



Faculté de médecine Toulouse Rangueil

Enseignement des techniques de réadaptation

Mémoire présenté en vue de l'obtention du

Certificat de Capacité d'Orthophoniste

**EvoLex, logiciel de tests de fluence verbale :
de l'amélioration de la reconnaissance vocale au test
en situation écologique auprès d'orthophonistes.**

Florence FUGIER – FRESNE et Magali SEGUI – DE LAPASSE

Sous la direction de :

Xavier DE BOISSEZON et Yann TANNOU

27 juin 2016

Remerciements

Merci à nos maîtres de mémoire Yann Tannou et Xavier de Boissezon, pour leur disponibilité, leurs conseils avisés, et leurs encouragements.

Merci aux chercheurs de l'IRIT, Maxime Le Coz, Julien Pinquier et Jérôme Farinas pour leurs éclairages, réflexions, compétences. Un remerciement particulier à Maxime, dont le savoir-faire, la réactivité et le dynamisme nous ont beaucoup aidées.

Merci à Ophélie Balland et Aurélie Courtade-Jouanicq pour leur enthousiasme et leurs conseils lors du passage de flambeau.

Merci à Muriel Mescam pour ses explications claires et son aide précieuse et généreuse.

Merci aux orthophonistes qui ont pris le temps de répondre à notre questionnaire et de nous apporter leur œil clinique et critique, et à ceux qui ont accepté de tester le logiciel avec des patients.

Merci aux patients d'avoir participé aux tests et d'avoir contribué à enrichir notre clinique.

Merci à nos proches, amis, famille, voisins, amis d'amis, qui ont acceptés d'être enregistrés, pour leur contribution très appréciée.

Merci à nos maris et à nos enfants pour leur grande patience et leur précieux soutien pendant ces 4 années.

Et merci à toi ma binôme, pour ce chemin partagé, ta présence dans les bons et les mauvais moments, ton dynamisme et ta bonne humeur !

Sommaire

. Introduction	1
. Partie théorique	3
I. Reconnaissance vocale	3
1. Quelques repères : définition et historique.....	3
2. Particularité de la parole.....	5
3. Fonctionnement de la reconnaissance vocale	13
4. Qu'est ce qui fait la difficulté de la reconnaissance de la parole ?	15
5. Quelques facteurs importants dans la conception d'une application impliquant la reconnaissance vocale.....	18
II. Fluence verbale	20
1. Définition de la fluence verbale.....	20
2. Stock lexical	21
3. Activités cérébrales impliquées dans l'évocation lexicale	25
4. Focus sur le clustering et le switching	33
5. Pathologies concernées par des perturbations de la fluence verbale	37
6. Mesures de la fluence	47
7. Facteurs influençant les scores aux étalonnages	48
8. Tests de fluence existants.....	50
. Problématique et hypothèses	54
I. Problématique.....	54
II. Hypothèses.....	55
. Méthodologie	57
I. Présentation du logiciel.....	57
1. EvoLex « version 2015 ».....	57
2. EvoLex « version 2016 ».....	64
II. Méthodologie utilisée pour l'amélioration de la reconnaissance vocale.....	66
1. Enregistrements de tâches de fluence et de textes lus.....	66
2. Reconnaissance vocale et correction des fichiers	70

III.	Tests en situation écologique	73
1.	Choix d'un questionnaire préalable	73
2.	Test et démonstration du logiciel auprès d'orthophonistes	77
.	Présentation et analyses des résultats.....	86
I.	Présentation des résultats	86
1.	Résultats de l'amélioration de la reconnaissance vocale	86
2.	Résultats des questionnaires et des tests en situation écologique	92
II.	Analyses des résultats.....	104
1.	Analyse des résultats sur la reconnaissance vocale.....	104
2.	Analyse des résultats du questionnaire sur les avis a priori des orthophonistes sur le logiciel	107
3.	Analyse des résultats des tests en situation écologique	109
.	Validation des hypothèses et discussion.....	111
I.	Validation des hypothèses.....	111
1.	Sur l'amélioration de la reconnaissance vocale depuis juin 2015	111
2.	Sur l'avis a priori des orthophonistes sur le logiciel.....	112
3.	Sur l'utilisation du logiciel en situation écologique	112
II.	Discussion	113
1.	Amélioration de la reconnaissance vocale.....	113
2.	Enquête sur les habitudes des orthophonistes en matière d'informatique et de tests de fluence.....	117
3.	Tests en situation écologique	119
4.	Perspectives du logiciel.....	120
.	Conclusion.....	121
.	Bibliographie	123
.	Table des annexes	127
Annexe 1	Epreuves et étalonnage de Cardebat.....	128
Annexe 2	Lettre de recrutement des participants	131
Annexe 3	Protocole de passation.....	132
Annexe 4	Textes lus « La bise et le soleil » et extrait du « Petit Prince ».....	139

Annexe 5	Questionnaire (enquête sur les pratiques professionnelles en matière d'informatique et de tests de fluence)	140
Annexe 6	Questionnaire visant à recueillir l'avis des orthophonistes après démonstration du logiciel.....	144

EvoLex, logiciel de tests de fluence verbale :
de l'amélioration de la reconnaissance vocale
au test en situation écologique auprès d'orthophonistes.

Mémoire soutenu par Florence FUGIER – FRESNE et Magali SEGUI – DE LAPASSE

Sous la direction de Xavier DE BOISSEZON et Yann TANNOU

Faculté de médecine Toulouse-Rangueil, le 27 juin 2016

Introduction de la problématique

Les études de Troyer et al. en 1997 (citées par Gierski et Ergis, 2004) et d'autres chercheurs depuis, ont permis de montrer le rôle prépondérant des fonctions exécutives dans la réalisation des tâches de fluence, notamment la mise en place de stratégies de recherche, la flexibilité mentale ou la planification. Les tests de fluence verbale existants ne permettent pas d'analyser ces processus, c'est pourquoi le logiciel EvoLex a été créé : il permet, après avoir enregistré le patient sur une tâche de fluence, de transcrire ses productions, puis de les analyser, en termes de temps d'initiation et de latence, de fréquence de mots et de regroupements / commutations. Cependant, la reconnaissance vocale n'était pas assez efficace dans la version 2015 pour pouvoir proposer ces analyses de façon pertinente. D'autre part, nous souhaitions savoir si ce logiciel pouvait intéresser les orthophonistes et dans quelle mesure.

Méthodologie

Hypothèse 1 : La reconnaissance vocale a pu être améliorée par les enregistrements de sujets sains, les corrections de fichiers et les améliorations techniques apportées au logiciel. Afin d'améliorer la reconnaissance vocale, nous avons procédé à des enregistrements de sujets sains sur 4 tâches de fluence (2 sémantiques et 2 formelles) et la lecture de 2 textes. Pour vérifier le taux de reconnaissance vocale, nous avons vérifié la transcription automatique d'un certain nombre de fichiers, les avons corrigés, puis après une phase de réapprentissage, en avons recorrecté afin d'en mesurer les effets. Nous avons fait de même sur des fichiers enregistrés avec des patients.

Hypothèse 2 : ce type de logiciel est susceptible d'intéresser les orthophonistes. Afin d'y répondre, nous avons rédigé un questionnaire - envoyé via un site internet dédié - pour connaître l'avis a priori des orthophonistes sur ce type d'outil, et par là-même, savoir quelles sont les habitudes d'utilisation des professionnels en matière d'informatique et de tests de fluence.

Hypothèse 3 : en situation écologique, le logiciel est pertinent et efficace, et apporte une plus-value par rapport aux passations des tests de fluence manuelles. Pour corroborer à cette hypothèse, nous avons procédé à une phase de pratique en situation écologique avec des patients et de présentation du logiciel auprès d'orthophonistes. Suite aux passations d'une épreuve de fluence sémantique (Animaux) et d'une de fluence formelle (Mots commençant par V) par les patients, nous avons recueilli l'avis des orthophonistes grâce à un questionnaire.

Résultats

Concernant la reconnaissance vocale, nous avons enregistré 86 sujets sains, ce qui a permis de fournir à la base de données du serveur 344 fichiers audio. La correction de 178 fichiers a permis de montrer que la reconnaissance vocale est passée de 30% (résultats des tests effectués lors de l'étude de faisabilité du logiciel en 2014-2015) à 63,41%. Cependant, les tests réalisés avec 10 patients ont montré de moins bons résultats: 59,66% pour les animaux et 28,02% pour les mots commençant par V.

Grâce au questionnaire envoyé en ligne, nous avons reçu 181 réponses d'orthophonistes, ce qui nous a permis de mettre en évidence leur intérêt pour les analyses quantitatives, mais surtout qualitatives (temps de latence, d'initiation, fréquence de mots, regroupements et commutations d'un regroupement à l'autre), comme l'ont signifié 66,85% d'entre eux.

Enfin, les tests réalisés avec 10 patients en situation écologique, et les démonstrations aux 10 orthophonistes ont permis de montrer que les analyses automatiques et chiffrées présentaient une réelle plus-value, surtout dans la perspective d'un logiciel intégrant également des tâches de dénomination et de génération de mots. L'amélioration de la reconnaissance vocale restant une condition à cet engouement.

Discussion et conclusion

Ainsi, nous avons pu en partie valider la 1^{ère} hypothèse, car les enregistrements de sujets sains et les corrections de fichiers ont permis une amélioration de la reconnaissance vocale. La sous-hypothèse concernant l'utilisation des lectures de texte pour l'amélioration de la

reconnaissance vocale n'a pas été validée car elles n'ont pas été efficaces dans l'amélioration du système. Concernant les enregistrements des patients, nous n'avons pas pu valider que la reconnaissance vocale fonctionnait aussi bien qu'avec les sujets sains.

La 2^{ème} hypothèse a pu être validée, et montre ainsi l'utilité de ce type d'outil dans la pratique clinique orthophonique.

La 3^{ème} hypothèse a été validée concernant la pertinence du logiciel, mais nous ne pouvons pas valider la sous-hypothèse concernant le gain de temps, les résultats n'étaient pas chiffrables de façon statistique, même si les avis sont plutôt positifs sur ce sujet également.

L'analyse de nos résultats a mis en perspective des limites à nos travaux. En effet, malgré le soin apporté aux enregistrements de sujets sains, et même si le taux de reconnaissance vocale a plus que doublé entre les versions 2015 et 2016 du logiciel, il n'est pas encore assez performant ; le gain de temps dépend fortement du nombre de mots reconnus correctement automatiquement par EvoLex. Des améliorations sont encore à apporter à ce logiciel, néanmoins, ce mémoire a montré que le prototype est déjà beaucoup plus fonctionnel que l'année dernière et nul doute qu'il deviendra à terme un outil performant.

Bibliographie

Gierski F., Ergis A.-M., Les fluences verbales : aspects théoriques et nouvelles approches. In L'année psychologique volume 104, 2004 : 331-353

Godefroy O. et le GREFEX, Fonctions exécutives et pathologies neurologiques et psychiatriques – évaluation en pratique clinique, Marseille : Solal, 2008

Haton J-P, Fohr D., Laprie Y., Cerisara C. Reconnaissance automatique de la parole - Du signal à son interprétation, Paris : Dunod, 2006

Mariani J., Reconnaissance de la parole, Traitement automatique du langage parlé, Hermes Science Publications, 2002

Seron X., Van Der Linden M., Traité de neuropsychologie clinique de l'adulte, Tome 1 – Evaluation. Louvain-La-Neuve (Belgique) : De Boeck, 2014

. Introduction

La tâche de fluence est fréquemment utilisée dans la pratique orthophonique, car elle donne rapidement une indication sensible sur les capacités langagières et surtout sur les fonctions exécutives du patient. Le projet de création d'un logiciel de reconnaissance vocale adapté aux tâches de fluence verbale (EvoLex) est né de plusieurs éléments. D'abord, à l'ère du numérique, le thérapeute a la possibilité de se dégager de contraintes matérielles et d'être ainsi plus disponible pour l'observation clinique. D'un autre côté, les résultats étalonnés apportés par les épreuves existantes de fluence verbale restent jusqu'à présent purement quantitatifs. Or, des recherches et études menées entre autres par Troyer et al. en 1997 (cités par Gierski et Ergis, 2004) ont mis en évidence des processus cognitifs et exécutifs particuliers, dont l'analyse donne des éléments intéressants sur le **fonctionnement cognitif du patient**.

Un logiciel comme EvoLex permettrait donc à la fois d'apporter des données quantitatives très précises et surtout des données qualitatives permettant de dégager un profil cognitif.

Ainsi, en 2014-2015, deux étudiantes en orthophonie, Ophélie Balland et Aurélie Courtade-Jouanicq, encadrées par Messieurs Xavier de Boissezon (Professeur en médecine physique et de réadaptation au C.H.U. de Rangueil, Toulouse) et Yann Tannou (orthophoniste à Muret), ont étudié la faisabilité d'un tel logiciel dans le cadre de leur mémoire de fin d'études. Lors de ce projet, elles ont travaillé en partenariat avec une équipe de l'IRIT (Institut de Recherche en Informatique de Toulouse, Université Toulouse III – Paul Sabatier) composée d'enseignants-chercheurs en informatique et d'étudiants, qui a permis la mise au point de l'outil informatique (interface, programmation, création de bases de données, premier travail sur la reconnaissance vocale, etc.). Elles ont notamment participé à la création de la maquette du logiciel. Elles ont constitué une base de données en alphabet phonétique informatique pour créer le dictionnaire indispensable au système de reconnaissance vocale. Elles ont également créé une base pour les analyses qualitatives en établissant des listes de mots fréquents, des regroupements (sémantiques, phonologiques, morphologiques et grammaticaux). Enfin, elles ont testé le logiciel auprès de sujets sains (45) et de patients (5) et ont présenté une étude de cas.

Avec ce mémoire, intitulé « Etude de faisabilité d'un logiciel de reconnaissance vocale adapté à des tâches d'évocation lexicale », soutenu en juin 2015, Ophélie Balland et Aurélie

Courtade-Jouanicq ont pu **démontrer la faisabilité du projet**, et l'intérêt de poursuivre les recherches en ce sens.

En revanche, elles ont noté plusieurs aspects techniques qui limitaient nettement la performance du logiciel. En premier lieu, la reconnaissance vocale ne s'est pas révélée assez efficiente ; il a ainsi été montré que seulement près de 30% des mots dits par le sujet étaient reconnus d'emblée par le logiciel. Cela s'avère insuffisant pour assurer des analyses quantitatives ou qualitatives pertinentes de façon automatique, et assurer un gain de temps à l'orthophoniste. De plus la transmission des fichiers audio et leur analyse n'étaient pas automatisées, la transcription ne pouvait être réalisée qu'en différé. Enfin, leurs dernières réserves concernaient la partie « Analyses des stratégies » qu'elles jugent perfectibles.

Leur conclusion était donc que le logiciel pourra aider l'orthophoniste faisant passer des tâches de fluences à ses patients, dans la mesure où, après des améliorations nécessaires (notamment la performance du système de reconnaissance vocale), il permettra de fournir des analyses quantitatives et qualitatives montrant les processus cognitifs en jeu chez le patient.

Dans la continuité de ce mémoire, nous avons **contribué à perfectionner le logiciel**, en travaillant sur l'amélioration de la reconnaissance vocale. D'autre part, nous avons souhaité connaître l'**avis a priori des orthophonistes** sur un logiciel comme EvoLex, et pour cela, nous avons fait une étude de l'utilisation de l'informatique et des tâches de fluence dans leur pratique. Enfin, nous avons **testé le logiciel en situation réelle**, pour recueillir l'avis des orthophonistes sur celui-ci.

Afin d'asseoir les fondements théoriques sous-tendus par l'élaboration du logiciel, nous présentons dans une première partie, les recherches et les données de la littérature sur la reconnaissance vocale et sur la fluence verbale. Puis, nous exposons notre problématique et nos hypothèses. Nous expliquons ensuite les méthodes mises au point pour valider nos hypothèses, nous présentons les résultats obtenus, et enfin, nous discutons ces résultats et validons nos hypothèses.

. Partie théorique

I. Reconnaissance vocale

Dans la mesure où l'amélioration du logiciel nécessitait au préalable l'amélioration de la reconnaissance vocale, il nous a paru nécessaire et intéressant de faire un point sur la question.

Si l'on interroge le grand public, il ressort que la reconnaissance vocale (ou reconnaissance automatique de la parole) est entrée dans le quotidien de beaucoup de personnes, que ce soit au téléphone quand une voix automatique enregistre nos réponses, sur notre GPS qui comprend où l'on veut se rendre, ou sur notre portable qui enregistre le sms que l'on dicte... Il est vrai que d'immenses progrès ont été réalisés surtout ces dernières décennies et que de nombreuses applications ont vu le jour. Cette reconnaissance semble donc fonctionner parfaitement dans l'esprit des gens. Mais si l'on creuse un peu la question, quand le dialogue s'avère plus complexe, que l'énonciation n'est pas très claire ou qu'il y a trop de bruit, bref quand les conditions ne sont pas stables et idéales... les interprétations par l'ordinateur sont beaucoup plus difficiles. De nombreux chercheurs travaillent encore sur la question pour améliorer la fiabilité et la robustesse des systèmes de reconnaissance vocale. Le vrai dialogue homme-machine n'est pas encore pour tout de suite... même si des robots japonais comme « Pepper » pourraient le laisser imaginer !

1. Quelques repères : définition et historique

1.1 Définition de la reconnaissance vocale

Commençons par préciser ce que l'on entend par reconnaissance vocale ou traitement automatique de la parole. C'est une technique informatique qui vise à reconnaître, dans une suite de signaux sonores, les mots prononcés par un locuteur et à les transcrire en texte exploitable par une machine. Elle appartient à la branche du traitement de la parole dans lequel on retrouve notamment le codage et la compression de la parole, la synthèse vocale (production d'un signal vocal), la reconnaissance du locuteur. De nombreuses disciplines s'y

entrecroisent tels que l'acoustique, l'électronique, la phonétique, la linguistique, la sémiologie, etc.

