurs de son, les ingénieurs en communications et les instructeurs de vol qui ont participé en donnant de leur temps, soit en tant que sujet participant à une expérimentation, où en tant qu'expert consulté à certaines étapes du projet. Merci à M. Yanik Guilbault et M. David Bowness pour leur support et leur disponibilité constante tout au long du projet. Enfin, j'ai une pensée toute spéciale pour M. Sébastien Malo, spécialiste en facteurs humains, pour son expertise dans le domaine, son support moral et son amitié.

Je remercie également mon directeur de recherche, M. Jean-Marc Robert, qui a su me transmettre sa passion pour l'ergonomie, qui m'a poussé à me dépasser et qui m'a inculqué les fondements de la méthodologie de recherche. Je remercie dans cette même pensée mon codirecteur de recherche, M. Michel C. Desmarais, pour sa générosité et son implication puisque c'est grâce à lui s'il y a eu partenariat avec la compagnie CAE pour la réalisation de ce projet.

Je remercie finalement ma copine, mes amis, mes parents et ma famille pour leurs encouragements si nombreux. Ils ont su me faire persévérer jusqu'à l'atteinte de mes objectifs.

RÉSUMÉ

L'entraînement des pilotes en simulateur de vol est très onéreux : entre 700 et 1 000 dollars de l'heure. C'est le prix à payer pour les transporteurs aériens qui n'ont d'autres choix que de former et de tenir leurs pilotes à jour. Augmenter l'efficacité de l'instructeur de vol en simulateur pour contrôler le simulateur à l'aide du poste d'opération lui donnerait plus de temps pour instruire et observer les pilotes pendant l'entraînement. Le poste d'opération a grandement évolué au fil du temps et plusieurs recherches ont porté sur l'amélioration de son utilisabilité. Cependant, peu de chercheurs se sont intéressés à l'intégration de la reconnaissance vocale comme moyen d'atteindre cet objectif d'utilisabilité. Plusieurs études menées dans les domaines militaire et de l'aviation montrent le potentiel de cette technologie et les bénéfices qu'elle peut apporter pour divers opérateurs. En vue d'évaluer la pertinence d'utiliser cette technologie au poste d'opération de l'instructeur en simulateur de vol, une analyse des facteurs humains en jeu lors de l'interaction avec cette technologie doit être faite.

Dans ce mémoire, nous effectuons diverses études au sein de la compagnie CAE Inc. (Montréal) dans le but de déterminer les facteurs humains qui ont un impact sur l'intégration de la reconnaissance vocale au poste d'opération de l'instructeur. Nous proposons d'abord deux profils d'instructeurs de vol, élaborés sur la base d'entrevues individuelles avec des instructeurs de vol. Nous avons identifié et analysé les tâches correspondant au travail de l'instructeur au poste d'opération.

Nous avons ensuite comparé les temps d'exécution des tâches entre une interaction manuelle et une interaction vocale. L'utilisation d'un modèle de performance humaine lors d'interactions manuelles avec une interface graphique, de la méthode d'analyse de tâche *KLM-GOMS*, des données sur

la vitesse de la parole provenant de la littérature et la réalisation d'une expérimentation sur la vitesse de la parole, nous permettent de comparer les deux modes d'interaction de manière peu coûteuse, puisqu'aucun instructeur n'est nécessaire. Cette comparaison montre que l'interaction vocale est plus rapide que l'interaction manuelle et qu'en moyenne un gain de 33% en temps d'exécution de certaines tâches est obtenu.

Nous avons réalisé une étude faisant appel à la technique du Magicien d'Oz et à la méthode Delphi pour recueillir les formulations naturelles des commandes vocales utilisées par six instructeurs de vol pour 30 énoncés décrivant un but à atteindre. Des facteurs qui influent sur la facilité d'apprentissage et d'utilisation des langages de commandes tels que la congruence, la hiérarchisation cohérente et un couplage fort ont été tirés d'études sur les langages de commandes textuelles. Les données recueillies dans notre expérimentation montrent entre autres que les instructeurs utilisent des verbes dans 73% des commandes (principalement en début de commande) et que l'ordre paramètre-valeur dans la formulation des commandes apparaît dans 89% des commandes (ordre valeur-paramètre pour les 11% restant). Les résultats de recherches menées dans le domaine des langages de commandes textuelles et les résultats obtenus nous ont permis de modéliser un langage de 78 commandes vocales qui doit être intégré au système de reconnaissance du poste d'opération de l'instructeur.

Nous avons ensuite analysé le contexte technologique de l'intégration du système de reconnaissance vocale au poste d'opération. Une analyse basée sur l'exactitude et l'efficacité de cinq engins de reconnaissance disponibles sur le marché permet de dégager le meilleur parmi ce lot. L'exactitude et l'efficacité de ce dernier sont de 97.5% et 95.8% lorsque l'engin est configuré à reconnaître 20 commandes, et de 77.5% et 55% lorsqu'il est configuré pour reconnaître 137 commandes. Ce dernier est par la suite testé à l'aide

d'enregistrements vocaux obtenus en milieu bruyant. Les résultats montrent que le meilleur engin voit son exactitude diminuer jusqu'à 40% lorsque le volume du bruit ambiant est au maximum (environ 88 dB), mais demeurer au-dessus de 65% lorsque le volume est diminué de moitié.

Nous avons ensuite comparé différents dispositifs d'activation/désactivation de la reconnaissance vocale qui sont nécessaires pour le poste d'opération. Les résultats montrent qu'une pédale ou un bouton-pression manuel est plus rapide qu'un bouton graphique à deux états présenté sur l'écran tactile ou un mini-levier. Cependant, un mot-clé dit avant une commande vocale est préféré par trois des cinq sujets ayant exécuté le scénario d'entraînement de vol.

Enfin, un prototype de système de reconnaissance vocale a été intégré à une interface graphique du poste d'opération et des tests d'utilisabilité ont été menés avec deux sujets afin de mesurer le taux de reconnaissance des commandes vocales et la satisfaction des utilisateurs. Nous avons observé deux erreurs de reconnaissance pour le premier sujet et huit pour le second sujet. La performance perçue par les sujets est de 32% et 8% respectivement. Les autres facteurs de satisfaction évalués soit : l'appréciation, la charge cognitive, l'ennui, le contrôle et la rapidité sont en général au dessus de 60%. Ces résultats montrent qu'une technologie vocale peut être intégrée au poste d'opération afin d'augmenter l'efficacité des instructeurs et révèlent que les sujets sont en général satisfaits du mode d'interaction proposé. Même si d'autres données devront être récoltées afin de quantifier plus exactement le gain d'efficacité en situations réelles d'entraînement, les études effectuées dans le cadre de ce mémoire apportent des données qui nous permettent de recommander l'intégration de ce mode d'interaction au poste d'opération de l'instructeur en simulateur de vol.

ABSTRACT

Flight simulator training is very expensive. Airlines must pay between 700 and 1 000 canadian dollars per hour to keep their pilots trained in a simulator. Increasing flight instructor's efficiency at controlling the operating station would leave him more time for instruction and observation tasks. Many studies have found ways to improve the operating station usability. However, few have seen the potential of voice recognition for this purpose. Studies in the military and avionic domains show that voice recognition technology has interesting benefits for various types of operators. In order to bring these benefits to the instructor's operating station, a human factors analysis is required for voice recognition to be integrated in this particular context.

This memoir shows the results of many studies made at CAE Inc. (Montreal) that establish the basis of human factors that impact on voice recognition at the instructor's operating station. We propose two flight instructor profiles built from individual interviews with five flight instructors. Studies that define optimized work of flight instructors at their operating station are used to create a model of their tasks. An analysis comparing execution time performances for manual and voice inputs is conducted. Speech rate data from various studies and a human performance model of manual inputs, based on the *KLM-GOMS* task analysis method, are used as a low cost method to compare these interaction modes. Results show 33% gain in task execution time using voice input. Then, a study that uses the Wizard of Oz technique and the Delphi method is conducted, giving us six instructors' natural voice commands for 30 tasks. A review of studies on text commands showed that congruence, consistent hierarchies and high coupling are factors influencing ease of learning and ease of use. From these results and the results obtai-

ned in our experiment, we build a language model of 78 voice commands for the operating station. Technologies needed for voice recognition at the operating station are then analysed. Five different voice recognition engines are compared based on accuracy and efficiency. The best engine found is then used for voice recognition analysis in noisy environment. Results show that no more than 40% accuracy can be reached at maximum noise volume while more than 65% is reached when noise is half reduced. We also compare different activation/deactivation devices that are essential for our context. Results show that a pedal or a push button are faster and more efficient than a software button or a toggle switch. However, a keyword said before a voice command is preferred by three out of five subjects that went through the training scenario to compare these devices. Finally, a voice recognition prototype is integrated at the operating station and usability tests are performed with two subjects. Voice recognition accuracy and user satisfaction are measured. Results show that the prototype still have some flaws and a user satisfaction questionnaire for voice systems indicates that subjects are generally satisfied. All the results show that the instructor's efficiency can be increased with this technology. However, more data needs to be gathered from real training situations in order to conclude about the real impact of voice recognition at the instructor's operating station. Human factors will always be important aspects to take into account in the process of integrating this new technology in the flight simulation training context.

TABLE DES MATIÈRES

DÉDICACE	iv
REMERCIEMENTS	v
RÉSUMÉ	vi
ABSTRACT	ix
TABLE DES MATIÈRES	xi
LISTE DES TABLEAUX	xvi
LISTE DES FIGURES	xviii
LISTE DES SIGLES ET ABRÉVIATIONS	xx
LISTE DES ANNEXES	xxi
INTRODUCTION	1
CHAPITRE 1 L'ENTRAÎNEMENT EN SIMULATEUR DE VOL ET LA RECONNAISSANCE VOCALE	7
1.1 Entraînement en simulateur de vol	7
1.1.1 Instructeurs	8
1.1.2 Poste d'opération de l'instructeur de vol	8
1.1.2.1 Fonctionnalités du poste d'opération	11
1.2 Reconnaissance vocale	19
1.2.1 Le fonctionnement des systèmes de reconnaissance vocale	20
1.2.2 Applications à des domaines connexes	22
1.2.2.1 Domaine avionique	22
1.2.2.2 Domaine militaire	23
1.2.3 Aspects contraignant l'intégration de la reconnais- sance vocale	27
1.2.3.1 Perturbation de la communication d'équipe	28
1.2.3.2 Limites de la mémoire et des tâches	28
1.2.3.3 Limites imposées par les standards	30
1.3 Sommaire	32

CHAPITRE 2	PROBLÉMATIQUE	SIMULATEUR-	
	INSTRUCTEUR ET OBJECTIF DE L'ÉTUDE		33
2.1	Problématique		33
2.2	Objectif		35
CHAPITRE 3	ANALYSE DES CARACTÉRISTIQUES DES INS-		
	TRUCTEURS DE VOL AU POSTE D'OPÉRATION		37
3.1	Élaboration d'un profil utilisateur		37
3.1.1	Méthodologie		38
3.1.1.1	Sujets		38
3.1.1.2	Outils utilisés		38
3.1.1.3	Procédures et consignes		39
3.1.2	Résultats		39
3.1.3	Description des profils		42
3.2	Sommaire		46
CHAPITRE 4	ANALYSE ERGONOMIQUE DES TÂCHES DE		
	L'INSTRUCTEUR DE VOL		47
4.1	Modélisation des tâches		47
4.2	Analyse des temps d'exécution des tâches selon deux modes d'interaction		49
4.2.1	Revue des études qui comparent interactions tactile et vocale		49
4.2.2	Méthodologie		51
4.2.2.1	Modèles prédictifs de performance humaine		51
4.2.2.2	Estimation des temps d'exécution à l'aide de la voix		54
4.2.2.3	Choix des tâches analysées		54
4.2.2.4	Choix des commandes vocales		55
4.2.3	Résultats d'analyse de l'interaction tactile		57
4.2.4	Résultats d'analyse de l'interaction vocale		60
4.2.5	Résultats de comparaison des interactions tactile et vocale		61
4.2.6	Discussion		64
4.3	Sommaire		64

CHAPITRE 5	ANALYSE DU LANGAGE DE L'INSTRUCTEUR	
	DE VOL	66
5.1	Définitions	67
5.2	Les langages de commandes	68
5.3	Modélisation du langage de commandes du poste d'opération	70
5.3.1	Méthodologie	70
5.3.1.1	Sujets	71
5.3.1.2	Outils utilisés	71
5.3.1.3	Tâche expérimentale	72
5.3.1.4	Procédures et consignes	75
5.3.2	Résultats	76
5.3.2.1	Constats	77
5.3.3	Description du modèle de langage de commandes vocales	83
5.4	Sommaire	86
CHAPITRE 6	ANALYSE DES TECHNOLOGIES DE RECON-	
	NAISSANCE VOCALE, DE L'IMPACT DU BRUIT	
	DU SIMULATEUR ET DE DISPOSITIFS D'ACTI-	
	VATION/DÉSACTIVATION	87
6.1	Analyse comparative de cinq engins de reconnaissance vocale	87
6.1.1	Méthodologie	88
6.1.1.1	Engins testés	88
6.1.1.2	Mesures d'exactitude et d'efficacité	89
6.1.1.3	Outils utilisés	91
6.1.1.4	Procédure	93
6.1.2	Résultats	94
6.1.3	Discussion	94
6.2	Analyse de l'impact du bruit du simulateur	97
6.2.1	Méthodologie	97
6.2.1.1	Outils utilisés	98
6.2.2	Résultats	100
6.2.3	Discussion	101
6.3	Analyse des dispositifs d'activation	102
6.3.1	Méthodologie	103
6.3.1.1	Sujets	103
6.3.1.2	Dispositifs utilisés	104
6.3.1.3	Tâches à exécuter	106

6.3.1.4	Procédures et consignes	106
6.3.2	Résultats	108
6.3.3	Discussion	109
6.4	Sommaire	109
CHAPITRE 7 CONSTRUCTION D'UN PROTOTYPE ET TESTS		
	UTILISATEURS	111
7.1	Analyse de l'intégration de la reconnaissance vocale au poste d'opération	111
7.1.1	Méthodologie	111
7.1.1.1	Sujets	112
7.1.1.2	Outils utilisés	112
7.1.1.3	Description du prototype	114
7.1.1.4	Tâches à exécuter	116
7.1.1.5	Procédures et consignes	117
7.1.2	Résultats	117
7.2	Sommaire	118
CONCLUSION		120
BIBLIOGRAPHIE		124
ANNEXES		135

LISTE DES TABLEAUX

Tableau 1.1	Aspects positifs et aspects négatifs de l'utilisation des technologies vocales	29
Tableau 3.1	Résultats d'analyse des entrevues effectuées auprès de cinq instructeurs	39
Tableau 4.1	Liste des commandes vocales analysées, divisées par tâches	55
Tableau 4.2	Vitesse de la parole selon diverses sources	60
Tableau 4.3	Temps d'exécution des tâches en secondes estimés à l'aide du logiciel CogTool pour deux versions de l'interface graphique (GUI 1, GUI 2) et les temps de production vocale théoriques (Voix 1) et empiriques (Voix 2)	61
Tableau 5.1	Exemple d'énoncé avec différentes formulations de commandes où le sujet doit indiquer 1 pour la commande préférée et 2 pour la seconde (deuxième exercice)	74
Tableau 5.2	Éléments de variation pour la construction de formulations de commandes	74
Tableau 5.3	Exemple de choix entre la formulation de commande préférée par le sujet et celle qui est préférée par l'ensemble des sujets	75
Tableau 5.4	Résultats où deux énoncés ont donné lieu à six formulations différentes par les sujets	82
Tableau 6.1	Engins de reconnaissance vocale retenus	88
Tableau 6.2	Moyenne des résultats des sujets pour les variables évaluées sur chaque dispositif	108

Tableau 6.3	Dispositifs d'activation/désactivation de système de reconnaissance vocale préférés par les sujets	109
Tableau 7.1	Pourcentage de satisfaction des sujets à l'égard du prototype testé et évaluée selon les six facteurs contenus dans le questionnaire.	118
Tableau A.1	Liste des tâches et sous-tâches retenues pour l'analyse	138
Tableau D.1	Liste des énoncés décrivant un besoin et à partir desquels le sujet devait composer un commande et la dire à haute voix	148
Tableau E.1	Caractéristiques des engins de reconnaissance vocale analysés	151
Tableau F.1	Scénario suivi par les sujets testant les dispositifs d'activation/désactivation	153

LISTE DES FIGURES

Figure I	Ligne du temps de l'évolution des simulateurs de vol et du poste d'opération	3
Figure 1.1	Vue de plan d'un simulateur	10
Figure 2.1	Concepts, techniques et analyses prévus dans cette étude	36
Figure 4.1	Modèles hiérarchiques des activités de l'instructeur de vol	48
Figure 4.2	Conception actuelle du poste d'opération des simula- teurs civils (GUI 1)	58
Figure 4.3	Nouvelle conception du poste d'opération (GUI 2) . .	58
Figure 4.4	Pourcentage du gain de temps moyen obtenu grâce à l'utilisation de commandes vocales	63
Figure 5.1	Résultats du deuxième exercice montrant, pour cha- cun des sujets, le nombre de fois (sur un total de 30 énoncés) qu'il a sélectionné la commande utilisée lors du premier exercice comme premier choix, comme se- cond choix, ou qu'il ne l'a pas sélectionnée.	80
Figure 5.2	Modèle de langage pour 78 commandes vocales. . . .	84
Figure 6.1	Niveau d'exactitude et d'efficacité des engins de re- connaissance testés à l'aide des six enregistrements et configurés pour reconnaître 20 commandes	95
Figure 6.2	Niveau d'exactitude et d'efficacité des deux meilleurs engins testés à l'aide des six enregistrements et confi- gurés pour reconnaître 137 commandes	96

Figure 6.3	Niveau d'exactitude de la reconnaissance de 20 commandes vocales avec l'engin E configuré pour reconnaître 137 commandes pour différents niveaux de bruit ambiant	100
Figure 6.4	Ensemble de dispositifs d'activation/désactivation testés par les sujets à l'aide d'un scénario de paramétrisation	105
Figure 7.1	Capture d'écran du prototype intégré au poste d'opération (bouton graphique identifié "Speech Recognition" en bas à droite dans l'application)	114
Figure 7.2	Rétroaction visuelle (trois dernières commandes reconnues) présentée au sujet lorsque le système de reconnaissance vocale est actif	115
Figure 7.3	Capture d'écran montrant l'effet de transparence du contrôle de rétroaction lorsqu'une partie d'une page se retrouve sous le contrôle	116
Figure B.1	CogTool - Fenêtre spécifique et les composantes graphiques accessibles	142
Figure B.2	CogTool - Représentation de la conception et des liens entre les fenêtres	143
Figure B.3	CogTool - Script décrivant la séquence d'actions à effectuer pour réaliser une tâche particulière	144
Figure B.4	CogTool - Liste de tâches et les temps calculés pour les deux interfaces comparées	145

LISTE DES SIGLES ET ABRÉVIATIONS

ACT-R	Adaptive Control of Thought - Rational
AOC	Air Operations Center
AWACS	Airborn Warning and Control System
BNF	Backus-Naur Form
FAA	Federal Aviation Administration (US government)
GPS	Global Positioning System
GSD	Glide Slope Display
GUI	Graphical User Interface
IOS	Instructor Operating Station
OACI	Organisation de l'aviation civile internationale

LISTE DES ANNEXES

ANNEXE A	SCÉNARIO D'ACTIVITÉS DE TRAVAIL POUR UN INSTRUCTEUR DE VOL	136
ANNEXE B	PRÉSENTATION DE COGTOOL	140
ANNEXE C	QUESTIONNAIRE D'ENTREVUE	146
ANNEXE D	LISTE DES ÉNONCÉS SOUMIS AUX SUJETS ET QUESTIONNAIRE POST-EXPÉRIMENTATION .	148
ANNEXE E	ENGINS DE RECONNAISSANCE VOCALE	151
ANNEXE F	SCÉNARIO ET QUESTIONNAIRE UTILISÉS DANS L'ÉTUDE DES DISPOSITIFS D'ACTIVA- TION/DÉSACTIVATION	153
ANNEXE G	SCÉNARIOS SERVANT AUX TESTS D'UTILISA- BILITÉ	156

INTRODUCTION

La vente de simulateurs de vol est un marché en expansion et très lucratif pour les quelques compagnies en place. La compagnie CAE Inc., par exemple, affichait pour l'année 2007 des revenus bruts de plus de 1.2 milliard de dollars canadiens ¹. Elle offre entre autres des services d'entraînement sur les simulateurs qu'elle possède dans différents centres de formation à travers le monde. Les compagnies aériennes n'ont donc pas nécessairement à acheter ces outils d'entraînement très onéreux. Les services d'entraînement coûtent entre 700 et 1 000 dollars de l'heure, selon le type d'appareil. C'est donc dire qu'à 1 000 dollars de l'heure, soutirer une minute d'entraînement pour les 18 000 heures d'entraînement annuelles que dispense CAE permet de faire des économies de 300 000 dollars par année.

Les outils d'entraînement de vol datent d'aussi loin que les avions eux-mêmes et pour cause : les entraînements sont critiques pour les pilotes. La ligne du temps et les images présentées à la figure I donnent un aperçu de l'évolution des simulateurs et plus particulièrement du poste d'opération de l'instructeur. Des recherches effectuées afin de rentabiliser ces entraînements ont été menées au début des années 90. Les études de Sylla et Ramesh (Ramesh et Sylla; Sylla et al.; Sylla et Ramesh) montrent que le temps efficace d'entraînement (I) est fonction du temps total disponible en simulateur (T) moins le temps de surcharge (H) induit par le processus d'entraînement et de simulation (équation 1). Leurs recherches tentaient ainsi de minimiser ce temps de surcharge afin d'augmenter le temps efficace d'entraînement. Un modèle d'optimisation du poste d'opération y est proposé.

¹http://www.cae.com/www2004/Investor_Relations/financialHighlights.shtml, Consulté le 15 janvier 2008

$$I = T - H \quad (1)$$

De plus, la conception du poste d'opération a fait l'objet de plusieurs rapports dans les années 80 (CAE Electronics Ltd.; Charles; Easter et al.; Elworth; Sanders; Schwartz; Warner), entre autres par le *Air Force Human Resources Laboratory* aux États-Unis, qui constatait alors des lacunes d'efficacité et d'utilisabilité. Ces guides de conception détaillent les contraintes d'affichage, de disposition, d'interactions et de fonctionnalités qui devaient alors être prises en compte afin d'obtenir de meilleurs postes d'opération. Plusieurs méthodes d'évaluation des postes d'opération ont de plus été développées. Ainsi, l'application des recommandations contenues dans ces rapports a grandement fait évoluer le poste d'opération pour l'amener à ce qu'il est aujourd'hui, comme le montre la figure I.

Dans un autre ordre d'idées, les technologies de reconnaissance vocale sont de plus en plus accessibles et intégrées dans les systèmes interactifs. La téléphonie est de loin le marché ayant le plus intégré ces technologies, créant des interfaces multimodales. Des entreprises de transport aérien comme American Airlines ont vu dans ces technologies la possibilité d'augmenter leurs services auprès des usagers en leur offrant des services d'information de vol automatisés (Srinivasan et Brown). Plus récemment, une voiture munie d'un système de reconnaissance vocale permettant de contrôler téléphone, radio et GPS était mise sur le marché (Meisel). Dans tous les cas, la reconnaissance vocale a permis aux utilisateurs de bénéficier d'un mode d'interaction supplémentaire avec le système.

Intégrer la reconnaissance vocale dans le domaine de la simulation n'est

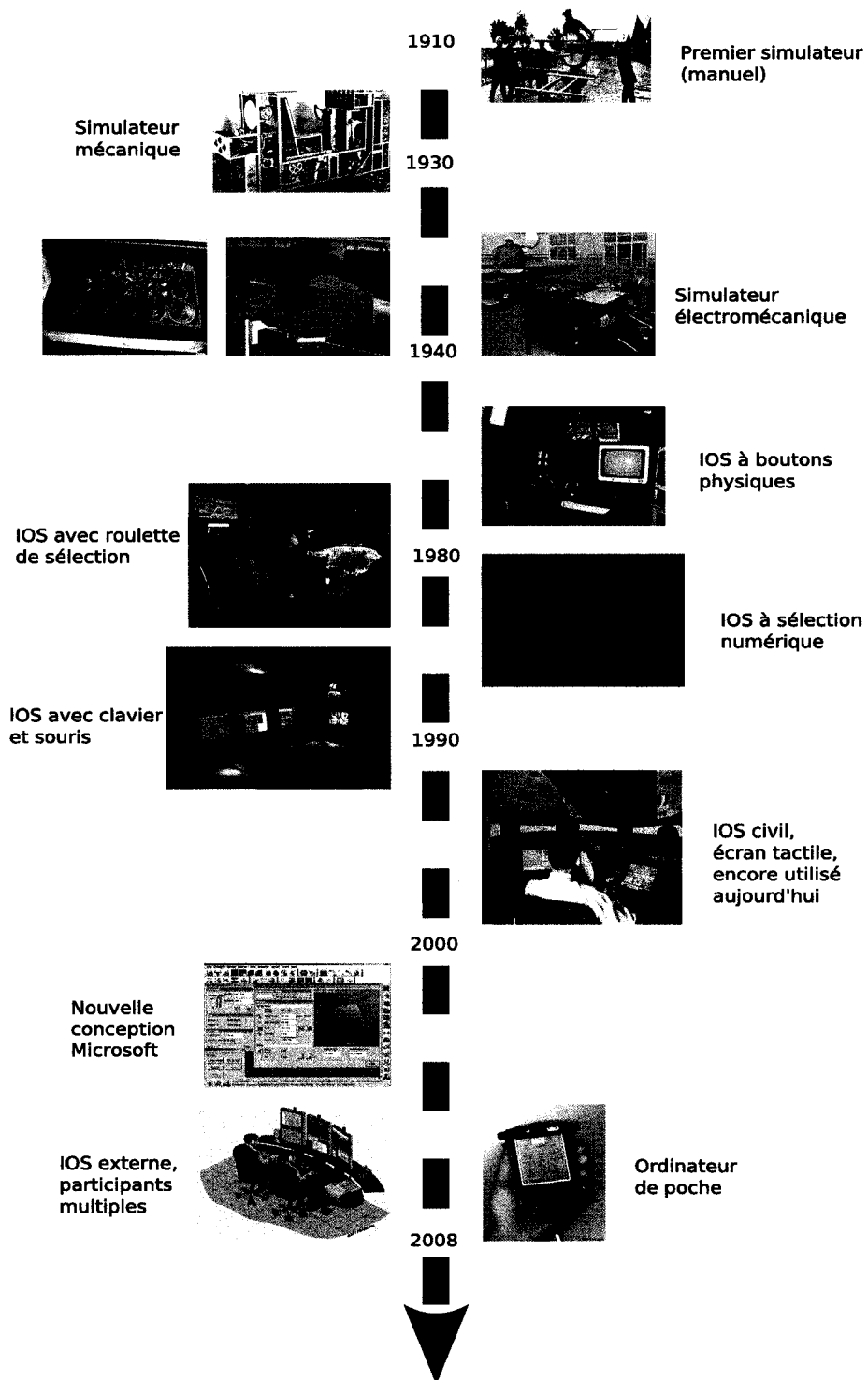


Figure I – Ligne du temps de l'évolution des simulateurs de vol et du poste d'opération

pas une tentative nouvelle. En effet, Fox et Weaver (1989) ont étudié l'intégration de ce mode d'interaction au poste d'opération de l'instructeur en simulateur de vol. L'interface du poste d'opération sur laquelle étaient basées les recherches était opérée à l'aide d'un clavier. L'instructeur devait entrer le numéro de la ligne du paramètre à modifier, la nouvelle valeur du paramètre et ensuite appuyer sur "Entrer". Un effort considérable était nécessaire aux concepteurs pour minimiser le nombre de touches permettant de changer de page. Les pages de paramètres qui ont été testées par cette étude portaient sur le contrôle d'une cible ennemie et de son armement, la communication radio, les dysfonctionnements et les conditions météorologiques. De plus, au moment de l'étude, il était impensable d'intégrer un système de reconnaissance totalement indépendant de l'utilisateur et ne nécessitant aucun entraînement. Cinq critères ont été proposés par les chercheurs en vue de produire un système répondant à leurs objectifs :

1. le système de reconnaissance vocale devait être plus facile à apprendre que l'interface existante ;
2. le temps nécessaire à l'instructeur pour entraîner le système devait être minimal ;
3. la grammaire et les commandes devaient être brèves tout en conservant l'aspect naturel de l'interaction ;
4. un système basé sur des échantillons de voix pour chaque utilisateur devait le rendre quasi-indépendant ;
5. le système devait être en mesure de reconnaître les mots isolés ainsi que la voix continue pour permettre à l'instructeur de vérifier que les commandes reconnues sont correctes.

L'objectif principal de Fox et Weaver (1989) était d'améliorer le temps

de réponse des instructeurs comme cela a déjà été fait avec les pilotes puisqu'ils doivent tous agir dans un contexte en temps réel. Un autre objectif était de réduire la complexité de l'interaction de l'instructeur avec le poste d'opération pour qu'il puisse se concentrer sur l'entraînement des pilotes. Plusieurs problèmes étaient prévisibles dans la réalisation de leur projet. Les auteurs mentionnent le développement d'une grammaire naturelle, l'impossibilité d'activer toutes les commandes à l'aide de la voix, le taux d'erreur du système de reconnaissance et la création d'un système difficile à entraîner. Les résultats de l'étude ne sont pas disponibles, mais les informations fournies donnent de bonnes indications sur les aspects importants à considérer dans l'intégration de la voix au poste d'opération de l'instructeur de vol. Puisqu'aucun simulateur ne possède encore une telle technologie, et ce, malgré l'évolution de l'interface du poste d'opération, il est toujours plausible que l'efficacité des instructeurs puisse être améliorée.

Afin de nous aider à cerner la problématique de l'intégration d'une technologie de reconnaissance vocale au poste d'opération, le chapitre 1 présentera une brève introduction sur l'entraînement en simulateur de vol et l'environnement du poste d'opération, de même qu'une revue de littérature sur le fonctionnement des systèmes de reconnaissance de la parole, leurs applications et les enjeux à considérer. Le chapitre 2 décrira ensuite la problématique et les objectifs de notre recherche et annoncera les différentes études que nous avons réalisées. Le chapitre 3 permettra de mieux cerner les caractéristiques des instructeurs à l'aide de profils utilisateurs. Le chapitre 4 présentera l'analyse des tâches de l'instructeur de vol en simulateur et comparera les temps d'exécution des tâches effectuées d'une part avec l'interface graphique et un écran tactile et d'autre part avec des commandes vocales afin de valider si l'efficacité peut être augmentée par ce nouveau mode d'interaction. Le chapitre 5 présentera un modèle de langage permettant aux

instructeurs de contrôler le poste d'opération par la voix. Le chapitre 6 présentera les résultats de trois analyses : la comparaison de cinq engins de reconnaissance vocale sur la base de l'exactitude et de l'efficacité, l'analyse de l'impact du bruit d'un simulateur sur le taux de reconnaissance vocale et enfin la comparaison de différents dispositifs d'activation/désactivation du système de reconnaissance vocale permettant de déterminer le plus approprié pour le contexte étudié. Le chapitre 7 présentera un prototype logiciel que nous avons développé intégrant les résultats obtenus dans les chapitres précédents et des tests avec utilisateurs visent à mesurer l'efficacité de système de reconnaissance vocale et la satisfaction des utilisateurs à l'égard du prototype.

CHAPITRE 1

L'ENTRAÎNEMENT EN SIMULATEUR DE VOL ET LA RECONNAISSANCE VOCALE

Les caractéristiques de l'entraînement en simulateur de vol sont peu connues. L'environnement dans lequel doit évoluer l'instructeur de vol est quelque peu différent de l'avion ou de l'hélicoptère réel. La première partie de ce chapitre introduit le travail de l'instructeur de vol et décrit les caractéristiques du poste d'opération avec lequel il est appelé à travailler.

Afin d'étudier l'intégration de la reconnaissance vocale au poste d'opération, il est nécessaire de bien comprendre les fondements de cette technologie. La seconde partie du chapitre identifie les aspects et les termes importants des systèmes de reconnaissance vocale. De plus, une revue de la littérature sur l'intégration de cette technologie dans divers domaines connexes est présentée. Enfin, les aspects mitigés de la reconnaissance vocale tels que les désavantages et les situations à éviter sont décrits afin mieux comprendre les contraintes qui se posent dans notre recherche.

1.1 Entraînement en simulateur de vol

Pour les pilotes, l'entraînement en simulateur de vol est similaire à l'entraînement réel. La fidélité des simulateurs d'aujourd'hui est telle qu'on se croit à l'intérieur de l'appareil lui-même. Les mises en situation possibles sont pratiquement illimitées et ne requièrent qu'un brin d'imagination et d'expérience de l'instructeur. Ce dernier doit cependant être en contrôle

de l'environnement du simulateur pour entrer et modifier les paramètres de différents scénarios de formation en vue de former les pilotes du mieux possible. Les sections suivantes décrivent quelques caractéristiques de l'instructeur de vol et de son travail, ainsi que le poste d'opération avec lequel il doit interagir pour mener à bien l'entraînement.

1.1.1 Instructeurs

L'instructeur de vol est un professionnel des plus polyvalents. Son but est d'enseigner à de futurs pilotes ou à des pilotes en exercice comment opérer un appareil, les règles qui s'y rattachent ainsi que la sécurité qui doit être constamment présente. Il doit être en mesure de piloter un appareil du côté droit comme du côté gauche et connaître tous les rudiments de vol. Il doit pouvoir communiquer ses connaissances et évaluer l'apprentissage fait par les pilotes. Au cours de l'entraînement, il a deux obligations : engager les étudiants dans des situations réelles, les observer et les corriger (Fox et Weaver).

Plusieurs documents réglementent le travail de l'instructeur de vol au même titre que celui des médecins puisqu'il porte une charge importante de responsabilités qui ont des conséquences majeures sur la qualité du travail des pilotes (FAA; FAA; FAA; FAA; FAA; FAA; Canada).

1.1.2 Poste d'opération de l'instructeur de vol

De nos jours, l'instructeur de vol est appelé à interagir avec son poste d'opération dans plusieurs contextes différents. Tout d'abord, il

peut contrôler la simulation à bord du simulateur avant ou pendant l'entraînement. Le poste de l'instructeur à bord du simulateur est conçu de manière à ce que l'instructeur puisse contrôler la simulation ainsi qu'observer les pilotes tout en demeurant assis à son poste. Le poste comprend entre deux et quatre écrans tactiles selon le type de simulateur et les préférences du client. Un ou plusieurs écrans sont installés sur un bras coulissant permettant à l'instructeur de regarder la simulation (vers les pilotes) et de la contrôler. Les autres écrans peuvent être positionnés l'un au-dessus de l'autre sur le côté du simulateur. Il est aussi possible pour l'instructeur d'utiliser un ordinateur de poche pour contrôler certaines fonctionnalités plus importantes du simulateur en n'étant pas devant un écran du poste d'opération. Dans certains cas, il y a aussi présence d'un dispositif de pointage pour contrôler les interfaces graphiques du poste d'opération en plus des écrans tactiles pour permettre une plus grande flexibilité, mais dans la plupart des simulateurs, les interactions ne sont que tactiles.

Le siège de l'instructeur est installé sur un rail, comme le montre la figure 1.1, lui permettant de se déplacer et de changer d'angle pour mieux interagir avec les pilotes ou avec son poste d'opération. Une entrée pour le casque de communication de l'instructeur est aussi accessible via le siège de l'instructeur.

L'instructeur a aussi accès à des boutons ou leviers pour contrôler quelques paramètres du simulateur comme l'arrêt d'urgence, les lumières à l'intérieur du cockpit, les mouvements du simulateur, la température, etc.

L'instructeur de vol a souvent accès à un poste d'opération à l'extérieur du simulateur. Ce poste lui permet de préparer la simulation, de contrôler les paramètres dynamiquement pendant l'entraînement de manière plus flexible

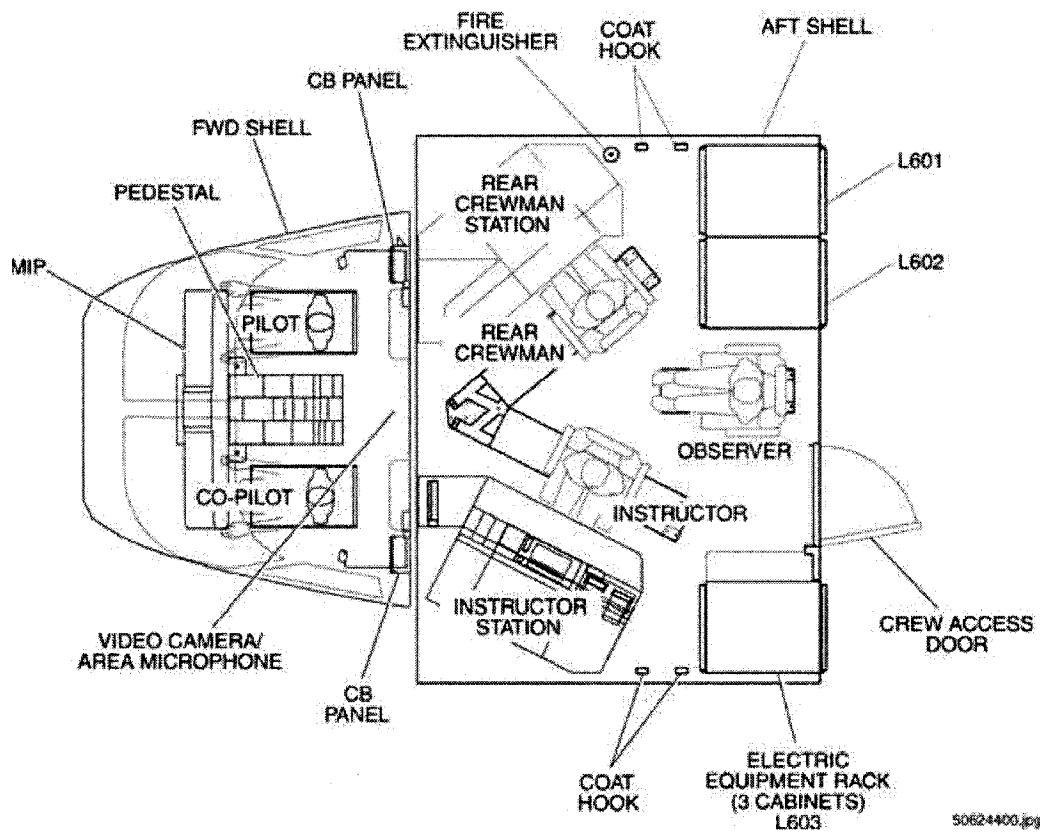


Figure 1.1 – Vue de plan d'un simulateur de vol (Source : CAE Inc. Chinook 47SD Full Mission Simulator - Instructor's Operating Manual - TPD 16969)

en étant à l'extérieur (sans mouvement ni bruit) et de contrôler diverses applications qui ne sont pas accessibles à l'intérieur du cockpit. Ce poste est muni d'un clavier et d'une souris, car les contraintes d'espace ne sont pas présentes comme dans le simulateur. Il est aussi possible de contrôler le simulateur ou d'autres entités de la simulation à l'aide d'une manette de contrôle. L'instructeur a accès au même système de communication qu'à l'intérieur du simulateur et peut donc communiquer avec les pilotes.

1.1.2.1 Fonctionnalités du poste d'opération

Les fonctionnalités du poste d'opération de l'instructeur en simulateur de vol ont été étudiées par (Ramesh et Sylla; Sylla et al.; Sylla et Ramesh). Ces fonctionnalités ont été classées par les auteurs en trois catégories : opération, instruction et gestion.

1.1.2.1.1 Fonctionnalités d'opération

Paramètres de vol - Les paramètres de vol sont les concepts physiques qui décrivent l'état de vol de l'appareil à travers différents instruments. L'instructeur est en mesure d'observer de son poste d'opération différentes valeurs d'altitude, de vitesse, d'orientation et autres que les cadrans et les moniteurs du poste de pilotage fournissent. De plus, l'instructeur peut visualiser les manipulations de l'appareil faites par les pilotes à l'aide de commandes telles que les freins, les injecteurs, etc. Par exemple, dans le cas d'hélicoptère où le transport de diverses charges est souvent présent, l'instructeur peut visualiser le centre de masse de l'appareil, les différentes distributions des charges ainsi que son poids total. Il lui est possible d'ajouter

des charges à l'appareil dynamiquement comme une charge externe au bout du treuil, un pilote supplémentaire ou encore de la glace sur l'appareil.

Paramètres des systèmes - Plusieurs systèmes sont présents dans un appareil et doivent être représentés graphiquement pour permettre de visualiser ou modifier leur état, ou de créer des défaillances. Par exemple, le système de carburant permet à l'instructeur d'ajuster le niveau de carburant dans l'appareil et de changer la configuration des réservoirs pour différents entraînements. Il lui est aussi possible de visualiser le taux d'essence consommé par les moteurs, la température des différents réservoirs, ainsi que le type de carburant à bord de l'appareil.

Le système électrique des appareils aériens est relativement complexe. Le poste de l'instructeur permet entre autres de visualiser le niveau de charge des batteries à bord ainsi que la tension des connexions électriques auxiliaires ou externes. Il est aussi possible de voir ou de modifier les panneaux de disjoncteurs de l'appareil, permettant ainsi de recréer des problèmes électriques en brisant certains circuits.

Les moteurs de l'appareil ont aussi leur lot de paramètres, c'est pourquoi plusieurs contrôles graphiques permettent de les représenter. L'instructeur peut visualiser la température et la puissance des moteurs, ainsi que la température et la pression de l'huile associée à chacun.

Les systèmes hydrauliques sont aussi représentés pour permettre à l'instructeur de connaître la pression, la température et la quantité du fluide contenue dans l'appareil.

Pause - Le poste de l'instructeur possède des fonctionnalités importantes pour le contrôle de l'environnement de simulation. Les contrôles permettant d'arrêter temporairement certaines parties de la simulation sont très utiles. L'instructeur peut ainsi changer certains paramètres sans que l'appareil soit en fonction, changer la position de l'appareil ou encore discuter avec les pilotes au milieu de l'entraînement. Un arrêt total de la simulation peut aussi être commandé, ce qui met en attente non seulement le simulateur, mais aussi les autres entités qui peuvent exister dans le monde virtuel. Enfin, plusieurs systèmes peuvent être arrêtés temporairement, ce qui signifie que leurs paramètres ne sont plus modifiés ou simulés et deviennent constants pour l'ensemble de la simulation (c.-à-d. un arrêt de la vitesse la rendrait constante et inhiberait toute accélération).

Reset - Une autre fonctionnalité importante du poste de l'instructeur permet de remettre instantanément les valeurs par défaut pour certains paramètres. Ces retours aux valeurs par défaut se font par exemple sur les températures de divers systèmes, la quantité de fluides comme le carburant ou l'huile, ainsi que plusieurs autres paramètres.

Disfonctionnements - Les disfonctionnements (traduction libre de "malfunctions") sont les fonctionnalités les plus importantes pour l'instructeur. Elles lui permettent de diriger l'entraînement des étudiants et de mettre ces derniers dans diverses situations peu communes. Un disfonctionnement est une défaillance dans une partie d'un système. L'instructeur peut choisir d'activer un disfonctionnement au moment présent ou le programmer pour plus tard, pour ainsi surprendre les pilotes et voir leur réaction dans cette mise en situation particulière. Les disfonctionnements possibles sont

très nombreux et variés : p.ex., le mauvais fonctionnement d'un moteur, une défaillance des communications radio, un problème du train d'atterrissage. C'est à l'instructeur de déterminer les combinaisons de dysfonctionnements et les moments propices pour les activer.

Visuel - Le système visuel du simulateur comprend plusieurs paramètres non seulement en ce qui a trait à l'activation/désactivation du matériel de projection graphique, mais plus spécifiquement au temps et à l'environnement de simulation. L'instructeur peut configurer le moment de la journée (jour, nuit, matin, soir) dans lequel il désire que les pilotes se retrouvent, les positions de la lune et du soleil, les conditions de surface au sol ou de l'océan, etc. Il lui est aussi possible de configurer l'intensité des différentes lumières des pistes d'atterrissage et de l'aéroport de référence. Enfin, l'interface graphique du système visuel permet de changer l'intensité des lumières dans le cockpit, de mettre les pilotes en situation de vision de nuit et de rendre l'horizon plus ou moins clair.

Navigation - Les paramètres de navigation comprennent entre autres les informations relatives aux aéroports et aux pistes d'atterrissage de références. Les conditions météorologiques, ainsi que la position et l'orientation d'une piste sont disponibles. L'instructeur peut de plus consulter les différentes fréquences utilisées par les contrôleurs aériens d'un aéroport en particulier et obtenir la liste des bornes radio de positionnement de la piste avec leur état et leur fréquence d'émission. L'instructeur ne peut sélectionner qu'un seul aéroport de référence à la fois pour lequel tous les paramètres sont importés d'une base de données. Chacun des aéroports possède un code OACI unique au monde qui permet à l'instructeur de l'identifier plus facile-

ment. Un autre code est ensuite requis pour identifier la piste d'atterrissage, ce qui nécessite une certaine connaissance des aéroports de la part de l'instructeur. Dans le cas d'un hélicoptère, il peut s'agir d'un site d'atterrissage (ou héliport) sur la terre ou sur un bateau.

Les bornes radio de positionnement des pistes d'atterrissage peuvent être désactivées par l'instructeur (un autre exemple de disfonctionnement). Ceci permet de rendre la tâche des pilotes plus compliquée et de voir leur façon d'opérer dans ce genre de situation. L'instructeur est aussi en mesure de consulter un graphique d'approche montrant les bornes radio et l'appareil en approche d'atterrissage ou en décollage pour en évaluer la pente et l'orientation.

Un contrôle graphique représentant un GPS permet de connaître la position de certains satellites de positionnement qui aident eux aussi à orienter les pilotes dans leur environnement. L'instructeur peut interroger ces satellites et les désactiver, ajoutant ainsi une fois de plus une certaine difficulté aux pilotes.

Communications - Les contrôles du système de communications permettent à l'instructeur par exemple d'activer ou de désactiver les liens de communication radio, d'ajuster le volume et de jouer le rôle d'un contrôleur aérien pour donner des instructions aux pilotes. Le système de communication comprend la simulation de fréquence radio pour simuler les communications externes de l'appareil ainsi qu'un système d'intercoms pour la communication interne. L'instructeur peut modifier le volume des différentes fréquences et modifier les liens de communications actifs. En vue de mettre les pilotes dans des situations les plus réelles possible, l'instructeur est appelé à communiquer avec les pilotes en jouant le rôle d'un tiers pour les informer

des conditions météorologiques ou aériennes par les communications radio. Enfin, il peut contrôler le niveau de bruit ambiant simulé pour reproduire les conditions à bord d'un appareil réel.

Météo - Les fonctionnalités relatives aux conditions météorologiques permettent à l'instructeur de changer les paramètres de l'environnement virtuel de simulation. L'instructeur peut enlever ou ajouter différents types de nuages, niveler la densité de brouillard, ajouter différents vents ou turbulences ou modifier les conditions de pluie ou de beau temps. Les conditions de pression atmosphérique, de nuages et de précipitation autour de l'appareil sont considérées en tout premier lieu par l'instructeur. Ce dernier peut paramétrer des couches de brouillard et nuages autour de l'aéroport de référence ou d'un point particulier dans l'environnement virtuel : la couche de brouillard près du sol, le premier et le second niveau de nuages, ainsi que la couche de nuages la plus élevée. L'instructeur est de plus en mesure d'ajouter des tempêtes à un endroit particulier dans l'environnement et de paramétrer sa configuration, son intensité, son altitude, etc. En plus des tempêtes, il peut activer certains événements environnementaux tels que des microsursauts près de l'aéroport de référence ou des cisaillements de vent autour de l'appareil.

Reposition - La fonction de reposition permet à l'instructeur de positionner l'appareil à un endroit particulier dans le monde virtuel de simulation. Il est possible de mettre l'appareil en situation de décollage ou d'atterrissage à diverses distances d'une piste de l'aéroport de référence, ou encore à une position absolue définie par la latitude, la longitude, l'altitude et la vitesse désirées de l'appareil.

L'instructeur a de plus la possibilité de commander au poste d'opération une transition de l'appareil dans une direction particulière. Contrairement à la reposition qui est instantanée, la transition a lieu dans le temps à une vitesse spécifiée par l'instructeur et sert à déplacer l'appareil graduellement à un endroit, selon un angle particulier.

Mouvement - En plus des paramètres de l'appareil et de l'environnement, l'instructeur a le contrôle sur la mise en marche et l'arrêt du simulateur. Tout comme les commandes pour le visuel ou le matériel audio de communication, une interface du poste d'opération permet d'activer et de désactiver le système de mouvement du simulateur, le chargement des contrôles, ainsi que la vibration.

Carte - Le poste d'opération de l'instructeur militaire permet de visualiser l'appareil dans le monde virtuel par le biais d'un plan topographique. L'instructeur peut accroître ou diminuer la résolution de la carte, déplacer l'appareil à une position particulière dans le monde virtuel et voir les autres joueurs dans la simulation.

1.1.2.1.2 Fonctionnalités d'instruction

Capture/Reprise - L'une des fonctionnalités d'instruction les plus intéressantes est la possibilité d'enregistrer l'état de la simulation à un moment précis. Cela permet par la suite à l'instructeur de revenir à cet état particulier pour permettre aux étudiants de refaire une partie de l'entraînement dans les mêmes conditions. Habituellement, la simulation comprend uniquement l'appareil dans lequel les pilotes s'exercent, mais dans le

cas où plusieurs entités (ou appareils) sont simulées en même temps dans le même monde virtuel, un enregistrement de l'ensemble de la simulation est nécessaire et une reprise permet à l'instructeur de revenir sur les autres éléments de la simulation.

Impression - L'instructeur peut au moment souhaité faire une impression de l'affichage qui paraît à son poste d'opération. Cela lui permet par la suite de visualiser par exemple les graphiques d'approche ou de décollage avec les étudiants.

1.1.2.1.3 Fonctionnalités de gestion

Exercices - Pour mieux soutenir la gestion et la préparation des entraînements en simulateur, le poste d'opération permet de créer à l'avance des exercices qui tiennent compte de l'environnement de simulation. Ces exercices sont présents plus spécifiquement dans les simulations militaires où plusieurs entités sont simulées en même temps. Un exercice permet à l'instructeur de positionner l'appareil à l'endroit qu'il le désire dans le monde virtuel, de créer et de paramétrer l'environnement de simulation avec les éléments nécessaires comme les conditions météorologiques, et aussi de planifier des événements devant se produire au cours de l'entraînement. L'édition d'un exercice se fait généralement avant l'entraînement sur un poste à l'extérieur du simulateur où il y a une souris et un clavier. Plusieurs applications peuvent venir compléter l'environnement graphique de base à l'extérieur du simulateur pour aider à la création d'exercices plus complexes. Ainsi, lorsque l'exercice est prêt, l'instructeur n'a qu'à le charger à bord du simulateur et commencer directement l'entraînement avec les paramètres de

simulation déjà entrés.

1.2 Reconnaissance vocale

La première technologie de reconnaissance vocale a été développée en 1952 sur un ordinateur analogique pour reconnaître les chiffres de 0 à 9 à l'aide d'un algorithme basé sur la reconnaissance de la voix d'une personne (Grasso et Finin). Depuis, plusieurs chercheurs se sont penchés sur ce mode d'interaction qui devient de plus en plus présent dans des systèmes qui nous entourent et des systèmes plus spécifiques. Les personnes atteintes de troubles de vision ou de cécité, ou qui ont des problèmes de lecture ou d'écriture, ou encore des troubles de mobilité des membres supérieurs ont grandement bénéficié de la technologie vocale (Klemmer et al.). Les interfaces standards ne sont pas adaptées à tous les utilisateurs de sorte que les technologies vocales viennent parfois combler une lacune importante dans l'interaction personne-machine qui ne pourrait l'être autrement. Le nombre d'utilisateurs d'ordinateurs ne faisant qu'accroître dans le monde, ceux qui désirent passer à l'ère de cette technologie sont donc aussi de plus en plus nombreux (James et Roelands). Dans certaines situations, on constate que la voix devient le médium le plus efficace pour interagir, comme lorsque l'utilisateur doit se déplacer alors que ses yeux et ses mains sont occupées à d'autres activités ou lorsqu'il doit concentrer son attention sur autre chose que le système. Bref, diverses raisons sont bonnes pour implanter et utiliser cette technologie, mais à quel prix ?

La reconnaissance vocale est un nouveau mode d'interaction, différent des modes conventionnels comme la souris et le clavier, qui offre de nouveaux

défis quant à sa facilité d'interaction et à sa diffusion auprès d'une plus grande population. La section suivante décrit le fonctionnement de cette technologie ainsi que les diverses applications auxquelles elle peut servir.

1.2.1 Le fonctionnement des systèmes de reconnaissance vocale

Certains systèmes de reconnaissance vocale doivent être entraînés à reconnaître la voix d'un locuteur particulier alors que d'autres sont indépendants du locuteur. Les résultats d'études montrent que l'entraînement des dispositifs à reconnaître la voix d'un utilisateur particulier est facile, ne nécessitant en moyenne que de 20 à 30 minutes, et que la performance en est grandement augmentée pour une gamme de commandes relativement simples. Un vocabulaire trop riche nécessiterait une plus grande effort de conception du dispositif pour qu'il soit en mesure de déterminer si deux commandes ou phrases ont la même sonorité (Gorniak et Roy). De plus, il est important que les systèmes soient entraînés dans l'environnement dans lequel ils seront intégrés. Cela permet d'adapter le système à cet environnement et améliore la distinction entre la voix et les autres bruits ambiants. Ensuite, certains systèmes permettent de parler de façon continue alors que d'autres exigent une pause entre chaque mot. Certains systèmes ne perçoivent qu'une certaine grammaire, d'autres des phrases ou des commandes bien spécifiques et d'autres se basent sur des modèles statistiques du langage naturel (Yankelovich et Lai). Dans tous les cas, on constate que plus les règles régissant les dispositifs de reconnaissance sont nombreuses, plus le taux de réussite est élevé. C'est donc dire qu'un système entraîné pour un utilisateur particulier qui ne reconnaît que quelques mots précis est

extrêmement fiable.

Tel que spécifié, le nombre de mots que reconnaît un dispositif est variable. Une application servant à la dictée de phrases ou à l'indexage multimédia est en mesure de reconnaître un vocabulaire allant jusqu'à 100 000 mots, alors qu'une application axée sur les transactions ou les commandes utilisera un vocabulaire plus orienté vers la tâche et comprenant une centaine de mots. Dans ce dernier cas, la grammaire permet la reconnaissance de certains mots ou de certaines phrases spécifiques formant une structure unifiée par des règles qui définissent ce que le dispositif est en mesure de reconnaître à un moment précis dans l'exécution d'une tâche (Srinivasan et Brown).

Il est inhabituel qu'un système se limite à l'interaction par la voix. À moins d'interfaces spécialisées pour les personnes atteintes d'un handicap, les systèmes comptent sur plus d'un mode d'interaction pour offrir une plus grande liberté à l'utilisateur. Les manipulations faites à l'écran à l'aide de la souris par exemple conviennent beaucoup mieux pour les actions simples effectuées sur des références graphiques visibles et limitées en nombre. À l'opposé, la reconnaissance vocale est plus performante dans les cas d'actions complexes, invisibles et en plus grand nombre (Grasso et Finin). La littérature fait mention d'interfaces multimodales comprenant plusieurs modes d'interaction. Puisque leurs diverses tâches varient en complexité, les utilisateurs sont enclins à utiliser des interfaces multimodales qui leurs permettent de sélectionner le mode le plus approprié. Un aspect intéressant cité par Yankelovich (1996) est que, dans certains domaines, il est avantageux d'avoir une correspondance entre l'interface graphique et les commandes vocales accessibles.

1.2.2 Applications à des domaines connexes

À notre connaissance, seul l'article de Fox et Weaver (1989) traite de l'intégration de la reconnaissance vocale au poste d'opération de l'instructeur en simulateur de vol spécifiquement. Par contre, dans d'autres domaines connexes, on a tenté d'intégrer cette technologie à des environnements, des contextes, ou des utilisateurs similaires. Les résultats de ces recherches peuvent nous informer sur le potentiel de cette technologie pour notre contexte particulier. Les sections suivantes décrivent les recherches faites dans les domaines avionique et militaire.

1.2.2.1 Domaine avionique

Swail et Kobierski (1997) rapportent l'une des nombreuses recherches qui ont été faites dans le domaine de l'avionique pour tester l'intégration de la reconnaissance vocale dans les cockpits d'avions ou d'hélicoptères. Cette recherche a en l'occurrence démontré les bénéfices de l'ajout de technologie vocale dans un cockpit d'hélicoptère Bell 412 pour contrôler le système de gestion de l'avionique. Le taux de reconnaissance du système fut d'environ 94.9% et requérait un minimum d'entraînement de la part de l'utilisateur.

Une étude menée par Starks et Morgan (1992) a démontré que la reconnaissance vocale était déjà à cette époque viable dans un environnement aussi défavorable pour ce mode d'interaction qu'un cockpit d'hélicoptère. Quatre cas de conditions de bruit ambiant ont été étudiés compte tenu des multiples possibilités qu'il est possible de rencontrer en opération : à l'intérieur du hangar avec le système électrique allumé, au sol avec les moteurs en attente à 70% du rotor principal, en survol stationnaire avec le

bruit de résonnance du sol et en vol à une vitesse de 60 à 80 noeuds. Les tests effectués en simulation en utilisant un vocabulaire associé aux commandes de vol électriques ou “fly-by-wire” (par exemple : chiffres de 0 à 9, “heading”, “forward”, “feet”, “cancel”, “enter”, etc.) ont permis d’atteindre après quelques améliorations près de 100% de taux de reconnaissance.

Plusieurs autres références traitant du succès de l’intégration de la reconnaissance vocale dans les cockpits d’avions ou d’hélicoptères pour les pilotes sont mentionnées dans (Weinstein).

1.2.2.2 Domaine militaire

Une étude de Weinstein (1995) basée sur des investigations auprès de plusieurs organisations militaires et gouvernementales aux États-Unis a révélé qu’il y avait plusieurs opportunités d’intégration de la reconnaissance vocale aux applications utilisées dans ces organisations. L’auteur mentionne que, malgré le fait que la robustesse aux bruits environnants et au stress des utilisateurs est nécessaire dans le domaine militaire, les applications sont dans la majeure partie des cas dédiées à des tâches particulières, ce qui restreint le vocabulaire et la grammaire associés à une application. Selon l’auteur, les prototypes réalisés avant les années 90 ne bénéficiaient pas d’une technologie aussi performante que celle qui existait lors de son étude, ce qui le portait à croire à cette époque que des applications opérationnelles en environnement réel seraient près d’obtenir du succès. À partir de plusieurs exemples dans des domaines appliqués comme l’armée de terre, les opérations navales, les cockpits d’avions ou d’hélicoptères, le contrôle du trafic aérien ou encore les applications d’investigations policières, Weinstein voit d’un bon oeil l’intégration de cette technologie dans des applications d’entraînement. Ces

applications permettraient de caractériser l'apport de ce mode d'interaction à des situations plus critiques et réelles.

Draper et al. (2003) ont fait une étude pour comparer l'interaction manuelle et la reconnaissance vocale pour le contrôle d'un véhicule aérien télécommandé. L'expérience au cours de laquelle différents types de tâches étaient exécutés par les sujets, a démontré de façon significative que l'utilisation de la voix diminue grandement le nombre d'étapes ainsi que le temps total pour accomplir une tâche particulière. Pour l'ensemble des tâches d'entrée de données en situation normale, d'alerte non critique, d'alerte critique et de requête d'information, la voix a permis de réduire d'environ 40% le temps requis pour compléter une tâche. Le nombre de tâches qui n'ont pas été complétées lors de l'expérimentation a de plus été moins grand avec l'interaction vocale. Le taux de succès du système de reconnaissance était de l'ordre de 95%. De plus, les sujets de l'expérience ont préféré l'interaction vocale au contrôle manuel. Draper et al. (2003) affirment toutefois que de plus amples recherches devraient être menées pour vérifier que l'interaction vocale n'interfère pas avec les opérations de communication radio ou intraéquipe.

Vidulich et al. (2006) ont travaillé pour le Air Force Research Laboratory et ont tenté de vérifier si les résultats de Draper et al. (2003) pouvaient s'appliquer à un système de contrôle et d'alerte aérienne (AWACS) pour un environnement de gestion et de contrôle de l'activité de combat aérien. Ce système, intégré à un avion, permet à une équipe de surveillance, une équipe d'armement et une équipe de support tactique de veiller à la gestion des activités de combat. Vidulich et al. (2006) ont pris comme hypothèses que le temps requis pour effectuer certaines tâches serait diminué en utilisant la technologie vocale et que les performances de partage de temps seraient

augmentées dû à l'utilisation d'un nouveau mode d'interaction. Ils croyaient de plus que l'augmentation de la charge cognitive observée dans d'autres recherches était due au problème de performance de la technologie qui existait il y a quelques années. Leurs hypothèses ont été confirmées : les temps d'exécution de tâches sont plus courts en utilisant la voix et les performances de partage de temps sont plus efficaces. Les opérateurs qui ont participé à cette étude ont de plus grandement préféré l'interaction vocale puisqu'eux-mêmes choisissaient ce mode d'interaction lorsque le choix leur était donné. L'évaluation de la charge cognitive en utilisant la méthode NASA-TLX a de plus permis de montrer que l'interaction vocale utilisait moins de ressources cognitives. L'expérience a de plus confirmé que l'opérateur est naturellement porté à s'adapter aux contraintes auxquelles il est soumis, puisqu'il a combiné la voix et l'interaction manuelle pour optimiser ses performances d'exécution de tâches. Vidulich et al. (2006) expliquent que pour tirer le plus de bénéfices possible de l'interaction vocale, une analyse approfondie doit être faite pour déterminer les tâches qui se prêtent le mieux à l'utilisation de la voix et le meilleur vocabulaire à concevoir pour le contexte d'utilisation.

L'article le plus révélateur sur les technologies de reconnaissance vocale est sans contredit celui de Williamson et al. (2005) qui fait le bilan de l'application des technologies de reconnaissance vocale dans deux projets de recherche militaire. Le premier projet concerne l'ajout d'interfaces vocales à un centre d'opération aérien (AOC) dans le style du système AWACS mentionné précédemment (à noter que les travaux de Williamson et al. (2005) ont été rédigés avant ceux de Vidulich et al. (2006) mais qu'ils font quand même mention des recherches en cours sur AWACS effectuées par ce deuxième groupe de chercheurs). Le but de cette recherche était de voir si le temps pour compléter certaines tâches s'améliorait en utilisant les com-

mandes vocales et aussi si l'ajout d'une interface vocale dans le processus de planification et de gestion de mission était intuitif et naturel. L'interface vocale fut d'abord prototypée pour permettre la navigation à travers les menus de l'interface graphique existante et aussi pour l'entrée de données par la voix. Au cours des premiers tests faits avec la nouvelle interface, 16 sujets ont utilisé ce système dans un contexte de simulation de mission. Les résultats ont révélé une augmentation de la rapidité d'entrée de données de 10.6% avec un taux de succès du système de reconnaissance de 97.4%, ainsi qu'une nette préférence des sujets pour ce mode d'interaction. Dans un désir d'améliorer le système, l'interface a été modifiée pour permettre la commande de plusieurs événements dans une même séquence de mots, ce qui a permis d'accroître encore la vitesse d'entrée de données. De plus, l'ajout d'un commutateur pour activer la reconnaissance à l'aide du pied de l'opérateur a permis de libérer les mains de ce dernier et d'augmenter le taux de reconnaissance en précisant au système le moment où les paroles de l'opérateur devaient être analysées. En somme, ces recherches ont démontré la capacité de la reconnaissance vocale à augmenter la rapidité d'utilisation d'un module de planification de mission aérienne. Les auteurs notent enfin que la diminution du niveau de frustration que peut apporter l'utilisation de technologies vocales se traduirait par une amélioration de la perception de la qualité du produit et une diminution de la fatigue d'utilisation. La deuxième recherche que Williamson et al. (2005) mentionnent est celle de Draper et al. (2003) citée plus haut : rappelons simplement les bénéfices de l'utilisation de la reconnaissance vocale pour la navigation et l'entrée de données concernant un véhicule aérien télécommandé à son poste d'opération. Enfin, cet article de Williamson et al. (2005) nous informe que la technologie de reconnaissance vocale commercialement offerte a atteint un niveau de maturité suffisant pour être intégrée de manière bénéfique dans

diverses situations de contrôle dans le domaine militaire.

Un document réalisé dans le cadre d'une étude sur l'utilisabilité d'un poste d'opération d'un simulateur de sous-marin a révélé plusieurs informations intéressantes concernant l'intégration de la reconnaissance vocale en contexte de simulation (Hays et al.). Le simulateur étant lui-même muni d'un tel dispositif, il est la preuve qu'un système de reconnaissance vocale indépendant de l'utilisateur peut être utilisé pour améliorer l'interaction de l'utilisateur. Les auteurs de ce document voient ainsi un potentiel intéressant dans l'intégration de cette technologie au poste d'opération de l'instructeur pour entre autres remplacer plusieurs actions faites avec la souris. Cela permettrait de plus à l'instructeur de s'éloigner de son poste pour mieux observer les étudiants. Les auteurs recommandent d'effectuer une analyse plus approfondie sur les différentes activités qui permettraient d'améliorer l'efficacité de l'instructeur et de l'ensemble du processus d'entraînement.

1.2.3 Aspects contraignant l'intégration de la reconnaissance vocale

Les recherches résumées ci-dessus nous montrent les bénéfices de la technologie de reconnaissance vocale dans divers environnements. Or, certains auteurs croient que l'intégration de cette technologie peut avoir des effets négatifs sur les utilisateurs ou sur l'environnement dans lequel elle est intégrée. Les sections suivantes décrivent ces aspects qui contraignent l'intégration de la reconnaissance vocale et qui devront faire l'objet d'attentions particulières dans notre étude.

1.2.3.1 Perturbation de la communication d'équipe

Leggatt et Noyes (1996) ont conclu, dans une étude sur l'intégration de la reconnaissance vocale dans un simulateur de véhicule de combat, que la communication entre les membres de l'équipe à bord du véhicule était affectée par l'interaction vocale. En effet, les commandes vocales données au véhicule ont eu des effets non seulement sur la personne donnant les commandes, mais aussi sur l'autre membre de l'équipage qui ne pouvait demander d'informations à son collègue pendant que ce dernier interagissait avec le système. Or, les échanges à l'intérieur de l'équipe, qui ont pour effet de mettre à jour les modèles mentaux des autres membres de l'équipe, ont diminué avec l'ajout d'un système de reconnaissance vocale. Pour Leggatt et Noyes (1996), dans des environnements où les communications rapides entre les membres d'une équipe sont nécessaires pour la performance du système, l'utilisation de la reconnaissance vocale devrait être questionnée.

1.2.3.2 Limites de la mémoire et des tâches

Cook et al. (1996) affirment que les interfaces vocales peuvent avoir des effets bénéfiques dans les cockpits. En augmentant le temps durant lequel les pilotes peuvent garder la tête levée, les interfaces vocales améliorent la performance et la sécurité puisque les manipulations du système et les informations reçues de celui-ci ne requerraient pas de baisser les yeux. Cependant, les auteurs ont démontré dans leur recherche que certains problèmes demeurent avec l'intégration de différentes modalités d'interaction. En effet, la mémoire et la charge de travail seraient des facteurs négatifs incontournables des applications vocales. Même en augmentant le taux de succès des technologies de reconnaissance, ces problèmes demeurent car ils ont pour

causes les limites immuables des opérateurs humains qui tentent d'interagir dans un environnement multimodal. Un aspect mentionné, qui découle des limites de capacités humaines, est le besoin de fournir aux opérateurs de l'information visuelle additionnelle, car il leur est impossible de se souvenir constamment de l'état dans lequel ils sont. Cependant, dans le cas des pilotes d'avion, l'ajout d'information visuelle est impensable sur une interface déjà surchargée.

Cook et al. (1996) dressent une liste des aspects positifs et négatifs de l'utilisation des technologies vocales (voir le tableau 1.1).

Tableau 1.1: Aspects positifs et aspects négatifs de l'utilisation des technologies vocales

Aspects positifs
- Offre une méthode d'apprentissage automatique - Meilleure opérabilité mains libres et tête libre - Utilisation des capacités intellectuelles inutilisées - Séparation de l'information et des sources d'information - Puissance grandissante des technologies de reconnaissance vocale - Cause une charge de travail modérée pour des tâches spécifiques
Aspects négatifs
- Une interface multimodale doit offrir une rétroaction plus adéquate pour l'ensemble des tâches - La réception d'information vocale peut correspondre à une plus grand charge mnémonique - Le changement entre les différentes modalités peut être lent et coûteux - La présentation d'information auditive peut prendre le dessus sur l'information visuelle - Le taux de reconnaissance peut diminuer lorsque l'utilisateur est en situation de stress - La restriction du vocabulaire peut être non-naturelle et coûteuse

Plusieurs études sont citées par Cook et al. (1996) qui apportent d'autres arguments contre les technologies vocales. Par exemple, les communications radio qui ne sont pas intégrées directement à la tâche courante des pilotes

d'avion seraient une cause possible d'accidents. Dans le même ordre d'idées, l'utilisation du téléphone cellulaire en automobile pendant la conduite serait aussi problématique. Ces exemples viennent appuyer l'affirmation selon laquelle le niveau d'attention nécessaire est plus élevé lorsqu'on utilise plusieurs modalités d'interaction. De plus, il appert que l'information auditive soit privilégiée par rapport à l'information visuelle. Premièrement, l'attention est dirigée à prime abord par les sons, et deuxièmement, le parcours du son est anatomiquement plus court que celui de l'information visuelle. La segmentation de l'information à travers plusieurs modalités pourrait aussi jouer un rôle négatif en créant une difficulté au niveau de l'intégration de l'information pour la prise de décision de plus haut niveau. En somme, c'est principalement la concurrence entre l'interaction vocale et visuelle pour les ressources cognitives et l'attention qui informent que l'interférence entre ces deux modes est un problème d'utilisabilité majeur.

Même si des études ont démontré les bénéfices de l'interaction vocale sur l'entrée au clavier pour un vocabulaire restreint, des contraintes causées par l'utilisation d'un vocabulaire plus ou moins naturel peuvent venir augmenter la charge cognitive. Cook et al. (1996) citent trois faits observés au cours de leurs études faites avec des pilotes qui viennent indiquer que la charge mnémonique de travail est un problème important : tout d'abord, certains opérateurs oublient le mode dans lequel ils sont ; ensuite, certains opérateurs doivent utiliser un bloc note pour noter les informations du système vocal ; enfin, certains opérateurs oublient certains paramètres de vol selon une fréquence qui pourrait être problématique. Il est donc important d'orienter le vocabulaire vers la tâche à accomplir pour diminuer l'effort cognitif.

1.2.3.3 Limites imposées par les standards

Le standard militaire MIL-STD-1472F fait mention de quelques aspects importants à considérer lorsqu'on intègre la reconnaissance vocale (DoD). Il est entre autres question de n'utiliser la reconnaissance vocale que dans des cas bien précis où, par exemple, les mains de l'utilisateur sont occupées,

lorsqu'il lui est nécessaire de se déplacer ou encore lorsque la tâche requiert son attention visuelle. Il est recommandé d'utiliser cette technologie lorsque le taux d'erreur est peu élevé, dans un environnement où la correction des erreurs est facilitée et où d'autres moyens d'interaction sont disponibles, rendant l'utilisation de la reconnaissance vocale minimale. Selon ce standard, la reconnaissance vocale ne devrait pas être utilisée pour des tâches liées à la description de position ou la manipulation d'objets. Les environnements où les utilisateurs sont soumis à un certain niveau de stress, qui sont bruyants ou qui ont une charge gravitationnelle supplémentaire, devraient aussi être prescrits. Le système de reconnaissance devrait nécessiter un minimum d'entraînement et ne permettre la reconnaissance que d'un nombre minimal de mots.