Depuis une trentaine d'années, des applications commerciales utilisant la reconnaissance vocale, de plus en plus nombreuses, voient le jour. Dans un premier temps, elles ne concernaient que la reconnaissance de mots isolés et/ou enchaînés, avant d'être capables de traiter des phrases en continu. Elles permettent aujourd'hui de réaliser des saisies de données, des commandes vocales, des transcriptions, de rechercher des informations, etc. Et on les retrouve principalement dans le vaste domaine des télécommunications mais également dans des domaines très variés comme les domaines bancaire, militaire, aéronautique, logistique, ou médical pour la partie qui nous intéresse.

1.2 Bref historique

Si l'on brosse un tableau rapide de l'évolution du traitement de la parole, on constate qu'elle est fortement liée à l'évolution de l'informatique et de l'électronique. L'augmentation des capacités de calcul et de stockage, la miniaturisation des composants, ont largement contribué à l'amélioration des systèmes. En 1952, Davis et al (cités par Haton, 2006) mettent au point un 1^{er} système qui permet de reconnaître des chiffres isolés, pour un locuteur, par un dispositif électronique câblé. Dans les années 60, le passage au numérique sur ordinateur entraîne une grande avancée technologique. De nombreuses recherches ont lieu dans les années 70 et les premières commercialisations voient le jour. En 1996 ce sont les premières machines à dicter en parole continue avec par exemple les systèmes, aujourd'hui bien connus, des sociétés IBM, Dragon... Les recherches s'orientent désormais principalement vers l'amélioration de la fiabilité et de la robustesse des systèmes notamment au travers de travaux sur les réseaux neuronaux et l'intelligence artificielle.

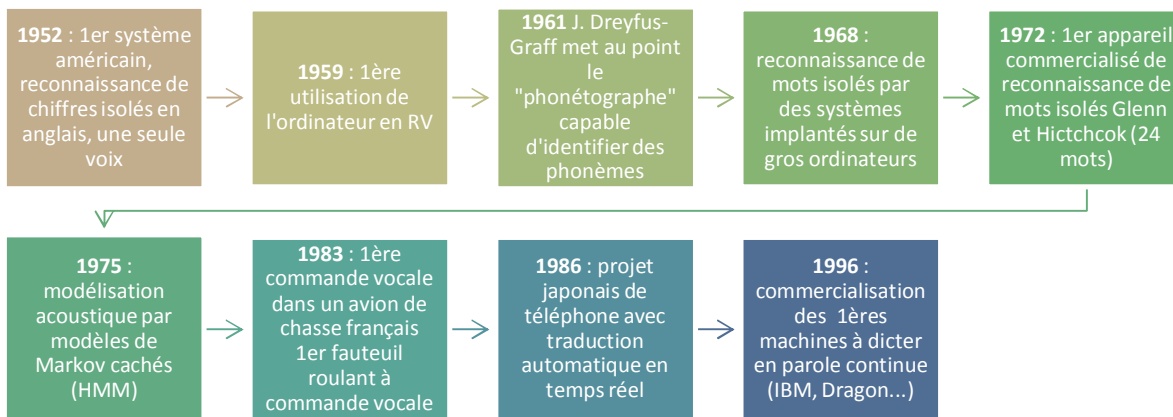


Figure 1 Quelques dates

2. Particularité de la parole

On peut se demander ce qui rend la parole si difficile à décoder ? Rappelons d'abord les caractéristiques de la parole, propre de l'humain.

2.1 Production de la parole

2.1.1 Principes

D'un point de vue **acoustique**, « *les ondes sonores sont des propagations de changement de pression, produits par les vibrations de particules du milieu ambiant : l'air atmosphérique pour les humains, l'eau pour les poissons* ». Le signal acoustique de parole résulte des variations de pression de l'air au niveau de la bouche d'un locuteur. Ce signal comporte à la fois des éléments linguistiques et des éléments suprasegmentaux comme la voix, la prosodie qui renseignent sur l'identité du locuteur, son état... R. Haton (2006) résume ainsi la chaîne de communication parlée :

- **Intention** : le locuteur transforme l'idée qu'il désire transmettre en une structure linguistique formée des mots adéquats
- **Articulation** : la structure linguistique est transformée en une suite de commandes motrices pour les organes articulateurs

- Emission : l'onde vocale est produite par le mouvement des muscles articulateurs et se propage dans l'air depuis la bouche du locuteur.
- Perception : l'onde vocale parvient aux oreilles d'un auditeur puis est transformée en impulsions nerveuses transmises au cerveau de façon à être comprise par l'auditeur.

Physiologiquement, la parole est le résultat d'un processus qui met donc en jeu une programmation (cerveau), une soufflerie (poumons), un phénomène de vibration (larynx), différents articulateurs (langue, lèvres, voile du palais, mâchoire) et des cavités de résonance.

La parole se fait normalement sur l'expiration. La pression de l'air provenant des poumons franchit le larynx et met en vibration (ou pas) les cordes vocales (fréquence fondamentale) avant de venir ensuite résonner et s'enrichir en harmoniques dans les cavités pharyngée, buccale, nasale... Un son est alors émis. La fréquence fondamentale varie en fonction de l'épaisseur, de la longueur et de la tension des cordes vocales. Elle est donc différente pour les enfants, les femmes et les hommes. Il en est de même pour les harmoniques qui dépendent notamment de l'anatomie des cavités de résonance.

En résumé, à une fréquence fondamentale notée F_0 qui est un signal périodique s'ajoutent des fréquences de résonance ou harmoniques notées $F_1 F_2 F_3...$ (appelées formants).

« La transformée de Fourier [...] permet de décomposer une onde, aussi complexe soit-elle, en une suite d'ondes élémentaires sinusoïdales différant par leurs fréquences, amplitudes et phases. » (Vaissière, 2011). On peut représenter une onde par un schéma temporel :

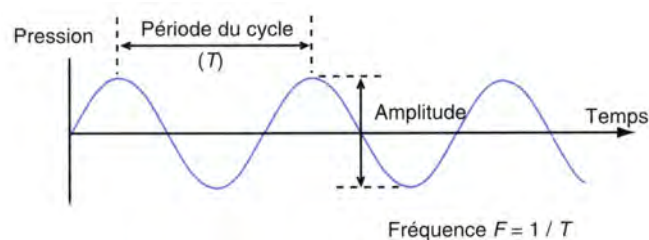


Figure 2 Onde sonore

On peut distinguer différentes propriétés acoustiques pour caractériser les sons (Vaissière, 2011) :

- La durée

- La fréquence fondamentale, notée F_0 (liée à la hauteur perçue, se mesure en Hz) et la forme du signal de source (liée à la qualité de la voix)
- L'intensité physique (en dB)
- La composition spectrale, en relation avec la répartition de l'énergie dans l'échelle des fréquences (formants pour les voyelles et la distribution de l'énergie dans l'échelle pour les bruits)
- L'aspect stationnaire ou dynamique (diphthongues, transitions...)

La parole est une onde sonore dont la particularité est d'être produite par un conduit vocal humain, contrairement aux « bruits » de la nature. Les sons de parole sont donc le produit d'un filtrage par un conduit vocal humain et sont interprétés comme tels par les auditeurs (Vaissière, 2011).

2.1.2 Les sons élémentaires du français

Ils peuvent se caractériser de 3 manières : par le voisement (vibration ou non des cordes vocales), le mode d'articulation et le lieu d'articulation. L'unité élémentaire est le **phonème**. C'est une unité abstraite qui peut se définir à travers l'idée de « paire minimale » : deux sons correspondent à deux phonèmes distincts s'ils peuvent définir une paire minimale dans la langue, comme dans « pain » et « main ». Le phonème est réalisé acoustiquement sous forme de phone. Les systèmes de reconnaissance de parole utilisent le phonème sous forme de phones, dipphones, triphones ou pentaphones. Ils peuvent également utiliser l'unité de la syllabe. Un phonème ne possède pas de traits invariants suffisants pour être utilisé tel quel.

Il existe trois types de phonèmes en français : les voyelles, les consonnes, les semi-voyelles.

2.1.3 Les voyelles

« Les voyelles correspondent aux sons du langage humain produits avec un libre passage de l'air dans les cavités supra-glottiques. Celles-ci constituent des cavités de résonance dont la configuration géométrique va déterminer des fréquences de résonance. La source laryngée est filtrée et donne naissance aux voyelles. » (Marchal, 2011)

Au nombre de 16 en français, elles peuvent être classées selon :

- leur nasalité : on parle de voyelles orales ou nasales, selon si le voile du palais est élevé et laisse passer l'air uniquement par la bouche (voyelles orales) ou s'il est abaissé et laisse passer l'air par les fosses nasales et par la bouche (voyelles nasales). Ainsi le [o] se distingue du [ɔ̃] dans les mots pot ou pont.

- leur lieu d'articulation : il correspond à la zone d'élévation maximale de la langue par rapport au palais. On distingue ainsi des voyelles antérieures, centrales, postérieures.
- leur aperture : voyelles ouvertes ou fermées
- la position et la forme des lèvres : voyelles étirées, [i], projetées [y] ou arrondies [u]

		Lieu d'articulation				
		antérieure		post.		
aperture	petite (fermée)	i	y	u	nasalité	
	Moyenne	e	ø	o		-
		mi-ouverte	ɛ	œ		ɔ
	grande (ouverte)	~ɛ	~œ	~ɔ		-
		a		ɑ		+
			~ɑ			
		-	+		labialité	

Figure 3 Tableau articulatoire des voyelles du français (Marchal, 2011)

Caractéristique acoustique de la voyelle

Le spectre fréquentiel d'une voyelle présente un ensemble de pics, ou formants, caractéristiques de cette voyelle. Elle peut être définie acoustiquement par la fréquence de ses 2 premiers formants F1 et F2, le 3^{ème} étant beaucoup moins discriminant. La nasalité introduit un déséquilibre, le 1^{er} formant perd de son intensité et modifie le spectre. R. Haton précise qu'en pratique, les caractéristiques acoustiques de la voyelle sont très variables. Cette variabilité provient principalement :

- de la physiologie du conduit vocal (différence homme, femme, enfant, notamment)
- de la coarticulation qui fait qu'en fonction du phonème qui suit ou précède la voyelle, elle sera prononcée différemment
- et de la latitude de réalisation dont dispose le locuteur sur le plan linguistique

2.1.4 Les consonnes

« Les consonnes correspondent aux sons de la langue dont la production est caractérisée par la présence d'un obstacle au libre écoulement de l'air dans les cavités supra-

laryngées. » (Haton, 2006) Elles ne peuvent exister seules et exigent la présence d'une voyelle.

On distingue deux types de consonnes selon si l'écoulement de l'air rencontre un obstacle partiel ou total :

- Les fricatives ou constrictives : le barrage à l'écoulement de l'air est partiel. Il existe trois sourdes [f][s][ʃ] et trois sonores [v][z][ʒ].
- Les occlusives : le barrage à l'écoulement de l'air est total. Elles peuvent être orales, nasales, voisées ou sourdes. On distingue ainsi des occlusives orales voisées [b][d][g], orales sourdes [p][t][k] et nasales voisées [m][n][ɲ][ŋ]

Ce sont des sons transitoires résultant de l'ouverture brusque du conduit vocal après son obstruction. Trois phases peuvent être observées :

- Un silence correspondant à l'occlusion complète du conduit vocal, sauf pour les occlusives voisées : la vibration des cordes vocales entraînent l'émission d'une faible énergie en très basses fréquences
- Une explosion résultant du relâchement de l'air comprimé par l'occlusion
- Un ensemble de transitions formantiques vers le son vocalique suivant

Les consonnes liquides [l] et [r] : elles proviennent d'une légère constriction du conduit vocal. Les réalisations acoustiques peuvent être très variées en fonction des locuteurs.

Mode d'articulation		Lieu d'articulation					
		Bi-labiale	Labio-dentale	Apico-dentale	Dorso-palatale	Dorso-vélaire	Dorso-uvulaire
Occlusive	Sourde	p		t		k	
	Nasale						
Sonore	Orale	b		d		g	
	Nasale	m		n		ɲ	
Constrictive							
Sourde	Orale		f	s		ʃ	
Sonore	Orale		v	z		ʒ	
Liquide				l			R

Figure 4 Tableau articuloire des consonnes (site <http://flenet.unileon.es/>)

2.1.5 Les semi-voyelles ou semi-consonnes

Elles sont au nombre de trois : [j][w][ɥ]. Et comme leur nom l'indique, elles se trouvent à la frontière entre la voyelle et la consonne. Elles se distinguent de voyelles par leur position non centrale dans la syllabe.

2.1.6 Le spectrogramme

Le spectrogramme permet de visualiser les caractéristiques acoustiques de la parole en trois dimensions ; les différentes fréquences de l'extrait sonore sont représentées en fonction de la durée. L'axe horizontal représente le temps, l'axe vertical les fréquences et le degré de noirceur (pour les représentations en noir et blanc), l'intensité.

Ainsi, sur un spectrogramme (Vaissière, 2011) :

- les sons voisés se repèrent grâce à la présence d'une barre de voisement dans les très basses fréquences et par la détection automatique de fréquence fondamentale ([b][d][g][v][z][ʒ]) les voyelles présentent des formants dans les basses et les moyennes fréquences, un sommet local d'énergie, et une barre de voisement
- les occlusives sourdes [p][t][k] et sonores [b][d][g] montrent une absence d'énergie dans les moyennes et hautes fréquences
- les fricatives sourdes [f][s][ʃ] et sonores [v][z][ʒ].se caractérisent par la présence d'un bruit continu
- les sonantes montrent des formants mais moins amples que les voyelles

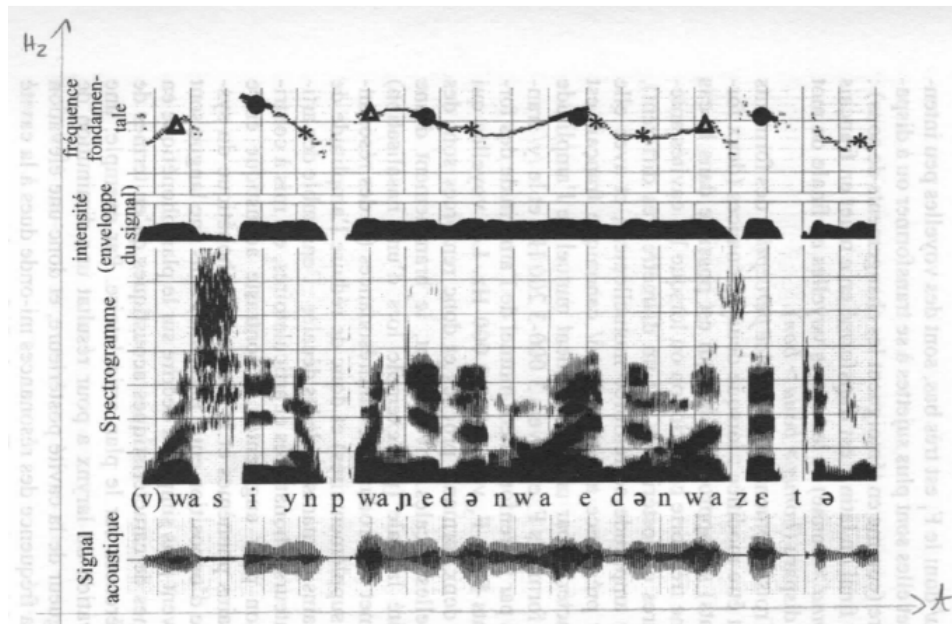


Figure 5 Courbe de fréquence fondamentale, enveloppe d'intensité, spectrogramme, transcription phonologique et signal du début de phrase « Voici une poignée de noix et de noisette... » (Vaissière, 2011)

Comme nous l'expliquerons plus loin de manière un peu plus détaillée, la parole est un signal extrêmement variable (variations intra-locuteurs, inter-locuteur...), et c'est bien toute l'habileté de l'oreille humaine que la machine essaie d'atteindre.

2.2 Perception de la parole

Un son est donc une sensation auditive engendrée par une onde acoustique générée par une vibration.

« *Tous les sons de la parole et bruits de la nature empruntent les mêmes voies auditives pour arriver au cerveau de l'auditeur, mais les tests psycho-acoustiques montrent que les sons de la parole ne sont pas perçus exactement de la même manière que les bruits.* » (Vaissière, 2011). C'est justement cette tâche de différenciation bruit/parole qui est complexe pour l'ordinateur. Contrairement à l'oreille humaine, la machine différencie difficilement le bruit de la parole (d'où la nécessité de très bonnes conditions d'enregistrement).

2.2.1 L'audition

Le système auditif humain a deux fonctions principales. Il permet d'une part de reconnaître un son, et de lui attribuer une signification. D'autre part de localiser la provenance de ce son

(grâce à l'analyse de la différence de niveau sonore et du temps d'arrivée de l'onde sonore (ou phase) entre les deux oreilles).

On peut distinguer deux systèmes auditifs (Haton, 2006) : un système périphérique (de l'oreille externe au nerf auditif) et un système central (des premiers neurones jusqu'au cortex).

➤ Anatomie et physiologie de l'oreille (site www.cochlea.eu) :

- L'oreille externe

Elle sert d'antenne acoustique en amplifiant et en transmettant le son à l'oreille moyenne. Le pavillon diffracte les ondes, pendant que le conduit auditif externe et la conque jouent un rôle de résonateur.

- L'oreille moyenne

Elle est principalement constituée d'un système tympano-ossiculaire qui a pour fonction de transmettre l'onde sonore à l'oreille interne. Il transforme pour cela des vibrations sonores aériennes en variations de pression dans les compartiments liquidiens de l'oreille interne (adaptation d'impédance).