Dans le même ordre d'idées, le standard de la FAA en facteurs humains (Human Factors Group of the FAA William J. Hughes Technical Center) fait référence aux systèmes de reconnaissance vocale comme étant utiles dans les situations où les performances visuelles et manuelles sont restreintes. De plus, il est mention de l'importance du choix du vocabulaire pour être phonétiquement distinctif tout en étant adapté à la terminologie associée à la situation. L'utilisateur devrait être en mesure de facilement annuler une action et de la reprendre.

Enfin, le standard qui présente les règles de conception d'interfaces dans le domaine nucléaire (O'Hara et al.) voit la reconnaissance vocale comme un mode d'interaction à n'utiliser qu'en dernier recours ou lorsque les mains de l'utilisateur sont restreintes dans leurs mouvements. Il est aussi mention de la nécessité d'intégrer un système non vocal pour activer ou désactiver le système de reconnaissance pour éviter que le dialogue entre utilisateurs soit assimilé à des commandes adressées au système.

1.3 Sommaire

Les instructeurs de vol en simulateur doivent travailler dans un environnement bien spécifique, à l'intérieur ou à l'extérieur du simulateur, et interagir avec un poste d'opération muni de fonctionnalités bien précises qui lui permettent d'entraîner les pilotes dans diverses situations simulées.

Les technologies intégrant la reconnaissance vocale sont de plus en plus nombreuses. Un dispositif de reconnaissance de la voix dépendant de l'utilisateur, entraîné à reconnaître un vocabulaire orienté vers la tâche et combiné à un autre mode d'interaction est à préconiser. Plusieurs applications de cette technologie dans les domaines de l'aviation et de la défense ont démontré son efficacité et ses avantages pour le travail de différents opérateurs. Cependant, il est important de tenir compte des limites relevées par certaines études comme l'impact négatif sur les communications interpersonnelles et sur la charge cognitive, et de retenir que les standards restreignent ce mode d'interaction à un environnement spécifique et propice.

CHAPITRE 2

PROBLÉMATIQUE SIMULATEUR-INSTRUCTEUR ET OBJECTIF DE L'ÉTUDE

À la lumière des informations recueillies dans la littérature et présentées au chapitre précédent, nous sommes en mesure de croire que l'intégration de la reconnaissance vocale est possible dans le contexte du poste d'opération de l'instructeur en simulateur de vol. Cependant, la difficulté ne semble pas être le développement d'une nouvelle technologie de reconnaissance vocale plus efficace, mais bien les contraintes entourant l'intégration d'une technologie déjà existante qui a fait ses preuves. Ce chapitre présente la problématique et les objectifs de cette recherche.

2.1 Problématique

Après l'analyse de la littérature récente sur l'intégration de systèmes de reconnaissance vocale, tout porte à croire que les problèmes technologiques mentionnés par Fox et Weaver (1989) ne sont plus aussi imposants qu'ils semblaient l'être au moment de leur recherche. De l'avis de plusieurs chercheurs dans le domaine, les aspects ergonomiques ou relatifs aux facteurs humains semblent à présent les défis à relever.

Weinstein (1990) mentionnait tout d'abord que les facteurs humains représentaient un élément crucial, mais souvent négligé, pour tirer bénéfice des technologies vocales. Les tâches de sélection, de conception du dialogue, d'élaboration des caractéristiques des utilisateurs, de conception des interfaces vocales, de modélisation des tâches dans leur environnement et enfin, d'analyse de performance, sont importantes dans les études sur les impacts réels.

Selon Larson (1995), il est important de choisir la bonne technologie et de cibler les tâches les plus appropriées pour l'utilisation de la voix. Il note aussi l'importance d'une bonne connaissance de l'utilisateur, des technologies et surtout des capacités du système développé pour intervenir en cas d'erreurs. Il conclut en affirmant qu'une analyse approfondie des requis, une collecte des données judicieuses, une analyse rigoureuse, une conception minutieuse ainsi que des tests sont essentiels pour faire ressortir la performance et l'utilisabilité d'une interface vocale.

Un auteur très cité dans le domaine des technologies de reconnaissance vocale, Micheal F. McTear, a aussi écrit sur les méthodologies de développement dans son article "Spoken Language Technology" (McTear). Cet article fait état des différents critères qui régissent le développement de systèmes à reconnaissance vocale. La conception basée sur la simulation comme dans le cas de la méthode du Magicien d'Oz, où une personne simule le rôle du système, questionnant et répondant à l'utilisateur comme s'il interagissait avec le système, permet de savoir comment l'utilisateur interagit avec l'application (McTear). L'auteur conclut en mentionnant que les meilleures pratiques dans ce domaine sont vues comme étant une spécialisation des meilleures pratiques de l'ingénierie logicielle, ce qui explique la convergence entre les différentes méthodologies que la recherche propose (McTear).

Kamm (1995) a écrit un article faisant état des considérations telles que l'interdépendance entre les tâches visées par une application vocale, les besoins et les attentes de l'utilisateur et la capacité de la technologie lorsqu'on songe à développer une application vocale. Il fait aussi mention de certaines stratégies pour alléger les contraintes du système et pour augmenter son utilisabilité. Le flux du dialogue peut par exemple être dirigé par des messages informatifs qui guident l'utilisateur dans son interaction avec le système. Une autre stratégie est de fournir suffisamment d'informations ou de confirmations pour que l'utilisateur soit en mesure de savoir si les interactions qu'il a avec le système sont acceptées et comprises par ce dernier ou s'il a commis des actions erronées.

Tous les aspects ergonomiques décrits par ces chercheurs n'ont jusqu'ici

fait l'objet d'aucune étude pour l'intégration de la reconnaissance vocale au poste d'opération et il serait hasardeux de concevoir un tel système sans connaître les facteurs humains à considérer. Il est primordial de connaître les contraintes relatives aux tâches réalisées par les instructeurs de vol au poste d'opération, tout comme les caractéristiques des instructeurs de vol qui sont les principaux utilisateurs de ce système. Le langage qui sera reconnu par le système nécessite aussi une attention particulière. Malgré les progrès technologiques touchant les engins de reconnaissance vocale depuis (Fox et Weaver), des tests sur les taux de reconnaissance s'avèrent toujours nécessaires pour trouver l'engin le plus approprié. Le contexte technologique ne se résume cependant pas aux engins et au vocabulaire reconnu. Le bruit ambiant sera un facteur important à considérer dans la décision d'utiliser ou non un système de reconnaissance. Un dispositif d'activation/désactivation du système de reconnaissance de la voix sera de plus nécessaire compte tenu de la complexité et du parallélisme des tâches des utilisateurs. Enfin, comme l'ont stipulé les auteurs cités ci-dessus, l'utilisabilité doit être mesurée à l'aide de tests adéquats servant à montrer l'efficacité et la satisfaction des utilisateurs pour contrôler le poste d'opération. Ces aspects ergonomiques nous amènent à adopter une démarche de conception centrée sur l'utilisateur dans notre étude.

2.2 Objectif

Comme les entraînements en simulateur sont très onéreux, il s'avère important que l'instructeur de vol puisse bénéficier du meilleur rendement possible de ces entraînements avec les pilotes. De ce fait, la présente étude a pour objectif principal d'augmenter, par rapport à la situation actuelle, l'efficacité de l'instructeur au poste d'opération à l'aide d'une technologie de reconnaissance vocale. Cela devrait lui permettre de consacrer plus de temps à la formation des pilotes et de diminuer le temps d'interactions avec le poste d'opération (Ramesh et Sylla).

Afin d'atteindre cet objectif, l'étude devra d'abord analyser les caractéristiques et les contraintes relatives à l'intégration d'un système de reconnaissance vocale au poste d'opération de l'instructeur de vol. L'ensemble des concepts et des techniques qui nous serviront dans cette recherche sont schématisés dans la figure 2.1.

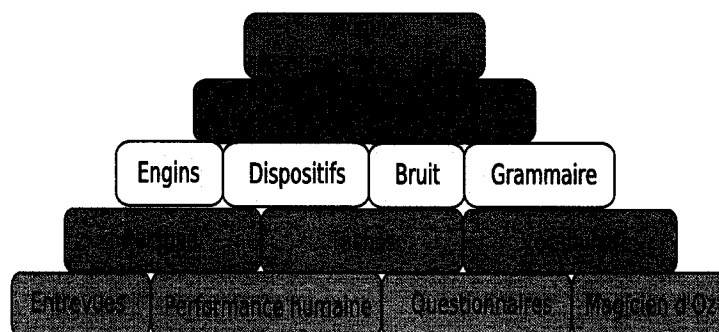


Figure 2.1 – Concepts, techniques et analyses prévus dans cette étude

Un profil d'utilisateurs sera d'abord élaboré à partir de données recueillies par des entrevues afin de mieux connaître les habitudes de travail et l'expertise des instructeurs de vol. Puis, un modèle des tâches de l'instructeur de vol sera construit. Une analyse de la performance des instructeurs à compléter leurs tâches sera effectuée afin de comparer le mode d'interaction actuel avec le mode d'interaction vocale. Une technique de recueil de données faisant intervenir les instructeurs de vol sera utilisée pour connaître les caractéristiques du vocabulaire qui servira de base à la construction du modèle de langage des commandes vocales. Cette analyse servira à améliorer les connaissances du langage opératif des instructeurs de vol. Ensuite, les caractéristiques du contexte technologique seront analysées. Les engins de reconnaissance disponibles, le bruit dans le simulateur et les dispositifs d'activation/désactivation feront l'objet d'expérimentations afin de cerner les outils technologiques à privilégier lors de l'intégration d'un système de reconnaissance de la voix. Enfin, nous allons construire un prototype expérimental intégrant un engin de reconnaissance vocale au poste d'opération. Des tests d'utilisabilité avec des utilisateurs permettront de mesurer l'efficacité du prototype, de connaître le niveau de satisfaction des utilisateurs et de caractériser l'impact de l'implantation d'un tel système au poste d'opération.

CHAPITRE 3

ANALYSE DES CARACTÉRISTIQUES DES INSTRUCTEURS DE VOL AU POSTE D'OPÉRATION

Peu de caractéristiques des instructeurs de vol sont connues. Leurs habitudes, leurs connaissances et leurs habiletés ont pourtant un impact déterminant sur la façon dont ils travaillent et interagissent avec le poste d'opération. Afin de mieux connaître ces caractéristiques, ce chapitre présente une étude que nous avons menée afin de dresser un portrait des instructeurs de vol.

3.1 Élaboration d'un profil utilisateur

La création d'un profil utilisateur est importante dans la conception des interfaces humain-machine (Hackos et Redish), et elle a une importance bien particulière lorsqu'il s'agit d'une interface vocale (Bradford). Les divers contextes dans lesquels les utilisateurs interagissent avec l'application nécessitent une connaissance approfondie de ces utilisateurs. Dans certains cas, des utilisateurs peuvent devoir interagir à voix haute avec les étudiants ou des collègues, ce qui peut créer de la confusion avec les commandes vocales. Il faut considérer dans ce cas leurs caractéristiques et la façon dont ils exécutent mentalement certaines actions. Dans d'autres cas, les émotions des utilisateurs peuvent avoir un impact sur le taux de réussite de la reconnaissance (Bradford). Pour aider à cerner efficacement les besoins et les attentes de l'utilisateur face à une interface vocale, il est important d'impliquer les utilisateurs dans le processus d'élaboration de l'application, et ce, dès le début et tout au long du cycle de vie de l'application (Yankelovich et Lai). Pour ces raisons, une démarche menant à l'élaboration d'un profil des instructeurs de vol est présentée dans cette section.

3.1.1 Méthodologie

Un profil d'utilisateur permet au concepteur de système de bien connaître les connaissances des utilisateurs, leurs expertises et leurs limites. Pour élaborer ce profil, on peut utiliser des entrevues (semi-structurées), qui demeurent une façon simple et peu coûteuse de recueillir des données (McTear).

Une fois les données recueillies, leur analyse permet de définir un ou plusieurs profils d'utilisateurs. Pour ce faire, une technique de conception interactive qui aide les concepteurs à focaliser la conception sur les besoins d'un utilisateur particulier que l'on nomme *persona* est utilisée. De plus, elle offre une base de communication partagée entre les membres de l'équipe de conception (Pruitt et Grudin). Les autres étapes de cette étude doivent ainsi tenir compte de ces profils.

3.1.1.1 Sujets

Les cinq sujets retenus sont des instructeurs de vol ou d'anciens instructeurs de vol qui connaissent l'entraînement en simulateur et le poste d'opération. Quatre des sujets ont été instructeurs sur des hélicoptères et un sujet a été instructeur sur des avions de chasse. Tous les sujets proviennent du domaine militaire, ont l'anglais comme langue maternelle et ont chacun plus de 20 ans d'expérience comme instructeur.

3.1.1.2 Outils utilisés

Le questionnaire présenté à l'annexe C nous a servi de guide pour le recueil de données. Il comporte différentes catégories d'informations à recueillir : le profil général des instructeurs de vol, les pratiques de travail utilisées au poste d'opération, les tâches reliées à l'instruction en simulateur de vol et l'impression des instructeurs face à l'intégration de la reconnaissance vocale au poste d'opération.

L'utilisation d'un enregistreur audio permet de recueillir l'ensemble de l'entrevue. Les bandes audio sont par la suite retranscrites textuellement afin de faciliter l'analyse.

3.1.1.3 Procédures et consignes

L'entrevue d'une durée d'environ 45 minutes s'effectue autour d'une table dans un environnement silencieux. Les enjeux de l'étude ainsi que les objectifs de l'entrevue sont d'abord exposés au sujet sur un document papier. Le sujet est informé que nous voulons obtenir des informations sur son travail d'instructeur et plus particulièrement sur ses interactions avec le poste d'opération. Il est informé qu'en tout temps, il peut refuser de répondre à une question, ou d'être enregistré, et qu'il peut arrêter l'entrevue pour une quelconque raison. L'entrevue débute et se termine lorsque toutes les questions ont été abordées.

3.1.2 Résultats

Les cinq bandes audio ont été retranscrites et analysées. Le tableau 3.1 présente les résultats des entrevues avec les instructeurs de vol.

Tableau 3.1: Résultats d'analyse des entrevues effectuées auprès de cinq instructeurs

Aspects	Résultats
Technologies	Un seul sujet utilise un ordinateur à la maison, les autres ne s'en servent qu'au travail
<i>suite sur la page suivante</i>	

Tableau 3.1: Résultats d'analyse des entrevues effectuées auprès de cinq instructeurs (suite)

Aspects	Résultats
Outils	Cartes d'approches imprimées, images de reproductions d'instruments, impressions météorologiques, présentations PowerPoint, outils informatiques d'aide à l'entraînement comme des bandes vidéo, des séquences d'événements, un ordinateur personnel de poche ou une manette de contrôle du simulateur. Un instructeur utilise un enregistreur audio ainsi qu'un logiciel de reconnaissance vocale pour retranscrire ses notes sur l'entraînement
Mode d'apprentissage	Seul ou avec un collègue en pratiquant à manipuler le poste d'opération. Cours de base existant selon certains.
Temps d'apprentissage	1-2 semaines selon trois instructeurs. Plus de 6 mois selon deux autres.
Manuel d'utilisation	Aucun manuel n'existait selon deux instructeurs. Trois autres se rappellent qu'un manuel existait ou était en conception, mais ne l'ont jamais utilisé.
Besoin d'assistance	En cas de scénarios complexes, division des responsabilités entre plusieurs instructeurs. En situation de familiarisation, demande d'assistance d'un collègue.
Scénarios d'entraînement	Constante réutilisation des anciens scénarios. Utilisés surtout pour les scénarios complexes. Tous s'accordent sur le fait qu'ils doivent constamment réagir aux actions des pilotes et ne peuvent donc être trop rigides.
Ergonomie du poste d'opération	Pas de table adéquate. Pas d'endroit pour poser crayon et papier. Dans le noir, sans lampe adéquate. Un instructeur mentionne des problèmes de stress au cou et au bras.
<i>suite sur la page suivante</i>	

Tableau 3.1: Résultats d'analyse des entrevues effectuées auprès de cinq instructeurs (suite)

Aspects	Résultats
Utilisation du poste d'opération	Trois instructeurs utilisent le plus d'écrans tactiles possible. Pourcentage du temps passé au poste d'opération par rapport à celui consacré aux autres tâches varie en fonction de l'expérience des pilotes, du type d'entraînement (plus élevé pour le vol aux instruments), du niveau d'interaction jouée par l'instructeur. Un instructeur évalue que 90% de son attention est donnée au poste d'opération. Utilisation fréquente des affichages suivants : alignement de descente (GSD), carte, vue furtive, répliques du cockpit, communications radio, dysfonctionnements.
Niveau de stress	Élevé lorsque : nouvel instructeur, simulation de guerre, instructeur aventureux.
Condition de bruit	Hélicoptère très bruyant. Casque d'écoute porté en tout temps. Un instructeur mentionne que le bruit n'est pas un problème et tend à le mettre à 20%.
Perception du poste d'opération	Tous les instructeurs affirment que l'interface du poste d'opération courant est correcte et nécessiterait probablement certains changements, mais qu'ils se sont adaptés avec l'expérience. Se soucient peu de l'apparence, en autant qu'il fait ce qu'il doit faire.
Situations frustrantes	Surviennent lorsque : trop de pages à ouvrir ou manque de logique, boutons trop petits, peu configurable, manque d'écrans, doit travailler avec deux doigts, écrans tactiles non calibrés, budget limité impliquant la suppression d'outils (caméra, vue furtive, réplique d'instruments, chaise supplémentaire).
Histoires	Boîtes de dialogue d'erreur interrompant la simulation. Nouvel instructeur ayant des problèmes avec le poste d'opération et manquant des actions des pilotes. Bras de positionnement d'un écran tactile trop flexible et heurtant un instructeur. Instructeur plus âgé peu habitué aux ordinateurs ayant de la difficulté à comprendre le poste d'opération. Entraînement en simulateur d'une mission réelle à venir en Bosnie.
<i>suite sur la page suivante</i>	

Tableau 3.1: Résultats d'analyse des entrevues effectuées auprès de cinq instructeurs (suite)

Aspects	Résultats
Fonctionnalités critiques	Trois instructeurs affirment que la carte est la fonctionnalité la plus critique, principalement en entraînement tactique.
Fonctionnalités moins importantes	Interfaces d'état et de maintenance du simulateur. Dépend beaucoup du type d'entraînement.
Arrêt d'une simulation	Selon quatre instructeurs sur cinq, uniquement si les pilotes ont tout faux et ne réussiront pas à atteindre un objectif de façon satisfaisante.
Session d'entraînement	Étapes selon tous les sujets : 1) S'assure que les pilotes connaissent bien l'entraînement. 2) Séance d'information ("briefing") du capitaine, puis de l'instructeur. 3) Entraînement en simulateur, utilisation de scénarios ou de leçons. 4) Séance de retour d'expérience ("debriefing") des pilotes, puis de l'instructeur.
Ajout de la reconnaissance vocale	Les cinq instructeurs interrogés aimeraient voir cette fonctionnalité et l'essayer. Un instructeur croit qu'à l'intérieur du simulateur ("onboard") c'est impossible et qu'à l'extérieur ("offboard") ce serait possible. Un autre affirme que les hélicoptères militaires sont trop bruyants. Trois instructeurs mentionnent qu'il est important de déterminer comment cela sera activé. Enfin, un instructeur souligne l'importance d'afficher le résultat du changement à l'écran.
Commandes utiles par la voix	Affichage de pages (2 instructeurs). Tout ce qui est possible (1 instructeur). Trafic et dysfonctionnements (1 instructeur). Macro effectuant plusieurs fonctions (1 instructeur). Aucune (1 instructeur).

3.1.3 Description des profils

La méthode de la *persona* est utilisée pour présenter des personnages fictifs intégrant les caractéristiques et les habitudes d'instructeurs de vol. Cette méthode, tel que stipulé précédemment, sert de base de référence pour la conception d'interfaces et facilite l'identification des fonctionnalités requises

et des adaptations particulières nécessaires. Les personas serviront à orienter le développement des applications du poste d'opération et à ajouter le nouveau mode d'interaction qu'est la reconnaissance vocale. Les choix de conception qui seront effectués plus loin concernant le système de reconnaissance vocale au poste d'opération devront ainsi répondre aux besoins et aux connaissances des personnages décrits.

Deux représentations fictives sont décrites, l'une pour un instructeur civil et l'autre pour un instructeur militaire, puisque ces deux classes d'instructeurs n'ont pas le même profil.

Simon, instructeur de vol pour la compagnie VolExpert



©Photographe : Lyn Baxter | Agence : Dreamstime.com

Simon a 37 ans et est instructeur depuis 5 ans. Il a volé sur plusieurs gros transporteurs aériens pendant 10 ans et est aujourd'hui qualifié pour enseigner dans les simulateurs de vol. Il fait environ deux séances d'entraînement sur les simulateurs par jour lorsqu'il est au centre d'entraînement. Il n'est pas un habitué des systèmes informatiques et préfère les sensations à bord des vrais appareils. Il considère que les simulateurs sont polyvalents, mais bien compliqués à utiliser. Il a de la difficulté à repérer les paramètres qu'il désire changer puisqu'il n'y est pas habitué. Pour Simon, il est plus facile d'utiliser le simulateur avec les scénarios préenregistrés que de configurer lui-même l'environnement virtuel. Il aimerait cependant être en mesure de mieux se servir du simulateur pour créer des scénarios plus complexes et mettre les pilotes dans des situations plus ardues. Dans ces situations, il doit faire appel à des collègues pour l'aider à créer de nouveaux scénarios. Simon connaît très bien les situations qu'il doit faire pratiquer aux pilotes et prend des notes constamment sur le déroulement de l'entraînement. Il garde son cartable de scénario sur ses genoux pour connaître exactement le déroulement du scénario en cours et donner les bonnes indications aux pilotes. Il prend le temps d'arrêter la simulation pour discuter de certaines situations avec les pilotes et leurs donner des conseils. Il connaît parfaitement les communications radio formelles de l'aviation et ne se gêne pas pour les appliquer pour mettre ainsi ses élèves en situation.

Frank, instructeur d'hélicoptère Chinook pour les forces armées



©Photographe : Xavier Marchant | Agence : Dreamstime.com

Frank est dans l'armée depuis 20 ans. Âgé de 52 ans, il a volé plus de 8 000 heures et est qualifié pour former les pilotes pour les missions tactiques et de vol de nuit. Il passe ses journées à la base à enseigner en salle de cours ou sur des appareils d'entraînement et fait un saut de temps en temps avec ses élèves au centre d'entraînement sur les simulateurs. Il connaît très bien les scénarios d'entraînement et est habile pour entrer les paramètres d'initialisation du simulateur. Il utilise efficacement tous les écrans tactiles qui sont disponibles en même temps. Il aime tenter de nouvelles configurations des paramètres du simulateur pour changer la simulation. Par contre, il est parfois distrait à regarder hors de l'appareil, les pilotes s'exécuter ou à configurer les paramètres du simulateur. Il connaît pourtant très bien l'interface du poste d'opération, mais a quelques problèmes à se souvenir de l'état du système. Il est très volubile et aime jouer le rôle des contrôleurs aériens ou des autres appareils lui-même pour rendre l'entraînement plus réaliste. De ce fait, il met le volume du bruit de la simulation au minimum pour pouvoir entendre et parler librement avec les pilotes. Frank ne prend pas beaucoup de notes pendant l'entraînement, préférant regarder les pilotes ou modifier les paramètres de la simulation et prendre des notes pendant les temps morts ou après l'entraînement.

3.2 Sommaire

Des entrevues semi-structurées ont été réalisées avec cinq instructeurs de vol afin de connaître les caractéristiques de leur travail et du milieu, leurs habitudes et leurs perceptions. Certains résultats, comme la non-utilisation d'un manuel, l'aspect collaboratif de l'apprentissage au poste d'opération, la forte utilisation du poste d'opération ou les situations frustrantes, sont intéressants pour guider la conception du poste d'opération. De plus, les données sur les doutes des instructeurs au sujet de l'intégration de la reconnaissance vocale seront utiles pour le reste de l'étude. Enfin, nous avons créé deux *personas* représentatives des instructeurs de vol auxquelles on pourra se référer dans la suite de l'étude lorsqu'il sera nécessaire de justifier un choix de modélisation sur la base des utilisateurs ciblés. Puisque l'objectif est de tester l'efficacité des instructeurs à l'aide de la reconnaissance vocale, ce chapitre a permis de caractériser les utilisateurs ciblés afin de guider la conception d'une telle interface. Le chapitre suivant abordera plus spécifiquement les tâches de ces utilisateurs.

CHAPITRE 4

ANALYSE ERGONOMIQUE DES TÂCHES DE L'INSTRUCTEUR DE VOL

Ce chapitre décrit les tâches de l'instructeur de vol au poste d'opération. De plus, nous présentons une analyse comparative entre l'interaction manuelle et l'interaction vocale pour ce qui est du temps d'exécution des certaines tâches de paramétrisation du simulateur. Cette analyse nous donne une meilleure estimation de l'efficacité du mode d'interaction que l'on désire intégrer par rapport au mode d'interaction actuel.

4.1 Modélisation des tâches

Le travail de l'instructeur de vol se divise en plusieurs étapes et comprend de nombreuses opérations. La figure 4.1 est une décomposition hiérarchique de ces étapes de haut niveau inspirée de Ramesh et Sylla (1990) qui se sont intéressés au travail de l'instructeur dans le but de développer un modèle pour optimiser le poste d'opération à la fin des années 80. Cette décomposition prend en compte toutes les étapes du travail de l'instructeur, autant à l'extérieur du simulateur qu'à l'intérieur pendant l'entraînement. De nombreuses opérations sont réalisées à l'extérieur du simulateur puisqu'une grande partie du travail de l'instructeur consiste en la préparation des entraînements et en la rédaction de documents.

La figure met en relief les étapes "Initialiser" et "Entraîner" puisqu'elles sont au centre de cette étude. C'est à travers l'optimisation de l'exécution des tâches relatives à ces étapes qu'il sera possible, selon l'hypothèse, d'augmenter l'efficacité des instructeurs.

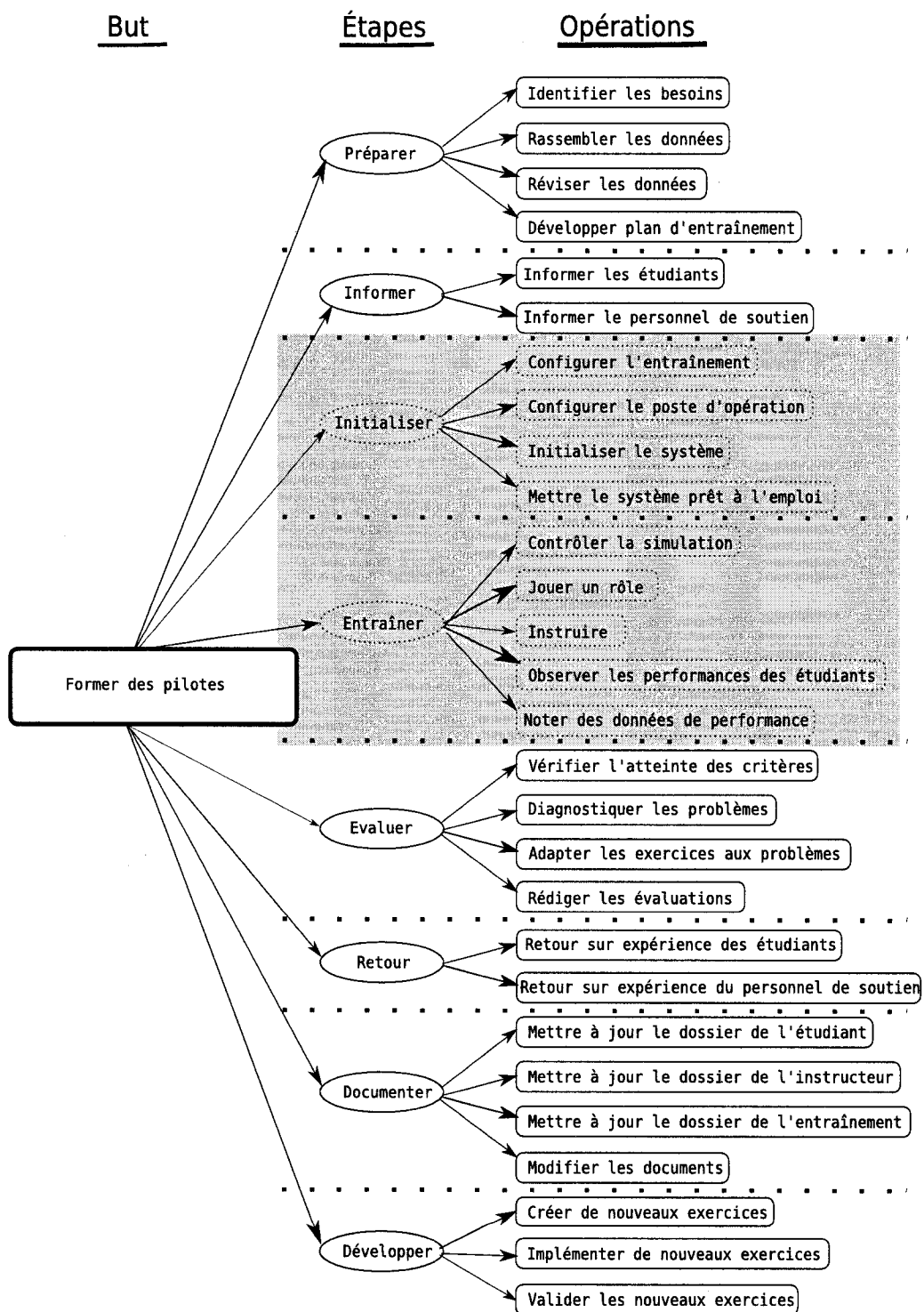


Figure 4.1 – Modèles hiérarchiques des activités de l'instructeur de vol

4.2 Analyse des temps d'exécution des tâches selon deux modes d'interaction

L'objectif de cette section est de déterminer si l'exécution d'une tâche est plus rapide par commande vocale ou par manipulation directe sur un écran tactile. La revue de littérature présentée dans la section suivante montre plusieurs résultats d'études qui ont comparé la commande vocale et l'interaction manuelle.

4.2.1 Revue des études qui comparent interactions tactile et vocale

L'étude de Tsimhoni et al. (2004) consistait à comparer l'interaction vocale avec l'interaction manuelle sur un écran tactile pour l'entrée de données (des adresses civiques) au cours de la conduite automobile. Un simulateur de conduite a été utilisé et 24 conducteurs différents ont participé à l'étude. La comparaison s'est effectuée entre l'interaction vocale avec des mots complets ou des caractères (lettres ou symboles) et la saisie de données sur un clavier virtuel présenté sur un écran tactile. L'étude a révélé que l'interaction vocale est plus rapide que l'interaction manuelle dans tous les cas étudiés de conduite avec courbe ou d'arrêt du véhicule. De plus, l'étude a démontré que le contrôle du véhicule était moins fluide lors de la saisie de données au clavier et que l'augmentation de la charge de travail lors de la conduite augmentait grandement le temps requis pour entrer une adresse au clavier. Les sujets de l'étude ont perçu la saisie de données au clavier comme étant plus difficile que la reconnaissance vocale. Notons enfin une recommandation des auteurs qui conseillent l'utilisation de la technique du Magicien d'Oz pour la simulation d'un système de reconnaissance vocale.

Lee et Lai (2005) ont réalisé une étude sur les préférences des utilisateurs pour la reconnaissance vocale ou l'entrée au clavier numérique (téléphonique) pour un système de gestion de messages et de calendrier.

L'étude a conclu que les utilisateurs préféraient l'utilisation du clavier pour des tâches linéaires ou directes (par exemple la réception de messages), et la reconnaissance vocale pour des tâches non-linéaires ou complexes (par exemple la recherche d'une entrée spécifique au calendrier). L'interaction par la voix a de plus été évaluée comme étant plus satisfaisante, amusante et facile que le clavier numérique.

Des études faites par Damper et ses collègues (Damper et Wood; Damper et al.) créent une certaine controverse à propos de la supériorité de l'interaction vocale par rapport à la saisie au clavier dans le cas d'applications de commande et de contrôle. Leur étude tente d'infirmer les résultats d'une étude un peu ancienne réalisée par Pooch (1980) qui révélait que l'interaction vocale était plus rapide que la saisie au clavier pour ce type d'applications. Or, pour des commandes saisies au clavier de quelques lettres seulement (par exemple "GTE" pour la commande "GO TO ECHO"), Damper et ses collègues ont observé, en ajustant la méthodologie de Pooch pour en éliminer certains biais, que l'interaction vocale était environ 10.6% plus lente que la saisie au clavier et était de loin plus propice aux erreurs. Damper a pris en compte dans sa méthodologie la notion de cycle de transaction, appelée de cette façon par Morrison et al. (1984). Cette notion consiste à considérer le nombre d'actions requises par l'utilisateur pour obtenir une certaine réponse du système, ce que Pooch n'avait pas considéré en prenant des chaînes de caractères trop longues qui requéraient plusieurs touches au clavier. Ainsi, dans leur première étude, Damper et Wood (1995) n'ont pas laissé le choix à l'utilisateur des commandes à entrer et ne lui demandaient pas d'exécuter d'autre tâche concurrente. Dans une deuxième étude (Damper et al.), les sujets devaient compléter la tâche primaire d'exécuter les commandes via la voix ou le clavier, et comme tâche secondaire, ils devaient retranscrire sur une feuille de papier certaines informations sans lien avec la tâche primaire. Les résultats ont ainsi révélé que la tâche secondaire était plus complétée lors de l'utilisation de l'interaction vocale et qu'en ne tenant pas compte des erreurs du module de reconnaissance vocale, l'interaction vocale était clairement supérieure. Le potentiel de l'utilisation de cette technologie est donc pour ces auteurs dans des situations de tâches concurrentes où il y a

une forte charge cognitive.

4.2.2 Méthodologie

Afin de vérifier si l'interaction avec le poste d'opération est plus rapide selon que l'on utilise un écran tactile ou la voix, l'analyse suivante permet d'estimer la vitesse à laquelle un utilisateur d'interface vocale est en mesure d'exécuter une tâche. Nous émettons l'hypothèse que pour l'ensemble des tâches étudiées (voir la section 4.2.2.3), effectuées par un instructeur de vol au poste d'opération dans un simulateur, l'interaction vocale est plus rapide que l'interaction manuelle sur un écran tactile. Cette hypothèse s'appuie sur les résultats des études revues dans la section précédente qui montrent la supériorité de l'interaction vocale pour des tâches similaires à celles qui sont étudiées. À noter que la syntaxe des commandes utilisées pour cette analyse est directement liée aux tâches, et que le vocabulaire utilisé est naturel et non optimisé. Or, un réaménagement pourrait éventuellement améliorer l'efficacité de l'interaction vocale. De plus amples détails sur cet aspect sont disponibles au chapitre 5.

4.2.2.1 Modèles prédictifs de performance humaine

La première partie de cette expérimentation est basée sur la méthode d'analyse de tâche *GOMS* (*Goals, Operators, Methods and Selection*) élaborée par Card et al. (1983). Les quatre éléments de cette méthode se décrivent ainsi :

Goals (Buts) : Les buts de l'utilisateur ou ce qu'il désire accomplir.

Operators (Actions) : Les actions représentent le plus bas niveau de l'analyse de la tâche. Elles peuvent représenter un changement à interpréter par le système, une saisie de données au clavier, ou encore une modification de l'état mental de l'utilisateur suite à la lecture d'une boîte de dialogue.