- L'oreille interne

Elle se compose de la cochlée, zone auditive de l'oreille interne, et du vestibule, zone de l'équilibre. La cochlée communique avec l'oreille moyenne par deux membranes appelées fenêtre ronde et fenêtre ovale. Une onde est transmise de la base à l'apex cochléaire. En fonction de la fréquence du son, l'onde n'est pas propagée jusqu'au même endroit (on parle de tonotopie cochléaire). Il est intéressant de noter que l'oreille répond de façon sélective en fonction de la fréquence d'un son.

L'onde induit ensuite un mouvement des cellules ciliées qui se transforme en signal électrique conduit par le nerf auditif jusqu'au cortex auditif, où il y a alors interprétation du son.

L'homme perçoit généralement les sons entre un seuil minimum de 0 à 40 dB jusqu'à un seuil de douleur de 120 dB. Les fréquences se répartissent du grave (fréquences basses) à l'aigu (fréquences hautes) et sont perçues entre 20 et 20 000 Hz.

Notons qu'avec le vieillissement, la perception diminue en précision, on assiste à un étalement des fréquences, contrairement à une machine qui conserve ses performances.

Chez l'homme, la perception de la parole correspond à un décodage du message vocal à partir de la représentation neuronale des sons. On peut distinguer trois grandes fonctions (Haton, 2006) :

- une discrimination de la parole par rapport au bruit même dans un environnement sonore intense
- la robustesse du décodage aux dégradations du signal
- et une invariance du décodage malgré de multiples variations inter-locuteurs ou, intra-locuteurs.

Plusieurs étapes sont donc impliquées dans la perception de la parole : auditive, phonétique, phonologique, lexicale, syntaxique et sémantique. Et même si différentes théories ont été avancées (ex. de la théorie motrice), différents modèles proposés (ex. modèle actif, passif) la perception de la parole demeure plus difficile à appréhender que sa production.

Toujours est-il, comme l'explique R. Haton (2006), que ces données et ces connaissances en matière de production, d'audition et de perception de la parole sont mises à profit dans la conception d'un système de reconnaissance de parole. Et qu'une meilleure connaissance des processus perceptifs et cognitifs mis en œuvre dans le cerveau contribuera sans doute beaucoup à améliorer les performances des systèmes automatiques.

3. Fonctionnement de la reconnaissance vocale

3.1 Principes

Pour enregistrer une phrase ou des mots, on utilise un microphone qui convertit le signal acoustique en signal électrique. Ce signal est ensuite numérisé et le bruit pertinent est séparé du bruit ambiant. Puis, il est paramétré par une technique d'analyse fréquentielle.

Un apprentissage automatique réalise une association entre les segments élémentaires de parole et les éléments lexicaux en faisant appel à une modélisation statistique (Modèles de Markov Cachés) et/ou par réseaux de neurones artificiels (ANN ou Artificial Neural Networks).

Enfin, le décodage ou reconnaissance se fait à partir d'algorithmes.

Pour résumer, nous pourrions reprendre le schéma ci-dessous utilisé par le site <https://synthesevoix.wordpress.com> :

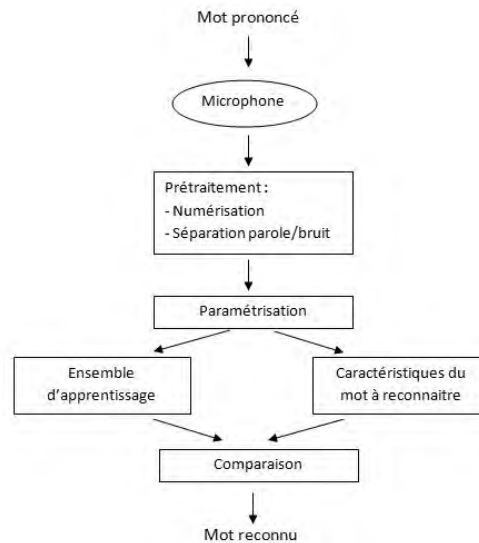


Figure 6 Principe de la reconnaissance vocale

3.2 Modèles

Ce système de reconnaissance s'appuie sur trois modèles principaux :

- Un modèle acoustique décrivant les entités à reconnaître et capable à partir du signal acoustique de donner la probabilité que le signal corresponde à chacun des phonèmes possibles de la langue cible
- Un modèle de prononciation ou lexique codant les mots du vocabulaire et donnant ainsi pour chaque mot la ou les prononciations possibles (au niveau phonétique avec des probabilités associées)
- Un modèle de langage décrivant la structure des phrases du langage et donnant pour chaque suite de mots sa probabilité dans le langage cible.

C'est la combinaison des 3 modèles qui permet de calculer les probabilités de chaque suite de mot en fonction du signal enregistré. La reconnaissance vocale correspond à la probabilité maximale.

3.3 Apprentissage et la reconnaissance de formes

« Une machine, tout comme l'être humain, ne peut reconnaître que les formes qu'elle a préalablement apprises. La qualité de l'apprentissage est cruciale pour les performances d'un système, indépendamment du modèle de décision choisi. » (Haton, 2006). Il y a donc d'abord une importante phase d'apprentissage. Les différentes entités sont mémorisées par le système sous diverses formes selon le type de reconnaissseur. Cela peut être des mots, ou pour des applications plus complexes, des unités phonétiques (phonèmes, demi-syllabes, syllabes, etc.). EvoLex fonctionne avec une reconnaissance de triphones. La référence est représentée par un modèle (Modèles de Markov Cachés) et conservée en mémoire. Il y a ensuite comparaison entre ce qui est dit et ce qui est mémorisé puis évaluation de la distance et décision sur la validation du mot.

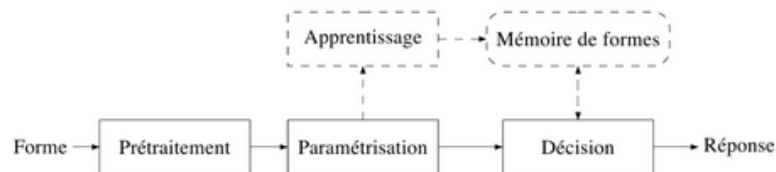


Figure 7 Schéma de principe d'un système de reconnaissance de formes (Haton, 2006)

4. Qu'est ce qui fait la difficulté de la reconnaissance de la parole ?

4.1 Quelques paramètres

Selon Mariani (1990), un ensemble de paramètres rendent la parole difficile à décoder :

- L'absence de séparateurs, de silences entre les mots, comparables aux blancs dans le langage écrit. Dans notre projet de logiciel de fluence, il s'agit de mots et non de parole, mais rien n'empêche le sujet de commenter ses productions entre deux énonciations de mots et d'autre part, le sujet peut prononcer plusieurs mots rapidement à la suite, sans pause.
- La grande variabilité de la parole :

- *Intra-locuteur* : inhérente au processus humain de production : variation du débit, modifications dues à l'émotion, au stress, à l'état de santé (voix enrourée, dysphonique, éraillée, soufflée), bégaiement...
 - *Inter-locuteurs* : timbres différents, voix masculines, féminines, voix d'enfants, d'adultes, de personnes âgées, accents régionaux, habitudes linguistiques...
- L'environnement acoustique qui est très important.
- La présence de bruit est néfaste. Cela complexifie la reconnaissance et peut engendrer une augmentation de la variabilité intra-locuteur par l'effet Lombard (dans une ambiance très bruitée, le locuteur a le réflexe de modifier sa voix) (Calliope, 1989).
 - Les conditions d'enregistrement d'autre part peuvent induire une variabilité, ne serait-ce que par l'utilisation du microphone, son type, son installation... D'où la très grande attention que nous avons portée lors de nos enregistrements à avoir des conditions les meilleures possibles - avec une pièce calme, le même microphone si possible dans la même position (au niveau du menton)... Sachant qu'il faudra que cela fonctionne aussi dans des conditions moins facilitatrices et plus écologiques.
- La coarticulation : cela rejoint la notion d'invariance du phonème. Un phonème n'est pas réalisé exactement de la même manière selon le phonème qui le précède ou qui lui succède. Le mouvement articulaire sera différent. Le [p] du mot « partir » n'est pas prononcé de la même manière que celui dans le mot « optimal ».
- La taille du vocabulaire : il est plus difficile de travailler avec un vocabulaire très étendu, comme dans la parole naturelle qu'avec un vocabulaire limité, comme dans notre projet pour lequel un dictionnaire a été élaboré recensant tous les noms d'animaux, de fruits, et les mots commençant par « R » et par « V ».
- La proximité phonétique des mots : plus les mots sont proches phonétiquement, plus leur différenciation sera complexe.
- Le langage : concerne la syntaxe, la sémantique voire la pragmatique. Les modèles de langage aident à prendre les meilleures décisions. Le seul aspect qui concerne EvoLex regarde les articles : un modèle a été créé pour prendre en compte « un », « une », « le », « la », « les » et « des ».

4.2 Différents types d'erreurs possibles

Toutes ces difficultés font que des erreurs peuvent apparaître lors de la transcription du message (Mariani, 1990). On note différents types d'erreurs :

- Erreur de substitution : un mot est reconnu à la place d'un autre
- Erreur d'insertion : un mot est reconnu alors que rien n'a été prononcé
- Erreur d'omission : rien n'est reconnu alors qu'un mot est prononcé

Pour mesurer ce taux d'erreur, deux mesures sont utilisées. Le « pourcentage de mots correct » qui mesure combien de mots ont été correctement reconnus, sans tenir compte des erreurs d'insertion. Et « l'acuité lexicale » qui prend en compte les trois types d'erreurs et peut donc conduire parfois à un score négatif.

4.3 Robustesse de la reconnaissance de la parole

Mariani (1990) conclut sur le fait que la qualité atteinte par les systèmes de traitement automatique de la parole est convenable mais que leur utilisation est « relative ». Si dans des conditions de « laboratoire », cela fonctionne bien, c'est beaucoup plus difficile dans des conditions réelles d'utilisation, avec des locuteurs réels. Haton (2006) souligne également que « *lorsque les conditions d'utilisation d'un système de reconnaissance diffèrent des conditions d'apprentissage, les taux de reconnaissance chutent fortement* ».

« Les auditeurs humains sont beaucoup moins sensibles à ces différences, ce qui laisse penser qu'il est possible d'améliorer les systèmes automatiques de ce point de vue, notamment en ce qui concerne le bruit environnemental (ou bruit additif), mais également les phénomènes qui modifient le signal émis (réverbération, canal de transmission, direction de propagation, etc.) ou les différences dues aux locuteurs eux-mêmes (prononciation, qualité de voix, etc.) ».

5. Quelques facteurs importants dans la conception d'une application impliquant la reconnaissance vocale

Mariani explique qu'un système de reconnaissance vocale ne peut traiter le signal de parole dans toute sa variabilité, et qu'une application résulte forcément d'un compromis entre différents paramètres. Nous dressons ci-dessous une liste qui résume les paramètres qui selon Mariani (2002) et Haton (2006) sont importants à définir lors de l'élaboration d'une application :

- Le mode d'élocution. La reconnaissance peut se faire en continu ou en isolé. Mariani souligne que de plus en plus de systèmes utilisent la reconnaissance en mode continu même pour traiter des mots isolés.
- Le style de voix. Le locuteur parle de façon régulière (type lecture) ou de manière spontanée.
- Le nombre de locuteurs. Il existe des systèmes indépendants du locuteur, dits multilocuteurs, qui fonctionnent même avec des locuteurs n'ayant jamais utilisé le système. Ou au contraire, d'autres systèmes sont dépendants du locuteur (dits unilocuteurs) et demandent un apprentissage préalable de l'utilisateur.
- Adaptation. Afin que le système s'adapte aux particularités vocales du locuteur, une phase d'apprentissage est nécessaire : le locuteur doit prononcer un certain nombre de phrases ou de phonèmes avant de pouvoir utiliser l'application pour la première fois. Cela dit, de plus en plus de systèmes sont adaptatifs en ligne, c'est-à-dire que l'apprentissage se fait en temps réel au fur et à mesure de l'utilisation de l'application.
- La taille et la difficulté du vocabulaire. De nombreuses applications ne nécessitent que des vocabulaires de taille petite (quelques dizaines de mots) à moyenne (quelques centaines de mots), mais dans certains cas le vocabulaire peut atteindre plusieurs dizaines de milliers de mots (dictée, transcription, accès à des bases d'information en langue naturelle).
- Le mode d'acquisition de la parole : la reconnaissance peut être effectuée en local sur un poste de travail ou à distance via une ligne téléphonique (réseau classique commuté ou, de plus en plus souvent, réseau GSM de téléphonie cellulaire). Les modèles acoustiques doivent être adaptés à ces différentes conditions.

- La prise de son : le microphone peut être soit de proximité (avec port d'un casque), intégré dans un combiné téléphonique, ou encore posé sur un bureau.
- Les contraintes imposées par la tâche : lorsque le vocabulaire croît en taille, il devient nécessaire d'introduire des contraintes syntaxiques et sémantiques de façon à limiter les combinaisons de mots autorisées.
- Les performances de reconnaissance. Mariani explique que le taux de reconnaissance attendu dépend de l'usage qu'il sera fait de l'application. Généralement il est censé atteindre au moins 95% au niveau du mot.
- La complexité de la tâche : *« les algorithmes de reconnaissance nécessitent une puissance de calcul assez importante. De ce fait, la mise en œuvre pratique d'un système est limitée par les ressources informatiques disponibles de façon à garantir un fonctionnement en temps réel »* (Haton, 2006).

Afin de clarifier les paramètres qui entrent en jeu dans le logiciel EvoLex qui nous intéresse ici, nous proposons le tableau suivant (RV = Reconnaissance vocale) :

Paramètres de reconnaissance vocale	Paramètres pour EvoLex
Mode d'élocution	Mots enchaînés
Style de voix	Manière spontanée
Nombre de locuteurs	Indépendant du locuteur (sans apprentissage préalable)
Adaptation	En temps réel
Taille et difficulté du vocabulaire	Plusieurs milliers de mots, mais restreints par la consigne de la tâche de fluence elle-même. Précisément sont recensés : 1489 animaux, 542 fruits, 5430 mots en « R » et 2819 mots en « V ». Mais au total, 105 000 formes fléchies.
Mode de traitement	Transcription, détection de silence...
Mode d'acquisition	Au moyen d'un micro-casque

Paramètres de reconnaissance vocale	Paramètres pour EvoLex
Contraintes de la tâche	Non (ne concerne pas les mots isolés)
Performance de la reconnaissance	En cours d'amélioration
Complexité de la tâche	La RV fonctionnera en très léger différé (traitement des fichiers à l'IRIT, puis rendus disponibles en ligne à l'utilisateur)

Figure 8 Paramètres d'EvoLex

EvoLex, dont quelques paramètres viennent d'être présentés, a été conçu afin de faire passer des tests de fluence verbale. Afin de comprendre les éléments importants à analyser lors du recueil des productions du patient, il est indispensable de faire un point concernant les données de la littérature sur la fluence verbale.

II. Fluence verbale

Nous avons choisi de faire un point sur la fluence en suivant la ligne directrice suivante : après une définition rapide de la fluence, nous verrons comment le stock lexical se constitue, et quels sont les processus anatomo-physiologiques et cognitifs impliqués dans la récupération de mots. Ensuite, nous préciserons quelles sont les pathologies concernées par des altérations de la fluence. Enfin, afin de montrer le niveau d'altération de la fluence chez un patient, nous aborderons l'évaluation de la fluence en tant que telle, avec les tests existants et les facteurs influençant les performances à ces tests.

1. Définition de la fluence verbale

Le dictionnaire d'orthophonie (Brin-Henry et collaborateurs, 2011) définit la fluence verbale comme étant la « *capacité d'un individu à évoquer et donner une série de mots à partir d'un champ sémantique (encore appelé champ lexical), ou d'un terme générique, ou du thème qui lui est proposé, exemples : énoncer le plus grand nombre de noms d'animaux, d'instruments de musique [...]* ». Dans le même ouvrage, l'évocation lexicale est définie par « *épreuve ou exercice consistant à donner oralement ou par écrit plusieurs mots se rapportant par*

exemple à un champ lexical déterminé (exemple : dire tous les noms d'animaux connus, les jours de la semaine...) ou consistant à trouver un synonyme, un antonyme. L'évocation permet de tester ou de travailler l'accès au stock lexical et sa constitution ». D'après ces définitions, on constate que les termes de fluence verbale ou d'évocation lexicale sont proches et seront donc utilisés indifféremment dans ce mémoire.

Dans le langage courant, les termes de « fluence » et de « fluidité » sont souvent assimilés, mais nous veillerons à ne pas les confondre. Pour de Weck et Marro, d'origine suisse, (2010) « *les troubles de la fluence verbale concernent la dimension suprasegmentale de la production verbale, et en particulier le rythme de la parole* », ce qui induit la notion de débit. Certains neuropsychologues québécois utilisent le terme de « fluidité verbale » dans le sens de « fluence » (Lussier, 1995). En France, le terme de « fluidité » se rapporte à une parole émise aisément, ou plus précisément selon Brin-Henry et collaborateurs (2011) à une « *émission orale quantitativement et qualitativement satisfaisante de mots et tournures syntaxiques donnant l'impression d'avoir été bien choisis, sans hésitations et avec un débit correct [...]. Chez une personne bègue par exemple, la fluidité verbale est presque toujours altérée. La fluence peut être quasi normale mais avec des achoppements, des hésitations, des blocages qui vont altérer le débit de la parole* ».