La granularité de ces actions est très flexible et laissée à la discrétion de l'analyste qui doit déterminer la granularité adéquate pour le public cible de son analyse.

Methods (Méthodes) : Les méthodes correspondent aux différentes possibilités offertes à l'utilisateur pour accomplir son but. Par exemple, pour fermer une application, l'utilisateur a la possibilité de cliquer sur l'icône de fermeture en haut à droite de la fenêtre ou encore d'appuyer sur "ALT-F4".

Selection (Sélection) : La sélection consiste à choisir une méthode parmi celles qui sont offertes à l'utilisateur pour accomplir un but. Cette sélection n'est pas faite au hasard et l'analyse faite avec *GOMS* cherche à prédire la méthode qui sera utilisée par l'utilisateur en tenant compte des connaissances et des compétences de l'utilisateur, ainsi que de l'état du système et du but recherché.

Cette méthode décrit l'interaction humain-ordinateur comme une suite des actions élémentaires de l'utilisateur avec le système dont on connaît les temps d'exécution. En utilisant ces actions élémentaires, il est possible de prédire le temps qu'un utilisateur expert prendra pour accomplir une certaine tâche sans avoir à observer et enregistrer l'utilisateur pendant qu'il effectue la tâche. Cette méthode considère deux parties à une action élémentaire : le temps d'acquisition de la tâche ou de préparation, et le temps d'exécution. Ainsi, l'analyse devient quantitative en se basant sur les temps que prennent en général ces actions. La méthode *KLM-GOMS (Keystroke-Level Model)* (Card et al.) a été développée dans cette optique et utilise des actions élémentaires telles que la préparation mentale, le positionnement des mains sur la souris ou le clavier, l'entrée d'une touche, un clic de souris, etc. Cela permet à un analyste d'évaluer différentes conceptions d'interfaces graphiques et de trouver celle qui donnera la meilleure performance pour un utilisateur expert. Il est cependant important de noter que l'analyse faite avec la méthode *GOMS* ne tient pas compte de la fatigue de l'utilisateur et des erreurs qu'il peut commettre.

L'outil *CogTool* (CogTool) permet d'automatiser l'évaluation de

différentes conceptions d'interfaces en utilisant la méthode *KLM-GOMS*. Cet outil, présenté à l'annexe B, est utilisé afin d'analyser deux versions du poste d'opération de l'instructeur de vol. La première version est utilisée dans plusieurs simulateurs de vols pour les lignes aériennes commerciales, alors que la seconde version est encore en développement et a été créée dans le but de maximiser la quantité d'information fournie à l'utilisateur. Cette dernière conception a de plus été développée à l'aide d'outils et d'environnements de développement graphique plus récents et fait appel à des concepts d'interfaces graphiques de Microsoft WindowsTM. L'analyse de ces deux interfaces permet de vérifier si la conception de l'interface graphique pour écran tactile a un impact sur l'efficacité de l'utilisateur et valide l'hypothèse de comparaison de l'interaction vocale et tactile, indépendamment de cette conception.

Cette méthode d'analyse a été utilisée dans plusieurs recherches. John et al. (2004) décrivent plusieurs applications de cette méthode qui démontrent sa validité. L'architecture logicielle du *Adaptive Control of Thought (ACT-R)* simule et prédit les performances humaines. Le groupe de recherche qui travaille sur ACT-R au département de psychologie de l'université Carnegie Mellon a créé entre autres des modèles de l'apprentissage humain, de la mémoire, de la résolution de problèmes et de la prise de décisions (ACT-R Research Group). Les résultats des modèles ont été comparés aux résultats de sujets effectuant les mêmes tâches. Le temps pour accomplir une tâche, la précision du résultat et certaines données neurologiques sont des mesures cognitives utilisées pour valider ces modèles. Le logiciel *CogTool* utilise cette architecture ACT-R et rend la prédiction de la performance humaine plus efficace (John et al.; Teo et John).

Les temps d'acquisition et d'exécution qui sont utilisés par le modèle de calcul de *CogTool* pour l'interaction avec un écran tactile nous sont malheureusement inconnus. Nous savons par contre que pour chacune des actions ou chaque groupe d'actions reliées, un temps de de préparation mentale de 1.2 seconde est ajouté puisque l'opérateur doit penser à l'action qu'il désire effectuer. De plus, un certain temps est alloué pour atteindre le contrôle

à utiliser et ensuite pour exécuter l'action d'appui sur le bouton. De plus amples détails sont disponibles dans (Card et al.; John et al.; Teo et John).

L'étude de Matessa et Remington (2005) considère un temps de 720 millisecondes pour l'interaction avec un écran tactile, soit 420 millisecondes pour repérer l'information, 50 millisecondes pour initialiser l'action et 200 millisecondes pour exécuter l'action sur l'écran. Ces données nous fournissent ainsi une idée du temps que peut mettre un utilisateur pour interagir avec un écran tactile.

4.2.2.2 Estimation des temps d'exécution à l'aide de la voix

La seconde étape de cette expérimentation consiste en l'estimation théorique du temps nécessaire pour exécuter une commande par la voix. Cette estimation prend en considération le nombre de mots que contiennent les commandes vocales. La vitesse théorique de production de la parole est établie sur la base des données recueillies dans la littérature. Un temps de 1.2 secondes est ajouté aux temps de production vocale calculés pour tenir compte de la préparation mentale des sujets avant chaque tâche.

En vue d'obtenir une estimation plus rigoureuse des temps d'exécution des tâches à l'aide des commandes vocales, une troisième étape consiste à recueillir et analyser des échantillons de voix humaines. Cela permet de mesurer les temps que prennent ces sujets pour dire certaines commandes pour ensuite calculer une moyenne pour chacune de ces commandes. Enfin, là aussi, un temps de 1.2 secondes est ajouté aux temps moyens recueillis pour les mêmes raisons que celles décrites ci-dessus et pour pouvoir comparer ces temps d'exécution avec les temps estimés pour les interactions tactiles.

4.2.2.3 Choix des tâches analysées

Le scénario d'activités qui représente un entraînement de vol pour une équipe pilote-copilote d'une ligne de transport commerciale est décrit à l'an-

nexe A. Ce scénario est inspiré des observations faites sur le terrain par une ergonome pendant une séance d'entraînement réelle. Ce scénario qui comprend l'initialisation de la simulation et quelques commandes en vol donne un bon exemple de ce qui se passe généralement pendant un entraînement en simulateur. Plusieurs autres tâches peuvent être accomplies par l'instructeur, mais uniquement celles décrites dans ce scénario sont utilisées pour évaluer l'efficacité des deux interfaces du poste d'opération.

4.2.2.4 Choix des commandes vocales

Les commandes vocales ont été extraites des différentes sous-tâches présentées à l'annexe A ; elles sont énumérées dans le tableau 4.1. Elles ont été prises à titre de référence pour l'analyse et ne constituent pas le résultat d'une analyse exhaustive du vocabulaire qu'utiliseraient les instructeurs de vol pour effectuer ces tâches dans un simulateur de vol. La combinaison de plusieurs commandes en une même phrase pourrait par exemple améliorer l'interaction vocale. Cette relation un-pour-un permet de considérer le pire cas d'interaction vocale tout en gardant l'aspect naturel des commandes. À noter que les manipulations via l'écran tactile se conforment aussi à cette relation et qu'aucun raccourci clavier, de même qu'aucune combinaison de tâches en une seule action n'est utilisée. Cette liste de commandes est ainsi utilisée à titre indicatif pour calculer les temps d'exécution des sous-tâches en utilisant la voix. Un raffinement de ces commandes est fait au chapitre 5.

Tableau 4.1: Liste des commandes vocales analysées, divisées par tâches

Tâches - Commandes
Ajuster le carburant – Set Fuel Weight to fifteen thousand
<i>suite sur la page suivante</i>

Tableau 4.1: Liste des commandes vocales analysées, divisées par tâches (suite)

Tâches - Commandes
Ajuster l'aéroport de référence - Set Reference Airfield to L I R F - Set Reference Runway to sixteen L Reposition - Reposition to Takeoff Activer la simulation - Unfreeze flight Ajuster les nuages - Set Cloud cover to scattered - Set Cloud visibility to zero point five - Set Cloud top to ten thousand - Set Cloud base to two hundred Ajuster le vent - Set Wind Direction to one eight zero - Set Wind Speed to forty five - Set Air Temperature to twenty Ajuster les conditions de la piste - Set Runway Roughness to three - Set Runway Contaminant to Dry Ajuster le visuel - Set Visual to Night
<i>suite sur la page suivante</i>

Tableau 4.1: Liste des commandes vocales analysées, divisées par tâches (suite)

Tâches - Commandes
Activer un dysfonctionnement – Set malfunction right engine flameout at twenty feet
Ajuster la vitesse de simulation – Set Slew Forward at three
Réinitialisation – Reset all systems – Reset all temperatures
Désactiver les dysfonctionnements – Clear malfunctions

4.2.3 Résultats d'analyse de l'interaction tactile

Le logiciel *CogTool* a été utilisé pour analyser deux conceptions différentes d'interfaces graphiques pour le poste d'opération de l'instructeur de vol dans un simulateur. L'utilisation de deux conceptions différentes permet de valider les résultats obtenus en comparant la moyenne des temps d'exécution via l'interaction tactile avec les temps d'exécution par commandes vocales. Pour les tâches décrites dans le scénario présenté à l'annexe A, nous avons utilisé des captures d'écran pour calculer les temps d'exécution de ces tâches avec les deux conceptions. Les figures 4.3 et 4.2 montrent les deux conceptions d'interfaces graphiques qui ont été analysées à l'aide du logiciel *CogTool*. Les tâches retenues pour cette analyse sont celles permettant d'exécuter les tâches du tableau A.1 à la page 138 (annexe A).

Les deuxième et troisième colonnes du tableau 4.3 présenté à la page 61

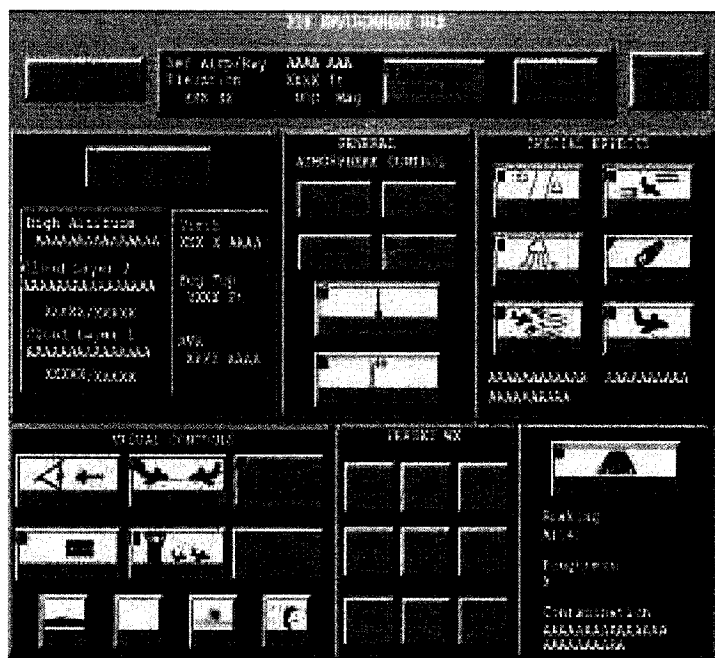


Figure 4.2 – Conception actuelle du poste d'opération des simulateurs civils (GUI 1)

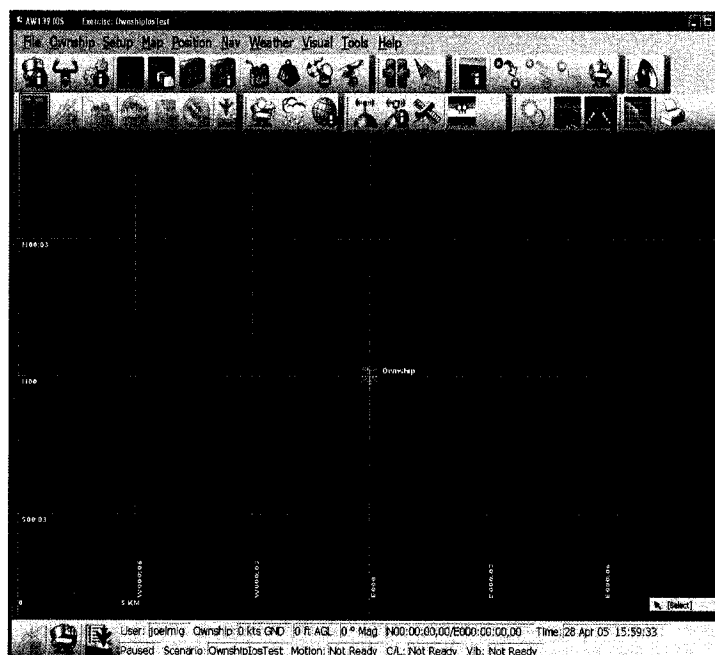


Figure 4.3 – Nouvelle conception du poste d'opération (GUI 2)

montrent les temps calculés pour les deux conceptions. On remarque que la conception montrée à la figure 4.2 est plus efficace que celle de la figure 4.3 pour la majorité des tâches. Plusieurs raisons semblent pouvoir expliquer ces résultats. Tout d'abord, les boutons de l'interface 4.2 sont plus gros et donc plus facilement accessibles lorsqu'on appuie sur l'écran tactile. La loi de Fitts permet de calculer ces temps d'exécution en fonction de la distance et de la grosseur des boutons. Il est important dans le cas des écrans tactiles de considérer cet aspect, puisque le contact avec l'interface n'est pas aussi précis qu'avec une souris. Les boutons de l'interface 4.3 respectent les standards décrits dans (DoD; Human Factors Group of the FAA William J. Hughes Technical Center) où la largeur minimale d'un bouton pour un écran tactile est de 16 mm. Or, pour un écran de 17 pouces et une résolution de 1024 par 768, les boutons ont exactement cette largeur minimale (plus petits encore pour une résolution plus grande). De plus, l'espace entre les boutons devrait être au minimum de 3 mm, ce que l'on ne retrouve pas dans plusieurs fenêtres de cette même interface. Enfin, l'interface 4.2 utilise des regroupements de contrôles sous des catégories générales offrant l'accès à plusieurs paramètres du simulateur sur une même fenêtre, alors que l'interface 4.3 offre plutôt une fenêtre par ensemble de paramètres. Il est donc nécessaire dans ce dernier cas d'ouvrir une nouvelle fenêtre chaque fois que l'utilisateur désire entrer de nouveaux paramètres pour un module particulier de la simulation. Il est cependant important de noter que nous avons recréé une même séquence d'actions pour les deux conceptions. Par exemple, dans l'interface 4.2 les fenêtres ne sont jamais fermées et l'utilisateur ne fait que superposer les fenêtres les unes aux autres. Or, les actions faites avec l'interface 4.3 vont aussi en ce sens, c'est-à-dire que la fermeture des fenêtres n'est jamais effectuée à la fin d'une séquence pour accomplir une tâche. Enfin, notons qu'un temps d'environ 9.2 secondes est nécessaire en moyenne pour exécuter une tâche via l'interaction tactile.

4.2.4 Résultats d'analyse de l'interaction vocale

L'estimation de la production des commandes vocales de l'instructeur de vol en simulateur a pu être effectuée selon deux méthodes différentes. Tout d'abord, nous avons estimé les temps pour produire une commande en utilisant la valeur théorique de 175 mots par minute issue de la littérature (voir le tableau 4.2). La colonne "Voix 1" du tableau 4.3 présente les différents temps obtenus à l'aide de cette méthode. Les commandes utilisées et présentées dans le tableau 4.1 ont en moyenne cinq mots, l'ensemble comprenant une commande minimale de deux mots et une commande maximale de huit mots. Les résultats montrent un temps moyen de 4.6 secondes pour commander une action en utilisant cet ensemble de commandes.

Tableau 4.2 – Vitesse de la parole selon diverses sources

Vitesse	Référence
150 à 200 mots par minute	Newell et al. (2003)
111 à 291 mots par minute	Yuan et al. (2006)
223 mots par minute (homme)	Liberman (2006)
125 à 150 mots par minute	Fulford (1992)
150 à 250 mots par minute	Rossi et al. (1981)
150 mots par minute	Minker et Bennacef (2004)
300 syllabes par minute	Wood (1973)

De plus, les temps de production des commandes ont été validés expérimentalement. Six sujets humains ont participé à l'étude en enregistrant leur voix pendant la lecture des commandes du tableau 4.1 sous format audio *wav*. Une analyse de l'oscillation sonore de ces enregistrements à l'aide du logiciel Adobe Audition® a permis de déterminer le temps de chacune des commandes. Un simple repérage manuel du début et de la fin de chaque commande permet d'obtenir cette mesure de temps. Une vitesse moyenne de 156 mots par minute a été estimée pour l'ensemble des commandes des six sujets. Les résultats des temps d'exécution mesurés à l'aide de cette méthode sont présentés dans la colonne "Voix 2" du tableau 4.3. En moyenne, un temps de 4.9 secondes a été nécessaire aux sujets pour dire une commande. Notons que pour les estimations théoriques et les estimations

Tableau 4.3: Temps d'exécution des tâches en secondes estimés à l'aide du logiciel CogTool pour deux versions de l'interface graphique (GUI 1, GUI 2) et les temps de production vocale théoriques (Voix 1) et empiriques (Voix 2) (suite)

Tâches	GUI 1	GUI 2	GUI Moy.	Voix 1	Voix 2	Voix Moy.	Moy. Diff. GUI-Voix
Ajuster le visuel	3.8	3.9	3.8	2.7	2.6	2.7	1.2
Activer un disfonctionnement	12.4	13.7	13.1	4.1	4.5	4.3	8.8
Ajuster la vitesse de simulation	7.3	7.8	7.5	3.1	3.3	3.2	4.3
Réinitialisation	5.3	5.4	5.4	4.8	4.1	4.4	0.9
Désactiver les disfonctionnements	1.8	1.9	1.9	2.0	2.6	2.3	-0.4
	8.9	9.6	9.2	5.2	4.6	4.9	4.4

Pour la majorité des tâches, la voix est plus rapide. L'interaction via l'écran tactile est plus rapide uniquement pour les tâches d'activation de la simulation et de désactivation des disfonctionnements. Cela est dû au fait que ces tâches ne requièrent aucune navigation dans l'interface puisque des boutons sont directement accessibles à l'utilisateur sur l'interface principale. Le fait que les commandes vocales associées ne comportent que deux mots ne vient pas jouer un rôle assez significatif pour que l'interaction vocale soit plus rapide. Les tâches nécessitant plus de manipulations sur l'interface graphique sont par contre beaucoup plus rapides avec la voix.

La figure 4.4 montre que la voix est en générale plus rapide que l'interaction via l'écran tactile. Un gain moyen de 33% en temps d'exécution de

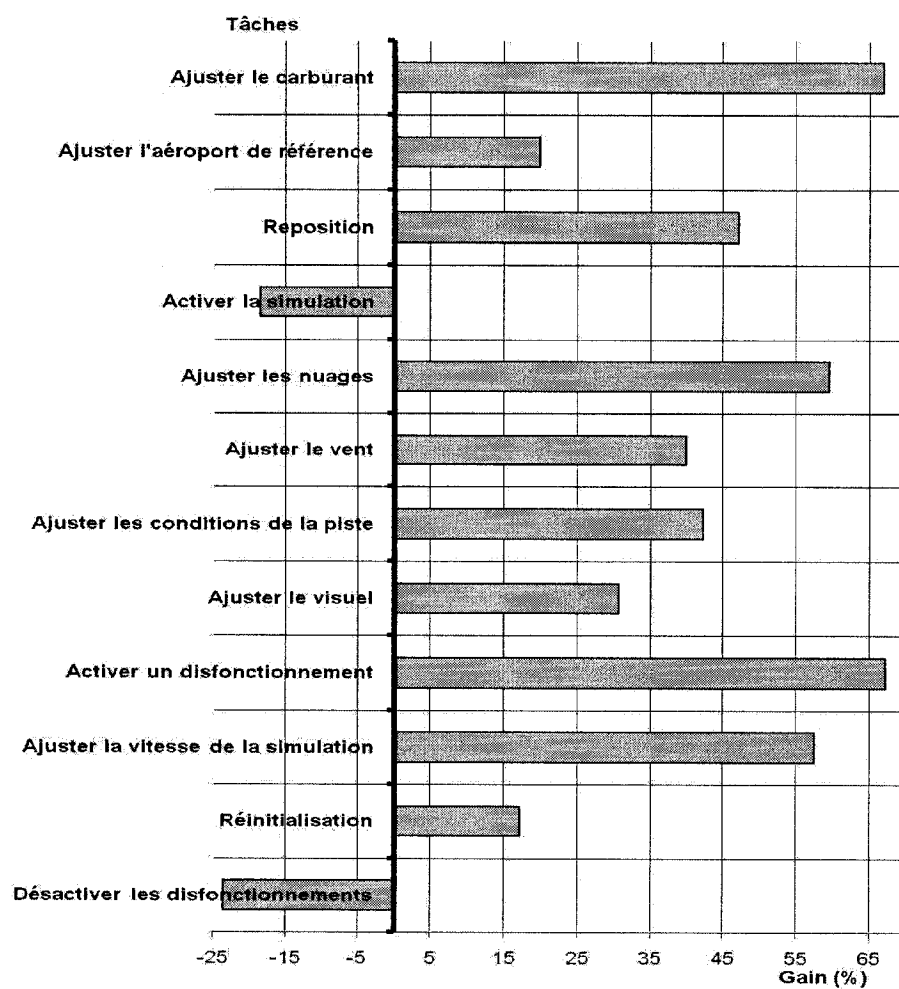


Figure 4.4 – Pourcentage du gain de temps moyen obtenu grâce à l'utilisation de commandes vocales

tâches pour l'ensemble des commandes est calculé. Le même résultat est observable en utilisant les estimations théoriques et les données empiriques. Le calcul du pourcentage de gain pour chacune des tâches a été effectué en comparant le temps de production de chaque commande vocale avec le temps le plus court d'exécution de la tâche avec les deux interfaces graphiques. L'interaction vocale n'est cependant pas adaptée à toutes les commandes, puisqu'il y a une perte de productivité pour certaines commandes qui sont plus rapides en utilisant l'écran tactile.

4.2.6 Discussion

Nous avons pu observer que la vitesse de production de la parole contenue dans la littérature, soit environ 175 mots par minute, est une bonne estimation de la production humaine. Nous avons pu remarquer à l'aide d'une étude empirique que cette estimation se rapproche grandement des temps réels de production mesurés dans des conditions normales pour différentes commandes vocales. De plus, l'analyse précédente nous a permis de comparer l'efficacité de l'interaction vocale avec celle de l'interaction manuelle sur écran tactile lorsqu'on effectue diverses tâches au poste d'opération de l'instructeur en simulateur de vol. Pour la majorité des tâches étudiées, la voix est le mode d'interaction le plus rapide. Toutefois, certaines tâches sont plus rapides avec l'écran tactile puisque le bouton graphique apparaissant à l'écran est directement accessible à l'utilisateur.

4.3 Sommaire

Le travail de l'instructeur de vol est complexe et comporte plusieurs activités que l'on peut découper en étapes et en opérations bien particulières. L'initialisation du simulateur et l'entraînement lui-même sont des étapes à améliorer en vue d'augmenter l'efficacité de l'instructeur.

De plus, les résultats de l'analyse de la performance de l'instructeur pour certaines tâches démontrent que l'intégration de la reconnaissance vocale au poste d'opération de l'instructeur de vol est profitable. Cela permettrait à ce dernier de gagner du temps dans la paramétrisation de la simulation et de conserver plus de ressources attentionnelles pour la formation. Une interface multimodale permettrait ainsi à l'instructeur d'optimiser ses interactions et de contrôler la simulation de manière plus intuitive. Une analyse approfondie du vocabulaire de travail qui devrait être utilisé pour la commande vocale dans un simulateur de vol est cependant nécessaire. Le chapitre 5 qui suit décrit cet aspect.

CHAPITRE 5

ANALYSE DU LANGAGE DE L'INSTRUCTEUR DE VOL

La voix ou la parole, bien qu'elle soit le mode de communication le plus naturel (Minker et Bennacef), n'est pas uniquement le résultat de fréquences sonores émises par l'Homme, mais aussi des aspects cognitifs qui s'y rattachent, tels que déterminer ce qui est juste à dire et comment le dire (Barber).

Le langage de commandes que l'instructeur de vol utilisera pour commander vocalement certaines actions du poste d'opération est une donnée importante de cette étude. Le vocabulaire ainsi que la syntaxe des commandes reconnues par le système de reconnaissance doivent être spécifiés de façon adéquate. Le temps d'apprentissage, la facilité d'utilisation et l'aspect naturel de l'interaction vocale sont les facteurs mis en jeu. En vue de fournir la meilleure interaction vocale possible, le langage devra être conçu en vue d'optimiser ces facteurs.

Aucun guide n'est disponible pour aider à la création de commandes vocales. Cependant, des articles concernant les langages de commandes textuelles ont été publiés dans les années 80. Plusieurs auteurs parlent de l'interface vocale en la comparant avec une interface de type "console" où une commande textuelle doit être écrite à l'aide du clavier pour être exécutée (Hakulinen et al.; Yankelovich). Dans les deux cas, le nom de la fonctionnalité à utiliser pour atteindre le but recherché est invisible et les limites de ce qui peut être commandé et de ce qui n'est pas supporté ne sont pas affichées à l'écran. Cela crée donc un besoin d'apprendre la base du fonctionnement du système avant de l'utiliser. L'approche "essai et erreur" qui est souvent utilisée par les utilisateurs d'interfaces graphiques n'est plus envisageable dans des circonstances où aucun indice n'est disponible. Il est donc raison-

nable de penser que les recherches entourant les paradigmes de commandes textuelles peuvent servir de guide à la création de commandes vocales.

Ce chapitre définit tout d'abord les termes se rapportant à l'analyse du langage de commande afin de créer une référence terminologique et d'éviter la confusion. Une revue de la littérature traitant des langages de commandes textuelles est ensuite présentée afin de déceler certains facteurs qui pourraient nous aider dans la création de commandes vocales optimales. Nous avons étudié comment six instructeurs de vol formulent des commandes vocales se rapportant à leur travail au poste d'opération du simulateur de vol. Enfin, un modèle de langage de commandes vocales est décrit et intègre les résultats de recherches sur les commandes textuelles et les résultats de l'étude menée auprès des instructeurs.

5.1 Définitions

Afin de poursuivre notre étude sur le langage de l'instructeur de vol, plusieurs termes doivent d'être définis. Les définitions qui suivent ont été tirées de la Banque de données terminologiques et linguistiques du gouvernement du Canada¹.

Langage : ensemble de caractères, de conventions et de règles employé pour transmettre de l'information.

Langage de commande : ensemble d'opérateurs de procédure, régi par une syntaxe, servant à spécifier les fonctions à exécuter par un système d'exploitation.

Vocabulaire : répertoire unilingue, bilingue ou multilingue de termes propres à un domaine technique ou spécialisé et dans lequel des notions sont définies ou expliquées.

¹Termium Plus®. <http://www.termiumplus.gc.ca/tpv2start/start.html?lang=fr>

Grammaire : dans un langage formel ou naturel, système générateur d'expressions bien formées à partir d'un ensemble de règles et d'un ensemble de symboles d'entrée.

Syntaxe : partie de la grammaire qui décrit les règles de combinaison des unités significatives (les mots, les syntagmes) en phrases.

Sémantique : ensemble de relations entre des primitives conceptuelles et leur référent, relativement indépendantes de la syntaxe.

Pragmatique : le fait de comprendre l'influence du contexte sur le sens transmis par une phrase.

5.2 Les langages de commandes

Dans les années 80, plusieurs chercheurs ont mené des études visant à évaluer l'impact des noms de commandes sur la facilité d'apprentissage et la facilité d'utilisation des systèmes informatiques pour les utilisateurs (Black et Moran; Carroll; Landauer et al.; Napier et al.; Rosenberg; Scapin). Carroll (1982) a mené trois expérimentations qui ont permis de montrer que, dans le contexte d'un scénario de manipulation d'un robot, des commandes congruentes les unes avec les autres favorisent l'apprentissage et l'utilisation. Ainsi, l'utilisation d'un même verbe pour désigner une action (c.-à-d. "MOVE FORWARD, MOVE BACKWARD") ou de mots d'une même famille (c.-à-d. "RIGHT, LEFT, PUSH, PULL") est à préconiser. De plus, à la lumière des résultats de ces expérimentations, cette congruence diminue les erreurs liées aux commandes complémentaires ou qui sont dites l'une après l'autre. Certains résultats amènent Carroll à affirmer que la hiérarchisation des commandes aide aussi la fluidité du dialogue avec le système (p.ex., "MOVE ROBOT FORWARD, MOVE ROBOT BACKWARD" plutôt que "ADVANCE, GO BACK").

Carroll a aussi cherché à savoir si la construction de commandes par l'uti-

lisateur plutôt que par le concepteur du système améliorerait les résultats. Il appert que les utilisateurs ont tendance à créer des commandes congruentes, mais créent des hiérarchies incohérentes. Ce sont plutôt les concepteurs de commandes qui réussissent à créer des hiérarchies cohérentes.

Landauer et al. (1983) ont demandé à des dactylographes de construire des commandes pour modifier un texte afin de le faire passer de l'état A vers l'état B. L'objectif de leur expérimentation était de déterminer si les mots employés par les utilisateurs pour décrire les tâches non informatisées aideraient à l'apprentissage des commandes informatisées. Leur hypothèse se basait sur le fait qu'un utilisateur naïf serait tenté de commander des tâches à un ordinateur de la même façon qu'il le ferait avec un collègue. Landauer et al. (1983) ont donc cherché à déterminer le vocabulaire que les dactylographes utilisaient pour décrire leur travail et à savoir si ce vocabulaire permettait de commander le système de traitement de texte. Les résultats montrent que les mots intuitifs ou familiers recueillis auprès d'utilisateurs potentiels ou d'experts dans un domaine particulier ne permettent pas de créer un vocabulaire de commandes adéquat puisque ces mots sont imprécis et diffèrent d'un contexte à l'autre ou d'un utilisateur à l'autre de sorte qu'ils n'améliorent pas la facilité d'apprentissage. Landauer et al. (1983) soulignent l'importance du caractère familier des mots recueillis auprès des utilisateurs, mais ils soutiennent que d'autres propriétés comme un couplage fort entre une commande et la fonctionnalité qu'elle contrôle sont nécessaires. Cependant, plusieurs mots familiers différents peuvent être adéquats pour une même fonctionnalité. Le choix de mots familiers avec une faible fréquence d'utilisation s'avérerait être un bon compromis. Ainsi, les concepteurs de langages de commandes seraient en meilleure position que les utilisateurs pour choisir les mots adéquats.

Napier et al. (1989) ont montré que l'utilisation d'un langage naturel, comprenant des mots ou des phrases familières pour l'interaction avec le logiciel de traitement de texte Lotus, améliorerait la performance des utilisateurs et était préférée de ceux-ci.

À la lumière de ces résultats, afin de faciliter l'apprentissage et l'utili-

sation du langage de commande et d'impliquer les utilisateurs dans le processus de modélisation Yankelovich et Lai (1999), il est pertinent de faire appel aux instructeurs de vol pour connaître la terminologie, la syntaxe et la sémantique du langage de commandes approprié. Cependant, les commandes naturelles que les instructeurs peuvent apporter ne devraient pas permettre d'établir à elles seules le vocabulaire qu'un système de reconnaissance devrait reconnaître. La congruence, la hiérarchisation cohérente ainsi qu'un couplage fort entre les commandes et leur fonctionnalité devraient être pris en compte lors de la modélisation finale du langage.

5.3 Modélisation du langage de commandes du poste d'opération

À partir de la littérature sur les langages de commandes, nous avons dégagé divers facteurs à considérer en vue de modéliser le langage de commandes des instructeurs au poste d'opération. Puisque les mots et la syntaxe du langage naturel des utilisateurs peuvent aider à la conception du système de reconnaissance, l'expérimentation suivante cherche à recueillir le vocabulaire et la syntaxe utilisés par les instructeurs de vol pour des tâches de paramétrisation du poste d'opération. Ils serviront de base à l'élaboration de la grammaire sémantique qui sera reconnue par le système de reconnaissance vocale.

5.3.1 Méthodologie

Afin d'obtenir le vocabulaire relatif aux tâches de paramétrisation du simulateur ainsi que la syntaxe des commandes, trois exercices sont présentés séquentiellement aux sujets. Ces derniers doivent déterminer la commande vocale qu'ils utiliseraient dans le but de changer la valeur ou l'état de certains paramètres d'une simulation de vol. Une rétroaction visuelle est de plus

offerte aux sujets afin de voir leurs réactions aux indices visuels de l'interface graphique suite à la paramétrisation vocale.

5.3.1.1 Sujets

Les six sujets retenus pour cette étude sont des instructeurs de vol ou d'anciens instructeurs de vol en simulateur. Ils ont plusieurs années d'expérience comme instructeur et connaissent l'entraînement en simulateur ainsi que le poste d'opération servant à contrôler la simulation. Tous parlent l'anglais comme langue maternelle. Deux sont âgés entre 40 et 49 ans, deux entre 50 et 59 ans et deux ont plus de 60 ans. Le nombre d'années d'expérience comme instructeur de vol varie de 4 à 28 ans. Un seul sujet n'a pas d'expérience comme instructeur, mais est un ancien pilote devenu spécialiste de tests en simulateur et du poste d'opération. Les données recueillies auprès de ce sujet ne devraient pas introduire de biais puisqu'il est un spécialiste en simulateur et qu'il connaît le rôle joué par l'instructeur à son poste d'opération. Trois sujets ont évolué dans le domaine militaire, un sujet connaît uniquement le domaine de l'aviation commerciale et deux ont de l'expérience dans les deux domaines.

5.3.1.2 Outils utilisés

La technique du "Magicien d'Oz" est très utilisée dans la conception des interfaces vocales puisqu'elle permet de simuler de telles interfaces sans implémenter de système (Fraser et Gilbert; Gustafson; Klemmer et al.; Lathrop et al.; McTear; Salber et Coutaz; Turunen). L'approche consiste à demander à un utilisateur d'interagir avec un prototype non fonctionnel ou semi-fonctionnel du système et de lui fournir une rétroaction provenant d'une personne qui joue le rôle du système et répond via le prototype. L'utilisateur voit ainsi la réponse de l'interface du poste d'opération lorsqu'il dicte des commandes vocales. Nous recueillons aussi des données sur l'appréciation du sujet face aux indices visuels que lui offre l'interface graphique. Dans cette

expérimentation, aucun système de reconnaissance vocale ne sera intégré au poste d'opération. Le sujet interagit vocalement avec le prototype, selon les tâches qui lui sont présentées sous forme de questionnaire papier, alors qu'un aperçu visuel des modifications apportées par ses interactions est simulé à l'écran.

Pour établir et évaluer le langage de l'instructeur de vol, il est judicieux d'examiner le langage de plusieurs instructeurs et le jugement que chacun d'eux porte sur le langage utilisé par leurs collègues. Les différences d'expérience et d'expertise entre les instructeurs peuvent mener à des divergences quant au choix de commandes à utiliser. À cet égard, la méthode Delphi (Linstone et Turoff) est utilisée puisqu'elle permet à des sujets ayant répondu une première fois de réévaluer leurs réponses en examinant les réponses des autres sujets et de converger ainsi vers un consensus.

Enfin, les commandes verbales sont enregistrées sur bande audio pour permettre une analyse future du vocabulaire et de la syntaxe utilisés, de même que des commentaires échangés.

5.3.1.3 Tâche expérimentale

Le scénario présenté à l'annexe A, auquel nous nous référons aussi dans la section 4.2.2.3 du chapitre 4 de ce mémoire, est utilisé pour définir la tâche expérimentale qui sera soumise aux sujets. Les tâches et sous-tâches de l'instructeur de vol proviennent des observations faites sur le terrain par une ergonome pendant une séance d'entraînement. La liste des énoncés qui ont été soumis aux sujets est disponible à l'annexe D. Les paramètres et les valeurs y ont été identifiés. Cette liste est utilisée pour les trois exercices suivants :

Exercice 1 - Dans la première partie de l'expérience, nous demandons au sujet de lire un énoncé qui décrit un besoin, de composer une commande pour satisfaire ce besoin et de l'évoquer à haute voix. Il lui est ensuite pos-

sible de voir à l'écran le résultat de la commande. Une liste des énoncés soumis aux sujets sous format papier est présentée à l'annexe D. Afin d'éviter de proposer des verbes ou des syntaxes spécifiques au sujet et de le biaiser dans la formulation de commandes vocales, nous lui avons présenté des énoncés qui ont toujours la forme "Vous avez besoin de..." suivis du nom du paramètre et de la valeur à lui associer (voir l'encadré ci-dessous). Cette stratégie permet de recueillir les commandes proposées par le sujet après compréhension du but de la tâche à accomplir. Cette méthode de recueil de données permet non seulement d'obtenir des commandes naturelles, mais aussi d'observer la formulation utilisée ainsi que le verbe qui commandent l'action.

You need aircraft mass to be 2 540.

Exercice 2 - La seconde partie de l'expérience est réalisée immédiatement après la première partie. Elle reprend les mêmes énoncés présentés à l'annexe D ; cette fois, cinq formulations de commandes sont fournies au sujet qui est appelé à choisir les deux formulations qu'il préfère dans l'ordre, en indiquant **1** dans la case pour la formulation préférée et **2** pour la seconde (voir le tableau 5.1) L'une des formulations est celle que le sujet a utilisée dans la première partie ce qui lui permet de la comparer avec d'autres formulations possibles et de remettre en question son choix. Les quatre autres formulations sont construites à l'aide des éléments de variation du tableau 5.2. Ces éléments de variation sont appliqués un ou deux à la fois pour la construction de commandes équivalentes. Nous cherchons ainsi à connaître l'importance qu'ont ces éléments pour les sujets. Ces derniers sont invités à commenter les raisons qui motivent leur choix d'une formulation plutôt qu'une autre, ou encore à proposer une formulation différente en conjugant certaines parties de celles présentées.

Afin de déterminer la formulation de commande préférée par les sujets, deux points sont attribués à celle choisie en premier, et un point à la formulation qui est choisie en deuxième. Pour chacun des 30 énoncés, la formulation

Tableau 5.1 – Exemple d'énoncé avec différentes formulations de commandes où le sujet doit indiquer **1** pour la commande préférée et **2** pour la seconde (deuxième exercice)

Énoncé	Formulations de commandes
You need aircraft mass to be 2 540.	☐ Set all up mass to two thousand five hundred forty. ☐ Set mass to two five four zero. ☐ Mass at two five four zero kilograms. ☐ Put mass at two thousand five hundred forty. ☐ Aircraft mass two five four zero.

Tableau 5.2 – Éléments de variation pour la construction de formulations de commandes

- | |
|--|
| 1) Ordre paramètre-valeur inversé
2) Mots-clés d'action différents.
3) Ajout des unités à la valeur.
4) Syntaxe différente des valeurs chiffrées.
5) Terme différent pour le même paramètre. |
|--|

ayant le plus de points est considérée pour la troisième partie comme étant le choix préféré des sujets.

Exercice 3 - Dans la troisième partie de l'expérience les mêmes sujets sont appelés à choisir entre deux formulations de commandes : celle qu'ils ont préférée dans l'exercice 2 et celle préférée par l'ensemble des sujets (exemple au tableau 5.3). Dans les cas où la formulation préférée par le sujet est la même que celle préférée des autres sujets, le second choix de formulation peut alors provenir d'un autre sujet ou être construit à partir des éléments de variation du tableau 5.2. Ce troisième exercice permet de faire choisir par le sujet la formulation de commandes qu'il préfère entre sa formulation préférée et celle des autres.

Tableau 5.3 – Exemple de choix entre la formulation de commande préférée par le sujet et celle qui est préférée par l'ensemble des sujets

Énoncé	Formulations de commandes
You need aircraft mass to be 2 540.	[] Set mass to two five four zero. [] Aircraft mass two five four zero.

5.3.1.4 Procédures et consignes

Un sujet à la fois participe à l'expérience. Le sujet lit sur papier l'introduction de l'expérience à laquelle il prend part. On l'informe que le but est de trouver la meilleure formulation possible pour les commandes vocales au poste d'opération et qu'il aura à suivre les indications écrites pour les trois exercices qu'on lui demande de compléter. On lui suggère d'utiliser un ton de voix naturel et un langage familier, tout en gardant en tête qu'il interagit avec le poste d'opération. On lui dit de ne pas se laisser influencer par la formulation des énoncés, c'est-à-dire par le vocabulaire et la syntaxe utilisés. Il a la possibilité de changer le nom du paramètre ou le type de valeur si cela lui convient. On lui conseille d'utiliser des commandes brèves et efficaces. Le sujet est encouragé à demander des précisions s'il a besoin

d'éclaircissements, et à émettre des commentaires s'il a des suggestions sur la présentation visuelle. Une rétroaction visuelle est présentée sur un écran d'ordinateur après chaque commande vocale.

Le sujet prend ensuite connaissance du premier exercice qui lui est proposé sur papier. L'expérimentation démarre lorsque le sujet est prêt à commencer le premier exercice et elle est enregistrée sur bande audio. Pour chaque formulation de commande, une rétroaction visuelle informe le sujet du changement provoqué par la commande. Cette rétroaction est un clip vidéo de l'interface du poste d'opération dont les changements de valeurs ont été préenregistrés.

L'expérimentateur note au fur et à mesure les formulations de commandes proposées par le sujet et prépare le second exercice. Une fois ce premier exercice terminé, il remet au sujet le second exercice imprimé. Pour chacun des 30 énoncés, le sujet est appelé à choisir la formulation de commande qu'il préfère parmi cinq choix possibles.

À la fin, le sujet remplit le questionnaire sur ses données biographiques et sur sa perception de l'utilisation de commandes vocales pour contrôler le poste d'opération.

Le troisième exercice de l'expérimentation a lieu une semaine plus tard quand tous les sujets ont terminé les exercices 1 et 2. Ici, pour chacun des 30 énoncés, le sujet note sa commande préférée parmi deux formulations : sa formulation préférée et celle qui est préférée par l'ensemble des sujets.

5.3.2 Résultats

Les résultats sont présentés sous la forme de constats. Pour chacun de ces constats, la provenance des résultats, l'analyse effectuée et l'impact en vue de l'élaboration d'un modèle de langage sont décrits.

5.3.2.1 Constats

1. Utilisation des mots d'action

Lors du premier exercice, 73% des 180 commandes construites par les sujets commencent par un verbe. Pourtant, la formulation des 30 énoncés ne comporte aucune action et spécifie plutôt le résultat attendu de la commande. Le verbe se retrouve toujours au début de la commande. Dans les cas où il n'y a pas de verbe, le sujet a, soit énoncé le nom du paramètre suivi immédiatement de la valeur, soit l'inverse. Par exemple, pour l'énoncé "8. *You need water on the runway*", moins de trois sujets ont utilisé un verbe dans leurs formulations. On retrouve entre autres "Wet runway" et "Water on runway" comme commandes vocales pour cet énoncé. Cela peut s'expliquer par la nécessité pour les sujets de créer une commande simple et naturelle évoquant uniquement l'état du paramètre, ou encore par le fait que ces paramètres sont familiers dans le discours opératif des instructeurs ou des pilotes. Le mot "Set" est utilisé pour 50% des commandes qui comportent un verbe. Parmi les autres verbes, il y a les mots "Select", "Reposition", "Clear", "Display" ou "Reset". Dans certains cas, par exemple pour les verbes "Reposition" et "Reset", le verbe est lié à l'effet de la fonctionnalité sur l'état du simulateur et est implicite dans l'énoncé décrivant l'objectif à atteindre. Le verbe fait donc partie de plusieurs formulations et devra être présent dans les règles de construction des commandes vocales retenues pour le modèle.

2. Ordre paramètre-valeur

Les données recueillies lors du premier exercice montrent que le nom du paramètre précède généralement la valeur qui lui est donnée. Ainsi, 89% des commandes recueillies correspondent à ce format syntaxique. Par exemple, pour mettre la quantité de carburant dans l'appareil à une certaine valeur, les sujets ont tous formulé leur commande sous la forme "Set fuel fifteen thousand", où le carburant précède la quantité. Il faut cependant prendre en considération le fait que la formulation des

énoncés (voir l'annexe D) présentait aussi cet ordre paramètre-valeur parce que cela nous semblait la façon la plus naturelle de présenter ces énoncés, ce qui a pu influencer les sujets. Pour 20 énoncés sur 30, on retrouve cet ordre paramètre-valeur. Il semble selon Carroll (1982) que cet ordre soit plus naturel. Notre modèle de langage devra tenir compte de cet ordre paramètre-valeur, qui est naturel pour les sujets.

3. Influence de la formulation des énoncés

Nous avons utilisé une stratégie visant à éviter le plus possible d'influencer les sujets dans leurs formulations. Tous les énoncés définissaient un besoin à satisfaire et étaient sous la forme "Vous avez besoin de...". Cependant, pour plusieurs énoncés (p.ex., "You need oil temperature to be at its default value."), la formulation de commande proposée par le sujet reprend le vocabulaire et la syntaxe pour désigner et placer le paramètre et la valeur : "Set oil temperature to default value". Cette influence est cependant compensée dans le second exercice où les sujets tendent à préférer une formulation plus directe (p.ex., "Reset oil temperature"). Le modèle de langage qui sera construit à partir des formulations des sujets ne devrait pas prendre en compte l'influence qu'a pu avoir la formulation des énoncés lorsqu'il est possible d'identifier un tel effet.

4. Divergence entre les sujets

L'analyse des données issues du premier exercice permet de voir que, malgré leurs antécédents similaires, les sujets ont un vocabulaire et des syntaxes de commandes qui leur sont propres. Les résultats montrent que pour seulement 7 énoncés sur 30, deux sujets ont exactement la même syntaxe. Plusieurs commandes diffèrent par des mots de liaisons comme "to" ou des verbes tels que "set" ou "select". L'inversion entre paramètre et valeur est souvent la cause d'une différence syntaxique entre les sujets. L'analyse montre aussi que les sujets n'ont pas tous la même façon de dire les valeurs numériques à haute voix : parfois ils décomposent chaque chiffre (p.ex., deux cinq zéro zéro), parfois en utilisant les centaines et les milliers (p.ex., deux mille cinq cent). Par

contre, chaque sujet est constant dans sa façon de dire les chiffres. Il est ainsi important de prendre en compte dans le modèle de langage qui sera élaboré les formulations de commandes qui tendent à se répéter ou à se ressembler d'un sujet à l'autre. Une structure rigoureuse et facile à retenir devra aussi être retenue afin d'aider les utilisateurs du système de reconnaissance à assimiler ce mode d'interaction.

5. Changements lorsque différents choix sont proposés

Les données recueillies dans le deuxième exercice de l'expérimentation révèlent qu'en moyenne 11 commandes formulées lors du premier exercice sur les 30 énoncés (c.-à-d. 37%) demeurent les commandes préférées du sujet après qu'il ait examiné les quatre autres formulations proposées. La figure 5.1 montre qu'à l'extrême le sujet 5 a préféré la commande qu'il a utilisée en premier dans 73% des cas, alors que le sujet 3 a conservé sa formulation initiale dans seulement 10% des cas. Les autres sujets ont en général préféré une nouvelle commande ou ont relégué la commande qu'ils ont utilisée dans le premier exercice au second rang. Cela démontre un désir de cohérence entre chaque commande et une orientation vers des commandes brèves.

6. Peu de changements au troisième exercice

Les résultats indiquent qu'il y a eu peu de changements entre le deuxième exercice et le troisième exercice pour ce qui est des commandes préférées par les sujets. Il y a eu des changements de préférence uniquement pour les énoncés 10, 13, 15 et 19, du tableau D.1 de l'annexe D. La stabilité des autres formulations s'explique par le fait que la commande préférée par l'ensemble des sujets lors du deuxième exercice se retrouve toujours parmi les deux choix présentés aux sujets dans le troisième exercice. En moyenne, les sujets ont choisi leur commande préférée lors du deuxième exercice pour la moitié des énoncés et ont choisi la commande préférée par l'ensemble du groupe pour l'autre moitié. Ces résultats ne permettent pas de prendre une décision définitive concernant la structure du modèle de langage à construire

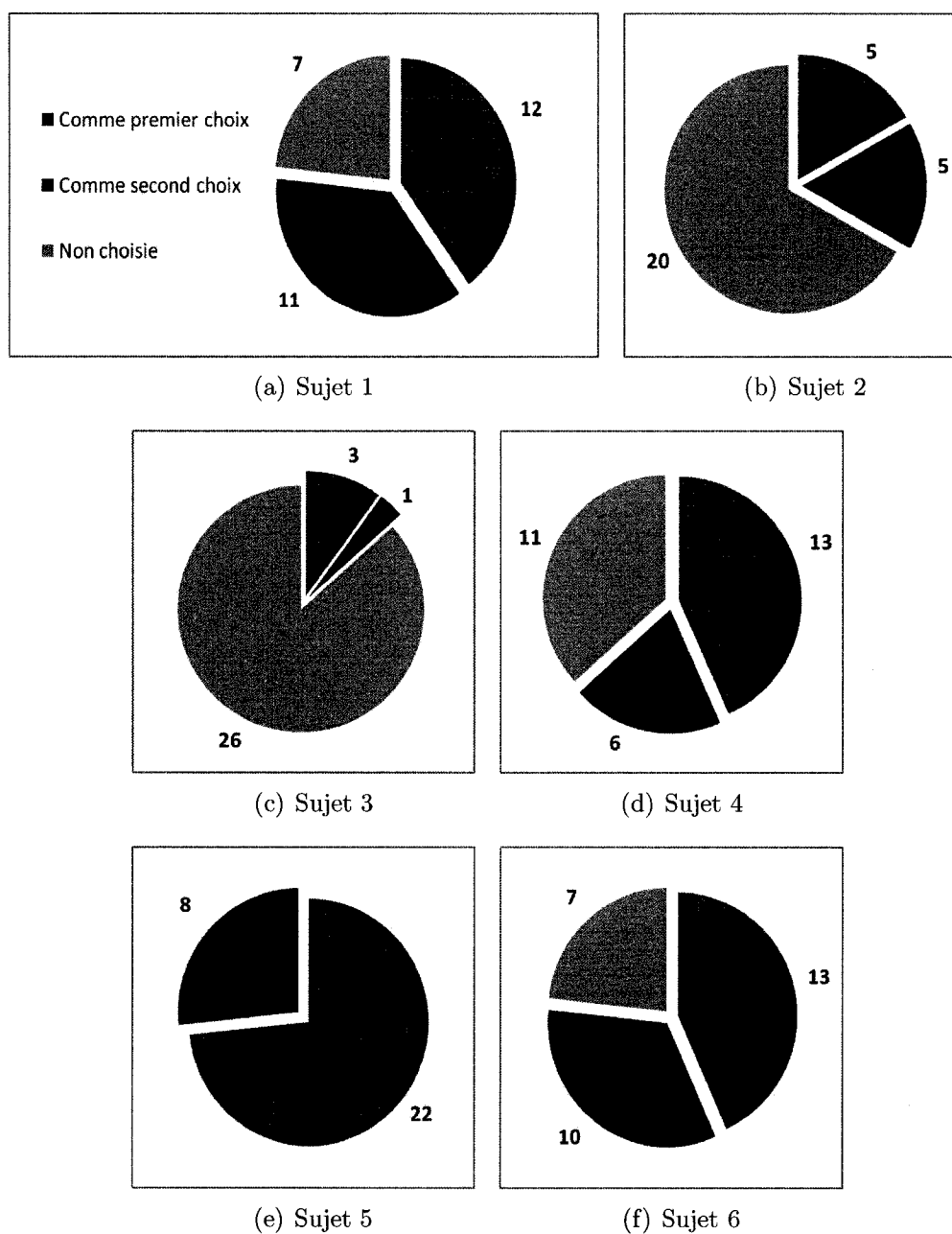


Figure 5.1 – Résultats du deuxième exercice montrant, pour chacun des sujets, le nombre de fois (sur un total de 30 énoncés) qu'il a sélectionné la commande utilisée lors du premier exercice comme premier choix, comme second choix, ou qu'il ne l'a pas sélectionnée.

à partir des formulations des sujets. Autant les préférences individuelles que les préférences de l'ensemble des sujets devront être prises en compte.

7. Cas d'aucune formulation commune

Pour deux énoncés (9 et 18), il n'est pas possible de faire ressortir une formulation plus qu'une autre après le troisième exercice, car six formulations différentes ont été proposées par les sujets. Comme le montre le tableau 5.4, ces formulations diffèrent de quelques mots seulement. La cohérence avec les autres commandes qui feront partie du modèle de langage devrait aider à déterminer la formulation la plus appropriée.

8. Classification des commandes pour la voix

L'analyse des données du questionnaire permet d'identifier les tâches que les instructeurs voient comme se prêtant bien à une commande vocale et celles qui ne leur semblent pas propices à une commande vocale. Les tâches entourant la météo ont été mentionnés par quatre sujets comme propices aux commandes vocales. Le repositionnement de l'appareil, la réinitialisation des systèmes, ainsi que la sélection de pages du poste d'opération sont aussi des éléments de tâches qui semblent propices aux commandes vocales selon deux sujets. Les dysfonctionnements ou la simulation de situations d'urgence sont quant à elles des fonctionnalités qui ne sont pas propices aux commandes vocales. L'effet de surprise ne serait plus présent puisque les pilotes pourraient entendre l'instructeur activer ces fonctionnalités vocalement. De plus, le nombre de dysfonctionnements et leur spécificité à chaque simulateur ou à chaque client viennent diminuer l'efficacité d'apprentissage des commandes disponibles.

Tableau 5.4: Résultats où deux énoncés ont donné lieu à six formulations différentes par les sujets

Énoncés	Formulations de commandes
9. Recirculation hasn't been used yet, you need 80% intensity of sand recirculation.	Sujet 1 : <i>recirculation on eighty percent sand</i> Sujet 2 : <i>eighty percent sand recirculation</i> Sujet 3 : <i>set recirculation sand intensity at eighty percent</i> Sujet 4 : <i>eighty percent sand recirculation on</i> Sujet 5 : <i>set recirculation to sand eighty percent</i> Sujet 6 : <i>activate sand recirculation at eight zero</i>
18. You need scattered clouds; visibility needs to be 0.5; cloud top needs to be 10 000 and cloud base 200.	Sujet 1 : <i>cloud level one scattered, top ten thousand, base two hundred, visibility five hundred meters</i> Sujet 2 : <i>cloud cover scattered, visibility zero point five, cloud top ten thousand, cloud base two hundred</i> Sujet 3 : <i>set two hundred scattered, tops ten thousand, visibility half a mile</i> Sujet 4 : <i>set scattered clouds, set visibility at zero point five, set cloud top ten thousand feet, set cloud base at two hundred feet</i> Sujet 5 : <i>set clouds scattered, set visibility zero point five nautical miles, set cloud top one zero thousand, set cloud base two zero zero</i> Sujet 6 : <i>select scattered clouds, set visibility at zero point five, set cloud top at ten thousand feet, set cloud base at two hundred feet</i>

9. Nécessité d'une rétroaction

La rétroaction visuelle présentée aux sujets suite à chaque commande vocale n'a pas semblé adéquate selon les commentaires recueillis par le questionnaire. Les sujets ont porté peu d'attention à la rétroaction, mais ils reconnaissent sa nécessité et quatre sujets proposent plutôt de faire apparaître à l'écran une liste des commandes reconnues par le système de reconnaissance. De plus, deux sujets s'accordent pour affirmer qu'une rétroaction ou une confirmation auditive serait intéressante et leur permettrait d'entendre l'effet de leur commande sur le système. Puisqu'ils portent déjà un casque d'écoute et qu'ils ont une certaine habitude des confirmations radio, cette solution leur semble adaptée au contexte. Enfin, les rétroactions devraient être désactivables pour permettre à l'instructeur qui a pris confiance en ce système de ne plus recevoir de confirmation.

5.3.3 Description du modèle de langage de commandes vocales

À la lumière des résultats décrits dans la littérature concernant les langages de commandes ainsi que de ceux recueillis au cours de cette expérimentation, nous sommes en mesure de modéliser les commandes qui serviront à la construction des grammaires de reconnaissance vocale. Il est important de rappeler que les variables dépendantes prises en considération sont : la facilité d'apprentissage et d'utilisation, et la perception subjective de l'aspect naturel des commandes. L'étude des langages de commandes textuelles nous amène à intégrer une congruence entre les commandes, une hiérarchisation cohérente et un couplage fort avec les fonctionnalités. Ainsi, le modèle retenu est schématisé à la figure 5.2.

Dans le modèle proposé, des verbes ont été intégrés et sont placés au début de la majorité des commandes. Il a semblé naturel pour les sujets de commencer leurs commandes avec des verbes (voir constat 1). Les commandes portant sur la météo et les communications radio font cependant

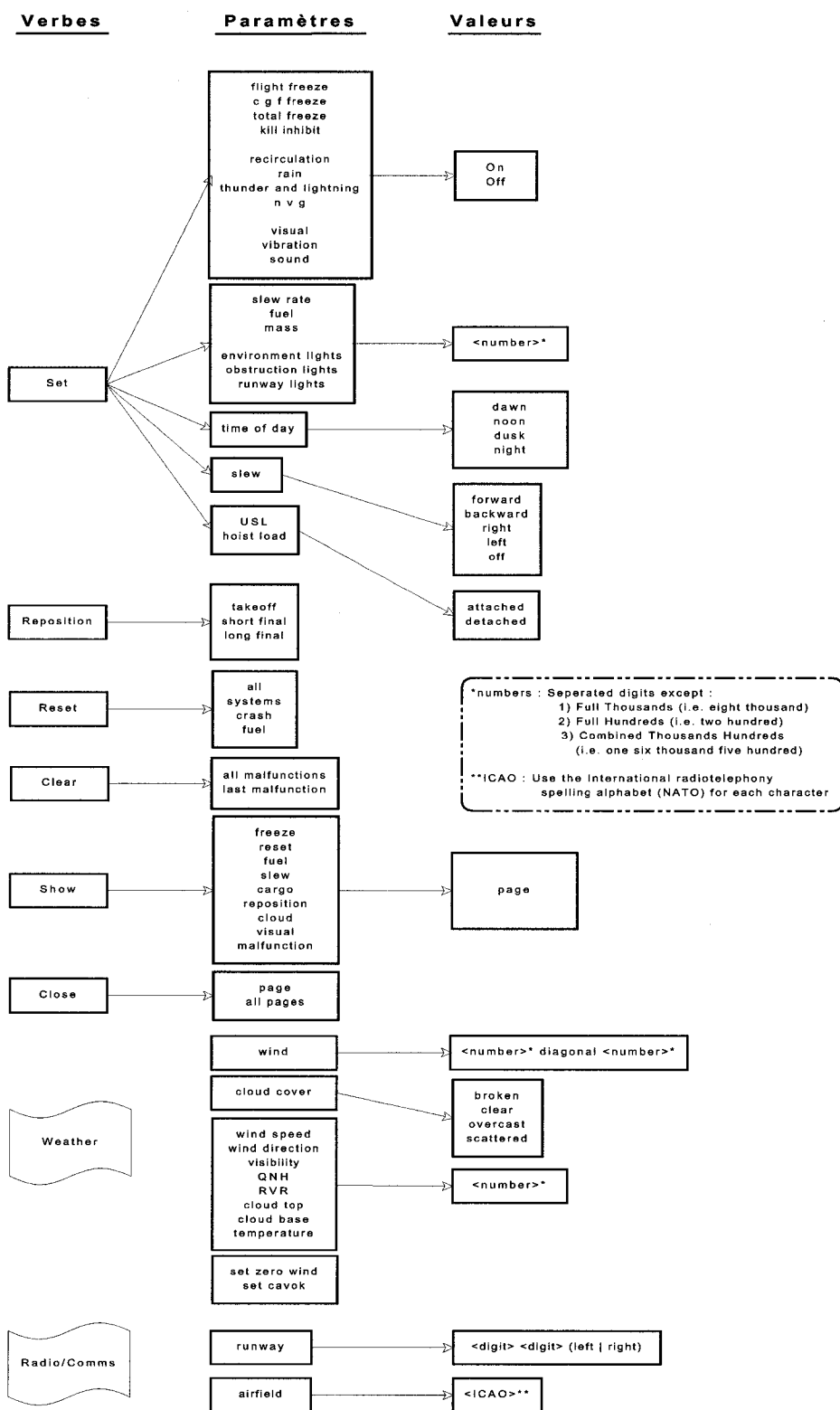


Figure 5.2 – Modèle de langage pour 78 commandes vocales.

exception. En effet, la formulation des commandes pour ces fonctionnalités provient du domaine du contrôle aérien avec lequel les instructeurs sont très familiers. De plus, la façon de formuler les nombres et les codes OACI² provient aussi de ce domaine et intègre des règles strictes décrites dans (Roos). Ainsi, tel que montré à la figure 5.2, les nombres doivent être décomposés en une suite de chiffres (c.-à-d. zéro à neuf), sauf pour les centaines et les milliers, alors que les codes OACI doivent être décomposés en une suite de lettres nommées à l'aide de l'alphabet international radiotéléphonique (c.-à-d. alpha, bravo, etc.). Il est important de conserver la formulation et le langage opératif déjà assimilés par ces professionnels afin de conserver l'aspect naturel et familier des commandes.

La structure des commandes est définie selon les facteurs identifiés dans la section 5.2, soit la congruence, la hiérarchisation cohérente et le couplage avec les fonctionnalités. La congruence est implémentée par la réutilisation de mêmes termes pour les paramètres de même type. Par exemple, les termes On/Off servent à modifier l'état actif/inactif de tous les paramètres qui reçoivent ces valeurs. Ensuite, la hiérarchisation cohérente est représentée par la règle selon laquelle le verbe précède le nom du paramètre qui précède la valeur à lui associer (voir constat 2). À noter qu'aucun mot de liaison comme "to" ou "at" n'a été introduit dans le modèle afin d'éviter les confusions (p.ex., "Set fuel two five thousand" est préférable à "Set fuel to five thousand"). De ce fait, les instructeurs devront éviter ces mots de liaison. Cependant, les instructeurs de vol sont déjà au fait de cette règle puisque les communications radio avec les contrôleurs aériens en tiennent compte. Enfin, les termes utilisés pour nommer les paramètres sont ceux identifiés par les sujets dans la section 5.3.2 et représentés dans l'interface graphique du poste d'opération.

Le nombre de commandes retenues dans le modèle est minimal. Les 78 commandes proviennent principalement de l'ensemble des fonctionnalités intégrées à un ordinateur de poche servant à contrôler avec quelques com-

²Les codes de l'Organisation de l'aviation civile internationale servent entre autre à identifier les aéroports à l'aide de quatre lettres (par exemple, l'aéroport Montréal-Trudeau, CYUL)

mandes principales le poste d'opération. Les commandes proviennent aussi de celles identifiées par les sujets comme étant intéressantes pour l'interaction vocale (voir constat 8). D'autres commandes pourraient être intégrées pourvu qu'elles respectent les règles décrites ci-dessus.

5.4 Sommaire

Le modèle de langage que nous avons élaboré en vue de l'intégration d'un système de reconnaissance vocale au poste d'opération provient des résultats de l'expérience menée et des résultats de recherches sur les commandes textuelles. Le parallèle peut être fait entre les commandes textuelles et les commandes vocales parce qu'elles font appel à la mémoire de l'utilisateur et qu'elles n'offrent pas d'indices visuels pour aider à la formulation exacte. Une revue de la littérature traitant des langages de commandes textuelles a permis de cibler des facteurs importants dans la modélisation de ces langages qui ont un impact positif sur la facilité d'apprentissage et d'utilisation, ainsi que la perception subjective de l'aspect naturel de l'interaction. La congruence, la hiérarchisation cohérente ainsi qu'un couplage fort entre les commandes et leurs fonctionnalités doivent être présents dans le langage de commandes. De plus, l'expérimentation effectuée avec des instructeurs a permis de connaître le vocabulaire et la syntaxe d'une trentaine de commandes. À partir des constats sur les formulations de commandes par les instructeurs et des facteurs importants dans la conception des langages de commandes textuelles, nous avons élaboré un modèle de langage de commandes vocales pour l'instructeur de vol pour des commandes ciblées. Ce modèle servira à l'élaboration d'une grammaire de reconnaissance intégrée au prototype décrit au chapitre 7. Le chapitre suivant analysera des engins de reconnaissance, l'impact du bruit ambiant dans l'environnement du simulateur et des dispositifs d'activation/désactivation.

CHAPITRE 6

ANALYSE DES TECHNOLOGIES DE RECONNAISSANCE VOCALE, DE L'IMPACT DU BRUIT DU SIMULATEUR ET DE DISPOSITIFS D'ACTIVATION/DÉSACTIVATION

Les caractéristiques des utilisateurs, des tâches, du milieu de travail et du langage de commandes doivent être prises en compte lorsqu'il s'agit d'intégrer une technologie de reconnaissance vocale au poste d'opération de l'instructeur de vol. Ce chapitre présente une analyse de différentes technologies à considérer. Cinq engins de reconnaissance sont d'abord comparés, et le meilleur engin est évalué en milieu bruyant. Cinq dispositifs d'activation/désactivation sont comparés puisqu'il faut un tel dispositif dans un système de reconnaissance vocale.

6.1 Analyse comparative de cinq engins de reconnaissance vocale

Plusieurs engins de reconnaissance vocale sont présentement disponibles sur le marché grâce à l'évolution constante des connaissances dans le domaine de la reconnaissance vocale. Ces engins diffèrent sur le plan de l'efficacité selon le contexte dans lequel ils sont utilisés, les options et les services offerts pour aider à leur intégration, le coût, etc. En vue de choisir un ou plusieurs engins de reconnaissance ayant les propriétés nécessaires pour le contrôle du poste d'opération de l'instructeur de vol, nous avons analysé et comparé cinq engins. Il faut un taux de reconnaissance minimal si l'on veut augmenter l'efficacité de l'instructeur. Nous avons analysé et comparé l'exactitude et l'efficacité de ces engins. Enfin, nous présentons un test de reconnaissance vocale où des commandes vocales préenregistrées sont soumises aux

différents systèmes de reconnaissance vocale que nous voulons comparer.

6.1.1 Méthodologie

Cette section décrit la méthodologie utilisée afin de comparer les engins de reconnaissance vocale. Les caractéristiques des engins utilisés sont d'abord présentées. Par la suite, des explications sur les données quantitatives retenues pour servir de base de comparaison entre les engins sont fournies. Enfin, la procédure servant à tester les engins est décrite.

6.1.1.1 Engins testés

Une fiche des caractéristiques des différents engins analysés est disponible à l'annexe E. L'accessibilité et la facilité d'intégration sont considérées lors du choix des engins à tester. Or, uniquement les engins du tableau 6.1 sont retenus et testés puisqu'ils sont accessibles gratuitement et sont facilement configurables en vue d'effectuer les tests décrits plus loin. Le lecteur peut se référer au tableau E.1 pour de plus amples détails sur les engins de reconnaissance.

Tableau 6.1: Engins de reconnaissance vocale retenus

Engin	Nom
A	Microsoft English Recognizer v5.1
B	Microsoft English (U.S.) v6.1 Recognizer
C	Microsoft English (U.S. Telephony) v7.0 Server
D	Nuance Vocon 3200 EDS 2.5
E	Nuance Vocon 3200 EDS 2.6

6.1.1.2 Mesures d'exactitude et d'efficacité

Les engins de reconnaissance vocale sont comparés sur la base de l'exactitude et de l'efficacité dans (Vinciguerra).

L'exactitude représente le rapport entre le nombre de mots reconnus par le système de reconnaissance et le nombre de mots total fourni au système. Ce rapport est souvent présenté sous forme de pourcentage. C'est l'un des paramètres utilisés pour comparer les engins de reconnaissance. Plus l'exactitude est près du 100%, plus il est capable de reconnaître les mots qui lui sont fournis pour l'analyse. Il peut y avoir des biais de comparaison si les mots qui sont fournis à deux engins différents ne sont pas les mêmes, ou si les contextes dans lesquels les engins sont testés diffèrent. Il est essentiel de tenir compte de ces remarques dans la procédure qui est mise en place pour tester et comparer les engins.

Certains systèmes de reconnaissance vocale permettent de traiter chacun des mots d'une commande dite par un utilisateur (par exemple dans le cas d'une dictée de texte). Dans ce cas, les mots non reconnus ne peuvent pas faire partie de l'information transmise au système et ce dernier doit composer avec l'information reconnue. Dans le contexte du contrôle du poste d'opération d'un simulateur de vol, seule une commande complète peut être exécutée. En effet, si uniquement une partie de la commande vocale est reconnue, commande composée généralement de plus de deux mots selon les résultats obtenus à la section 5.3, alors le système n'est pas en mesure d'exécuter cette commande. Pour illustrer cette contrainte, prenons la commande suivante :

Set fuel to ten thousand kilograms

Dans cette commande, si l'engin est en mesure de reconnaître que l'utilisateur cherche à modifier le carburant (*set fuel*), mais qu'il ne reconnaît pas la quantité demandée, il ne pourra pas effectuer la commande. Idéalement, l'engin pourrait demander à l'utilisateur de répéter une partie

de la commande (dans l'exemple, la valeur), pour ensuite corriger ou terminer l'exécution de la tâche. Cependant, dans le cadre de cette étude, seules les commandes totalement reconnues sont considérées dans le calcul de l'exactitude des engins. L'exactitude est calculée comme le nombre de commandes reconnues sur le nombre total de commandes fournies à l'engin.

$$Exactitude = \frac{C_r}{C} * 100 \quad (6.1)$$

Avec

C : nombre de commandes

C_r : nombre de commandes reconnues

En ce qui a trait à l'efficacité, ce critère peut être défini de manière subjective par le gain ou la perte d'efficacité que l'utilisateur ressent lors de l'exécution de ses tâches avec l'utilisation de cet engin. Ainsi, le calcul de cette valeur considère qu'une commande reconnue avec succès par le système est bénéfique pour l'efficacité de l'utilisateur, que la non-reconnaissance d'une commande n'a pas d'impact sur son efficacité avec ce système, alors que la reconnaissance d'une commande autre que celle émise par l'utilisateur a un impact négatif sur cette efficacité. La formule suivante permet de calculer cette donnée et traduit cet énoncé :

$$Efficacité = \left[\frac{(1 * C_r) + (0 * C_{nr}) + (-1 * C_{fr})}{C} \right] * 100 \quad (6.2)$$

Avec

C : nombre de commandes

C_r : nombre de commandes reconnues

C_{nr} : nombre de commandes non reconnues (*miss*)

C_{fr} : nombre de commandes faussement reconnues (*false alarms*)

Le résultat est multiplié par 100 pour avoir un pourcentage où 100% correspond à la situation où toutes les commandes seraient reconnues. Avec cette méthode de calcul de l'efficacité, un résultat négatif démontre qu'il y a plus de fausse reconnaissance que de commandes correctement reconnues, ce qui indique un grave problème avec le dispositif. Cette méthode diffère de celle présentée par Vinciguerra (2002) qui utilise plutôt la formule suivante :

$$Efficacité = \left[\frac{(1.5 * C_r) + (1.0 * C_{nr}) + (0.5 * C_{fr})}{C} \right] * 100 \quad (6.3)$$

Or, dans le contexte de notre étude, puisque seules les commandes complètes sont considérées, une fausse commande reconnue réduit l'efficacité puisque l'instructeur doit prendre conscience de l'erreur, corriger le paramètre erroné et rappeler à nouveau la commande. Un engin reconnaissant plus de fausses commandes qu'un autre aura donc une valeur d'efficacité faible. Les commandes non reconnues sont quant à elles considérées comme neutres puisqu'elles n'induisent pas de charge supplémentaire pour l'utilisateur.