On peut donc imaginer qu'une bonne capacité d'évocation lexicale aidera à obtenir une bonne fluidité du discours. D'autre part, la fluence et la fluidité ne se mesurent pas de la même façon. La première concerne la recherche de mots isolés et s'évalue donc par un nombre de mots trouvés dans un temps imparti, alors que la fluidité, qui s'inscrit dans le discours, peut se mesurer par la longueur d'un énoncé par exemple.

Notons que la fluence verbale recouvre deux composantes : la fluence sémantique ou catégorielle (recherche par champ sémantique ou catégorie) et la fluence formelle ou littérale ou phonémique (recherche de mots commençant par un son ou par une lettre).

2. Stock lexical

2.1 Comment se constitue-t-il ?

Le stock lexical ou lexique mental se forme au fur et à mesure de l'acquisition du langage et tout au long de la vie. En effet, dès la naissance, le bébé est plongé dans un mode sonore, dont la parole, tout d'abord perçue comme un flot continu de sons, est de plus en plus segmentée au fur et à mesure de la connaissance et reconnaissance de formes acoustiques

désignant des objets ou des actions. Ainsi « *la reconnaissance et la production des premiers mots indiquent qu'il existe une conscience que les formes sonores ont un sens* » (de Boysson-Bardies, 1996). Par exemple, le mot « chat » dit par la voix grave de papa ou par la voix aiguë de la grande sœur, et entendu dans différents contextes linguistiques, affectifs, kinesthésiques (« oh, regarde le joli chat beige de mamie, il est tout doux » ou « attention, ce chat peut te griffer et te faire mal ! »), aura plus tard la même définition. L'enfant a ainsi **généralisé et catégorisé le mot** « chat » : il entre dans la catégorie des animaux (hyperonyme), puis des animaux à 4 pattes, poilus, domestiques, etc. qui sont autant de sous-catégories. On voit ainsi que le lexique mental se construit à partir d'indices acoustiques, phonologiques et sémantiques, et grâce à un processus élaboré de catégorisation. De plus, avec le temps et les expériences, un mot pourra être catégorisé selon différents critères : l'orange est un fruit mais aussi une couleur, l'avion peut être classé dans ce qui vole, de même que le pinson, alors que l'avion est un produit manufacturé et l'oiseau un animal. Bref, les réseaux sémantiques vont se croiser, se complexifier ; cela rendra le langage d'autant plus riche (polysémie, nuances, etc.). Sur le plan phonologique, Thibault et al. (2010) expliquent qu'il y a une forte corrélation entre les résultats à l'épreuve de fluence phonémique et l'apprentissage de la lecture ; le traitement phonémique se renforce avec l'acquisition des correspondances grapho-phonémiques.

Il est admis que les premiers mots acquis sont des mots concrets, que ce soient des noms ou des verbes. En effet, Marmion (2015) explique que les noms tels que « doudou » ou des verbes tels que « manger » sont acquis en premier, du fait de leur caractère concret et tangible. Les mots représentant des objets ou des personnes non visibles directement par l'enfant ou des phénomènes abstraits ne seront acquis que plus tard, au cours de diverses expériences et confrontations avec autrui. Les études CDI (*Communicative Development Inventories*), dont parle Marmion, déterminent trois étapes dans l'acquisition du vocabulaire : « *l'expansion des noms d'objets [...], suivi de l'augmentation des prédicats (verbes et adjectifs) [...], enfin des mots grammaticaux* ». De Boysson-Bardies (1996) confirme ce point en précisant que « *l'apprentissage des verbes dépend, en partie au moins, de la possibilité de comprendre des phrases* ». L'acquisition des verbes est donc fortement liée à l'acquisition de la syntaxe.

On voit ainsi que la constitution du stock lexical commence très tôt et se fait de façon « organisée », et par ailleurs, Crowe (1998) fait remarquer que plus le lexique est organisé, meilleures seront les performances en recherche de mots.

2.2 Pérennisation du stock lexical : rôle des mémoires

Il convient ici de préciser le rôle de la mémoire, ou plutôt des mémoires, en particulier épisodique et sémantique qui permettent la pérennité du stock lexical. Le modèle MNESIS (Modèle NEOStructural InterSystémique de la mémoire humaine) présenté ci-dessous, développé par Eustache et Desgranges (2008), permet de spécifier le rôle de chaque mémoire dans l'acquisition du lexique. Avant tout, un peu de vocabulaire : la **mémoire procédurale** concerne l'apprentissage par entraînements, qu'ils soient cognitifs (apprendre à calculer mentalement, par exemple), verbaux (pouvoir réciter une procédure de sécurité), ou moteurs (apprendre à skier), et concerne donc principalement la mémoire des gestes.

La **mémoire de travail**, dont la structure modélisée par Baddeley (modèle de 1986) a été conservée ici, met en évidence des processus phonologiques (verbaux) et visuo-spatiaux qui sont pilotés par l'administrateur central (focalisation sur les tâches à accomplir). La mémoire de travail est reliée à la mémoire à long terme via le buffer épisodique, qui permet de faire passer une information dans un sens ou un autre : une information à court terme peut être stockée à long terme (le code de l'entrée de l'immeuble, par exemple), de même qu'une information tirée de la mémoire à long terme peut être réactivée dans la mémoire de travail si elle a besoin d'être modifiée (doubler les quantités de la recette des crêpes, par exemple).

En ce qui concerne les **mémoires à long terme**, on peut assister à un processus « montant » : une information sensorielle est enregistrée dans la mémoire perceptive (image d'un arbre), puis cette image acquiert un sens pour entrer dans la mémoire sémantique (c'est un pommier), puis elle va jusqu'à la mémoire épisodique si on rattache à cette information un lieu, un moment précis se rapportant à notre histoire personnelle (la première fois que j'ai vu un pommier, c'était dans le jardin de mon grand-père). Le processus peut également être « descendant » par le biais des reviviscences de souvenirs (**mémoire épisodique**), dont le sens peut changer lorsqu'on y repense (mémoire sémantique), de même que les perceptions que l'on en a (mémoire perceptive).

On comprend ainsi que plus les expériences et les interactions avec autrui sont nombreuses et diversifiées, et plus le stock lexical pourra augmenter, en quantité, en nuances et en réseaux.

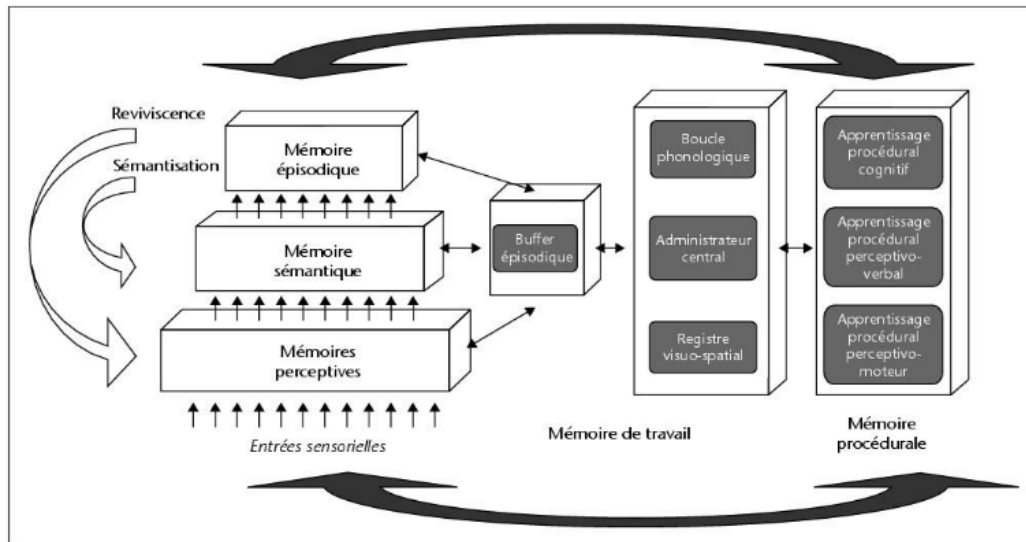



Figure 9 Modèle des interactions entre les différentes formes de mémoire (MNESIS)

Eustache et Desgranges précisent néanmoins que chez les enfants avant trois ans, l'apprentissage de la signification des objets s'encode dans la mémoire sémantique tout en ne gardant que très peu de souvenirs épisodiques. La mémoire de travail avec l'association de la boucle phonologique et du calepin visuo-spatial permet d'associer un son [ʃa] à un objet  vu. Concernant les mémoires à long terme, les entrées sensorielles (toucher doux, vue des formes de l'animal, odeur) sont stockées dans la mémoire perceptive, et enfin, l'association des percepts et du mot permettent un encodage dans la mémoire sémantique.

Lorsque l'enfant grandit, et d'autant plus à l'âge adulte, l'apprentissage de nouveaux mots fait appel à ces processus, mais la mémoire épisodique peut également avoir un rôle de fixation de l'information, par la prise en compte du contexte d'apprentissage, plus personnel, rattaché à un lieu et à un instant précis dans la vie du sujet. Cependant, Eustache et Desgranges dans leur article précisent qu'un patient amnésique peut apprendre une nouvelle langue tout en oubliant les circonstances dans lesquelles il a fait cet apprentissage ; on dira alors qu'il y a formation de connaissances (mémoire sémantique) sans formation de souvenirs (mémoire épisodique).

Bonin (2008) précise que le lexique mental d'un adulte contient entre 50 000 et 100 000 mots (pour un adulte cultivé). On comprend ainsi que la constitution du stock lexical se fait tout au long de la vie.

3. Activités cérébrales impliquées dans l'évocation lexicale

La production de mots a été modélisée par de nombreux auteurs. Le modèle le plus communément validé et utilisé est celui de Caramazza et Hillis de 1990. Il se réfère à la production de mots isolés, ce qui nous intéresse particulièrement dans le cadre de notre mémoire.

Dans leur modèle, ces auteurs précisent que le sujet, lorsqu'il veut dire un mot dans la catégorie des animaux va chercher dans son **système sémantique**. Cependant, on peut supposer que dans le cadre de la fluence phonémique, le sujet va se référer au stock phonologique d'entrée, sans forcément passer par le système sémantique.

Puis dans le **lexique phonologique de sortie**, le sujet retrouve les formes phonologiques des mots (phonèmes et ordre des phonèmes). Enfin, la mémoire tampon phonologique permet de structurer les praxies à effectuer pour dire le mot correctement.

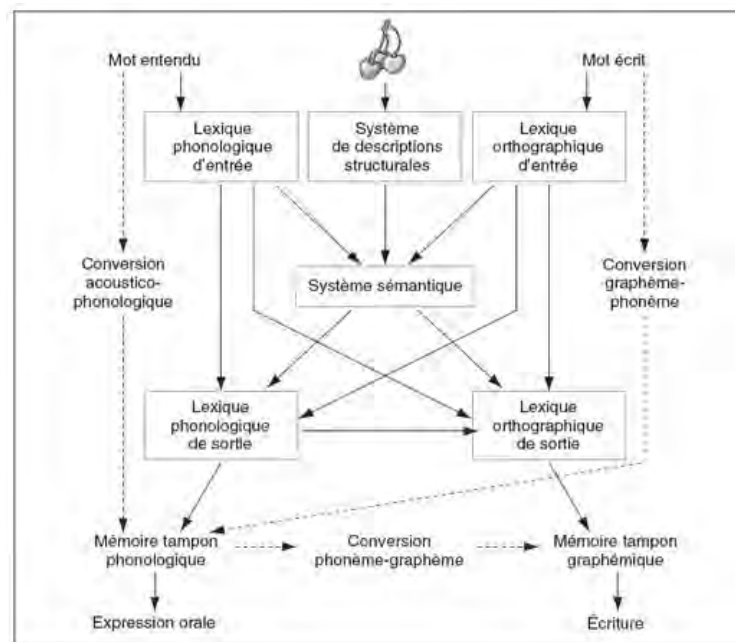


Figure 10 Modèle sériel de production du mot selon Caramazza et Hillis (tiré du site <http://pontt.net/>)

Bonin (2008) explique qu'un mot est produit environ entre 600 et 1200 millisecondes après que le dessin représentant le mot a été présenté (en tâche de dénomination, par exemple). Il explique que la facilité et la rapidité de production viendraient de plusieurs facteurs : non ambiguïté du mot (un objet = un mot), l'âge d'acquisition des mots (plus le mot est acquis tôt, plus il est facile à récupérer), la fréquence d'exposition (d'utilisation) et la fréquence lexicale

(les mots fréquents sont traités plus rapidement). Il ajoute que le coût cognitif de recherche de mots augmente au fur et à mesure de la baisse de fréquence des mots produits, ce qui rejoint Crowe (1998), comme nous le voyons plus loin.

Tout ce processus est également sous-tendu par des activités cognitives et exécutives que nous allons détailler ici.

3.1 Anatomophysiologie des activités cérébrales sous-tendues par les tâches de fluence

Hirshorn et Thompson-Schill (2006) précisent qu'une bonne performance en fluence requiert des processus de sélection lexicale, d'encodage phonétique, de mémoire de travail et de contrôle exécutif. Elles font remarquer que lors de tâches de fluences, l'activation d'un vaste réseau de régions corticales et sous-corticales est visible en imagerie.

3.1.1 Rôle du lobe temporal

La recherche de mots dans une catégorie donnée (animaux, par exemple) fait intervenir les régions frontotemporales et/ou pariétales (d'après Weill-Chounlamountry et ses collaborateurs in Mazaux et al., 2014). Des études, notamment celles de Martin et al. (cités par Gierski et Ergis, 2004), montrent qu'une **lésion temporale gauche entraînerait une altération plus importante en fluence sémantique** qu'une lésion temporale droite. Par ailleurs, Troyer et al. (1998) précisent que « *la fluence sémantique est souvent associée à un test du fonctionnement du lobe temporal* », mais ils précisent également, que les études sont souvent contradictoires. En effet, ils ont pu montrer une réduction du nombre de mots générés en fluence sémantique que ce soit chez les patients avec lésions frontales ou temporales, ce qui montrerait que cette performance ne serait pas liée uniquement au fonctionnement du lobe temporal. De plus, l'altération des résultats était significative pour les 2 tâches de fluence (phonémique et sémantique) chez des patients ayant subi une lobectomie temporale suite à des épilepsies ; cela montre d'un autre côté que la fluence phonémique est également sous-tendue par l'activité du lobe temporal.

D'autre part, Jokeit et al. (cités par Gierski et Ergis, 2004) ont montré que « *les lésions du lobe temporal peuvent produire un déficit catégorie-spécifique dépendant de l'hémisphère lésé* », car ils ont observé que les résultats en fluence sémantique pour les animaux étaient inférieurs si la lésion temporale était à droite, alors que les résultats étaient moins bons pour les catégories des outils et pour les recherches de mots se référant à des attributs visuels (forme, couleur, taille, etc.) lorsque la lésion était à gauche. D'autres auteurs (Zarino et al.,

2014) ont pu montrer que les performances en fluence sémantique pour la catégorie des animaux étaient corrélées à l'activation bilatérale du lobe temporal inférieur et du gyrus fusiforme (gyrus temporo-occipital inférieur) ; alors que la dénomination de produits manufacturés impliquait plutôt d'autres parties du lobe temporal (en particulier le gyrus postéro-médial, le gyrus inférieur bilatéral et la région médiale gauche) ainsi que la région prémotrice gauche.

Certains auteurs, dont Caramazza et al. (cités par Zarino et al., 2014) partent du principe que différentes sous-catégories d'items vivants (animaux, fruits, légumes) sont générés par des circuits neuraux séparés et spécialisés ; ils expliquent ceci par le fait que ces items ont des rôles différents dans la survie humaine (animal = danger, fruits et légumes = nourriture). D'autres auteurs expliqueraient ces différences par les différentes représentations que l'individu peut avoir : des informations visuelles, auditives et langagières permettraient la représentation d'animaux sauvages, alors que ce seraient des informations olfactives, tactiles et relatives au goût qui aideraient à la représentation de fruits, de légumes ou de fleurs. Cela signifierait que chaque catégorie sémantique serait liée à une partie anatomique du cerveau.

3.1.2 Rôle de l'hippocampe

D'autre part, lorsqu'on objective une lésion de l'hippocampe, associée à une lésion temporale, on constate (études de Gleissner et Elger, cités par Gierski et Ergis, 2004) que les **performances en fluence sémantique sont plus altérées**, qu'en cas de lésion temporale seule. Or, l'hippocampe est la structure sous-corticale temporale impliquée dans l'encodage et la récupération d'informations stockées en mémoire. Sachant que le lobe temporal (antérieur) est le siège de la mémoire sémantique, ces deux informations expliqueraient la difficulté à accéder au stock sémantique, due à une difficulté de récupération et à une atteinte du stock sémantique.

3.1.3 Rôle du lobe frontal

Communément, « *la fluence phonémique est souvent associée à un test du fonctionnement du lobe frontal* » (Troyer et al, 1998). En effet, ces auteurs ont montré que le nombre de mots générés en fluence phonémique par les patients atteints de lésions frontales était moins important que pour les patients atteints de lésions temporales. D'autres études, dont celle de Corcoran et Upton (cités par Gierski et Ergis, 2004) ont eu les mêmes résultats. Toutefois, Troyer et al. (1998) émettent une réserve, observant que des patients avec des lésions du lobe pariétal avaient également de moins bonnes performances en fluence phonémique.

Sur une tâche de fluence formelle, on observe une **corrélacion entre de meilleures performances et l'activation du lobe frontal gauche**, mais pas du droit (étude de Wood et al. de 2001 citée par Gierski et Ergis, 2004). Cependant, on observe également une différence d'activation frontale en fonction du temps : le cortex frontal inférieur gauche est plus actif pendant les 30 premières secondes, et le cortex frontal médian gauche s'active plus dans les 30 dernières secondes. Cela augurerait de la mise en place de stratégies différentes en début et en fin de test, ces dernières relevant d'un processus moins « routinier » que les productions des premières secondes, et étant sous-tendues par le cortex médio-frontal gauche.

D'un autre côté, comme pour les atteintes temporelles, l'influence d'un hémisphère ou de l'autre dans les lésions frontales n'est pas négligeable. Hirshorn et Thompson-Schill (2006) font remarquer que **le lobe préfrontal gauche est la région la plus impliquée dans les tâches d'évocation lexicale**. Des études (citées par Gierski et Ergis, 2004) ont montré grâce à l'imagerie fonctionnelle que la partie postérieure du lobe préfrontal inférieur gauche participait à la réalisation des deux tâches d'évocation lexicale. Ces études ont également montré que le cortex préfrontal inférieur droit s'activait lors de tâches de fluence sémantique, mais peu en fluence phonémique. Cela s'expliquerait par le fait que cette région corticale est plus impliquée dans la stratégie de catégorisation.

On voit ainsi que même si les lobes frontal et temporal ont un rôle majeur dans la réalisation des tâches d'évocation lexicale, en formant une alliance sémantique et exécutive, la localisation cérébrale à elle seule ne peut expliquer le fonctionnement d'un individu sur les tâches de fluence. Troyer et al. (1998) précisent bien que, parmi leur échantillonnage, il peut y avoir des patients avec des atteintes diffuses, en plus de lésions focales, ce qui aiderait au fonctionnement cérébral dans les tâches de fluence. La plupart des auteurs cités ici, même s'ils ont pu mettre à jour des liens entre localisation cérébrale et processus, sont conscients que des recherches plus approfondies doivent être menées pour préciser la physiologie cérébrale sur les tâches de fluence. Plus les recherches avancent, et plus il devient évident qu'on a plutôt affaire à des **réseaux neuronaux, impliquant plusieurs régions cérébrales**, et donc à des activations diffuses. Encore reste-t-il à montrer lesquelles et si ces découvertes peuvent être généralisées à la majorité des individus.

	Fluence sémantique	Fluence formelle
Lobe temporal	<p>Altération plus importante en cas de lésion de l'hémisphère gauche</p> <p>En cas de lésion à gauche, altération plus importante sur les catégories des outils et des attributs visuels (couleurs, formes, etc.)</p> <p>Si lésion à droite, la catégorie des animaux est plus altérée</p>	Altération des performances en cas de lésion temporale isolée
Lobe frontal	<p>Baisse des performances en cas de lésion frontale</p> <p>Activation de la partie postérieure du lobe préfrontal droit</p>	<p>Altération plus importante qu'en cas de lésion temporale seule</p> <p>Rôle primordial de la partie préfrontale gauche pour de meilleures performances</p>
Hippocampe	Altération plus importante en cas de lésion hippocampique associée à lésion du lobe temporal , que si lésion temporale seule	
Autres	Altération plus importante en cas de lésion temporale associée à une lésion du gyrus fusiforme (temporo-occipital inférieur)	Altération importante en cas de lésion pariétale associée

Figure 11 Résumé des altérations dans les tâches de fluence en fonction des lésions cérébrales

3.2 Processus cognitifs et exécutifs

De nombreux processus cognitifs sont en jeu lors de ces tâches de fluences verbales ; en effet, une recherche active et efficace de mots nécessite, entre autres, un fonctionnement optimum de la mémoire, de l'attention, mais aussi des fonctions exécutives.

3.2.1 Langage

Bien que les tâches de fluence verbale, comme leur nom l'indique, font référence à une compétence verbale, il est maintenant admis qu'elles testent moins le langage que les fonctions exécutives. En effet, Robinson et al. (cités par Seron et Van Der Linden, 2014), ont montré qu'une altération aux tâches de fluence n'était que peu corrélée à un trouble du

langage de type trouble de la dénomination, que l'on retrouve par exemple chez les patients frontaux. Les tâches de dénomination permettent en effet de tester l'état du stock lexical, alors que les tâches d'évocation lexicale évaluent l'accès à ce stock.

3.2.2 Mémoire

La **mémoire** est bien entendu sollicitée lors des tâches de fluence verbale. Comme nous l'avons vu dans l'acquisition du lexique, les mémoires **épisodique** et **sémantique** ont un rôle important dans la constitution du stock lexical, mais c'est principalement la mémoire **sémantique** qui est sollicitée dans les tâches d'évocation lexicale.

Selon Lecouvrey, Desgranges et Eustache (in Dubois et al., 2015), la mémoire sémantique « *est définie comme la mémoire des mots, des concepts, des connaissances que nous avons sur le monde et sur soi, affranchies de leur contexte spatial et temporel d'acquisition* ». Elle peut être explorée à partir de différentes épreuves, notamment des épreuves d'évocation de mots sur critère sémantique. Un score faible en fluence catégorielle peut avoir plusieurs significations : une difficulté d'accès au lexique (plutôt due à des déficits exécutifs) ou une perturbation des représentations sémantiques, qui signent plutôt une altération de la mémoire sémantique. En effet, Hahn et Guichart-Gomez (in Dubois et al., 2015) précisent que le dysfonctionnement d'accès au lexique est imputable à une atteinte frontale, alors que l'atteinte du stock des connaissances sémantiques est plutôt en rapport avec les régions temporales. C'est pourquoi un score faible aux épreuves de fluences sémantiques ne peut être considéré isolément et son analyse doit être combinée avec d'autres tests afin d'avoir une idée précise du déficit : atteinte exécutive ou atteinte sémantique ?

3.2.3 Attention

Une des fonctions cognitives impliquées dans la tâche de fluence est l'**attention**. La fonction d'attention est « *une focalisation sélective sur des aspects spécifiques de notre environnement ou la concentration lors d'opérations mentales et des processus mentaux spécifiques* » (Weill-Chounlamountry et collaborateurs, in Mazaux, 2014). C'est une structure hiérarchisée en plusieurs composantes : la vigilance, la fonction d'alerte, l'attention soutenue, l'attention sélective et l'attention divisée. L'hémisphère gauche serait le siège de l'attention divisée et sélective, alors que l'hémisphère droit concernerait plutôt la vigilance et l'attention soutenue. Un déficit attentionnel est souvent présent chez les patients aphasiques ; le traitement du langage (en expression et en compréhension) demande des **ressources attentionnelles importantes**, d'autant plus lorsque celui-ci est altéré du fait de lésions cérébrales. On comprend ainsi qu'un déficit en attention soutenue (rester concentré pendant 2 minutes sur une tâche) et/ou sélective (focalisation sur une tâche) entraîne une

difficulté pour chercher les mots selon la consigne. Nous avons ainsi pu entendre, lors des enregistrements de sujets sains en tâches de fluence, des modalisations telles que « heu, bon, il faut que je me reconcentre », ou alors, des commentaires émis après les tests tels que « c'est très difficile de rester attentif pendant les 2 minutes, j'ai dû me forcer à me recentrer parfois ».

3.2.4 Fonctions exécutives

Allain et Le Gall (in Godefroy et al., 2008) précisent qu' « *en neuropsychologie humaine, la notion de “fonctions exécutives” fait référence à des “fonctions de direction” permettant, lors de la réalisation d'une tâche, la définition du but ou des objectifs à atteindre, d'une stratégie pour y parvenir, le contrôle de sa mise en œuvre et des résultats. Elles correspondent donc à des fonctions de haut niveau (des fonctions de contrôle), impliquées dans de nombreuses formes d'activités cognitives* ». Ainsi les fonctions exécutives incluent différents processus, tels que la mise en route de l'activité mentale (initiation), la sélection entre des informations pertinentes et non pertinentes (inhibition), la formulation de plan d'actions (planification), la mise en place de stratégies, le changement de stratégies en cas d'échec (flexibilité mentale), entre autres.

Au niveau anatomique, d'après Luria (cité par Allain et Le Gall in Godefroy et al., 2008), la région prémotrice participe à la **dynamique de l'action**. Lors de lésions de cette région, le patient sait ce qu'il doit faire, a envie de le faire, mais le plan d'action n'aboutit pas (réponses ébauchées mais non abouties, autocorrections d'erreurs, etc.). Il explique que la région dorso-latérale (partie antérieure du cortex latéral frontal) est impliquée dans la **décision d'action, la planification et le contrôle de cette action**. Ainsi, si cette région est lésée, les sujets ont des difficultés à construire un discours, à résoudre un problème. Enfin, la région médio-basale (structures internes du lobe frontal incluant le système limbique et le cortex orbitaire) serait, quant à elle, responsable du **maintien de l'activité tonique et de la synthèse des informations** en provenance du milieu interne. D'après Luria, des lésions de ces structures entraîneraient un déficit de conduite de l'activité par une **désinhibition** importante alliée à un déficit de la sélectivité des informations non pertinentes. De plus, on note des altérations de l'affect, avec notamment des troubles des conduites sociales ou des modifications des caractères.

3.2.5 Rôle de la mémoire de travail

Certains auteurs, dont Goldman-Rakic (cité par Allain et Le Gall in Godefroy et al., 2008) situent la mémoire de travail, et en particulier l'**administrateur central**, comme faisant partie des fonctions exécutives, dans la mesure où une difficulté à faire le lien entre les événements précédant un changement de critère et la consigne serait due à un « *déficit des*

mécanismes qui permettent la mise en mémoire de travail des représentations et leur mise à jour » (dans le cas des tests de type Wisconsin Card Sorting Test qui visent à évaluer la flexibilité mentale). D'autre part, dans le cas des tâches de fluence, dans la mesure où la mémoire de travail permet de garder la consigne à l'esprit et de la mettre à jour, en se remémorant les mots déjà produits, on peut considérer qu'elle entre dans la définition des fonctions exécutives donnée plus haut par Allain et Le Gall.

La mémoire de travail a été modélisée par Baddeley en 1986 (voir descriptif dans le paragraphe 2.2), et ce modèle est jusqu'à aujourd'hui largement admis et utilisé. Pour lui (cité par Allain et Le Gall in Godefroy et al., 2008), à propos de l'administrateur central, « *son implication est nécessaire dans les épreuves mettant en jeu la flexibilité, dans les conduites des doubles tâches, l'attention sélective et l'activation de la mémoire à long terme* ».

D'autre part, Weill-Chounlamounry et ses collaborateurs (Mazaux, 2014) précisent le rôle fondamental de la boucle phonologique et du buffer épisodique. La **boucle phonologique** permet de stocker temporairement des informations verbales et le modèle articulatoire. Ces traitements permettent aux lexiques (d'entrée et de sortie) de former des « modules » entrant dans la mémoire tampon (ou **buffer épisodique**).

On comprend donc bien le lien privilégié entre mémoire de travail et langage, et en particulier lors d'épreuves de fluence, qui impliquent flexibilité (pouvoir changer de stratégies si nécessaire), attention sélective, informations phonologiques et articulatoires et activation de la mémoire à long terme (notamment la mémoire sémantique).

C'est ainsi, lors des enregistrements de tâches de fluence, qu'on a pu entendre des commentaires tels que « est-ce que j'ai déjà dit "girafe" ? » ou encore « ah non, pas les légumes, il faut que je donne des noms de fruits ». Ces illustrations montrent que la mémoire de travail est très impliquée dans la réussite aux tâches de fluence.

3.2.6 Mise en place de stratégies

Le rôle prépondérant des fonctions exécutives dans les tâches de fluence est également décrit par Weill-Chounlamounry et ses collaborateurs (Mazaux et al., 2014). Selon eux, « *il est établi [que les tâches de fluences verbales] dépendent des fonctions exécutives selon qu'elles soient catégorielles ou phonologiques* ».

Ces mêmes auteurs font toutefois une différence entre l'évocation lexicale formelle et sémantique, en précisant que les tâches de fluence à partir d'une lettre n'étant pas une tâche habituelle, elles demandent la mise en place de **stratégies**. Par exemple, certains

sujets que nous avons enregistrés mettaient en place des stratégies phonémiques, en disant à haute-voix « je n'ai plus de mots en RA, j'essaie en RE ».

En ce qui concerne l'évocation lexicale sémantique, la difficulté peut venir d'un déficit d'accès à la mémoire sémantique et/ou d'un dysfonctionnement de fonctions exécutives. Selon Weill-Chounlamounry et ses collaborateurs (Mazaux et al., 2014), le problème viendrait de l'altération de la capacité à trouver des éléments superordonnés (félin, par exemple) en comparaison à des éléments subordonnés (chat, tigre, etc.) qui demandent en effet une hiérarchisation d'éléments sémantiques ; les éléments superordonnés ont moins de traits sémantiques communs (félin, reptiles, etc.) que les éléments subordonnés (chat et tigre ou serpent et lézard). Les patients aphasiques ont plus de facilité à trouver des éléments subordonnés, du fait d'une altération des fonctions exécutives. Afin d'évaluer si le déficit provient d'un défaut de mise en place de stratégies de recherche, on peut proposer des indices de récupération (sous-catégories d'animaux en fluence sémantique ou ébauche phonologique, par exemple). Ainsi, le patient est aidé dans sa stratégie et trouve plus de mots (Seron et Van Der Linden, 2014).

D'autres auteurs (Mayr, 2002) expliquent que les tâches de fluence sémantique impliquent d'autres composantes exécutives : notamment, la mise à jour des critères de recherche, et l'arrêt ou l'initiation des processus de récupération simple.

Le rôle des fonctions exécutives est donc prépondérant dans la recherche de mots, en particulier la mise en place de stratégies, qu'il convient de détailler plus spécifiquement ici.

4. Focus sur le clustering et le switching

Lorsque nous avons enregistré des patients sains réalisant des tâches de fluence, nous nous sommes aperçues que les sujets mettent en place des stratégies pour trouver le plus de mots possibles. On observe ainsi que certains font des **regroupements**. Par exemple, dans la recherche de noms d'animaux, ils disent des noms d'animaux domestiques (chien, chat, poisson rouge), puis des noms d'animaux de la ferme (mouton, chèvre, cheval, etc.), puis de la savane (lion, hyène, éléphant, etc.), etc. Dans le cas des épreuves de fluence formelles, certains sujets procèdent par regroupements phonologiques avec d'abord des mots commençant par RA puis RE puis RI, ou par rimes finales. Certains regroupements sémantiques peuvent être observés également, par exemple route, rouler, reculer, reconduire.

Les phénomènes de regroupement et de commutation d'un regroupement à l'autre ont été modélisés entre autres par Troyer et al. en 1997 (cités par Gierski et Ergis, 2004) et ont été nommés respectivement *clustering* et *switching*.

Gierski et Ergis (2004) rapportent que les données quantitatives des tests de fluence « *ne rendent pas suffisamment compte de la complexité des processus cognitifs en jeu dans ce type de tâches. Pour ces raisons, Troyer et al. en 1997 ont proposé une nouvelle méthode d'analyse des fluences verbales basée sur les aspects qualitatifs* ». En effet, Troyer et al. (1998) expliquent que « *le nombre de mots générés dans les tests de fluence verbale n'est pas forcément en lien avec les lésions d'une partie du cerveau en particulier. Nous avons montré que deux composantes cognitives importantes dans la performance en fluence sont le clustering et le switching* ». Nous avons pu constater que les corrélations entre les performances aux tâches de fluence et les lésions cérébrales ne sont pas si simples, comme nous l'avons vu précédemment. Troyer et al. ont donc tenté de trouver un autre moyen de clarifier la nature précise du déficit, notamment par une **analyse qualitative plus approfondie**, portant sur deux processus : le « clustering » et le « switching ». Ils précisent que de bons résultats lors de tâches de fluences verbales impliquent la production de clusters de mots reliés phonémiquement ou sémantiquement, et une fois un cluster (sous-catégorie) terminé, le passage à un autre.

4.1.1 Définitions

Le processus de regroupement (**clustering**) peut être réalisé selon un critère phonologique ou sémantique. Un cluster est composé d'au moins deux mots liés sémantiquement ou phonologiquement (commençant par les deux mêmes syllabes). Les regroupements phonologiques vont permettre de faire des groupes de mots commençant par la même lettre ou le même son ou la même syllabe (rare, rapide, ravi), ou rimant entre eux (rempli, repli, rétabli). Les regroupements sémantiques concernent des sous-catégories (éléments subordonnés) ayant des caractéristiques communes (pie, rouge-gorge, aigle pour les oiseaux), et impliquent donc une capacité de classification et de hiérarchisation. L'analyse des clusters de classes grammaticales (noms, verbes), qui seraient corrélés aux lobes frontal (verbes) et temporal (noms), est également intéressante. De même que les clusters « morphologiques » dont les mots sont reliés par des morphèmes communs (cane, caneton, canard, ou voir, vu, vision, par exemple).

Le **switching** permet de passer d'un regroupement à l'autre, entre deux mots isolés (non liés sémantiquement ou phonologiquement), d'un regroupement à un mot isolé ou l'inverse.

4.1.2 Rôle du switching et clustering dans les tâches de fluence

Troyer et al. (1997, cités par Gierski et Ergis, 2004), font également remarquer que pour qu'un grand nombre de mots soient produits en tâche de fluence, il faut une **combinaison entre des clusters de taille conséquente et des switchings rapides** entre chaque regroupement. Plus précisément, de meilleures performances en évocation lexicale ont été observées si les clusters sont de grande taille, s'il y a un nombre important de clusters et de switches et si le nombre de clusters et de switches sont proches ; cela montrerait qu'il y a peu de switches avec un seul mot, et que chaque switch correspond au passage d'un cluster à l'autre.

Ils ont constaté d'autre part, que dans les tâches de fluences formelles, les regroupements sont de plus petites tailles, et qu'il y a plus de switchings. Contrairement aux tâches de fluence sémantique, où les regroupements sont plus importants, mais les switchings moins nombreux, comme on le voit sur ces schémas.