En somme, l'exactitude et l'efficacité calculées sont des mesures quantitatives de la capacité des différents engins permettant ainsi de les comparer.

6.1.1.3 Outils utilisés

Six enregistrements audio sont utilisés provenant de six sujets différents, soit cinq hommes et une femme, âgés entre 25 et 50 ans. Pour nos tests, les différences de voix dues au sexe du sujet ne nous importent pas. C'est plutôt la variation pour un même sujet d'un engin à l'autre qui est étudiée. Quatre sujets parlent le français comme langue maternelle et les autres l'anglais. Pour chaque enregistrement, 20 commandes (voir tableau 4.1) sont dites à haute voix par le sujet de façon naturelle et dans un environnement calme

de bureau. Aucun sujet ne connaît les commandes à enregistrer. Le sujet a devant lui la liste écrite des commandes à dire. Le microphone Sennheiser PC110 qui est utilisé pour enregistrer les sujets a les caractéristiques suivantes :

- Fréquence de réponse : 80 - 15000 Hz
- Mode : Omnidirectionnel
- Sensibilité : -36 dB
- Impédance : $\approx 2k\Omega$

Tous les tests sont effectués en soumettant les enregistrements audio en format *WAVE* aux systèmes de reconnaissance A à E. Les caractéristiques des enregistrements utilisés sont les suivantes :

- Format : ACM Waveform (.wav)
- Filtre : PCM (*Pulse-code modulation*)
- Fréquence : 11025 Hz
- Mode : Mono
- Compression : 16 bits

Tous les enregistrements sont ajustés à l'aide du logiciel Adobe®Audition®2.0 pour correspondre à ces caractéristiques.

Les trois premiers engins de reconnaissance (A, B et C) sont entraînés pour reconnaître la voix d'un même individu, soit l'expérimentateur. Puisque les sujets peuvent faire varier l'exactitude et l'efficacité de ces engins en entraînant plus ou moins adéquatement les engins, les tests sont effectués sans entraînement préalable pour les différents sujets. Pour les deux derniers engins, aucun entraînement n'est nécessaire. L'intégration de la reconnaissance vocale au poste d'opération de l'instructeur de vol nécessite le moins d'entraînement possible du système par l'utilisateur tel que mentionné dans la section 2.1. Ainsi, les résultats prennent en considération qu'aucun entraînement n'est fait sur ces engins.

6.1.1.4 Procédure

Pour comparer de façon rigoureuse l'exactitude et l'efficacité de différents engins de reconnaissance vocale lors de tests, plusieurs facteurs doivent être contrôlés. Tout d'abord, le contexte et l'environnement dans lesquels les tests sont effectués ne doivent pas varier d'un système à l'autre. L'entraînement du système à reconnaître la voix d'une personne particulière, le bruit ambiant et le microphone utilisé sont des aspects qui doivent demeurer inchangés pour chaque système à tester.

De plus, les caractéristiques de la voix des sujets participant aux tests (ex., voix d'homme vs de femme, force, fréquence, timbre, rapidité, tonalité, clarté, etc.) peuvent influencer sur les résultats. Des différences telles que l'âge ou la langue maternelle, l'accent, l'articulation, la taille des lèvres, le fait d'être enrhumé ou non, la perte d'une dent ont un impact sur la voix. Ces diverses caractéristiques sont responsables du fait que le taux de réussite de la reconnaissance vocale peut varier d'un individu à l'autre. Certains systèmes doivent donc être entraînés pour avoir une meilleure performance. Un aspect culturel ou sociologique vient aussi jouer un rôle entre les individus. Des personnes de nationalités différentes peuvent parler la même langue, mais l'accent peut souvent être bien différent et influencer la reconnaissance.

Ainsi, pour éviter que ces différents facteurs viennent introduire un biais dans les résultats de tests, des enregistrements audio servent de signal d'entrée commun aux engins de reconnaissance testés. Tous les enregistrements audio doivent contenir les mêmes commandes vocales, dites cependant par des sujets différents. Pour chaque enregistrement, 20 commandes, soit celles mentionnées dans le tableau 4.1, sont dites à haute voix. Un même environnement peu bruyant (c.-à-d. un environnement calme de travail individuel) est utilisé pour l'enregistrement de tous les sujets. Une analyse en environnement bruyant est effectuée plus loin à la section 6.2.

Pour tester les engins de reconnaissance, nous dirigeons une source audio préenregistrée vers l'interface programmable de l'engin qui doit l'analy-

ser. Avant d'effectuer les tests, chaque engin est configuré avec une liste de commandes. Ces commandes font partie du vocabulaire que l'engin est en mesure de reconnaître. Enfin, les engins doivent être configurés afin d'écrire les résultats de reconnaissance dans un fichier texte que nous pouvons par la suite analyser. Ce fichier doit comprendre la liste des commandes reconnues par le système. Connaissant les commandes faisant partie de l'enregistrement audio, il est ainsi possible de déterminer si l'engin a reconnu toutes les commandes et s'il a faussement reconnu certaines commandes. L'analyse se fait par la suite en calculant l'exactitude et l'efficacité de chaque engin. Ces résultats nous donnent ainsi des mesures servant à les comparer.

6.1.2 Résultats

La figure 6.1 montre l'exactitude et l'efficacité des engins A à E configurés pour reconnaître 20 commandes enregistrées par les sujets.

À la lumière de ces résultats, les deux meilleurs engins D et E ont été sélectionnés et configurés pour reconnaître 137 commandes. La figure 6.2 montrent l'exactitude et l'efficacité de chacun des engins : l'exactitude de l'engin D passe de 91.7% à 78.3%, celle de E passe de 97.5% à 77.5%, l'efficacité de D passe de 87.5% à 58.3% et celle de E passe de 95.8% à 55%.

6.1.3 Discussion

Les enregistrements audio ont permis de créer un banc d'essai consistant et fiable pour tester les engins. L'analyse des données sur l'exactitude et l'efficacité révèle que le meilleur engin de reconnaissance vocale est l'engin E. Cet engin peut atteindre un niveau d'exactitude de 97.5% pour un ensemble de 20 commandes et de 77.5% pour un ensemble de 137 commandes. Il est important de noter que chacun des enregistrements audio n'a pas obtenu le même taux de reconnaissance que les autres. Cela peut être dû à l'accent ou au timbre de voix du sujet. En plus de nous amener à proposer un engin en

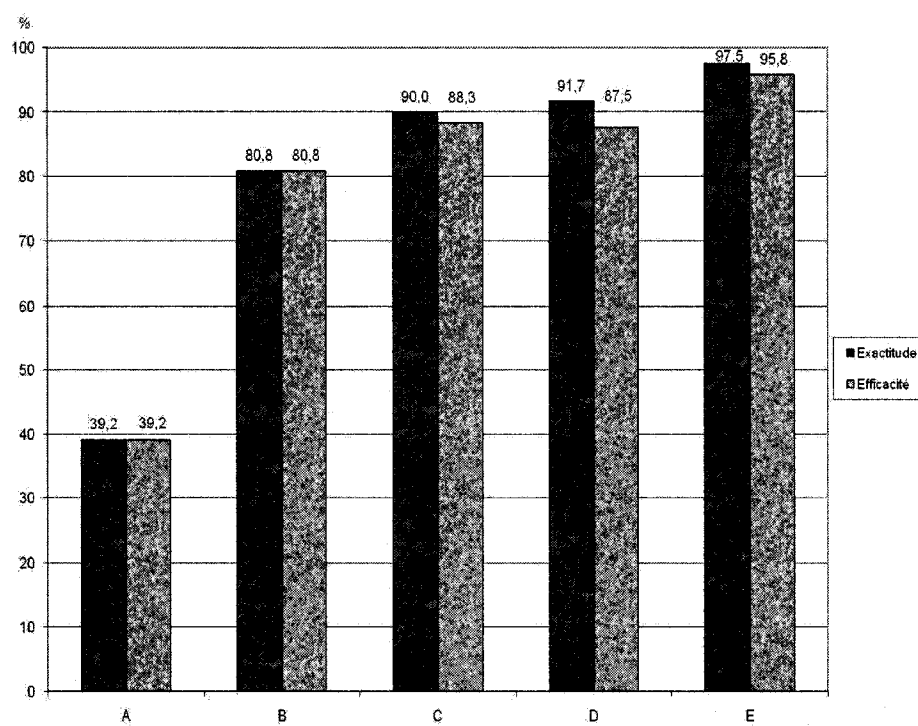


Figure 6.1 – Niveau d'exactitude et d'efficacité des engins de reconnaissance testés à l'aide des six enregistrements et configurés pour reconnaître 20 commandes

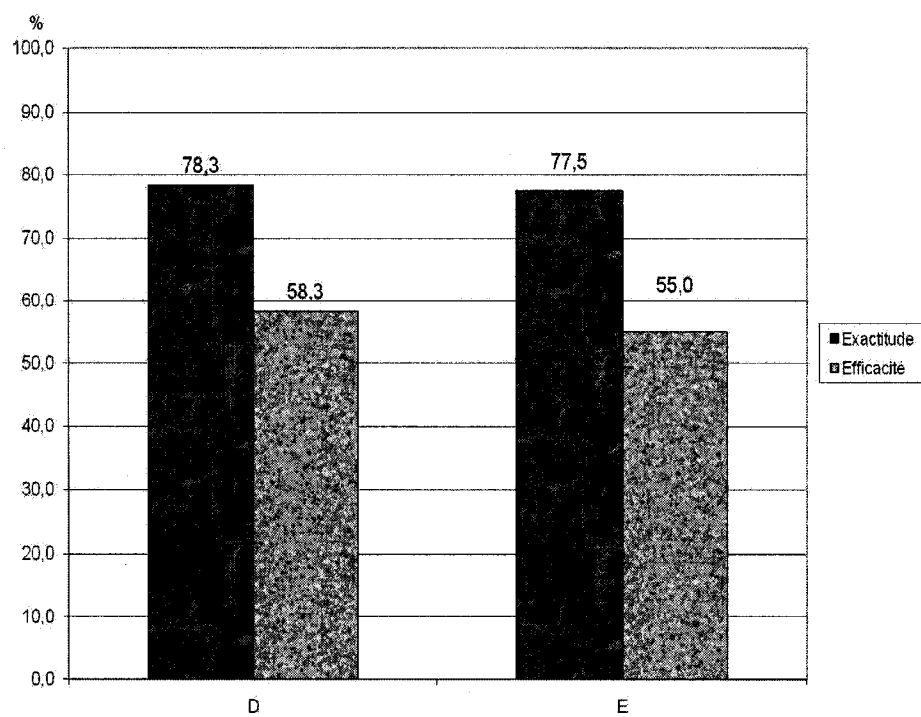


Figure 6.2 – Niveau d'exactitude et d'efficacité des deux meilleurs engins testés à l'aide des six enregistrements et configurés pour reconnaître 137 commandes

particulier, ces résultats montrent les possibilités de ce type d'engins pour faire reconnaître des commandes au poste d'opération. Malgré le fait que les engins n'étaient configurés que de façon minimale, le nombre de commandes composant la liste s'avère être une variable importante à considérer dans l'intégration d'un tel engin. Il est nécessaire de rechercher la configuration optimale de l'engin, par exemple à l'aide d'un dictionnaire phonétique plus adéquat, par l'entraînement de l'engin ou par la structure du vocabulaire. Tous ces paramètres pourraient ainsi influencer grandement sur l'intégration de l'engin choisi.

6.2 Analyse de l'impact du bruit du simulateur

Le bruit ambiant dans un simulateur est très élevé puisque les ingénieurs de son tentent de reproduire le plus fidèlement possible les conditions de bruit réel. À titre d'exemple, le bruit ambiant d'un hélicoptère peut atteindre plus de 100 dB, de sorte que toutes les personnes dans le simulateur doivent porter un casque de communication. Nous avons mené une expérience afin de déterminer l'impact du bruit sur un système de reconnaissance vocale intégré au simulateur : la méthodologie est décrite dans la section suivante.

6.2.1 Méthodologie

Comme l'affirme Gong (1995) dans son rapport sur la reconnaissance vocale en milieu bruyant, la principale cause de dégradation du taux de reconnaissance est la différence entre le milieu d'entraînement du module de reconnaissance et l'environnement réel. Cependant, pour respecter le second critère de (Fox et Weaver) ("le temps nécessaire à l'instructeur pour entraîner le système devait être minimal"), nous préconisons un engin de reconnaissance ne nécessitant pas d'entraînement. Pour l'évaluation du taux de reconnaissance dans des conditions bruyantes, il est possible d'effectuer des

tests dans le milieu lui-même, avec plusieurs sujets, en enregistrant les tests et en évaluant le taux de reconnaissance. Cette méthodologie est coûteuse en ressources et l'environnement réel n'est pas toujours accessible. Des auteurs comme Yamada et al. (2006) ont tenté d'appliquer une certaine distorsion à une voix de synthèse dans le but d'évaluer efficacement et à moindre coût le taux de reconnaissance en milieu bruyant.

Puisque le bruit à l'intérieur d'un simulateur est artificiellement généré, une superposition de ce bruit à des enregistrements de commandes vocales peut être faite. Il est ensuite possible de réintroduire ces fréquences sonores dans le module de reconnaissance afin d'observer les effets du bruit sur le taux de reconnaissance des commandes. La méthodologie fait ainsi appel à une technique moins coûteuse en ressources.

Pour effectuer l'expérience, il faut connaître les paramètres du bruit et de l'enregistrement vocal qui seront utilisés. Tout d'abord, pour obtenir les limites d'un système de reconnaissance vocale dans un simulateur bruyant, l'environnement le plus hostile est utilisé, soit un enregistrement du rotor et des moteurs d'hélicoptère à plein régime. Ensuite, comme le niveau sonore normal de la voix se situe autour de 60 dB (Marsh), c'est le niveau de l'enregistrement vocal utilisé. Différents échantillons de voix et de bruit superposés sont constitués en faisant varier de 0 à 100% (entre 0 et 88 dB) le niveau de bruit présent dans un simulateur.

6.2.1.1 Outils utilisés

Les échantillons de voix et de bruit superposés sont faits dans un laboratoire de son, qui est un environnement fermé et insonorisé. Les deux haut-parleurs suivants sont utilisés : le premier est pour la voix et le second est pour le bruit ambiant ; ce dernier est placé à une distance de deux mètres du premier et est dirigé vers un mur pour provoquer un son indirect et résonnant comme dans un simulateur.

Haut-parleur 1 :

- Modèle : JBL 4206
- Intervalle de fréquence : 42 Hz - 21kHz (-10dB)
- Sensibilité : 87 dB SPL
- Impédance : 8Ω
- Puissance : 75 W

Haut-parleur 2 :

- Modèle : JBL 4602
- Intervalle de fréquence : 50Hz - 15kHz (-10dB)
- Sensibilité : 103 dB SPL
- Impédance : 8Ω
- Puissance : 150 W

Le microphone Plantronics DSP-400 décrit ci-dessous est fixé devant le premier haut-parleur à une distance similaire à celle qui existe entre le microphone et la bouche d'une personne.

- Technologie : Diminution du bruit ("Noise Canceling")
- Type : Boom
- Fréquence de réponse : 100Hz - 10kHz
- Mode : Mono
- Sensibilité : -38 dB

Le bruit ambiant est un enregistrement de bruit simulé dans les conditions décrites plus tôt, soit un rotor et deux moteurs d'hélicoptère à plein régime. Ce bruit a un niveau maximal de 88 dB à la fréquence de 200 Hz. La voix correspond à un enregistrement décrit à la section 6.1.2 utilisé pour l'analyse des engins de reconnaissance. L'enregistrement du bruit simulé de l'hélicoptère ainsi que l'enregistrement vocal sont calibrés à l'aide d'un signal sonore normal de 94 dB. Le logiciel Adobe®Audition®2.0 est utilisé pour la superposition des enregistrements ainsi que pour l'acquisition. Pour des raisons de confidentialité, aucune image de l'installation ne peut être présentée.

6.2.2 Résultats

Les enregistrements ont été recueillis à intervalle de 18 dB ou à des augmentations du volume du bruit d'environ 30%. Les pourcentages de bruit ambiant représentés dans la figure 6.3 correspondent à différents niveaux de bruit que l'instructeur peut mettre à l'intérieur du simulateur. Un volume de 100% est le maximum de bruit que l'instructeur peut mettre. Un niveau supplémentaire de 141% a été ajouté afin de montrer la tendance.

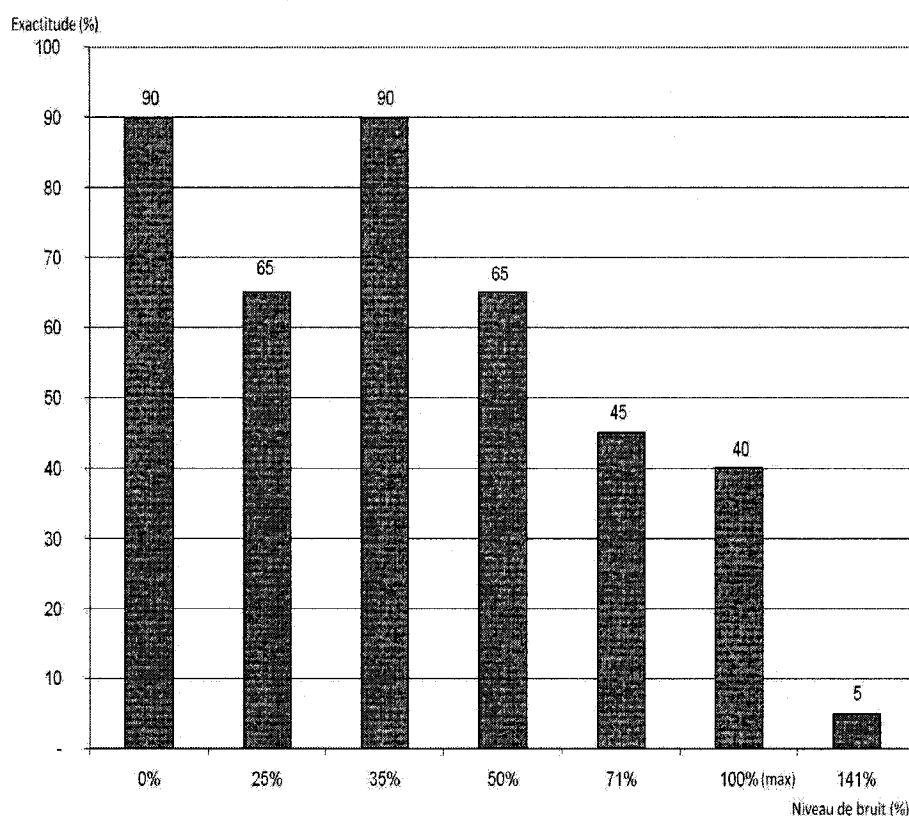


Figure 6.3 – Niveau d'exactitude de la reconnaissance de 20 commandes vocales avec l'engin E configuré pour reconnaître 137 commandes pour différents niveaux de bruit ambiant

Les enregistrements recueillis à l'aide du microphone et des conditions de superposition décrites ci-dessus ont été analysés à l'aide de l'engin de reconnaissance E décrit à l'annexe E. Cet engin s'avère être le plus performant selon les résultats de la section 6.1.2. À noter que 137 commandes

composaient la grammaire et pouvaient être reconnues par l'engin de reconnaissance.

Les résultats d'exactitude obtenus (voir la section 6.1.1 pour des détails sur la procédure de calcul) montrent qu'en situation de bruit ambiant maximal (soit 100%), l'engin est en mesure de reconnaître 40% des 20 commandes faisant partie de l'enregistrement vocal. Lorsque le bruit passe sous les 50%, l'engin s'avère beaucoup plus efficace et permet de reconnaître en moyenne près de 80% des commandes. La baisse d'exactitude observable à 25% du volume peut être due au condition de test avec l'engin difficilement contrôlable. Il est difficile de s'assurer que l'engin de reconnaissance est dans un état parfait pour recevoir le signal sonore et qu'il pourra reconnaître les mêmes commandes qu'à l'essai précédent. Un plus grand nombre de commandes pourrait aussi aider à avoir une granularité plus fine sur l'exactitude de l'engin à un niveau de bruit ambiant donné.

6.2.3 Discussion

La mauvaise performance de l'engin en situation bruyante peut être due à plusieurs facteurs. Tout d'abord, l'engin n'a pas été entraîné et n'a donc pas été soumis à l'environnement avant d'effectuer l'expérimentation. Des tests informels effectués avec l'engin ont montré qu'il tend à s'adapter à la voix de l'utilisateur et à l'environnement. Ensuite, la qualité de l'enregistrement peut avoir eu un impact sur les résultats. Il est cependant important de noter que selon les informations recueillies au chapitre 3, un instructeur est porté à réduire le bruit à l'intérieur du simulateur afin de mieux instruire les pilotes. En considérant ce facteur, la performance de l'engin en milieu bruyant s'avère passable. Enfin, il faut noter qu'une fois intégré, l'engin tendra à s'ajuster à l'utilisateur et à l'environnement compte tenu de sa propriété d'auto-adaptation.

6.3 Analyse des dispositifs d'activation

L'intégration d'un système de reconnaissance vocale au poste d'opération de l'instructeur en simulateur de vol nécessite un dispositif d'activation/désactivation du système pour les raisons suivantes :

1. L'instructeur peut être appelé à jouer le rôle d'un contrôleur aérien (principe d'interférence entre les tâches).
2. L'instructeur peut avoir à discuter avec les étudiants pour leur enseigner ou pour discuter de certaines conditions de la simulation (principe d'interférence entre les tâches).
3. L'instructeur peut vouloir désactiver totalement le système de reconnaissance pour ne travailler qu'avec l'écran tactile (préférence personnelle).
4. Selon le standard militaire 1472F(DoD), tout système de reconnaissance intégré à un équipement doit pouvoir être désactivé lorsqu'il n'est pas utilisé (standard ergonomique).

Pour permettre à l'instructeur de se sentir en contrôle du système de reconnaissance et pour éviter les erreurs d'interaction, un dispositif d'activation/désactivation doit être intégré. Boff et Lincoln (1988) ont comparé différents dispositifs. Les résultats montrent que lorsque la vitesse d'activation est primordiale, un bouton à enfoncer ("push") avec faible résistance doit être utilisé, alors qu'un bouton à plusieurs états ("toggle") ayant une plus forte résistance est nécessaire lorsque les erreurs ou les manipulations commises par inadvertance doivent être évitées. Nous émettons l'hypothèse dans le contexte du poste d'opération qu'un bouton à enfoncer sera plus adéquat puisqu'il permettra une interaction directe, rapide et ponctuelle nécessaire au contrôle par commandes vocales.

La section qui suit présente une étude comparative de différents dispositifs d'activation/désactivation qui peuvent être utilisés dans le contexte du poste d'opération de l'instructeur en simulateur de vol. Cette analyse permettra de tester l'hypothèse définie ci-dessus et de déterminer les dispositifs à considérer pour la suite de l'étude.

6.3.1 Méthodologie

L'expérimentation suivante cherche à ordonner différents dispositifs d'activation/désactivation selon leur facilité d'utilisation et les préférences des utilisateurs dans un contexte d'interaction avec le poste d'opération de l'instructeur.

Différents dispositifs sont présentés aux sujets à qui on demande d'effectuer diverses tâches de paramétrisation d'un simulateur fictif en vue de recueillir leurs impressions sur ces dispositifs ainsi que des données sur les erreurs et les problèmes qu'ils peuvent causer dans ce contexte spécifique. Les données recueillies sur le nombre d'erreurs commises par les sujets lorsqu'ils manipulent le dispositif ou essaient d'interrompre leur scénario permettront de quantifier l'efficacité d'utilisation de chaque dispositif, tout comme le temps total de l'exécution du scénario.

6.3.1.1 Sujets

Les cinq sujets mâles qui participent à l'étude sont des ingénieurs qui connaissent le contexte d'utilisation du poste d'opération et les tâches qui y sont effectuées. Quatre d'entre eux sont droitiers, l'autre gaucher. Trois ont entre 30 et 39 ans alors que les deux autres sont âgés de plus de 50 ans. Tous ont conçu ou ont travaillé avec l'interface graphique du poste d'opération et connaissent les termes relatifs aux paramètres de la simulation. Le choix de cette catégorie de sujets se justifie par le coût et le manque de temps des instructeurs de vol. De plus, les instructeurs disponibles ne sont pas experts du poste d'opération retenu. Il est donc plus approprié de demander aux concepteurs de l'interface graphique de participer à l'expérimentation puisqu'ils sont familiers avec le poste d'opération.

6.3.1.2 Dispositifs utilisés

Cinq dispositifs sont testés à tour de rôle par les sujets. Ils ont été choisis pour leur facilité d'intégration dans un contexte d'entraînement en simulation, parce qu'ils requièrent peu d'adaptation ou sont déjà présents dans le simulateur et à la portée de l'instructeur. À noter qu'aucun dispositif n'est intégré au poste d'opération réel puisque l'effort nécessaire est trop important et qu'une simple analyse comparative est visée. Il n'y a donc aucun effet engendré par la manipulation des dispositifs.

1. Le premier dispositif consiste en un mot-clé, choisi par le sujet, qui doit être inséré avant chaque phrase ou commande donnée au poste d'opération. La désactivation du système de reconnaissance n'est pas nécessaire dans ce cas, puisque le système ne reconnaîtra que les phrases précédées de ce mot. Ainsi, une commande est reconnue en général par un système de reconnaissance lorsque l'opérateur commence sa phrase par le mot-clé et finit lorsqu'une reconnaissance est effectuée par le système ou lorsque le système détecte une pause (silence) après la commande. Nous prenons pour acquis que la commande est valide et reconnue en tout temps, dès que le mot-clé est inséré par le sujet, indépendamment de la syntaxe de la commande.
2. Le second dispositif correspond à un bouton graphique (figure 6.4(a)) sur lequel on doit appuyer sur l'écran tactile pour que le système de reconnaissance s'active, et appuyer de nouveau pour qu'il se désactive. Le bouton graphique fait 3 cm x 2 cm, suivant ainsi les recommandations pour une taille minimale de 2.2 cm de diamètre faites par Boff et Lincoln (1988). Ce dispositif offre une rétroaction au sujet puisqu'il est gris lorsque désactivé et devient vert lorsqu'activé.
3. Le troisième dispositif est un mini-levier à deux positions, fixé au bras du siège, qui doit être mis à la position élevée pour activer le système et remis en position normale pour le désactiver. Pour cette expérimentation, la chaise du sujet est en position fixe pour que ce dernier soit à une distance adéquate de l'écran tactile et le mini-levier

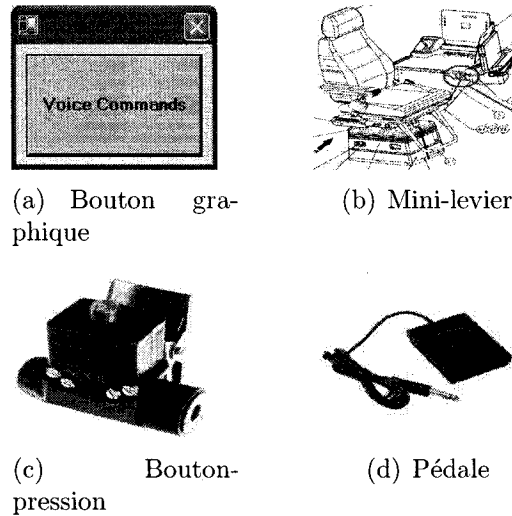


Figure 6.4 – Ensemble de dispositifs d'activation/désactivation testés par les sujets à l'aide d'un scénario de paramétrisation

est fixé sous cet écran, reproduisant ainsi la disposition montrée sur la figure 6.4(b).

4. Le quatrième dispositif est un bouton-pression qui s'ajoute au bout du fil du casque d'écoute et qui peut être attaché à la ceinture de l'opérateur (figure 6.4(c)). Les sujets n'ont qu'à l'attacher à leur ceinture pour cette expérimentation.
5. Le dernier dispositif est une pédale (figure 6.4(d)) placée aux pieds du sujet, sur laquelle il doit appuyer pour enclencher l'activation et la désactivation de la reconnaissance.

Dans les deux derniers cas, le bouton-pression et la pédale, il est nécessaire que le sujet tienne le bouton ou la pédale dans une position enfoncée pour paramétrer le poste d'opération et qu'il relâche le bouton ou la pédale lorsqu'il n'utilise pas la reconnaissance vocale. En ce qui a trait au mini-levier et au bouton graphique, ils sont actifs tant et aussi longtemps que le bouton graphique est actionné ou que le mini-levier est en position relevée.

6.3.1.3 Tâches à exécuter

On demande aux sujets d'exécuter un scénario de paramétrisation du poste d'opération. Ce scénario, décrit à l'annexe F, est inspiré d'une procédure de test dans un simulateur et permet de savoir si certains paramètres fonctionnent correctement. Certains ajouts ont été faits au scénario afin de forcer les sujets à interrompre l'entrée de données pour effectuer une autre tâche : premièrement, les "jeux de rôle" qui sont exécutés par l'instructeur pour simuler l'interaction entre les pilotes et la tour de contrôle ou les agents de bord, et deuxièmement, "l'entrée de notes" qui est nécessaire pour que l'instructeur puisse rapporter les erreurs survenues pendant l'entraînement. Ces ajouts permettent d'observer le comportement des sujets lorsqu'ils brisent momentanément la tâche de paramétrisation et qu'ils manipulent (activent/désactivent) le dispositif. Le scénario met ainsi en jeu quatre types d'interactions différentes soit, la modification de paramètres via l'écran tactile, la paramétrisation à l'aide d'une commande vocale (au choix du sujet), la prise de notes avec crayon et papier, et le "jeu de rôle" ou la parole.

6.3.1.4 Procédures et consignes

Aucun système de reconnaissance n'est utilisé. La fluidité est requise pour cette expérimentation. Les commandes sont inventées par les sujets. De cette façon, l'apprentissage des commandes ne vient pas biaiser les résultats et la configuration d'un système de reconnaissance selon les préférences de chaque sujet n'est pas requise.

Un sujet à la fois participe à l'expérience qui a lieu à une station de tests du poste d'opération avec un écran tactile. Le sujet reçoit les informations écrites sur l'expérience à laquelle il prend part. On informe le sujet que le but de l'expérimentation est d'obtenir des informations sur le dispositif le plus approprié pour activer un système de reconnaissance dans un contexte d'interaction avec le poste d'opération. On lui explique que l'exercice consiste

à modifier des paramètres du poste d'opération avec la voix et avec l'écran tactile en suivant un même scénario d'entraînement pour différents dispositifs d'activation/désactivation. On invite le sujet à se familiariser avec le scénario et à poser des questions s'il en ressent le besoin. Les quatre types d'interaction (manuelle, voix, entrée de notes, jeu de rôle) lui sont expliqués. On informe le sujet que la voix et l'interaction manuelle servent à modifier les valeurs des paramètres tel qu'indiqué, que l'entrée de notes se fait en écrivant sur une feuille de papier blanc les informations requises, et que, dans le jeu de rôle, il doit dire à haute voix l'information décrite dans le scénario. On lui explique enfin qu'il doit désactiver le dispositif lorsqu'il fait le jeu de rôle.

Pour tous les dispositifs à tester, le sujet est assis à une même distance de l'écran tactile du poste d'opération. La séance expérimentale est enregistrée sur bande audio pour analyse future.

L'expérience démarre lorsque le sujet est prêt à commencer avec le premier dispositif. Le scénario est répété avec les cinq dispositifs différents. Chaque sujet a une minute avant l'exécution du scénario pour se familiariser avec le dispositif qu'on lui demande de tester. Il est important de noter que les dispositifs sont testés dans un ordre différent pour chaque sujet afin de contrebalancer l'impact de la familiarisation avec le scénario sur le temps d'exécution.

Les erreurs de manipulation (c.-à-d. omission d'activations) et les problèmes de fluidité (c.-à-d. interruption momentanée du scénario) dans l'exécution des tâches par le sujet sont comptabilisés par l'expérimentateur. Le temps total de l'exécution du scénario avec chacun des dispositifs est mesuré. À la fin, le sujet remplit le questionnaire présenté à l'annexe F sur ses données biographiques et sur sa perception de la facilité de manipulation et de l'efficacité des différents dispositifs utilisés.

6.3.2 Résultats

Le tableau 6.2 montre la moyenne des résultats des sujets pour chaque dispositif. Ces résultats portent sur le nombre d'erreurs de manipulation, le nombre de problèmes de fluidité dans l'exécution de la tâche, les évaluations de la facilité de manipulation et l'efficacité, ainsi que le temps d'exécution du scénario.

Tableau 6.2: Moyenne des résultats des sujets pour les variables évaluées sur chaque dispositif

Dispositifs	Nb. d'erreurs de mani- pulation	Nb.de problèmes de flui- dité	Niveau de fa- cilité	Niveau d'effi- cacité	Temps d'exécution (minutes)
Mot clé	1	0	96%	92%	4 :33
Bouton graphique	1	3	64%	56%	4 :41
Mini-levier	0	5	64%	60%	3 :50
Bouton- pression	1	4	72%	64%	3 :38
Pédale	0	3	76%	80%	3 :32

La pédale, suivie du bouton-pression sont les dispositifs les plus rapides, et le bouton graphique, le plus lent. Le mini-levier est le dispositif ayant causé le plus de problèmes de fluidité dans l'exécution des tâches. Selon les résultats présentés sous forme de pourcentage dans le tableau 6.2, le mot clé est le dispositif le plus facilement manipulable et le plus efficace. Comme le montre le tableau 6.3, le mot-clé est le dispositif préféré des sujets. Trois sujets ont préféré le mot-clé, un la pédale et un le bouton-pression.

Tableau 6.3: Dispositifs d'activation/désactivation de système de reconnaissance vocale préférés par les sujets

Sujets	Premier choix	Deuxième choix
Sujet 1	Bouton-pression	Pédale
Sujet 2	Pédale	Mot clé
Sujet 3	Mot clé	Bouton-pression
Sujet 4	Mot clé	Pédale
Sujet 5	Mot clé	Pédale

6.3.3 Discussion

Nous pouvons confirmer l'hypothèse selon laquelle un dispositif avec faible résistance est préférable dans un contexte de manipulations rapides et fréquentes. La préférence des utilisateurs pour le mot-clé peut s'expliquer par le fait qu'il nécessite une moins grande concentration et devient rapidement un automatisme chez le sujet. De plus, il peut être mobile et utiliser ses mains à d'autres escients. Cependant, des tests plus poussés devraient être faits pour vérifier la réceptivité des instructeurs de vol à constamment insérer un mot-clé avant leurs commandes, ainsi que pour déterminer le type de mots à utiliser pour éviter que le système reconnaisse des paroles qui ne s'adressent pas à lui.

6.4 Sommaire

Les analyses présentées dans ce chapitre ont permis de cerner le contexte technologique entourant l'intégration d'un système de reconnaissance vocale au poste d'opération de l'instructeur en simulateur de vol. Tout d'abord, l'engin E s'avère le plus efficace pour reconnaître des commandes vocales servant à paramétrer un simulateur de vol. Ensuite, les tests effectués en milieu

bruyant ont démontré que ce facteur pouvait avoir un impact négatif important sur la performance d'un engin de reconnaissance. Cependant, puisque les utilisateurs ont l'habitude de réduire le bruit et que les tests n'ont pas été effectués avec une configuration optimale (p.ex., le dictionnaire phonétique et l'engin pourraient être adaptés au contexte et au vocabulaire de l'entraînement en simulateur), l'engin peut être efficace dans un tel milieu. Enfin, un dispositif physique comme une pédale ou un bouton-pression s'avère adéquat lors de l'exécution de tâches reliées à l'entraînement en simulateur. Un mot-clé inséré avant la commande est le dispositif préféré par les sujets ayant participé à l'expérimentation, mais une analyse plus poussée de ce dispositif particulier serait nécessaire.

Les technologies analysées dans ce chapitre peuvent à présent servir à créer un prototype. Il nous reste à vérifier à l'aide de tests d'utilisabilité si leur intégration est possible et adéquate. Le chapitre suivant intégrera l'ensemble des connaissances acquises jusqu'à présent dans ce mémoire afin de tester l'utilisabilité d'un système de reconnaissance vocale pour contrôler le poste d'opération.