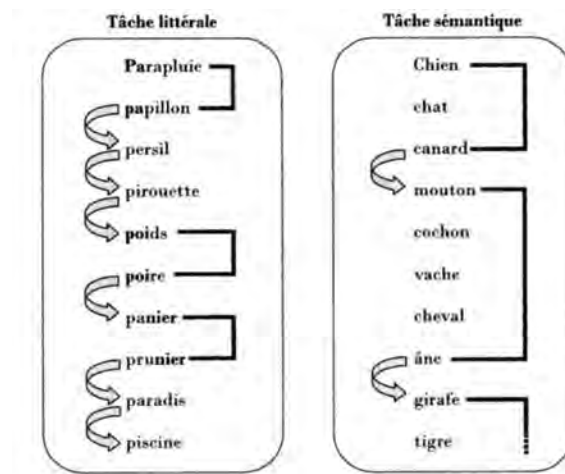


Figure 12 Exemples de clusters et switches en fluence phonologique et sémantique

De plus, Troyer et al, (1998), affirmaient que, lors des tâches de fluence sémantique, les recherches de mots intra-clusters (à l'intérieur des clusters) étaient du ressort d'un processus sémantique, alors que le switching était plutôt sous-tendu par les fonctions exécutives. Seron et Van Der Linden (2014) confortent cette idée en précisant que la génération de clusters aurait un caractère plus automatique que le switching, qui demande des processus de stratégies de recherche et de flexibilité mentale, plus coûteux cognitivement. Cependant, Mayr (2002) affirmait que les fonctions exécutives étaient nécessaires à tout acte de recherche de mots, pour le clustering et le switching, et que d'autre part, le switching serait sous-tendu par un processus sémantique.

4.1.3 Physiologie du switching et du clustering

Les analyses en termes de clustering et de switching ont permis de montrer des processus différents en fonction des lésions cérébrales observées.

Les études menées par Troyer et al. (1998) sur des tâches de fluences littérales (f, a, s) et sémantiques (animaux) ont montré un déficit de regroupements uniquement sur la fluence sémantique, et non sur les tâches de fluence formelle ; les patients avec une **lésion temporelle gauche** **faisaient des regroupements plus petits** que ceux avec une lésion temporelle droite. D'autre part, une réduction du nombre de switchings a été retrouvée dans les deux types de fluences, en particulier chez les adultes âgés et les patients avec des atteintes frontales médiales supérieures et dorsolatérales gauches ; nous avons vu plus haut que d'après Luria, cette partie du cortex est dévolue à la décision d'action, la planification et au contrôle de cette action.

Dans leur étude de 1997, Troyer et al. (cités par Gierski et Ergis, 2004) avaient également montré que les déficits en switching étaient plus importants en fluence phonémique qu'en fluence sémantique ; ce qui serait dû à une plus grande nécessité **d'avoir recours à du switching du fait d'une déficience de la stratégie de clustering** en comparaison avec la fluence sémantique. Cela suggérerait que le switching est fortement dépendant du lobe frontal. Ces observations sont à mettre en lien avec des persévérations, des déficits d'initiation comportementale ou une atteinte de la flexibilité des processus de recherche/récupération que l'on observe chez des patients avec ces lésions.

Dans cette même étude de 1997 rapportée par Mayr (2002), Troyer et al. précisent que, sur une tâche de fluence sémantique, le nombre de switches a tendance à diminuer en fonction de l'âge, alors que la taille des clusters ne diminue pas ; cela s'expliquerait par le vieillissement cérébral qui entraîne une altération des fonctions exécutives et peu du stock sémantique. Cela prouverait que le **switching est une composante exécutive**. Le switching impliquerait l'initiation, la flexibilité mentale, et, comme le suggèrent Troyer et al. (1998), la prise en compte de l'information indiquant que la recherche dans un cluster est terminée. Ils ont montré que le **switching était altéré dans les deux tâches de fluences pour les patients avec des lésions frontales**. D'autre part, le switching en fluence phonémique est normal chez les patients avec des lésions temporelles. Mais ils font également remarquer que le switching en fluence sémantique est altéré chez ces mêmes patients. Ils pensent que cela serait dû à une atteinte plus diffuse impliquant les connexions entre les régions frontale et temporelle.

D'autre part, une étude de Mayr et Kliegl (citée par Mayr, 2002) avec la réalisation de tâches de fluence sémantique avec alternance de catégories a montré que les temps de **switching entre les clusters et entre les catégories dépendaient d'une composante sémantique véritable**, car les temps de recherche entre les catégories les plus difficiles étaient plus longues qu'entre les catégories plus faciles. Hirshorn et Thompson-Schill (2006) ont montré à travers deux expériences que le gyrus frontal inférieur gauche était fortement impliqué dans les activités de switching, en particulier dans les tâches de fluence sémantique. Elles émettent cependant une réserve en spécifiant que des recherches complémentaires doivent être réalisées pour bien différencier les contributions du lobe frontal et celui du lobe pariétal en switching.

4.1.4 Critère temporel et fréquence de mots

Pour compléter ces analyses en termes de clustering et de switching, d'autres auteurs ont approfondi ces résultats, notamment concernant le **temps passé à trouver des mots à l'intérieur d'un cluster, et le temps entre les clusters**, Mayr et Kliegl (Mayr, 2002) précisant que le temps inter-clusters est long, alors que le temps de recherche entre les mots à l'intérieur d'un cluster est plus court. D'autre part, il a été montré (Mayr, 2002) que les adultes âgés étaient plus lents que les adultes jeunes. Une autre étude (Fitzgerald, cité par Mayr, 2002) a montré que les adultes âgés accédaient à un nouveau cluster plus vite que les jeunes adultes, mais que les temps de recherche à l'intérieur des clusters étaient plus longs.

Concernant la prise en compte du critère temporel, Crowe (1998) a mis en évidence le fait que lors des 15 premières secondes, le sujet produit des **mots fréquents**. Il a également constaté que la fréquence des mots diminuait significativement entre les 15 et les 30 premières secondes et la fin du test. Il explique que cela serait dû au temps nécessaire pour mettre en place une stratégie de recherche efficace. Ou autrement dit, la recherche devient plus difficile à mesure que la fréquence des mots diminue.

On voit ainsi que de nombreux processus cognitifs et exécutifs entrent en jeu pour que l'accès au mot du sujet soit fluide. Il est donc logique que de nombreuses pathologies soient concernées par une altération de la fluence, si un ou plusieurs processus sont atteints.

5. Pathologies concernées par des perturbations de la fluence verbale

D'après les définitions du dictionnaire d'orthophonie, l'évocation lexicale est principalement utilisée dans les bilans auprès de patients aphasiques (en particulier d'origine vasculaire).

C'est très largement le cas, puisque toutes les batteries de tests évaluant les patients atteints d'aphasies proposent un test de fluence. Or, la capacité d'accès au lexique peut être altérée dans bien d'autres pathologies. Nous ferons donc ici un inventaire non exhaustif des principales maladies dans lesquelles on retrouve une altération de la fluence, en particulier les maladies dont les patients pris en charge par les orthophonistes sont atteints.

Ainsi, nous détaillerons les pathologies suivantes : les aphasies d'origine vasculaire, les pathologies neurodégénératives, le traumatisme crânien, les pathologies psychiatriques et les troubles des apprentissages.

5.1 Fluence et aphasie d'origine vasculaire

Brun (Mazaux et al., 2007) précisent que « *les perturbations de la fluence sont un des marqueurs les plus fiables de classification des aphasies* ». Les aphasies sont actuellement classées en deux grands groupes : les aphasies fluentes et les aphasies non fluentes.

Les aphasies sont caractérisées par une distorsion importante du discours qui perturbe la communication. Cela peut aller du mutisme (absence d'expression orale) à un discours peu informatif du fait d'un manque du mot ou d'un agrammatisme importants. On retrouve pour les principales l'aphasie globale, de Broca et transcorticale motrice.

5.1.1 Fluence et aphasie globale

L'aphasie globale se caractérise par une anomie majeure avec un mutisme initial qui évolue vers des stéréotypies qui pénalisent une quelconque communication. On peut voir dans les cas d'aphasies globales des **épreuves de fluence complètement chutées**, et les seules productions du patient sont des répétitions d'une même syllabe ou d'un même mot ou d'un segment de phrase dépourvu de sens (stéréotypies).

5.1.2 Fluence et aphasie de Broca

Dans le cas des aphasies de Broca, le discours est altéré par un **manque du mot sévère**, un agrammatisme fréquent (phrases sans verbes ou verbes non conjugués) et des troubles praxiques qui perturbent l'intelligibilité de la parole. Le manque du mot se mesure avec une épreuve de fluence, dont le résultat renvoie à une difficulté d'accès au stock lexical. En outre, on note des **paraphasies**, qui peuvent être phonémiques ([təva] pour cheval), sémantiques (tabouret pour chaise) ou verbales (parapluie pour table). « *La sévérité du syndrome est intimement liée à la notion de fluence : à une réduction massive de la fluence correspondra une aphasie sévère* » (Chomel-Guillaume, 2010).

5.1.3 Fluence et aphasie transcorticale motrice

L'aphasie transcorticale motrice, quant à elle, se caractérise principalement par un défaut d'incitation spontanée et d'initiation de la parole. La **fluence est très perturbée**, du fait principalement de cette asponanéité verbale et de l'anomie. Le patient est incapable de générer une phrase et de la formuler dans son intégralité.

5.2 Fluence et pathologies neurodégénératives

Nous détaillerons ici les principales pathologies neurodégénératives dont l'altération de la fluence verbale est au premier plan, auxquelles nous pouvons être confrontés dans nos cabinets d'orthophonie pour une rééducation du langage.

Une pathologie neurodégénérative est une pathologie d'origine neurologique, dont l'évolution est progressive, plus ou moins rapidement. Dans le cas de dégradations cognitives progressives, l'autonomie est remise en cause ; le patient ne peut plus habiter seul et a besoin d'aides humaines et/ou techniques pour la vie quotidienne (y compris la nuit dans certains cas). On parle alors de **démence**, ou plutôt de démences au pluriel, car une maladie neurodégénérative entraîne une démence différente d'une autre maladie neurodégénérative. La démence a donc une origine organique, comme le précise Dubois (Dubois et al., 2015) « *la démence, qui résulte de l'atteinte de ces mêmes fonctions [supérieures, actuellement appelées cognitives], est une maladie du cortex cérébral* ». Récemment, le terme « démence » (DSM-4) a été remplacé par le terme « **trouble neurocognitif majeur** » (DSM-5), dont la définition est en fait plus large. En effet, d'après la définition du DSM-4, un déclin cognitif substantiel dans plusieurs domaines permettait de poser un diagnostic de démence, alors qu'avec la définition du DSM-5, un seul critère peut suffire. D'autre part, les troubles de la mémoire ne sont plus essentiels pour porter le diagnostic. Cependant, le terme de « démence » étant toujours largement utilisé, nous le conserverons dans ce mémoire.

Parmi les pathologies neurodégénératives dont l'altération des fluences verbales est une des caractéristiques, nous détaillerons ici les pathologies corticales dont la maladie d'Alzheimer et les aphasies primaires progressives, la SLA, les pathologies sous-corticales dont la maladie de Parkinson, la paralysie supranucléaire progressive, la maladie à corps de Lewy, la maladie de Huntington, et la sclérose en plaque.

5.2.1 Fluence et maladie d'Alzheimer (MA)

La maladie d'Alzheimer se caractérise physiologiquement, par l'apparition de plaques séniles (accumulation extracellulaire du peptide β -amyloïde) et la dégénérescence neurofibrillaire

(lésions de taupathie intraneuronales), ces deux marqueurs étant retrouvés dans le liquide céphalorachidien. D'autre part, on assiste, en imagerie cérébrale à une atrophie hippocampique, des structures temporales (interne et inférieure), mais aussi pariétales et cingulaires. Les deux hémisphères sont concernés, généralement de façon symétrique.

Les premières manifestations de la MA sont la perturbation de la mémoire épisodique, du fait notamment de l'atteinte hippocampique. Puis, viennent des perturbations cognitives telles que des altérations du langage (manque du mot, perturbations praxiques, des fonctions exécutives et visuospatiales). Petit à petit, le comportement est perturbé, ce qui conduit à des pertes de l'autonomie du patient.

D'après Mosca et Godefroy (Godefroy et al., 2008), se référant à des études de Henry, Crawford et Philipps de 2004, on constate une altération plus importante de la fluence sémantique qu'alphabétique, notamment aux stades pré-démontiel et léger de la maladie (atteinte de la mémoire sémantique). A un stade plus avancé de la maladie, on observe une **baisse des performances aux tâches d'évocation lexicale alphabétique**. D'après Lecouvrey, Desgranges et Eustache (in Dubois et al., 2015) les patients atteints de la maladie d'Alzheimer auraient des difficultés à « *opérer des regroupements sémantiques (ou « clusters »)* ». Ces difficultés seraient plus de l'ordre d'une diminution des capacités exécutives des sujets, et notamment de mises en place de stratégies.

Au plan neuropsychologique, Mosca et Godefroy (Godefroy et al., 2008) soulignent que plusieurs épreuves permettent de poser un diagnostic précoce de la maladie d'Alzheimer y compris avant l'apparition des symptômes : les principaux étant les tests de mémoire, d'inhibition (Stroop), de flexibilité mentale (TMT-B), et de fluences verbales. Comme le précisent Zarino et al. (2014), les épreuves d'évocation lexicale catégorielle sont plus sensibles que les épreuves de fluence formelle dans la détection précoce de la maladie d'Alzheimer. Il y a un consensus autour du fait qu'une **altération plus importante en fluence catégorielle**, signe le développement de la maladie d'Alzheimer et cela précocement (entre 3 et 5 ans avant l'apparition des premiers symptômes).

5.2.2 Fluence Aphasie primaire progressive (APP)

Contrairement à l'aphasie d'origine vasculaire, l'aphasie primaire progressive a la particularité d'être due à une lésion corticale focale. On distingue trois formes d'APP (dont les localisations cérébrales sont présentées sur l'image ci-dessous tirée de Dubois et al., 2015), pour lesquelles une évaluation de la fluence verbale est préconisée.

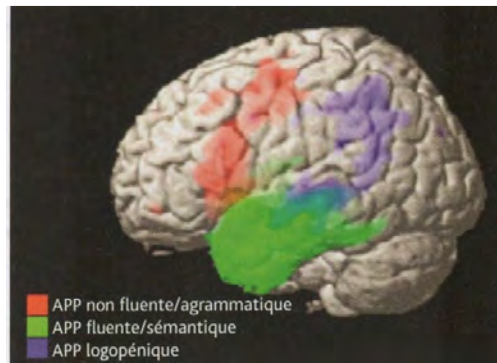


Figure 13 Visualisation des lésions corticales dans les aphasies primaires progressives (Dubois et al., 2015)

➤ Variante non fluente/agrammatique

La forme agrammatique est liée à une atrophie focalisée sur le cortex frontal postéro-inférieur. Le langage est perturbé du fait d'un déficit en compréhension, mais aussi, en production, de paraphrasies phonémiques, d'erreurs syntaxiques, et l'évolution de la maladie conduit à des troubles exécutifs et attentionnels. L'atteinte exécutive engendre une **altération plus importante de la fluence formelle**.

➤ Variante fluente / démence sémantique

Les patients atteints de démence sémantique se plaignent de « *trouble de la mémoire des mots* » (Didic et Felician, in Dubois et al., 2015). Outre des tâches d'appariement sémantique, de dénomination, entre autres, les résultats aux **tâches de fluence verbale catégorielle sont faibles**, en comparaison à la fluence formelle. Cette altération est due à un trouble progressif de la mémoire sémantique, et non à un syndrome dysexécutif. Didic et Felician, (Dubois et al., 2015) insistent sur le fait que la démence sémantique est une altération des savoirs et non du langage.

➤ Variante logopénique

La forme logopénique est associée à une atteinte du cortex temporal externe postéro-supérieur empiétant sur le cortex pariétal inférieur, englobant l'aire de Wernicke. Didic et Felician, (Dubois et al., 2015) évoquent un manque du mot caractéristique (« *mot sur le bout de la langue* »). Les concepts sémantiques sont conservés au début de la maladie, mais les patients sont gênés par le fait qu'ils ne retrouvent plus la forme sonore ou écrite du mot. Le manque du mot est très important, de ce fait, les **épreuves de fluence ne sont pas forcément préconisées** ; en revanche, les épreuves de dénomination mettent en avant les difficultés telles que des paraphrasies phonémiques et sémantiques.

5.2.3 Fluence et SLA

La sclérose latérale amyotrophique (SLA) est due à une dégénérescence touchant les motoneurones centraux et périphériques, qui entraîne en quelques années le décès des patients par une paralysie musculaire progressivement globale. Au plan anatomique, on constate une atteinte dorsolatérale du lobe frontal. D'après Portet et al. (in Dubois et al., 2015), de nombreuses études montrent des anomalies aux tests neuropsychologiques chez les patients atteints de SLA, mais sans évoquer forcément de démence. En effet, on retrouve des troubles mnésiques (en particulier épisodique), exécutifs (difficultés de conceptualisation et de planification) et de façon moins systématique, des troubles du langage. Dans ces cas précis, on notera une **diminution de la fluence verbale**.

5.2.4 Fluence et pathologies neurodégénératives sous-corticofrontales

La particularité des maladies neurodégénératives par atteinte sous-corticale (noyaux gris centraux en lien avec le lobe frontal) est que les perturbations touchent le contrôle moteur, et les fonctions cognitives et exécutives. En effet, des lésions au niveau des boucles entre le striatum, les pallidums, le thalamus et le cortex frontal provoquent d'une part une akinésie (défaut d'initiation du mouvement) et les tremblements au repos (défaut de contrôle des mouvements musculaires) et d'autre part, des troubles cognitifs, exécutifs et/ou comportementaux plus ou moins importants en fonction des pathologies. La fluence est perturbée dans les cas de maladie de Parkinson, la paralysie supranucléaire progressive et la maladie de Huntington, que nous allons détailler ici.