CHAPITRE 7

CONSTRUCTION D'UN PROTOTYPE ET TESTS UTILISATEURS

Tous les aspects ergonomiques influençant l'intégration d'un système de reconnaissance vocale au poste d'opération ont été analysés et les technologies qui sont adéquates pour le contexte du poste d'opération ont été ciblées. Il reste à présent à incorporer ces informations à un premier prototype et à voir si les analyses effectuées permettent à l'utilisateur d'interagir de façon satisfaisante par la voix avec le poste d'opération. Ce chapitre présente le prototype que nous avons construit ainsi que les tests d'utilisabilité menés avec un premier système de reconnaissance vocale au poste d'opération.

7.1 Analyse de l'intégration de la reconnaissance vocale au poste d'opération

Dans le but de connaître le niveau de satisfaction des instructeurs à l'égard de l'utilisation d'un prototype de système de reconnaissance vocale intégré au poste d'opération, nous avons réalisé des tests d'utilisabilité qui sont présentés dans la section qui suit.

7.1.1 Méthodologie

Les mesures d'utilisabilité sont primordiales dans le domaine des interactions humain-machine. Elles permettent de comparer différentes conceptions entre elles et de valider l'amélioration d'une conception particulière. Un système est utilisable lorsqu'il permet à l'utilisateur de réaliser sa tâche avec

efficacité, efficience et satisfaction dans le contexte d'utilisation spécifié¹. Dans le cas d'un système de reconnaissance vocale, bien que les mesures objectives de la performance soient importantes, elles ne sont pas suffisantes, puisqu'elles ne permettent pas de connaître le niveau de satisfaction des utilisateurs (Hone et Graham). De plus, le succès de l'intégration d'un tel système est lié à la facilité d'utilisation de ce système, ce que les mesures de performance ne révèlent pas. La méthodologie utilisée s'appuie donc sur ces prémisses et utilise une mesure de la performance d'un prototype du système de reconnaissance à utiliser au poste d'opération, ainsi qu'une mesure de satisfaction des utilisateurs à interagir avec ce prototype dans le cadre d'un scénario de paramétrisation.

7.1.1.1 Sujets

Les cinq sujets étaient invités à participer à l'expérimentation, mais uniquement deux étaient disponibles. Ces deux instructeurs de vol ou d'anciens instructeurs de vol en simulateur de sexe masculin et âgés entre 50 et 60 ans parlent l'anglais comme langue maternelle. Ils ont plusieurs années d'expérience comme instructeur et connaissent l'entraînement en simulateur ainsi que le poste d'opération servant à contrôler la simulation.

7.1.1.2 Outils utilisés

Nous avons développé un prototype de système de reconnaissance vocale intégré au poste d'opération sur la base des résultats expérimentaux précédents, relatifs au vocabulaire de commandes et à la rétroaction (section 5.3.3). L'intégration d'un dispositif d'activation/désactivation y est omise pour simplifier le prototype ainsi que les tests. L'engin de reconnaissance qui s'est avéré le plus exact et le plus efficient (voir la section 6.1.2) est utilisé.

¹Norme ISO 9241-11 : Exigences ergonomiques pour travail de bureau avec terminaux à écrans de visualisation (TEV) – Partie 11 : Lignes directrices relatives à l'utilisabilité, http://www.iso.org/iso/fr/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883

Les ressources ne permettent cependant pas d'intégrer le prototype dans l'environnement réel du simulateur. Une station de tests est utilisée. Le prototype est détaillé à la section 7.1.1.3 suivante.

En plus de comptabiliser les erreurs du système de reconnaissance, un outil d'évaluation subjective de la satisfaction des utilisateurs est utilisé : un questionnaire de satisfaction post-expérimentation, développé par Hone et Graham (2000). Ce questionnaire a été bâti en suivant une approche empirique de développement de questionnaires du domaine de la psychométrie. Ces chercheurs ont tenté de créer un instrument valide et fidèle servant à mesurer six facteurs pouvant traduire l'expérience d'un utilisateur, soit la performance perçue, l'appréciation, la charge cognitive, l'ennui, le contrôle et la rapidité². Deux itérations de conception du questionnaire ainsi que les résultats obtenus auprès d'un grand nombre de répondants révèlent que les trois premiers facteurs sont fiables (Hone et Graham). Les trois autres facteurs nécessitent un travail supplémentaire afin de valider leur fiabilité. Ce questionnaire est approprié pour mesurer l'utilisabilité du prototype intégré au poste d'opération, et ce, malgré le fait que l'évaluation de ce questionnaire ne soit pas terminée. Aucune autre technique valide de mesures subjectives n'a pu être identifiée.

Le questionnaire est composé de 34 questions disponibles dans (Hone et Graham) et une échelle de notation comparative à sept niveaux est utilisée. Cette échelle permet de mesurer la direction et l'intensité de la satisfaction du sujet.

Échelle :

- *Strongly Agree* (Fortement en accord)
- *Agree* (En accord)
- *Slightly Agree* (Légèrement en accord)
- *Neutral* (Neutre)
- *Slightly Disagree* (Légèrement en désaccord)
- *Disagree* (En désaccord)

²Traduction libre des termes des auteurs (Hone et Graham) : *accuracy, likeability, cognitive demand, annoyance, habitability, speed*

– *Strongly Disagree* (Fortement en désaccord)

7.1.1.3 Description du prototype

L’engin de reconnaissance vocale E a été intégré au poste d’opération. Le contrôle graphique, présenté en bas à droite sur la capture d’écran 7.1, a été construit en langage de programmation C# pour faciliter l’intégration avec l’interface du poste d’opération. Ce contrôle permet d’activer ou de désactiver l’engin de reconnaissance vocale, d’afficher l’état de l’engin et de fournir une rétroaction à l’utilisateur.



Figure 7.1 – Capture d’écran du prototype intégré au poste d’opération (bouton graphique identifié “Speech Recognition” en bas à droite dans l’application)

Si l’engin est inactif, le bouton graphique (identifié “Speech Recognition”) demeure bleu et la zone de texte vide. Lorsque l’engin est actif, ce même bouton graphique devient rouge tel que le montre la figure 7.2. Les trois dernières commandes reconnues par le système sont affichées dans cette

zone de texte. À noter qu'une confirmation auditive est aussi perceptible dans le casque d'écoute du sujet. La commande reconnue est répétée à l'aide d'un module de synthèse vocale à partir du texte de cette commande. Pour aider le sujet dans sa tâche, la page contenant le paramètre modifié est affichée automatiquement. Il peut ainsi s'assurer que la valeur a été entrée correctement et la corriger en cas d'erreur en rappelant la commande ou en modifiant le paramètre manuellement.



Figure 7.2 – Rétroaction visuelle (trois dernières commandes reconnues) présentée au sujet lorsque le système de reconnaissance vocale est actif

Le contrôle graphique du système de reconnaissance est toujours en premier plan, donc toujours visible pour le sujet. L'effet de transparence montrée à la figure 7.3, où on peut voir la page sous le contrôle graphique du système de reconnaissance, permet au sujet d'observer la rétroaction visuelle sans qu'il y ait interférence avec les pages du poste d'opération qui sont primordiales à la tâche principale de paramétrisation.

La grammaire sous-jacente à l'engin de reconnaissance est de type hors contexte et a la forme Backus-Naur où les symboles terminaux sont représentés par les mots pouvant être reconnus par le système. L'encadré ci-dessous donne un exemple de la syntaxe utilisée. Les mots “<” “>” sont les symboles non terminaux utilisés pour former les règles de production. Cette forme est imposée par l'engin de reconnaissance utilisé.

```

...
< S > :
set fuel < number > {method="Fuel"; param = <
number >;}
...

```

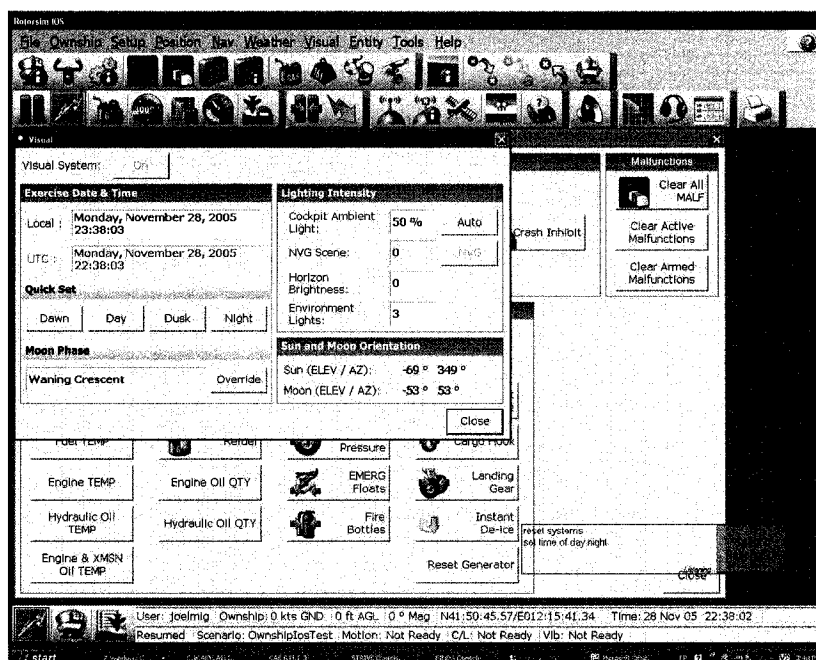


Figure 7.3 – Capture d’écran montrant l’effet de transparence du contrôle de rétroaction lorsqu’une partie d’une page se retrouve sous le contrôle

Les mots réservés *method* et *param*, dont la valeur est reçue par le module implémenté, permettent d’appeler la fonctionnalité adéquate en lui fournissant la valeur appropriée du paramètre à modifier. Dans le cas d’une valeur numérique comme dans l’exemple, le mot-clé *param* prendrait la valeur “2000” si l’utilisateur a énoncé la commande “set fuel two thousand”. Ces ajouts à la grammaire permettent de découpler, d’un point de vue logiciel, la fonctionnalité de la syntaxe de la commande associée.

7.1.1.4 Tâches à exécuter

Le scénario exécuté par ces sujets est divisé en deux exercices (voir l’annexe G). Tout d’abord, un scénario de paramétrisation d’un simulateur est présenté aux sujets. Ce scénario fictif est inspiré d’une procédure de test d’un simulateur qui sert à déterminer si certains paramètres fonctionnent correctement. Les motivations derrière ce choix de scénario sont la familiarité des

sujets avec ce type de description et le nombre important de manipulations à effectuer avec le poste d'opération. Ensuite, nous demandons au sujet de créer lui-même un scénario d'entraînement lui permettant d'interagir avec le poste d'opération. Ce deuxième exercice donne le contrôle au sujet tel qu'en situation réelle.

7.1.1.5 Procédures et consignes

Un seul sujet à la fois participe à l'expérimentation. Les explications sur les tests d'utilisabilité auxquels il participe lui sont présentées sur papier. On informe le sujet que le but de l'expérimentation est d'observer son comportement et de recueillir ses impressions sur l'utilisation de commandes vocales pour contrôler le poste d'opération. On lui demande de se familiariser avec la formulation des commandes vocales et la rétroaction. La figure 5.2 présente la liste hiérarchique des commandes vocales divisées par catégorie de commandes et par construction syntaxique. Le sujet peut consulter cette liste en tout temps sur une feuille grand format posée près de lui. On lui demande d'utiliser un ton de voix naturel tout en gardant à l'esprit que le système doit reconnaître les commandes. Un temps de 10 minutes lui est ensuite alloué pour se familiariser avec les commandes et tester le système.

Le sujet prend ensuite connaissance du scénario à exécuter présenté sur papier. Il exécute les tâches qui sont requises dans le scénario de l'annexe G. L'expérimentateur comptabilise les erreurs du système de reconnaissance.

À la fin de l'expérimentation, le sujet répond au questionnaire sur papier.

7.1.2 Résultats

Deux instructeurs de vol seulement se sont prêtés à l'expérimentation. L'expérimentation a porté sur 30 commandes et a duré environ 10 minutes pour chaque sujet. Le nombre d'erreurs de reconnaissance est de 2 pour le

premier sujet (ou 7%) et de 8 pour le second sujet (ou 27%). Il est important de souligner que le second sujet a eu des problèmes avec une commande particulière, ce qui a influencé le taux d'erreurs. Après investigation, un problème a été décelé et réglé dans la grammaire de reconnaissance.

Les résultats de l'analyse du questionnaire de satisfaction sont présentés dans le tableau 7.1.

Tableau 7.1 – Pourcentage de satisfaction des sujets à l'égard du prototype testé et évaluée selon les six facteurs contenus dans le questionnaire.

Facteurs	Sujet 1	Sujet 2
Performance perçue	32%	8%
Appréciation	67%	41%
Charge cognitive	77%	63%
Ennui	60%	60%
Contrôle	68%	61%
Rapidité	79%	79%

Plus un pourcentage dans le tableau 7.1 est élevé, plus le sujet a été satisfait de ce facteur durant le test. Les résultats démontrent qu'en général le niveau de satisfaction est assez bon, bien que la performance perçue soit plutôt faible. Cela peut être dû, dans un premier temps, au problème dans la grammaire, et dans un deuxième temps, à un manque de connaissance du modèle de langage supporté. Le temps d'apprentissage du modèle était relativement court et les sujets n'ont pas ressenti le besoin de connaître adéquatement le langage à utiliser puisqu'ils pouvaient s'y référer. Cependant, les sujets ont tenté à quelques reprises de donner des commandes sans toutefois analyser correctement le modèle, ce qui a provoqué des erreurs de reconnaissance.

7.2 Sommaire

Un système de reconnaissance vocale a été intégré à l'interface graphique du poste d'opération. Le meilleur engin de reconnaissance résultant de l'ana-

lyse du chapitre 6 ainsi que le modèle de langage construit au chapitre 5 ont été intégrés. Des tests d'utilisabilité avec des instructeurs de vol ont permis de déceler certaines lacunes dans le prototype et d'observer que la satisfaction globale d'une telle interaction est assez bonne. Les tests permettent toutefois de constater qu'une meilleure configuration de l'engin serait nécessaire et qu'une période d'apprentissage du langage est primordiale afin que les instructeurs de vol soient davantage satisfaits de ce mode d'interaction.

CONCLUSION

Puisque l'entraînement en simulateur de vol coûte très cher, il s'avère important d'augmenter l'efficacité du travail de l'instructeur au poste d'opération pour qu'il puisse consacrer une plus grande part de son temps à l'interaction avec le pilotes plutôt qu'à l'interaction avec le poste d'opération. Plusieurs résultats de recherches indiquent que la reconnaissance vocale est une technologie qui a fait ses preuves dans des contextes connexes à l'entraînement en simulateur. La fiabilité de cette technologie est un facteur de moins en moins limitatif puisque la technologie a grandement évolué et que l'on retrouve de plus en plus de systèmes qui l'intègrent, et ce, dans diverses environnements. De plus, la qualité de l'interaction des utilisateurs avec cette technologie est primordiale selon plusieurs auteurs. Nous nous sommes donc attardés dans ce mémoire à définir et analyser les facteurs humains qui sont en jeu en vue d'améliorer l'efficacité du travail de l'instructeur à l'aide d'une technologie de reconnaissance vocale.

Une analyse des caractéristiques et des habitudes de travail des instructeurs de vol a d'abord été effectuée sur la base de données recueillies par entrevue individuelle auprès de cinq instructeurs. Cela nous a permis en utilisant la méthode de la *persona* de définir deux personnages représentatifs des instructeurs de vol. Nous avons ainsi une base commune de connaissances sur les utilisateurs du poste d'opération qui nous permet de mieux orienter les décisions de conception du poste d'opération en répondant aux besoins et aux attentes des instructeurs.

Les tâches des instructeurs de vol ont été analysées, plus particulièrement celles qui sont effectuées dans les étapes d'initialisation du simulateur et d'entraînement. La performance des utilisateurs a été analysée selon un modèle prédictif de la performance humaine et la méthode d'analyse de tâche *KLM-GOMS*. Cette analyse a permis de comparer, de manière peu coûteuse, l'interaction manuelle sur l'écran tactile du poste d'opération avec l'interaction vocale. Les résultats ont montré que la voix était plus rapide que l'écran tac-

tile pour la majorité des tâches analysées. Un gain moyen de 33% est obtenu, ce qui nous permet de conclure que ce mode d'interaction diminue le temps d'exécution pour plusieurs tâches et qu'une augmentation de l'efficacité est possible.

L'analyse des tâches des instructeurs a permis d'identifier plusieurs commandes vocales correspondant aux fonctionnalités des différents systèmes à contrôler ; cependant, on ne connaît pas le vocabulaire et la syntaxe des commandes qui sont les plus appropriés pour l'interaction vocale. Une analyse du langage des instructeurs de vol a été conduite en utilisant la méthode du Magicien d'Oz et la méthode Delphi qui ont permis de faire participer plusieurs. Cela a permis de connaître le langage naturel utilisé par ces derniers et le degré d'accord entre les sujets. De plus, nous nous sommes inspirés des résultats de recherches faites dans le domaine des langages de commandes textuelles pour améliorer le modèle de langage et construire une syntaxe qui offre une certaine facilité d'apprentissage et d'utilisation. Il est important de noter que le gain en temps d'exécution de plusieurs tâches analysées tend à augmenter avec les commandes que l'on retrouve à présent dans le modèle conçu.

Les technologies qui doivent être prises en compte dans l'intégration de la reconnaissance vocale au poste d'opération ont aussi fait l'objet d'analyses. Ainsi, cinq engins de reconnaissance ont été comparés sur la base de l'exactitude et de l'efficacité, ce qui nous a permis de sélectionner le meilleur engin et de construire une méthode de tests qui pourra éventuellement servir à analyser d'autres engins disponibles. À l'aide de cet engin de reconnaissance, nous avons par la suite effectué des tests en milieu bruyant puisque l'environnement principal dans lequel le système doit être intégré, soit le simulateur, recrée les conditions de bruit réelles d'un cockpit d'avion ou d'hélicoptère. Une étude faite dans un laboratoire de son a permis de recueillir des données sur le taux de reconnaissance vocale avec le bruit du simulateur en arrière-plan. Les résultats ont démontré qu'au-delà de 50% du volume du bruit ambiant (ou 44 dB), l'exactitude de l'engin est inférieure à 50% ce qui est faible. Il faut cependant rappeler que l'analyse du travail

des utilisateurs a révélé que la plupart des instructeurs de vol ne placent pas le volume au niveau maximum et tendent à le réduire à environ 25%. Enfin, l'intégration d'un système de reconnaissance vocale nécessite un dispositif d'activation/désactivation. Une étude a permis de montrer que les dispositifs de type mot clé, bouton-pression ou pédale sont préférables au contexte d'entraînement puisqu'ils sont plus efficaces et préférés des sujets. Ces résultats donnent une meilleure perspective sur les technologies qui devraient être mises en place dans l'environnement réel de simulation pour augmenter l'efficacité des instructeurs.

Les résultats des études évoquées ci-dessus ont été intégrés à un prototype de système de reconnaissance vocale au poste d'opération. Ce prototype a permis d'effectuer des tests d'utilisabilité avec des instructeurs de vol qui ont mis le système à l'épreuve et ont donné leur niveau de satisfaction à l'égard d'un système de reconnaissance vocale intégré au poste d'opération pour définir les paramètres du simulateur au cours d'un scénario de travail représentatif. Un questionnaire a été utilisé pour recueillir le niveau de satisfaction des sujets à l'égard du prototype de système de reconnaissance vocale. Nous avons remarqué que certains réglages du système doivent être faits pour permettre une meilleure performance de l'engin de reconnaissance, mais que dans l'ensemble le niveau de satisfaction des utilisateurs est assez bon.

L'analyse des facteurs humains entourant l'intégration d'un système de reconnaissance vocale au poste d'opération de l'instructeur en simulateur de vol permet de conclure que cette technologie a un effet bénéfique sur l'efficacité de l'instructeur. Cependant, certaines données sont encore manquantes.

Travaux futurs et perspectives

Certaines données n'ont pas pu être recueillies au cours de cette recherche. Elles sont importantes pour la suite du projet d'intégration de la reconnaissance vocale au poste d'opération.

Tout d'abord, le nombre de tests effectués avec les instructeurs a été restreint à cause de leur manque de disponibilité. Il serait important de poursuivre ces tests d'utilisabilité et de tenter de faire intervenir un plus grand nombre de sujets.

La fréquence des tâches effectuées au cours d'un entraînement est une donnée cruciale pour calculer le temps gagné avec l'interaction vocale. De plus, il serait important de cibler certains entraînements typiques qui guideraient ces évaluations de performance et les tests d'utilisabilité. Il est impossible pour le moment de calculer le gain de temps apporté par l'utilisation de la voix par rapport à l'écran tactile et de calculer les économies réalisées. L'activation d'un journal d'actions effectuées au poste d'opération dans un simulateur et servant à des entraînements réels permettrait de recueillir les données nécessaires sur la fréquence, la durée et la séquence de tâches effectuées par les instructeurs.

Les analyses décrites dans cette recherche n'ont pu être faites dans un environnement réel de simulation compte tenu du coût important des simulateurs, du temps d'entraînement des pilotes et du manque de disponibilité des instructeurs. Il serait intéressant d'intégrer le système de reconnaissance vocale au poste d'opération dans un simulateur et de recueillir des données sur le temps d'exécution des tâches, la charge mentale, la facilité d'apprentissage du langage de commandes vocales et la satisfaction des instructeurs à interagir par la voix lors d'un entraînement.

Cette recherche s'est attardée uniquement à l'efficacité de l'instructeur de vol à l'intérieur du simulateur. Il serait intéressant de poursuivre les recherches selon différentes avenues. Par exemple, le poste d'opération à l'extérieur du simulateur s'avère être une opportunité d'intégration intéressante puisque l'environnement de travail est beaucoup moins bruyant. De plus, d'autres applications sont utilisées par l'instructeur de vol lorsqu'il est à l'extérieur du simulateur. La voix pourrait ainsi servir de mode d'interaction pour contrôler d'autres joueurs simulés dans le cas d'une simulation d'entraînement tactique militaire, ou encore pour préparer les exercices et les scénarios d'entraînement en dictant à voix haute l'état de la simulation

selon les phases de l'entraînement.

Il existe aussi des dispositifs d'entraînement de pilotage individuel à moindre coût. Ils permettent aux pilotes de s'entraîner par eux-mêmes sans l'intervention onéreuse de l'instructeur en simulateur. Les pilotes pourraient aussi bénéficier de l'interaction vocale pour modifier les paramètres de leur dispositif d'entraînement, évitant ainsi d'interrompre leurs tâches principales d'entraînement.

Finalement, il serait pertinent d'étudier l'opportunité de remplacer l'ordinateur de poche utilisé par les pilotes de tests lors de la certification des simulateurs par l'interaction vocale. Les pilotes de tests doivent habituellement manoeuvrer l'appareil tout en contrôlant les paramètres de la simulation afin de s'assurer que tout fonctionne correctement. Un certain temps d'apprentissage serait nécessaire, mais cela permettrait d'offrir un plus grand nombre de fonctionnalités aux pilotes qui sont jusqu'à présent réduits à quelques boutons importants compte tenu des contraintes d'affichage sur l'écran d'un ordinateur de poche.

BIBLIOGRAPHIE

ACT-R RESEARCH GROUP (2007). "Act-r". Website. <http://act-r.psy.cmu.edu/>.

BABER, C. (1997). *Beyond The Desktop, Designing and Using Interaction Devices*. Academic Press, Californie, USA.

BLACK, J. B. et MORAN, T. P. (1982). "Learning and remembering command names". *Proceedings of the 1982 conference on Human factors in computing systems*, New York, NY, USA, pages 8-11. ACM.

BOFF, K. R. et LINCOLN, J. E. (1988). *Engineering Data Compendium of Human Perception and Performance*, chapitre 12.4. Wright-Patterson A.F.B., Harry G. Armstrong Aerospace Medical Research Laboratory, Ohio.

BRADFORD, J. H. (1995). "The human factors of speech-based interfaces, a research agenda". *SIGCHI Bulletin*, **27**(2), 61-67.

CAE ELECTRONICS LTD. (1981). "Instructor-simulator interface design". Technical report, Air Force Human Resources Laboratory.

CANADA, T. (2004). *Flight Instructor Guide - Aeroplane*.

CARD, S. K., MORAN, T. P., et NEWELL, A. (1980). "The keystroke-level model for user performance time with interactive systems". *Communication*. ACM, **23**(7), 396-410.

CARD, S. K., MORAN, T. P., et NEWELL, A. (1983). *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, NJ.

CARROLL, J. M. (1982). "Learning, using and designing command paradigms". *Human Learning*, **1**(1), 31-62.

CHARLES, J. P. (1988). "Instructor/operator station (ios) design guide". Technical report, Air Force Human Resources Laboratory.

COGTOOL (2006). "The cogtool project website". Website. [http://www.cs.cmu.edu/~sim\\$bej/cogtool/index.html](http://www.cs.cmu.edu/~sim$bej/cogtool/index.html).

COOK, M. J., CRANNER, C., FINAN, R., SAPELUK, A., et MILTON, C.-A. (1996). *Engineering Psychology and Cognitive Ergonomics Volume One - Transportation Systems*, chapitre 15 - Memory load and task interference : hidden usability issues in speech interfaces, pages 141-149. Ashgate.

DAMPER, R. I., TRANCHANT, M. A., et LEWIS, S. M. (1996). "Speech versus keying in command and control : Effect of concurrent tasking". *International Journal of Human-Computer Studies*, **45**(3), 337-348.

DAMPER, R. I. et WOOD, S. D. (1995). "Speech versus keying in command and control applications". *International Journal of Human-Computer Studies*, **42**(3), 289-305.

DOD (1999). *MIL-STD-1472F, Design Criteria Standard - Human Engineering*. Department of Defense, Washington, D.C.

DRAPER, M., CALHOUN, G., RUFF, H., WILLIAMSON, D., et BARRY, T. (2003). "Manual versus speech input for unmanned aerial vehicle control station operations". *Proceedings of the Human Factors & Ergonomics Society's 47th Annual Meeting*, pages 109-113.

EASTER, A. W., KRYWAY, J. T., OLSON, W. R., PETERS, S. M., et SLEMON, G. K. (1986). "Development of instructor support feature guidelines". Technical report, Air Force Human Resources Laboratory.

ELWORTH, C. (1981). "Instructor/operator display evaluation methods". Technical report, Air Force Human Resources Laboratory.

FAA (1999). *Aviation Instructor's Handbook*. Federal Aviation Administration.

FAA (2002). *Flight Instructor - Practical Test Standards for Airplane*. Federal Aviation Administration, Department of Transportation, Washington, DC.

FAA (2003a). *Federal Aviation Regulations - Part 121 - Operating requirements : Domestic, flag, and supplemental operations*. Federal Aviation Administration, Department of Transportation.

FAA (2003b). *Federal Aviation Regulations - Part 135 - Operating requirements : Commuter and on demand operations and rules governing persons on board such aircraft*. Federal Aviation Administration, Department of Transportation.

FAA (2003c). *Federal Aviation Regulations - Part 141 - Pilot schools*. Federal Aviation Administration, Department of Transportation.

FAA (2003d). *Federal Aviation Regulations - Part 61 - Certification : Pilots, flight instructors, and ground instructors*. Federal Aviation Administration, Department of Transportation.

FOX, R. et WEAVER, T. (1989). "A voice-operated instructor's station". *AIAA Flight Simulation Technologies Conference and Exhibit*, Boston, MA, pages 304–309. American Institute of Aeronautics and Astronautics.

FRASER, N. M. et GILBERT, G. N. (1991). "Simulating speech systems". *Computer Speech and Language*, **5**, 81–99.

FULFORD, C. P. (1992). "Systematically designed text enhanced with compressed speech audio". *Proceedings of Selected Research and Development Presentations at the Convention of the Association for Educational Communications and Technology and Sponsored by the Research and Theory Division*.

GONG, Y. (1995). "Speech recognition in noisy environments : A survey". *Speech Communication*, **16**, 261–291.

GORNIAK, P. et ROY, D. (2003). "Augmenting user interfaces with adaptive speech commands". *Proceedings of the 5th international conference on Multimodal interfaces*, pages 176–179.

GRASSO, M. A. et FININ, T. W. (2005). "Task integration in multimodal speech recognition environments". *Crossroads - The ACM Student Magazine*, **n.d.** [www.acm.org/crossroads/xrds3-3/taskint.html], Consulté le 25 janvier 2006.

GUSTAFSON, J. (2002). *Developing Multimodal Spoken Dialogue Systems Empirical Studies of Spoken Human Computer Interaction*. PhD thesis, Department of Speech, Music and Hearing, KTH, Stockholm.

HACKOS, J. T. et REDISH, J. C. (1998). *User and Task Analysis for Interface Design*. Wiley Computer Publishing.

HAKULINEN, J., TURUNEN, M., SALOMEN, E.-P., et RÄIHÄ, K.-J. (2004). "Tutor design for speech-based interfaces". *Proceedings of the 2004 conference on Designing interactive systems : processes, practices, methods, and techniques*, pages 155–164.

HAYS, R. T., SEAMON, A. G., et BRADLEY, S. K. (1997). "User-oriented design analysis of the vesub technology demonstration system". Technical Report 97-013, Naval Air Warfare Center Training Systems Division, Orlando, FL.

HONE, K. S. et GRAHAM, R. (2000). "Towards a tool for the subjective assessment of speech system interfaces (sassi)". *Natural Language Engineering*, **6**(3/4), 287–305.

HONE, K. S. et GRAHAM, R. (2001). "Subjective assessment of speech-system interface usability". *Proceedings of Eurospeech 2001, 7th European*

Conference on Speech Communication and Technology, volume 3, Aalborg, Denmark, pages 2083–2086.

HUMAN FACTORS GROUP OF THE FAA WILLIAM J. HUGHES TECHNICAL CENTER (2003). *Human Factors Design Standard*. Federal Aviation Administration.

JAMES, F. et ROELANDS, J. (2002). “Voice over workplace (vowp) : Voice navigation in a complex business gui”. *Proceedings of the fifth international ACM conference on Assistive technologies*, pages 197–204.

JOHN, B. E., PREVAS, K. C., SALVUCCI, D. D., et KOEDINGER, K. R. (2004). “Predictive human performance modeling made easy”. DYKSTRA-ERICKSON, E. et TSCHELIGI, M., editors, *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, pages 455–462. ACM.

KAMM, C. (1995). “User interfaces for voice applications”. *Proceeding of the National Academy of Sciences of the USA - Paper edition*, volume 92, pages 10031–10037.

KLEMMER, S. R., SINHA, A. K., CHEN, J., LANDAY, J. A., ABOOBAKER, N., et WANG, A. (2000). “Suede : A wizard of oz prototyping tool for speech user interfaces”. *Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 1–10.

LANDAUER, T. K., GALOTTI, K. M., et HARTWELL, S. (1983). “Natural command names and initial learning : a study of text-editing terms”. *Commun. ACM*, **26**(7), 495–503.

LARSON, J. A. (1995). “Ten guidelines for designing a successful voice user interface, speech technology magazine”. Consulté le 12 février 2006.

LATHROP, B., CHENG, H., WENG, F., MISHRA, R., CHEN, J., BRATT, H., CAVEDON, L., BERGMANN, C., HAND-BENDER, T.,

PON-BARRY, H., BEI, B., RAYA, M., et SHRIBERG, L. (2005). "A wizard of oz framework for collecting spoken human-computer dialogs : An experiment procedure for the design and testing of natural language in-vehicle technology systems".

LEE, K. M. et LAI, J. (2005). "Speech versus touch : A comparative study of the use of speech and dtmf keypad for navigation". *International Journal of Human-Computer Interaction*, **19**(3), 343–360.

LEGGATT, A. et NOYES, J. (1996). *Engineering Psychology and Cognitive Ergonomics Volume One - Transportation Systems*, chapitre 51 - Speech recognition technology : implications for crew communication in armoured fighting vehicle, pages 451–457. Ashgate.

LIBERMAN, M. (2006). "Sex and speaking rate". Website. [http://itre.cis.upenn.edu/\\$\sim\\$myl/languagelog/archives/003423.html](http://itre.cis.upenn.edu/$\sim$myl/languagelog/archives/003423.html).

LINSTONE, H. A. et TUROFF, M. (1975). *The Delphi Method, Techniques and Applications*. Addison-Wesley.

MARSH, A. (1999). "Online course on acoustics, speech intelligibility". http://www.kemt.fei.tuke.sk/Predmety/KEMT320_EA/_web/Online_Course_on_Acoustics/index_acoustics.html.

MATESSA, M. et REMINGTON, R. (2005). "Reusable templates of human performance in space shuttle procedures". Technical report, NASA Space Human Factor Engineering.

MCTEAR, M. F. (2002). "Spoken dialogue technology : Enabling the conversational user interface". *ACM Computing Surveys*, **34**(1), 90–169.

MEISEL, B. (2007). "Ford offers sync using microsoft auto platform and nuance speech technology". *Speech Strategy News, Speech Technology in Business and Communications*, **Février**, 5.

MINKER, W. et BENNACEF, S. (2001). *Parole et dialogue homme-machine*. Éditions Eyrolles, Paris.

MINKER, W. et BENNACEF, S. (2004). *Speech and Human-Machine Dialog*. Kluwer Academic Publishers, Boston.

MORRISON, D. L., GREEN, T. R., SHAW, A. C., et PAYNE, S. J. (1984). "Speech-controlled text-editing : effects of input modality and of command structure". *International Journal of Man-Machine Studies*, **21**(1), 49–63.

NAPIER, H. A., BATSELL, R. R., GUADANGO, N. S., et LANE, D. M. (1989). "Impact of a restricted natural language interface on ease of learning and productivity". *Commun. ACM*, **32**(10), 1190–1198.

NEWELL, A. F., CARMICHAEL, A., GREGOR, P., et ALM, N. (2003). *The Human-Computer Interaction Handbook*, chapitre Information Technology for Cognitive Support, pages 464–481. Lawrence Erlbaum Associates.

O'HARA, J., BROWN, W., LEWIS, P., et PERSENSKY, J. (2002). *Human-System Interface Design Review Guidelines*. U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington, DC.

POOCK, G. K. (1980). "Experiments with voice input for command and control : using voice input to operate a distributed computer network". Technical Report NPS55-80-016, Naval Post-graduate School, Monterey, CA.

PRUITT, J. et GRUDIN, J. (2003). "Personas : practice and theory". *DUX '03 : Proceedings of the 2003 conference on Designing for user experiences*, New York, NY, USA, pages 1–15.

RAMESH, R. et SYLLA, C. (1990). "A framework for the optimal design of instructor/operator stations in flight simulators". *IEEE Transactions on Systems, Man, and Cybernetics*, **20**(3), 571–581.

ROOS, W. (1987). *calling TOWER, manuel de radio aéronautique, VFR+CVFR*, aéro-club de suisse edition.

ROSENBERG, J. (1982). "Evaluating the suggestiveness of command names". *Proceedings of the 1982 conference on Human factors in computing systems*, New York, NY, USA, pages 12–16. ACM.

ROSENBERG, J. (1983). "A featural approach to command names". *CHI '83 : Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, New York, NY, USA, pages 116–119. ACM.

ROSSI, M., DI CRISTO, A., HIRST, D., MARTIN, P., et NISHINUMA, Y. (1981). *Intonation de l'acoustique a la semantique*. Institut de phonetique d'Aix-en-provence.

SALBER, D. et COUTAZ, J. (1993). "Applying the wizard of oz technique to the study of multimodal systems". *Proceedings of East/West Human Computer Interaction*, pages 219–230.

SANDERS, C. D. (1981). "Task analytic techniques : Application to the design of a flight simulator instructor/operator console". Technical report, Air Force Human Resources Laboratory.

SCAPIN, D. L. (1982). "Computer commands labelled by users versus imposed commands and the effect of structuring rules on recall". *Proceedings of the 1982 conference on Human factors in computing systems*, New York, NY, USA, pages 17–19. ACM.

SCHWARTZ, N. F. (1977). "Display and speech devices for simulator instructor/operator station applications". Technical report, Air Force Human Resources Laboratory, Advanced Systems Division, Wright-Patterson Air Force Base, Ohio 45433. Final Report, June 1975 - June 1976.

SRINIVASAN, S. et BROWN, E. (2002). "Is speech recognition becoming mainstream?". *Computer*, **2**, 38–41.

STARKS, D. R. et MORGAN, M. J. (1992). "Integrating speech recognition into a helicopter". *Proceedings of ESCA Tutorial and Research Workshop on Speech Processing in Adverse Conditions*, Cannes-Mandelieu, France.

SWAIL, C. et KOBIERSKI, R. (1997). "Direct voice input for control of an avionics management system". *Proceedings of the American Helicopter Society 53rd Annual Forum*.

SYLLA, C., RAMASWAMY, R., et BRADY, R. (1988). "Design optimization of flight simulator's instructor/operator station". *Computers and Industrial Engineering*, **15**(1), 418–427.

SYLLA, C. et RAMESH, R. (1990). "Modeling instructor training preparation in a simulation flight training program". *Proceedings of the 12th annual conference on Computers and industrial engineering*, Elmsford, NY, USA, pages 362–366.

TEO, L. et JOHN, B. E. (2006). "Comparisons of keystroke-level model predictions to observed data". *Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems*, volume 2 of *Work-in-progress*, pages 1421–1426. ACM.

TSIMHONI, O., SMITH, D., et GREEN, P. (2004). "Address entry while driving : Speech recognition versus a touch-screen keyboard". *Human Factors*, **46**(4), 600–610.

TURUNEN, M. (2004). "Speech application design and development". Department of Computer Sciences, University of Tampere.

VIDULICH, M. A., NELSON, W. T., et BOLIA, R. S. (2006). "Speech-based controls in simulated air battle management". *International Journal of Aviation Psychology*, **16**(2), 197–213.

VINCIGUERRA, B. (2002). "A comparaisn of commercial speech recognition components for use with the project54 system". Mémoire de maîtrise, University of New Hampshire.

WARNER, H. D. (1987). "Instructor/operator station design : Handbook for aircrew training devices". Technical report, Air Force Human Resources Laboratory.

WEINSTEIN, C. J. (1990). "Opportunities for advanced speech processing in military computer-based systems". *Human Language Technology Conference / Proceedings of a workshop on Speech and natural language*, pages 433 – 452.

WEINSTEIN, C. J. (1995). "Military and government applications of human-machine communication by voice". *Proceedings of the National Academy of Sciences of the United States of America*, pages 10011–10016. The National Academy of Sciences.

WILLIAMSON, D. T., DRAPER, M. H., CALHOUN, G. L., et BARRY, T. P. (2005). "Commercial speech recognition technology in the military domain : Results of two recent research efforts". *International Journal of Speech Technology*, 8(1), 9–16.

WOOD, S. (1973). "Speech tempo". Technical Report 9, Working Papers, Phonetics Laboratory, Department of General Linguistics, Lund University.

YAMADA, T., KUMAKURA, M., et KITAWAKI, N. (2006). "Performance estimation of speech recognition system under noise conditions using objective quality measures and artificial voice". *IEEE Transactions on Audio, Speech and Language Processing*, 14(6), 2006– 2013.

YANKELOVICH, N. (1996). "How do users know what to say?". *Interactions*, 3(6), 32–43.

YANKELOVICH, N. et LAI, J. (1999). "Designing speech user interfaces". *CHI '99 extended abstracts on Human factors in computing systems*, pages 124–125.

YUAN, J., LIBERMAN, M., et CIERI, C. (2006). "Towards integrated understanding speaking rate conversation". Department of Linguistics, Linguistic Data Consortium, University of Pennsylvania, USA.

ANNEXE A

SCÉNARIO D'ACTIVITÉS DE TRAVAIL POUR UN INSTRUCTEUR DE VOL

La période d'entraînement débute tout d'abord par une session de briefing dans une salle aménagée à cet effet. Pendant environ une heure, l'instructeur révise les objectifs et les détails de l'entraînement à faire. Il utilise une présentation PowerPoint ainsi qu'une carte montrant un diagramme d'approche d'un aéroport comme supports visuels pour expliquer différents éléments concernant le décollage et l'atterrissage. Ensuite, il discute en détail de l'entraînement et des différentes parties qui le composent. Il informe les pilotes des conditions météorologiques qui seront en vigueur pendant les exercices. Il questionne par la suite les pilotes sur les différentes procédures qu'ils devront suivre pendant les parties de l'entraînement. Il utilise un tableau blanc avec des crayons de différentes couleurs pour expliquer l'atterrissage en détail et les erreurs communes à ne pas faire. Enfin, il utilise également une réplique d'un cockpit sur une affiche sur le mur pour pointer les différents instruments.

L'entraînement dans le simulateur débute à la suite de cette session d'explications. L'entraînement est planifié pour durer trois heures séparées en parts égales pour que chaque étudiant joue le rôle du pilote à son tour, pendant que l'autre est le copilote. Une pause de cinq minutes est prévue entre ces deux parties d'une heure et demie.

L'instructeur débute en s'asseyant au poste derrière les pilotes et installe son matériel sur la table à sa gauche ou sur ses genoux. Il place son siège de façon à être confortable et bien voir

les pilotes travailler devant lui. Il positionne les écrans tactiles pour bien les atteindre, celui de droite un peu plus bas que celui de gauche. Il prépare les feuilles d'informations météorologiques qu'il distribuera pendant l'entraînement aux pilotes pour jouer le rôle de l'imprimante qui fournit généralement les informations provenant des centres de contrôle aérien. Il ouvre son cartable de scénarios et de plans de vol à la page de l'entraînement en cours. Il installe ensuite son casque de communication pour pouvoir interagir avec les pilotes via le système d'intercom.

L'instructeur débute son interaction avec le poste d'opération en entrant la quantité d'essence dans l'appareil selon les données prescrites dans son scénario. Il entre l'aéroport de référence et la piste de décollage qui servira pour le début de l'entraînement. Il commande une reposition de l'appareil en bout de piste prêt pour le décollage et désactive la fonction pause de la simulation pour que la reposition s'effectue. Il entre les données concernant les nuages et la visibilité à basse altitude. Il change la direction et la vitesse du vent, ainsi que la température extérieure, toujours selon les conditions inscrites dans son scénario. Il modifie les conditions de la piste selon les contraintes auxquelles l'entraînement doit soumettre les pilotes. Après avoir discuté avec les pilotes et vérifié qu'ils sont prêts, il active le mouvement du simulateur et met en fonction tous les systèmes de simulation. Il modifie le moment de la journée pour mettre la simulation de nuit.

Pendant que les pilotes suivent la procédure de décollage de l'appareil, l'instructeur observe et joue le rôle d'un contrôleur aérien pour informer les pilotes de certaines situations ou pour fournir des informations comme les données météorologiques. Il prépare le mauvais fonctionnement du moteur droit et commande à la simulation d'activer ce problème à une certaine altitude. Il attend ensuite que le mauvais fonctionnement se

produise et prend en note les réactions des pilotes. Une fois la procédure terminée, il arrête temporairement la simulation pour pouvoir recharger les systèmes de l'appareil et remettre les valeurs normales. Il réinitialise donc les systèmes ainsi que toutes les températures de l'appareil. Enfin, il efface le mauvais fonctionnement activé et prépare l'appareil pour un nouvel objectif de l'entraînement. Au cours d'une nouvelle partie du vol, l'instructeur augmente la vitesse de la simulation pour que cette dernière arrive à un endroit particulier de l'entraînement. Pour cette raison, l'entraînement ne dure pas le temps que prendrait le vol réel normalement.

L'instructeur poursuit donc l'entraînement jusqu'à ce que tous les objectifs soient atteints ou que le temps alloué dans le simulateur soit écoulé.

Le tableau A.1 énumère les différentes tâches et sous-tâches décrites dans le scénario précédent et retenues pour l'analyse.

Tableau A.1: Liste des tâches et sous-tâches retenues pour l'analyse

Tâches et sous-tâches
Ajuster le carburant
Mettre 15000 kilogrammes d'essence dans l'appareil
Ajuster l'aéroport de référence
Spécifier l'aéroport LIRF comme référence pour l'entraînement
Spécifier la piste de décollage 16L comme référence pour l'entraînement
Reposition
Repositionner l'appareil en position pour le décollage
Activer la simulation
Ajuster les nuages
Mettre la couche nuageuse dispersée
suite sur la page suivante

Tableau A.1: Liste des tâches et sous-tâches retenues pour l'analyse
(suite)

Tâches et sous-tâches
- Ajuster la visibilité de la couche nuageuse à 0.5 - Mettre la fin de la couche nuageuse à 10000 pieds - Mettre le début de la couche nuageuse à 200 pieds - Ajuster le vent - Ajuster la direction du vent à 180 degrés - Ajuster la vitesse du vent à 45 noeuds - Ajuster la température de l'air à 20 degrés - Ajuster les conditions de la piste - Mettre la condition rugosité de la piste à 3 - Mettre les conditions de surface de la piste à sèche - Ajuster le visuel - Mettre la simulation de vol de nuit - Activer un dysfonctionnement - Activer un dysfonctionnement du moteur droit à 20 pieds - Ajuster la vitesse de simulation - Mettre l'avance rapide à 3 - Réinitialisation - Réinitialiser tous les systèmes de l'appareil - Réinitialiser toutes les températures de l'appareil - Désactiver les dysfonctionnements

ANNEXE B

PRÉSENTATION DE COGTOOL

Le projet *CogTool* a été développé à l'université Carnegie Mellon, à Pittsburgh, dans les laboratoires de l'institut d'interaction humain-ordinateur (CogTool). Ce logiciel permet de calculer les temps que prendrait un utilisateur expert pour effectuer certaines tâches avec une interface particulière. Le logiciel intègre des modèles permettant de calculer les interactions avec une souris, un clavier, un écran tactile ainsi que toutes les manipulations faites à l'aide d'un ordinateur personnel de poche comme la technique d'entrée de données *Graffiti*.

La première étape consiste à insérer dans le logiciel les images représentant l'interface conçue. Il peut s'agir de captures d'écran ou de prototypes graphiques, comme le montre la figure B.2, qui permettent de représenter les interactions que l'utilisateur aurait avec le système. Une fois les images intégrées dans *CogTool*, il est nécessaire d'ajouter les composantes graphiques qui seront utilisées pour les tâches à analyser (figure B.1). Plusieurs types de composantes graphiques sont disponibles, allant du simple bouton sur lequel on appuie jusqu'à la boîte de texte, en passant par le champ de reconnaissance *Graffiti*. Les composantes graphiques sont représentées par un champ coloré qui montre la région qui est accessible par l'utilisateur. Nous pouvons ici supposer que la taille du champ doit correspondre à la taille de la composante graphique représentée dans l'image, car la loi de Fitts doit être considérée lors du calcul des temps d'exécution des tâches. Une fois les composantes graphiques disposées, il suffit de lier les fenêtres entre elles en insérant une transition entre une composante graphique et la fenêtre qui est appelée par une action effectuée à l'aide de cette composante, comme dans la figure B.2. Si l'action ne fait pas appel à une nouvelle fenêtre par contre, il faut lier la transition de la composante à la fenêtre courante et l'action sera effectuée en laissant l'interface dans le même état. Lorsque toutes les

transitions ont été ajoutées pour permettre l'exécution des tâches que l'on désire représenter, il faut créer un script d'exécution de la tâche. Pour ce faire, il faut indiquer la fenêtre de départ de la tâche et sélectionner dans l'ordre les actions que doit effectuer l'utilisateur pour réaliser la tâche (voir figure B.3). Une fois le script d'une tâche terminé, le logiciel calcule automatiquement le temps nécessaire pour effectuer la tâche selon la séquence d'actions élémentaires inscrites. Il est important de noter que lors de l'entrée des actions dans le script, le logiciel insère automatiquement un temps pour penser à l'action à effectuer, ainsi qu'un temps pour repérer la composante graphique avec laquelle interagir. Finalement, la figure B.4 montre le résultat des calculs effectués pour les différentes tâches. Nous pouvons par la suite répéter les manipulations avec une nouvelle conception et ainsi comparer les deux interfaces en termes de temps de réalisation des différentes tâches.

Le logiciel permet aussi d'exporter les conceptions vers une série de pages HTML dynamiques dans lesquelles il est possible de naviguer d'une fenêtre à l'autre comme l'utilisateur le ferait avec l'interface réelle. C'est donc un moyen intéressant de créer des prototypes semi-fonctionnels rapidement et qui nous informent en plus de l'efficacité de l'interface créée.

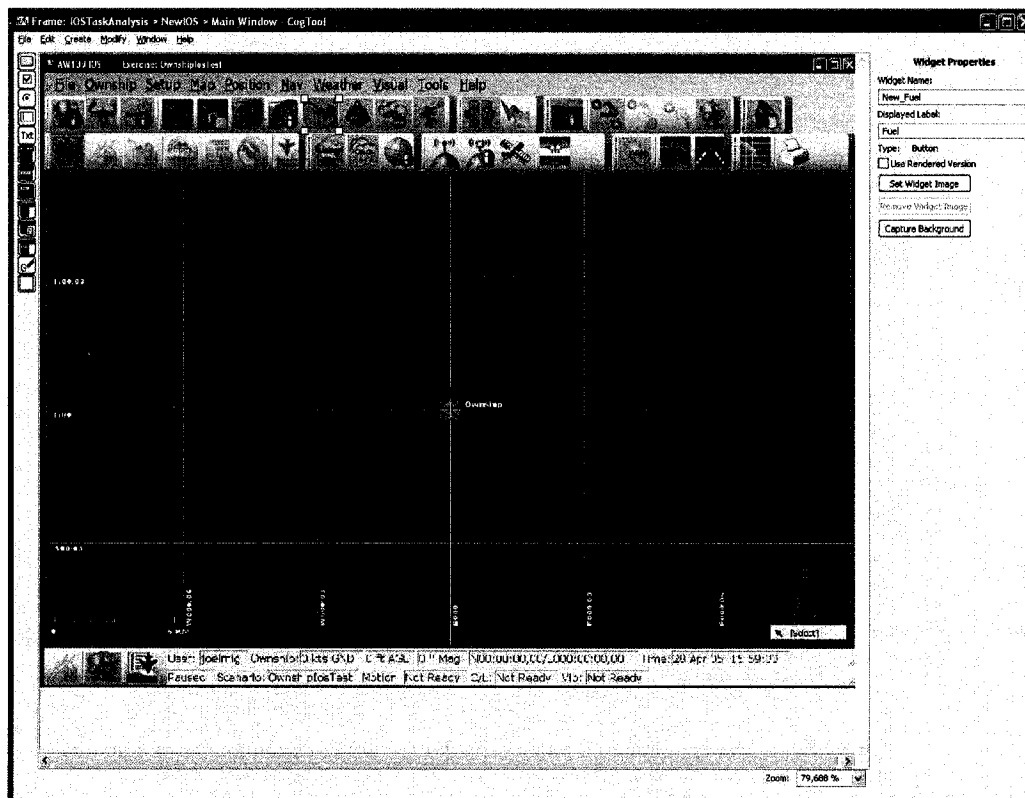


Figure B.1 – CogTool - Fenêtre spécifique et les composantes graphiques accessibles

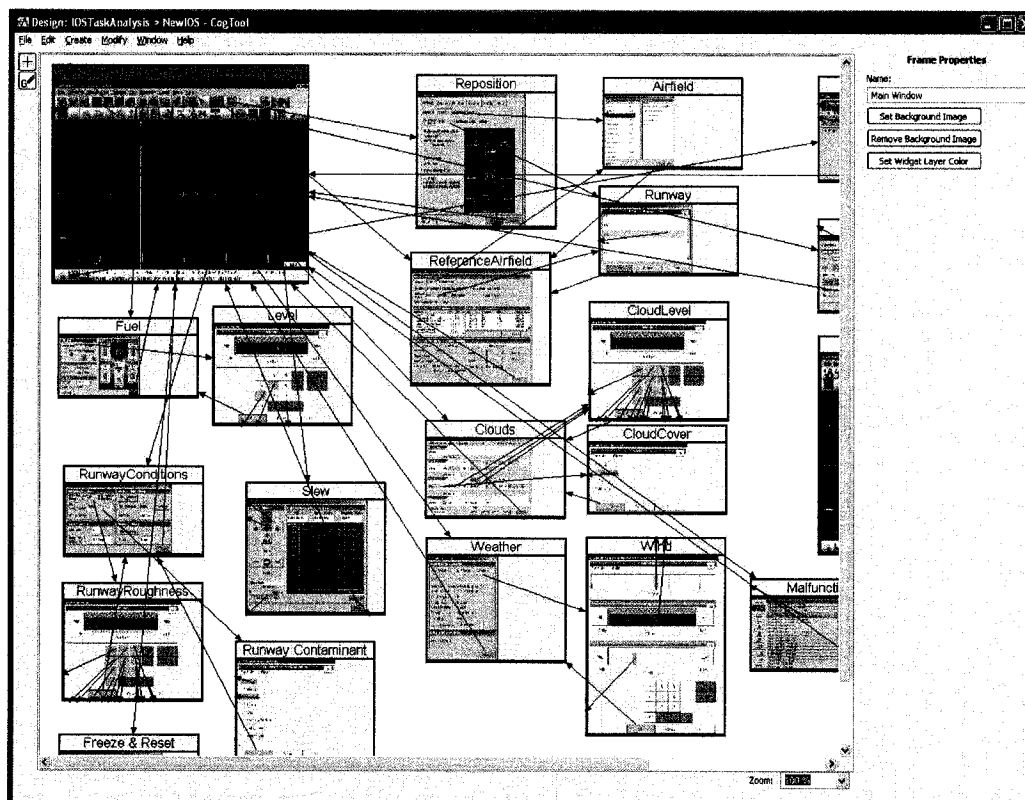


Figure B.2 – CogTool - Représentation de la conception et des liens entre les fenêtres

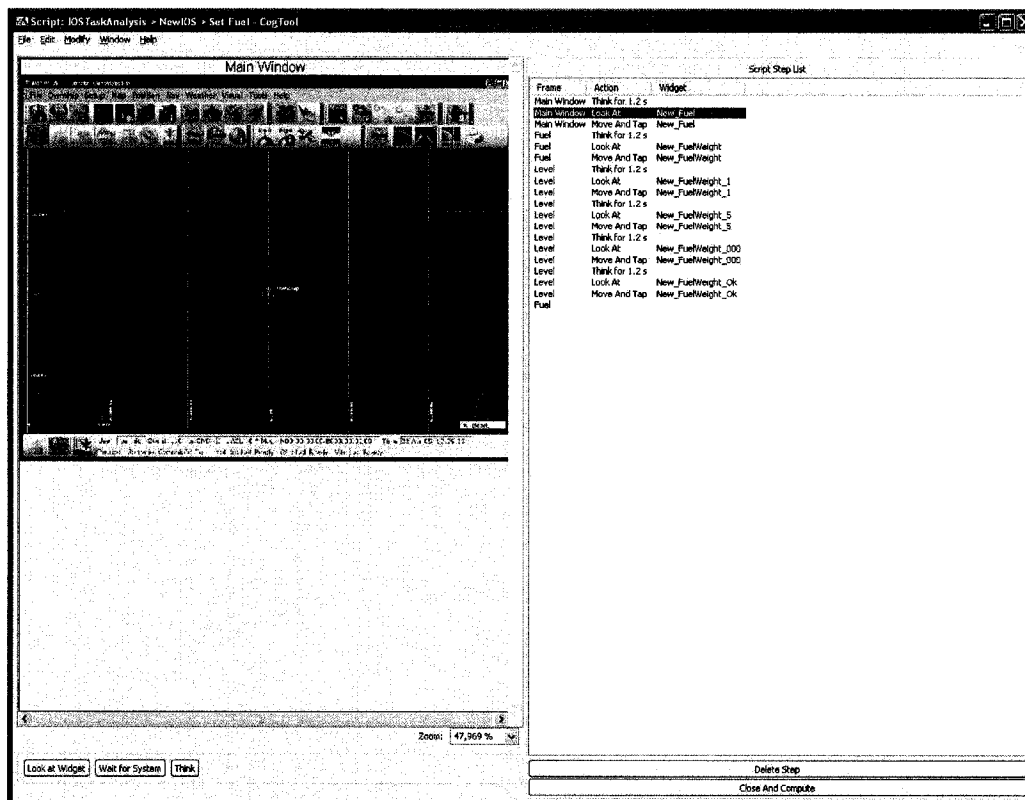
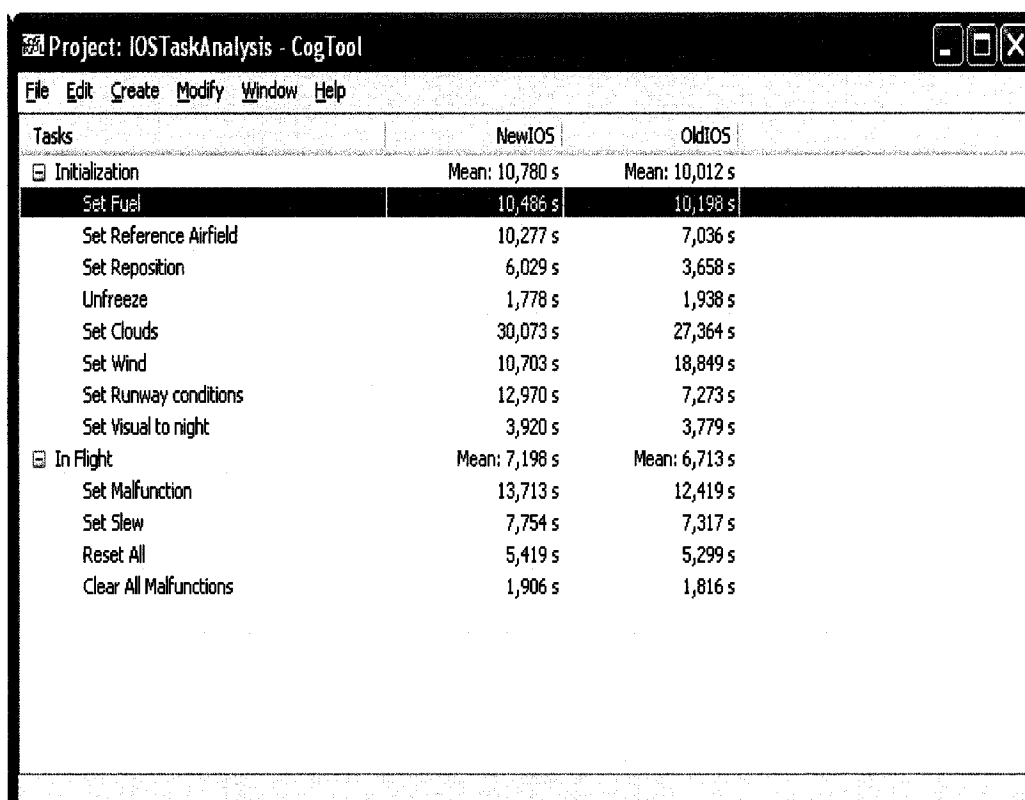


Figure B.3 – CogTool - Script décrivant la séquence d'actions à effectuer pour réaliser une tâche particulière



The screenshot shows the 'Project: IOSTaskAnalysis - CogTool' window. It features a menu bar with 'File', 'Edit', 'Create', 'Modify', 'Window', and 'Help'. Below the menu is a table with three columns: 'Tasks', 'NewIOS', and 'OldIOS'. The table is organized into two main sections: 'Initialization' and 'In Flight', each with a sub-total row showing the mean time for each scenario.

Tasks	NewIOS	OldIOS
Initialization	Mean: 10,780 s	Mean: 10,012 s
Set Fuel	10,486 s	10,198 s
Set Reference Airfield	10,277 s	7,036 s
Set Reposition	6,029 s	3,658 s
Unfreeze	1,778 s	1,938 s
Set Clouds	30,073 s	27,364 s
Set Wind	10,703 s	18,849 s
Set Runway conditions	12,970 s	7,273 s
Set Visual to night	3,920 s	3,779 s
In Flight	Mean: 7,198 s	Mean: 6,713 s
Set Malfunction	13,713 s	12,419 s
Set Slew	7,754 s	7,317 s
Reset All	5,419 s	5,299 s
Clear All Malfunctions	1,906 s	1,816 s

Figure B.4 – CogTool - Liste de tâches et les temps calculés pour les deux interfaces comparées

ANNEXE C

QUESTIONNAIRE D'ENTREVUE

1. What is your background ? (commercial, military, transport ; tactical, flight maneuvers ; aircraft types)
2. What is your knowledge of technologies ? Do you like using them ?
3. What other applications are you using ?
4. Is English your primary language ?
5. How do you learn how to use the IOS ?
6. Do you ever have read the IOS manual or use any documentation ?
7. Do you use any other tools while doing training ?
8. How long do you practice to manipulate the IOS before doing pilot training ?
9. How often do you use the IOS in one training session ?
10. Do you spend more time using the IOS or doing pilot training tasks (speaking, looking, evaluating, etc) ?
11. In which contexts do you use the IOS ?
12. For which purposes do you use the IOS ?
13. Under which stress conditions do you have to use the IOS ?
14. What are the environment conditions (quiet, noisy, relax, stress) while using the IOS ?
15. Do you need assistance while using the IOS ?
16. How do you perceive the IOS ?
17. What is your biggest concern about how the IOS actually work ?
18. Is there any changes you would like to make to actual IOS ?
19. Which are the most critical/important features or functions ?
20. Which are the less important features or functions ?

21. Can you remember of any training incident/story related specifically to IOS use?
22. Could you describe your tasks in a typical training session?
23. Could you list all the high level tasks instructor-related you can think of? (ex : briefing, preparing training, etc.)
24. Do you think speech recognition could help you do some particular tasks?
25. For which tasks you think speech recognition could be helpful?
26. Can you think of any problem it could introduce?

ANNEXE D

LISTE DES ÉNONCÉS SOUMIS AUX SUJETS ET QUESTIONNAIRE POST-EXPÉRIMENTATION

Liste des énoncés soumis aux sujets

La liste des énoncés à partir desquels les sujets devaient composer une commande à haute voix sont présentés dans le tableau D.1. Pour aider le lecteur à repérer l'information importante, les paramètres de la simulation sont identifiés par un soulignement, alors que les valeurs à insérer pour ces paramètres sont encadrées. Lors de l'expérimentation, les sujets ne bénéficiaient pas de cette emphase qui est présentée ici uniquement pour aider le lecteur.

À noter que pour certains énoncés, le paramètre n'est pas identifié (par exemple pour l'énoncé 22). Dans ces cas, le paramètre est implicite dans l'énoncé.

Tableau D.1: Liste des énoncés décrivant un besoin et à partir desquels le sujet devait composer un commande et la dire à haute voix

Énoncés
1. There is currently <u>no</u> <u>sound</u> , you need sound at a <u>maximum</u> level.
2. You need <u>aircraft mass</u> to be <u>2 540</u> .
3. You need <u>ice accumulation</u> to be <u>stopped</u> .
4. <u>Microburst</u> hasn't been used yet, you now need a microburst of intensity <u>15</u> .
5. You need the <u>hydrostatic mode</u> .
6. You need to be <u>positioned</u> at <u>short final</u> .
7. You need <u>oil temperature</u> to be at its <u>default value</u> .
<i>suite sur la page suivante</i>

Tableau D.1: Liste des énoncés décrivant un besoin et à partir desquels le sujet devait composer un commande et la dire à haute voix (suite)

Énoncés
8. You need <u>water</u> on the <u>runway</u> .
9. <u>Recirculation</u> hasn't been used yet, you need <u>80%</u> intensity of <u>sand</u> recirculation.
10. You need an <u>ECU Fail 2</u> <u>malfunction</u> .
11. You need the <u>cabin heater</u> <u>preset</u> for the last malfunction you just requested.
12. You need to <u>see</u> the <u>clouds</u> page on the IOS.
13. You need <u>fuel</u> to be <u>15 000</u> .
14. You need <u>airfield</u> to be <u>LIRF</u> .
15. You need <u>runway</u> to be <u>16 L</u> .
16. You need to be <u>aligned</u> for <u>takeoff</u> .
17. You need <u>flight</u> to be <u>unfrozen</u> .
18. You need <u>scattered</u> <u>clouds</u> ; <u>visibility</u> needs to be <u>0.5</u> ; <u>cloud top</u> needs to be <u>10 000</u> and <u>cloud base</u> <u>200</u> .
19. You need <u>wind speed</u> of <u>45</u> and <u>direction</u> <u>180</u> .
20. You need a <u>temperature</u> of <u>20</u> .
21. You need a <u>dry</u> <u>runway</u> .
22. You need to fly during <u>night</u> .
23. You need at <u>20 feet</u> the right engine to <u>flameout</u> .
24. You need to <u>slew</u> at <u>3</u> towards the <u>front direction</u> .
25. You need <u>all systems</u> to be at their <u>default value</u> .
26. You need <u>all temperatures</u> to be at their <u>default value</u> .
27. You need <u>no</u> <u>malfunction</u> .
28. You need to <u>see</u> the <u>fuel</u> page on the IOS.
29. You need <u>300 more</u> on <u>aircraft mass</u> .
30. You need <u>player 2</u> (Bell 412) to <u>head</u> <u>270</u> at a <u>speed</u> of <u>10</u> .

Questionnaire post-expérimentation

Thank you for filling the form below the best you can.

Age : 20-29 ____ 30-39 ____ 40-49 ____ 50-59 ____ 60+ ____

Number of years as instructor : ____

Last training session as instructor was (year) : ____

[_] Commercial or [_] Military ?

Aircraft types : _____

After doing these exercises, what are the tasks you would preferably accomplish using voice commands ?

After doing these exercises, what are the tasks you wouldn't accomplish using voice commands ?

Any comments about the visual feedback the IOS was providing ?

ANNEXE E

ENGINS DE RECONNAISSANCE VOCALE

Tableau E.1: Caractéristiques des engins de reconnaissance vocale analysés

Engin	Nom	Spécifications	Langage	Outils	Requis	Commentaires
A	Microsoft English Recognizer v5.1	Continuous, Speaker-dependent (10 minutes training)	C++, VB, C#	Grammar Compiler, Reco (Emulator)	Aucun	Gratuit
B	Microsoft English (U.S.) v6.1 Recognizer	Continuous, Speaker-dependent (10 minutes training)	ASP.NET	Debugging Console, Telephony	VS.NET, IIS, IE6 sp1, .NET Framework	Gratuit
C	Microsoft English (U.S. Telephony) v7.0 Server			Application Simulator, VS.NET Grammar Editor, VS.NET Embedded Emulator, VS.NET Pronunciation Editor	1.1 sp1, Microsoft Enterprise Instrumentation framework	

suite sur la page suivante

Tableau E.1: Caractéristiques des engins de reconnaissance vocale analysés (suite)

Engin	Nom	Spécifications	Langage	Outils	Requis	Commentaires
D	Nuance Vocon 3200	Continuous,	C(C++)	Audio Data	Python	\$\$
	EDS 2.5	Boost accu-		Collector, Sound	2.2.3, Jedit	
E	Nuance Vocon 3200	racy frequent		File Analyzer,	4.2	\$\$\$
	EDS 2.6	words,		Speech Verifier,		
		Phoneme-		User Dictionary		
		based,		Editor, Gram-		
		Speaker-		mar Creator		
		independent		Tool, Grammar		
				Compiler Tool,		
				Dictionary		
				Compiler Tool,		
				Spelling Tree		
				Compiler Tool,		
				Context Verifier		
				Tool, Vocabulary		
				Verifier Tool,		
				Confusability		
				Tool, Batch Tool,		
				Scoring Tool		
F	Sphinx 3.3.6	Continous,	C++		Aucun	Gratuit
		Speaker-				Difficulté
		independent				d'intégration
G	Nuance 9 Speech Re- cognizer	Speaker- independent, VoiceXML, W3C gram- mar format, Multilingue	Web	Vbuilder (Phrases vers GRXML)	Construction d'un ser- veur Web	\$\$\$

ANNEXE F

SCÉNARIO ET QUESTIONNAIRE UTILISÉS DANS L'ÉTUDE DES DISPOSITIFS D'ACTIVATION/DÉSACTIVATION

Scénario

Tableau F.1: Scénario suivi par les sujets testant les dispositifs d'activation/désactivation

#	Type d'interaction	But de la tâche
1	-Touch-	You need rescue winch fitted.
2	-Touch-	You need emergency floats fitted.
3	Voice	You need to be full of fuel.
4	Voice	You need fuel to be unfrozen.
5	Voice	You need AC weight to be 2800 kg.
6	-Touch-	You need reference airfield to be LIRF.
7	Notes =	You need to write down current time.
8	Role-play	"FLIGHT 341, are you ready to copy clearance?"
9	Voice	You need to fly during night.
10	Voice	You need cloud base to be 3000 feet.
11	Voice	You need cloud top to be 5000 feet.
12	Voice	You need visibility to be 15 km.
13	-Touch-	You need surface wind to be 240/25.
14	Notes =	You need to mark that pilot 1 is responding late.
15	Voice	You need sea state to be 5.
<i>suite sur la page suivante</i>		

Tableau F.1: Scénario suivi par les sujets testant les dispositifs d'activation/désactivation (suite)

#	Type d'interaction	But de la tâche
16	Voice	You need QNH to be 1013 mb.
17	Voice	You need temperature to be 10 degrees.
18	Role-play	"FLIGHT 341 is cleared to Fiumicino airfield, surface wind 240/25. Contact Departures 119.5 airborne passing 2000 feet."
19	-Touch-	You need to increase weight by 300 kg.
20	Voice	You need to decrease weight by 300 kg.
21	Notes =	You need to note that pilot 2 is having problem controlling stability.
22	Voice	You need to have 1200 kg of fuel.
23	Voice	You need the hoist cable to lock at position (malfunction)
24	Role-play	"FLIGHT 341 is cleared to commence handover procedures. Not below FL 250. Advise complete."
25	Voice	You need flight to be frozen.
26	-Touch-	You need all systems to be at their default value.
27	Voice	You need no malfunction.
28	Voice	You need to be positioned for takeoff.
29	Notes =	You need to write down that you are restarting the exercise.
30	Role-play	"FLIGHT 341 you are cleared for takeoff."

Questionnaire post-expérimentation

Check the appropriate answer for each characteristic.

Male ____ Female ____

Right-handed ____ Left-handed ____

Age : 20-29 ____ 30-39 ____ 40-49 ____ 50+ ____

Mark the activation devices based on ease of use and how you felt comfortable.

	Very easy		Can manage		Not easy
Activation word	☐	☐	☐	☐	☐
Soft button	☐	☐	☐	☐	☐
Chair switch	☐	☐	☐	☐	☐
Belt switch	☐	☐	☐	☐	☐
Foot switch	☐	☐	☐	☐	☐

Mark the activation devices based on effectiveness in the context of simulation training.

	Very easy		Can manage		Not easy
Activation word	☐	☐	☐	☐	☐
Soft button	☐	☐	☐	☐	☐
Chair switch	☐	☐	☐	☐	☐
Belt switch	☐	☐	☐	☐	☐
Foot switch	☐	☐	☐	☐	☐

Mark by 1 and 2 the activation devices you preferred first and second.

Activation word ____

Soft button ____

Chair switch ____

Belt switch ____

Foot switch ____

ANNEXE G**SCÉNARIOS SERVANT AUX TESTS
D'UTILISABILITÉ****Scenario 1**

Full internal fuel, AUM 2800 kgs

Shut down at Fiumicino Intl Rome (LIRF) - Runway 16L Ready for
takeoff

TOD : Night

Weather : Overcast base 3000 feet, tops 5000 feet, Visibility 15 km

Surface wind : 240/25

QNH 1013 mb

Temp : +10°C

N V G

Flight Freeze OFF

Running Engine flamout malfunction

End malfunction

Scenario 2

Create your own landing training scenario using the current ownship status