➤ Fluence et maladie de Parkinson

Dans le cas de la maladie de Parkinson, les lésions des noyaux gris centraux associées à un hypométabolisme et à un amincissement cortical, en particulier des zones temporales, pariétales et occipitales, conduisent à une démence. Czernecki, Michon et Dubois (Dubois et al., 2015) indiquent qu'une épreuve de **fluence phonologique** est systématiquement proposée pour établir un diagnostic de maladie de Parkinson, et que le score est considéré comme pathologique s'il est inférieur à 9 mots en 1 minute. Czernecki, Michon et Dubois (Dubois et al., 2015) relatent qu'une étude menée sur 5 ans par Williams-Gray et ses collaborateurs en 2009, montre que la difficulté de reproduction de figures et une **diminution de la fluence sémantique** (impliquant les régions temporales) « *s'avéraient de bons prédicteurs [de la démence]* ». D'autre part, Troyer et al. en 1998 ont montré un **réduction du nombre de switches** chez les patients déments, que ce soit en fluence littérale ou catégorielle.

➤ Fluence et Paralyse supranucléaire progressive (PSP)

Verny (Dubois et al., 2015) explique que le type de troubles cognitifs rencontrés chez les patients atteints de paralyse supranucléaire progressive (ou maladie de Steele-Richardson) sont très similaires à ceux trouvés chez les patients atteints de maladie de Parkinson. Le diagnostic différentiel peut se faire notamment par le constat d'une atteinte oculomotrice propre à la PSP et d'autre part, par la sévérité du syndrome dysexécutif. Dujardin (Godefroy et al. 2008) précise que pour ces patients, la formation de concepts est altérée et qu'elle est en partie responsable de **faibles scores aux épreuves de fluences phonémique et sémantique**, avec une forte tendance à la **persévération**.

➤ Fluence et démence à corps de Lewy

Les corps de Lewy sont des corps cellulaires qui s'invitent dans d'autres éléments : neurone, noyau de nerf, par exemple. La présence de ces corps de Lewy est généralement associée à une perte neuronale, et peut provoquer des troubles cognitifs, avec notamment, comme dans la maladie de Parkinson, un **déficit notoire aux épreuves de fluence formelle**.

➤ Fluence et maladie de Huntington

Cette maladie d'origine génétique entraîne une dégénérescence des neurones striataux ; nous retrouvons donc chez ces patients des manifestations comportementales initialement, puis motrices (mouvements involontaires de type choréiques), psychiatriques et cognitives, d'évolution progressive. Au niveau cognitif, les patients ont des difficultés attentionnelles majeures, et sur le plan exécutif, une perturbation de la flexibilité mentale, de l'organisation et de la planification. Aux **épreuves de fluence, le score est déficitaire** (Dujardin cité par Godefroy et al., 2008). Comme dans la maladie de Parkinson, les sujets déments présentent des **déficits en switching**.

5.2.5 Fluence et sclérose en plaque (SEP)

La SEP est une maladie due à une inflammation du système nerveux qui entraîne une démyélinisation des axones, évoluant par pallier. Les atteintes cognitives sont semblables aux pathologies sous-corticales avec des troubles attentionnels, de la mémoire de travail, de la vitesse de traitement, des fonctions exécutives et de la mémoire (en particulier verbale et visuospatiale). Le langage est peu altéré, même si Stoquart-Elsankari, Roussel et Godefroy (Godefroy et al., 2008) précisent qu'une **altération des fluences verbales** est un signe précoce de la maladie, voire « *un marqueur de dysfonctionnement exécutif le plus sensible dans la SEP [études de Wiwhart et Sharpe en 1997 et de Zakzanis en 2000]* ». Cependant, il existe une importante hétérogénéité des profils de patients atteints de SEP. Certains auteurs

dont Henry et Beatty en 2006, cités par Stoquart-Elsankari, Roussel et Godefroy (Godefroy et al., 2008) émettent l'hypothèse que le déficit aux tests de fluence serait dû au ralentissement cognitif plutôt qu'à un dysfonctionnement exécutif.

5.3 Fluence et traumatisme crânien

Azouvi et ses collaborateurs (Godefroy et al., 2008) indiquent que l'une des plus fréquentes causes de syndrome dysexécutif est le traumatisme crânien. Si le syndrome dysexécutif peut prendre diverses formes et divers degrés de sévérité, le trouble cognitif qui en résulte est souvent global. Les lésions retrouvées le plus fréquemment sont focales (en particulier à la base des lobes frontaux et à la pointe des lobes temporaux), ou diffuses (lésions axonales par étirement ou cisaillement de la substance blanche).

Même si une modification du comportement ou de la personnalité est une des principales conséquences du traumatisme crânien (étude du cas Phineas Gage, qui a eu le crâne perforé par une barre à mine en 1848), on note également des troubles exécutifs de type désinhibition, troubles de la planification ou de la flexibilité mentale, mais également un trouble attentionnel et de la mémoire de travail. En ce qui concerne la fluence verbale, les **performances sont faibles** avec une tendance à donner des **réponses persévératives**, surtout dans le cas de traumatismes crâniens sévères.

5.4 Fluence et pathologies psychiatriques

Certaines pathologies psychiatriques sont également concernées par des troubles de la fluence, en particulier la schizophrénie et la dépression.

5.4.1 Fluence et schizophrénie

Le DSM-5 et le CIM-10 donnent parmi les symptômes caractéristiques de la schizophrénie, outre les idées délirantes ou les hallucinations, une diminution des performances et des relations sociales, une apathie, ou encore une altération du langage, avec une désorganisation du discours (DSM-5) et un discours incohérent, hors de propos et pauvre (CIM-10). Robert et Michel (Godefroy et al., 2008) précisent que les troubles cognitifs sont présents chez environ 80% des patients, avec des déficits d'inhibition, de flexibilité mentale, de planification et de résolution de problèmes, et une faiblesse dans les épreuves de fluences verbales. Selon l'étude de Bockart citée par ces auteurs, le **déficit serait plus important en fluence sémantique qu'en fluence formelle**. Chez ces patients, il a été

relevé une **réduction du nombre de switches en fluence littérale**, et d'autre part, une **réduction du nombre de switches et de la taille des clusters en fluence sémantique**. Cependant, les profils sont très hétérogènes, et les performances sont donc très liées aux atteintes cognitives : plutôt exécutives et/ou plutôt lexicales.

5.4.2 Fluence et dépression

En ce qui concerne la dépression, les troubles exécutifs sont principalement dus à des difficultés de concentration ou à prendre des décisions, une baisse de l'affect, une anxiété, entre autres. Nous notons parfois un ralentissement général (que ce soit au niveau psychomoteur ou cognitif). Robert et Michel (Godefroy et al., 2008) font remarquer que lors des épreuves de fluence verbale, les performances de ces patients sont plus faibles que chez des sujets contrôles, et la **fluence sémantique est plus altérée que la fluence formelle**. Ils précisent que les baisses de productions peuvent être dues à un trouble de la flexibilité mentale. D'autre part, il a été observé (études citées par Gierski et Ergis, 2004) que, bien que le stock sémantique soit préservé, l'évocation lexicale sémantique est perturbée ; cela serait dû à une **réduction du nombre de switches** alors que la taille des clusters serait normale. Il s'agirait donc ici d'un déficit exécutif plutôt que sémantique.

5.5 Fluence et troubles des apprentissages

Thibault et al. (2012), dans la batterie EXALANG, font remarquer qu'un résultat faible dans les épreuves d'évocation lexicale phonémique ne sera pas forcément un indicateur d'un trouble spécifique du traitement phonémique, mais pourra être dû à une difficulté d'évocation ou de catégorisation.

Ces mêmes auteurs (2009) faisaient également remarquer que l'épreuve de fluence sémantique permettait de « *mesurer la fonction d'évocation, tout en appréciant les ressources lexicales de l'adolescent. Au plan qualitatif, il sera intéressant de noter de quelle manière le sujet a procédé* ».

Lors de passation de bilans de langage (oral et/ou écrit) par des patients présentant des troubles des apprentissages, on note souvent une **faiblesse aux épreuves de fluence verbale**, en particulier lors de diagnostic des troubles spécifiques tels que la dyslexie ou la dysphasie.

5.5.1 Fluence et dyslexie

Dehaene (2007) explique que chez les sujets dyslexiques, au moins deux régions sont sous-activées dans le cerveau au moment de la lecture : le cortex temporal latéral et une région temporale inférieure (voie visuelle ventrale). Or, une des fonctions de la région temporale latérale est le traitement des informations phonologiques du langage parlé ; ce serait une anomalie dans la mise en place du réseau neuronal au cours du développement qui expliquerait donc les troubles phonologiques chez les enfants dyslexiques. L'auteur indique que l'on peut observer de nombreuses ectopies (groupes de neurones ayant migré au-delà de leur position normale dans les couches du cortex, faisant ainsi apparaître des « bosses » par endroits), notamment dans les régions linguistiques et visuelles essentielles à la lecture (particulièrement frontal inférieur, gyrus temporal et région temporale inférieure). D'autre part, Dehaene (2007) précise que la littérature fait état d'autres types d'anomalies montrées dans des études sur les cerveaux de patients dyslexiques (diminution de la matière grise dans la région frontale inférieure gauche, entre autres). Touzin dans Rousseau et al. (2013) précise que la **lenteur d'évocation** est en effet une des caractéristiques de la dyslexie, et préconise un travail spécifique sur la rapidité d'évocation avec des listes de mots, par exemple dans un champ sémantique, en faisant des liens entre le visuel, le **sémantique et la phonologie** afin de renforcer l'accès au stock lexical, mais aussi avec un travail de catégorisation.

5.5.2 Fluence et dysphasie

Concernant la dysphasie, Lussier et Flessas (2009) écrivent que « *bien que la dysphasie soit la résultante d'une atteinte neurologique, les techniques d'imagerie cérébrale actuelles ne permettent généralement pas de détecter un site lésionnel anatomique spécifique* ». Des hypothèses sont avancées, notamment concernant des malformations au niveau des microarchitectures cérébrales, mais les recherches ne sont pas encore concluantes (peu d'études sur la dysphasie de l'enfant). Les IRM ont pu parfois mettre en évidence des anomalies morphologiques dans les régions péri-sylviennes gauches, une absence (ou une inversion) d'asymétrie du planum temporale, une diminution volumétrique du cortex préfrontal gauche ou du pariétal supérieur droit. On comprend ainsi pourquoi un des marqueurs linguistiques et comportementaux de la dysphasie est le **trouble de l'évocation lexicale**, comme le précisent Gérard en 1991 et Mazeau en 2003 (Rousseau et al. 2013). Même s'il existe différents profils de dysphasies, il est souvent retrouvé un score très échoué aux épreuves de fluences verbales, tant sémantiques que phonologiques, qui relèvent d'une difficulté « *d'accès au lexique qui n'arrive pas spontanément au niveau de conscience pour être immédiatement disponible* » (Lussier et Flessas, 2009). Elles notent cependant que, malgré des scores en dessous de la moyenne dans les deux modalités, les épreuves de

fluence sémantique sont souvent mieux réussies que la fluence phonologique. Elles expliquent que cette différence est vraisemblablement due à de meilleures capacités d'évocation dans le registre visuel.

Nous le voyons, de nombreuses pathologies d'étiologies diverses (dégénératives ou non, acquises ou développementales) peuvent être concernées par des altérations de la fluence verbale. De plus, en fonction des pathologies, les résultats aux tâches d'évocation lexicale peuvent être différents, la fluence catégorielle étant plus ou moins touchée que la fluence formelle. Il est donc indispensable de pouvoir mesurer ces performances, et nous allons donc décrire les types de mesure, les tests et les étalonnages existants.

6. Mesures de la fluence

Pour chercher un mot, on peut procéder grâce à différents indices : le premier son ou la première lettre, le champ lexical (animaux) ou la fonction (tout ce qui sert à faire la cuisine) par exemple. Il est donc généralement proposé dans les tests de fluence de chercher des mots avec des indices principalement sémantiques et formels, en un temps donné. Les tests d'évocation lexicale sont conçus pour faire produire au patient le plus de mots possibles répondant à des consignes strictes et les principes des conditions de passation sont généralement les suivants :

- L'épreuve est **limitée dans le temps** (2 minutes pour le test de Cardebat, 90 secondes pour le MT-86 ou le BDAE, 4x15 secondes pour le test d'Isaacs, par exemple)
- Le patient doit **produire le plus de mots possibles**, soit dans une **catégorie** donnée (animaux, fruits, meubles, villes, etc.), soit commençant par une **lettre** donnée (P, R, V, entre autres)
- **Consignes particulières** : généralement, le patient doit éviter de répéter et ne doit pas donner de noms de la même famille (cane et canard, par exemple), ni de noms propres

La fluence sémantique permet d'évaluer la capacité à trouver des mots dans une catégorie sémantique précise, par exemple les animaux. La fluence formelle permet quant à elle, d'évaluer la capacité à chercher des mots commençant par une lettre. Les résultats obtenus à ces 2 épreuves par un même sujet sont en général différents, avec des scores en fluence sémantique supérieurs à la fluence formelle, mais il peut y avoir des différences de scores

entre deux catégories sémantiques ou deux lettres. Gierski et Ergis (2004) précisent qu' « *on considère classiquement les tâches sémantiques comme étant plus simples à réaliser par le sujet que les tâches littérales* », en se référant à l'étude effectuée par Cardebat et al. en 1990. Ces résultats peuvent s'expliquer par plusieurs facteurs.

7. Facteurs influençant les scores aux étalonnages

Lors de l'établissement des étalonnages de ces tests de fluence verbale, des études ont permis de montrer les facteurs expliquant les différences de scores entre les individus (études réalisées sur des sujets sains). On voit ainsi que l'âge, le sexe et le niveau socio-culturel jouent un rôle dans les performances, mais on note aussi des différences de productions entre la fluence sémantique et formelle.

7.1 Formelle / sémantique

Marquier-Morin et Lahille (1989) ont montré que les performances globales sont fortement **corrélées au nombre d'items présents dans chaque critère** : il y a plus de mots qui commencent par P que par V, et les résultats des tests sont meilleurs avec la lettre P qu'avec la lettre V ; de même pour les animaux par rapport aux meubles. D'autre part, lors des épreuves dont le stock d'items est important, la production des réponses se distille régulièrement dans les deux minutes allouées. Alors que pour les épreuves dont le nombre d'items est plus limité, la quasi-totalité des réponses est donnée lors des 30 premières secondes. En effet, les mots produits dans les premières dizaines de secondes sont le résultat d'une **recherche automatique** (mots fréquemment utilisés et disponibles rapidement), alors que pour trouver les mots moins fréquents, l'individu doit entamer une **recherche plus contrôlée** avec la mise en place d'une stratégie éventuellement. Cela expliquerait qu'un stock lexical réduit entraînerait un processus de recherche contrôlée moins efficace, et donc un score plus faible.

D'autres facteurs peuvent expliquer les différences de performance entre la fluence sémantique et formelle car, si les 2 fluences testent les capacités exécutives et la capacité d'accès au stock sémantique, les **circuits neuronaux sont différents** (comme il a été démontré au paragraphe 3.1).

7.2 Sexe

En analysant le tableau des étalonnages de Cardebat (Annexe 1), on voit distinctement que les performances des femmes dans la catégorie des fruits sont meilleures que celles des hommes, quels que soient le niveau socio-culturel et l'âge. En revanche pour les animaux, les résultats sont nuancés : les performances des hommes sont meilleures chez les hommes avant 65 ans de niveau socio-culturel 1, par rapport aux femmes, mais les résultats s'inversent chez les femmes niveau socio-culturel 2 de plus de 50 ans par rapport aux hommes. Pour ce qui concerne la fluence formelle, on constate que pour le niveau socio-culturel 1, les hommes ont de meilleures performances que les femmes, alors que c'est l'inverse chez les individus de niveau socio-culturel 2.

D'autre part, Zarino et al. (2014) ont montré avec leur étude que les hommes obtenaient de meilleurs scores dans la catégorie des marques de voitures, alors qu'elles n'ont pas montré de différences significatives pour les catégories des animaux et des fruits.

7.3 Age

Rousseau et Gagnol (Mazaux et al., 2014) indiquent que les **performances en fluence sémantique et formelle se dégradent avec l'âge**. Lorsqu'on observe les tableaux des scores au test de Cardebat, nous voyons qu'en fluence sémantique, après une amélioration des performances pour les catégories 50-65 ans, on assiste à une baisse significative chez les sujets les plus âgés, à un niveau plus bas que les 30-45 ans. On pourrait donc penser que l'âge (jusqu'à 65 ans) est un facteur favorisant, car le sujet a d'autant prolongé son contact avec les mots. En revanche, le ralentissement du fonctionnement cérébral après 65 ans se traduit par une **baisse d'efficacité des processus de recherches automatique et contrôlée ainsi qu'un manque d'habileté à changer de stratégie (flexibilité mentale)**.

D'autre part, le **nombre d'erreurs**, quelle que soit l'épreuve, augmente avec l'âge, en particulier la répétition de mots, ce qui pourrait s'expliquer par une diminution de la performance de la mémoire de travail.

7.4 Niveau socio-culturel

Le niveau socio-culturel a également un rôle important. En effet, d'après l'étude de Ruff et al. (cités par Gierski et Ergis, 2004), les connaissances verbales plus étendues permettent de sélectionner des mots à partir d'un échantillonnage plus important ; on constate que globalement, les **performances des sujets sains de niveau socio-culturel élevé sont meilleures.**

Lors de la réalisation d'une normalisation d'un test de fluence sur les substantifs et les verbes (Boulangue et Martin, 2014), les étudiantes ont noté des **différences significatives en nombre de mots produits** entre le niveau socio-culturel 1 (jusqu'à la terminale) et 3 (au moins bac + 3), et entre le niveau 2 (du baccalauréat jusqu'à bac + 2 inclus) et 3, mais pas entre les niveaux 1 et 2.

Rousseau et Gatignol (Mazaux et al., 2014) rapportent qu'une étude menée sur 350 sujets contrôles a démontré un effet significatif du niveau d'études sur les performances dans les deux tâches de fluence verbale. Ratecliff et al. (cités par Gierski et Ergis, 2004) rapportent, quant à eux que des effets plus importants du niveau d'études ont été notés sur la fluence littérale que sémantique. Cependant, les **performances en fluence sémantique seraient très dépendantes de variables autobiographiques**, de l'exposition à tel ou tel champ sémantique et des expériences individuelles (voyages, profession, etc.).

Des études sur la fluence verbale chez des sujets brésiliens (de langue portugaise) et suédois (cités par Zarino et al., 2014) ont montré que le niveau d'études avait un **impact sur l'amplitude des clusters et le nombre de switches**. Ces résultats sont à rapprocher du mémoire de Boulangue et Martin (2014) montrant des différences concernant le nombre et la taille des clusters : les sujets sains de niveau d'études supérieures à la licence produisent plus de clusters, et de taille plus grande. Concernant les switchings, il a également été retrouvé des différences significatives entre les sujets de niveau socio-culturel 1 et 3, et 2 et 3.

8. Tests de fluence existants

Les tests de fluence verbale sont largement utilisés, que ce soit isolément, ou dans des batteries. En effet, ils permettent d'avoir rapidement et facilement un repère fiable (moyenne, écart type).

En France, les **tests isolés** les plus utilisés sont le test de Cardebat et le Set Test d'Isaacs (réétabli en français par Thomas-Antérion). Le test de Cardebat est cependant plus complet, car il permet d'évaluer les capacités d'accès au lexique sémantique (fluence sémantique avec les noms d'animaux, de fruits et de meubles) et formel (fluence littérale ou formelle) avec les mots qui commencent par les lettres P, R et V. Alors que le test d'Isaacs ne propose que l'évaluation de la fluence sémantique, même si le nombre de catégories est plus important (animaux, fruits, meubles, articles de supermarché). Le test de Cardebat, malgré un étalonnage ancien (1990) et des tranches d'âge manquantes (entre 45 et 50 ans et de 65 à 70 ans), reste néanmoins une référence encore aujourd'hui.

Les tests de fluence sont également proposés dans des **batteries complètes**, orientées vers des pathologies ciblées : le MT86 ou le GREFEX, utilisés dans les pathologies neurologiques (AVC, ou maladies neurodégénératives), BCcogSEP (inspirée de la BRB-N, validée en français pour la SEP), L2MA pour les pathologies du langage écrit ou oral (dyslexies, dysphasies), ou encore la BREF utilisée plus spécifiquement par les neuropsychologues pour les fonctions exécutives. Balland et Courtade-Jouanicq (2014) ont largement décrit et comparé les tests de fluence présents dans des batteries. En voici un récapitulatif (non exhaustif) :

Langage oral	Langage écrit	Neurologie
L2MA (phonémique et sémantique)	L2MA (phonémique et sémantique)	BEC96 (sémantique)
ELOLA (sémantique)	BALE (phonémique)	BCcogSEP (phonémique et sémantique)
EDA (sémantique)	DENO48 (phonémique et sémantique)	GREFEX (phonémique et sémantique)
		GREMOTS (verbes, phonémique et sémantique)
		MT86 (sémantique)
		HDAE (phonémique et sémantique)
		MEC (phonémique, sémantique et évocation libre)
		BREF (phonémique)
		TLE (fluence induite)

Figure 14 Récapitulatif non exhaustif des batteries comprenant des épreuves de fluence verbale

De plus, des tests de fluence se retrouvent dans des **batteries informatisées**, dont voici les principales :

LANGAGE ORAL / LANGAGE ECRIT	NEUROLOGIE
BILO (fluence de noms et verbes)	i-MEC (P, vêtements et épreuve libre)
ALOE (fluence phonémique)	BIA (fluence sémantique et phonémique)
EXALANG (fluence phonémique et sémantique)	

Figure 15 Bilans informatisés existants sur le marché français

Généralement, les tests de fluence inclus dans les batteries sont plus courts (1 minute 30 dans le MT-86 au lieu de 2 pour Cardebat) et proposent une seule lettre ou une seule catégorie ; ils permettent d'avoir une idée rapide des déficits, qui peuvent ensuite être analysés plus finement grâce à d'autres épreuves (attention, inhibition, etc.) proposées dans les mêmes batteries.

Ces tests de fluence verbale sont utilisés afin d'avoir des **résultats quantitatifs** et qualitatifs. Les mots produits sont soit écrits par l'orthophoniste ou le professionnel qui fait passer l'épreuve dans une grille (tests manuels) et les résultats sont calculés à la main, soit l'orthophoniste clique pour définir si le mot est correct ou non dans le cas d'épreuves informatisées, et le nombre de mots est comptabilisé par l'ordinateur. On peut ainsi connaître le nombre de mots corrects produits (correspondant à la consigne), le nombre d'erreurs (non-respect de la consigne), ou encore le nombre de mots répétés. Ces résultats sont alors comparés à l'étalonnage, afin de calculer l'écart par rapport à la moyenne d'âge et de niveau socioculturel. On a ainsi un état des lieux initial et une base de comparaison pour suivre l'évolution de la prise en charge du patient au long cours.

Sur le **plan qualitatif**, à ce jour aucun test existant ne propose d'analyses étalonnées. Or, comme nous l'avons vu dans les chapitres précédents, il est intéressant de connaître la stratégie utilisée par le patient (fréquence des mots, types et nombre de regroupements, utilisation de sous-catégories (subordonnés), les mots dérivés, laps de temps entre les réponses, etc.). Ces informations sont tout aussi importantes que les mesures quantitatives, mais ne sont pas quantifiées ni étalonnées à ce jour. Du point de vue de la mise en œuvre du projet thérapeutique, on comprend aisément leur utilité afin de s'appuyer sur ce qui fonctionne, proposer des exercices pour développer ce qui fonctionne moins, et mieux cibler

les fonctions exécutives à travailler (manque d'inhibition, d'initiation, de mise en place de stratégies, par exemple).

De façon générale, les étalonnages disponibles à ce jour sur les tests de fluence verbale, en particulier celui de Cardebat ou d'Isaacs ont été réalisés il y a plus de 20 ans. Zarino et al. (2014) avec leur nouvel étalonnage, font par ailleurs remarquer des modifications des résultats par rapport aux normes plus anciennes. Elles pensent que ces différences sont probablement dues aux changements intervenus dans le monde (allongement de l'espérance de vie, mondialisation, etc.) ces 20 à 30 dernières années ; il est donc absolument nécessaire de **réactualiser les étalonnages** de ces tests.

Enfin, si l'on souhaite que les tests de fluence fournissent des analyses les plus intéressantes et exhaustives sur le fonctionnement exécutif du patient, il conviendrait de contrôler les points suivants :

- Nombre de mots produits : corrects, nombre et types d'erreurs
- Nombre de mots produits par intervalles de temps
- Fréquence des mots en fonction du temps
- Temps d'initiation de la tâche
- Temps de latence (entre et à l'intérieur des clusters)
- Clusters : types (sémantiques, phonologiques, grammaticaux, morphologiques), taille, nombre
- Switches : nombre, mots isolés
- Paraphasies (sémantiques, phonologiques, verbales), néologismes, etc.
- Facilitations (ébauche phonologique, sémantique)

. Problématique et hypothèses

I. Problématique

La faisabilité du logiciel EVOLEX a été démontrée l'an dernier grâce au mémoire d'Aurélié Courtade-Jouanicq et Ophélie Balland soutenu en juin 2015 et intitulé *Etude de faisabilité d'un logiciel de reconnaissance de vocale adapté à des tâches d'évocation lexicale*. Le logiciel a été créé et modifié en équipe avec des enseignants informaticiens-chercheurs de l'IRIT et suite aux premiers tests, il s'avère que la reconnaissance vocale (nombre de mots correctement identifiés) n'est efficace qu'à environ 30%. Les pistes d'amélioration proposées par les étudiantes sont de plusieurs ordres :

- amélioration de la reconnaissance vocale, grâce à l'adaptation de modèles acoustiques, au codage temporel, à l'utilisation de modèles de langage incluant les déterminants et prenant en compte la grammaire, à la normalisation des conditions de passation
- améliorations ergonomiques, avec notamment l'affichage de la consigne à l'écran
- amélioration des analyses, avec la prise en compte de la fréquence des mots pour le lexique de chaque épreuve.

Le but de ce logiciel est de pouvoir apporter au thérapeute des analyses tant quantitatives que qualitatives des productions du patient. Mais ces analyses se calculent à partir des mots identifiés par le logiciel. Or, si le taux de reconnaissance vocale est trop faible, nous perdons le bénéfice de ces analyses faites automatiquement, et par là-même une plus-value essentielle du logiciel. Il semble ainsi évident que l'amélioration de la reconnaissance vocale est la pierre angulaire de la poursuite de la mise au point du logiciel.

Nous avons orienté notre travail sur deux axes principaux. Le premier axe consiste à contribuer à l'augmentation du taux de reconnaissance vocale. Le second, plus écologique, consiste à connaître les avis a priori des orthophonistes sur ce type de logiciel, puis à effectuer des tests en situation réelle avec des patients, afin de voir quels seraient les atouts et les points faibles du logiciel.

Nous nous sommes donc interrogées sur les points suivants : comment pouvons-nous participer à l'amélioration de la reconnaissance vocale ? Comment savoir si les analyses quantitatives et qualitatives effectuées par le logiciel sont pertinentes pour une pratique orthophonique quotidienne ? Le logiciel apporte-t-il une réelle plus-value aux orthophonistes ? Quelles sont les perspectives de développement du logiciel ?

II. Hypothèses

L'étayage théorique, allié à ce qui a déjà été réalisé sur le logiciel ainsi que les interrogations de notre problématique, nous conduisent vers les hypothèses suivantes.

- **Hypothèse 1** : la reconnaissance vocale a pu être améliorée depuis juin 2015

Nous avons vu que sans une reconnaissance vocale performante, le logiciel n'a plus de raison d'être. Il nous a semblé intéressant de poser des sous-hypothèses pour distinguer les différents facteurs pouvant jouer un rôle dans cette amélioration. Et dans un second temps, nous nous sommes demandées si cette reconnaissance vocale serait aussi efficace en situation écologique. Ce questionnement peut se poser dans deux domaines : l'un plutôt technique, et l'autre plutôt d'ordre écologique.

- Hypothèses techniques :
 - **H1-1** : la reconnaissance vocale a pu être significativement améliorée par les enregistrements de « La bise et le soleil » et d'un extrait du « Petit Prince »
 - **H1-2** : l'amélioration du système d'exploitation du logiciel et les enregistrements ont permis une amélioration de la reconnaissance vocale
 - **H1-3** : la correction des fichiers permet l'amélioration de la reconnaissance vocale par réapprentissage
- Hypothèse écologique :
 - **H1-4** : la reconnaissance vocale fonctionne bien avec des patients.

- **Hypothèse 2** : ce type de logiciel est susceptible d'intéresser les orthophonistes
- **Hypothèse 3** : en situation écologique, le logiciel est pertinent et efficace, et apporte une plus-value par rapport aux passations des tests de fluence manuelles

- **H3-1** : le logiciel apporte un gain de temps aux orthophonistes
- **H3-2** : les orthophonistes « testeurs » pensent que le logiciel apporte des analyses quantitatives et qualitatives pertinentes pour la clinique (aide au diagnostic et à la définition des axes thérapeutiques)

. Méthodologie

Le cœur de notre mémoire repose ainsi sur l'amélioration de la reconnaissance vocale et sur le test en situation écologique du logiciel. Afin de répondre à notre problématique et à nos hypothèses, nous avons mis en place des méthodes spécifiques pour les deux grands domaines sur lesquels nous avons travaillé. Mais avant de détailler notre méthodologie, il est essentiel de présenter le logiciel en cours d'élaboration.

I. Présentation du logiciel

1. EvoLex « version 2015 »

EvoLex est né de la volonté de créer un outil permettant d'affiner et de faciliter les analyses des productions du patient lors des tâches de fluence. Une première version a été présentée en juin 2015, permettant la passation des tests et l'analyse des données.

Cette première version a permis de poser les bases de l'interface, des fonctionnalités du logiciel, des éléments demandés par les orthophonistes pour les analyses et l'ergonomie.

1.1 Principes de fonctionnement

Le **patient est enregistré avec un microphone** pendant la réalisation des tâches de fluence, et en même temps, l'orthophoniste appuie sur A si la réponse est correcte ou sur P si la réponse est jugée incorrecte (suivant les consignes du test de Cardebat, voir Annexe 1).

Le logiciel contient des « dictionnaires » correspondant à chaque épreuve de fluence : un lexique des animaux, des fruits, les mots du dictionnaire pour les mots commençant par R et par V. Tous les mots de la base de données ont été écrits en phonétique informatique, en tenant compte des différentes façons de prononcer un même mot. Par exemple, le mot « rebattre » est transcrit :

rebattre	rr eu bb aa tt rr	rr eu bb a tt rr ee	rr bb aa tt rr	rr bb aa tt rr ee
----------	-------------------	---------------------	----------------	-------------------

Figure 16 Ecran de correction et validation de la transcription

fluence permettent de se rendre compte des stratégies de recherche de mots, ce qui enrichit l'observation clinique, le diagnostic et fournit une base dans le cadre du suivi du patient.

1.2.1 Analyses quantitatives

Ces analyses correspondent aux **normes de Cardebat** et contiennent :

- Le nombre et le pourcentage de mots produits répondant strictement à la consigne
- Le nombre et le pourcentage de répétitions
- Le nombre et le pourcentage de mots de la même famille (canard, cane, caneton, par exemple)
- Le nombre et le pourcentage d'erreurs (mots ne correspondant pas dans la consigne)

Ces résultats sont mis en rapport avec les étalonnages de Cardebat et cela permet d'avoir des résultats en termes d'écart type. Ils sont présentés comme sur le schéma suivant :



Figure 18 Résultats quantitatifs selon les normes de Cardebat

1.2.2 Résultats qualitatifs

Comme nous l'avons vu sur un plan théorique, les stratégies mises en place par le patient pour trouver le maximum de mots sont intéressantes à observer et à pouvoir quantifier afin d'avoir une base de comparaison et un profil à un instant T. Les indicateurs à étudier ont été regroupés en 3 grandes catégories : temps de latence et fréquence des mots, clusters et switches, et processus cognitifs.

➤ Temps de latence et fréquence de mots

Le **temps de latence** est un indicateur important : en effet, nous avons vu que dans certaines pathologies (maladie de Parkinson entre autres), **l'initiation de la tâche** est très longue. D'autre part, nous avons également noté que les temps de latence pouvaient augmenter avec l'âge.

Concernant la **fréquence des mots**, il a été montré que dans les 15 premières secondes du test, les patients trouvaient des mots fréquents, puis petit à petit, des mots moins fréquents, qui demandent plus de temps et la mise en place de stratégies. Ces analyses sont présentées avec le diagramme suivant :



Figure 19 Graphique de présentation des temps de latence et fréquence de mots

L'échelle de 0 à 100 en abscisse représente le pourcentage de fréquence des mots prononcés, et en ordonnée se trouve l'échelle du temps.

➤ Clusters et switches

Une recherche efficace de mots se caractérise par des clusters de grande taille et des switches rapides. Les analyses suivantes permettent de **visualiser les regroupements** réalisés par le patient et les liens entre les mots (sémantique, phonémique, morphologique, grammatical).

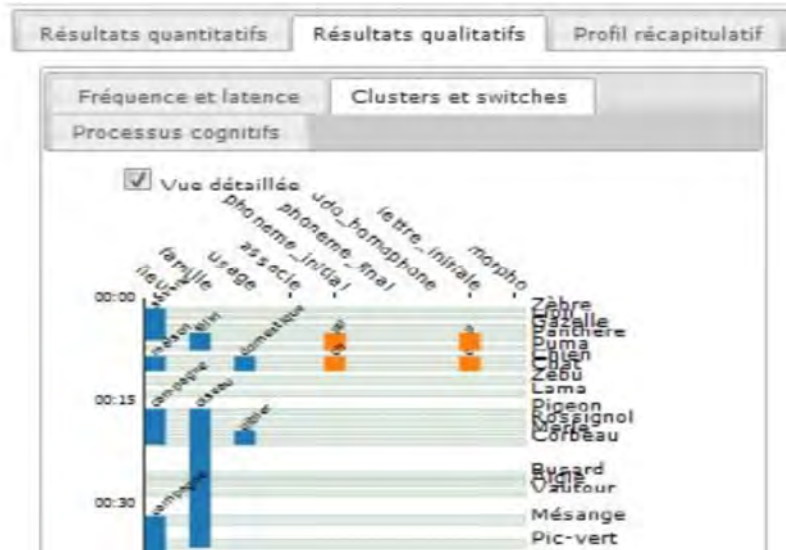


Figure 20 Vue détaillée des regroupements effectués par le patient

Sur ce schéma, les regroupements matérialisés en bleu montrent un lien sémantique, et les clusters matérialisés en orange, un lien formel (phonémique ou alphabétique).

Les analyses qualitatives sont ensuite synthétisées et représentées par des schémas, avec les résultats chiffrés suivants :

- Le nombre et le pourcentage de mots regroupés et isolés
- Les types de regroupement effectués (sémantique, morphologique, etc.)
- Les processus de stratégies de recherche, avec le nombre de clusters et de switches

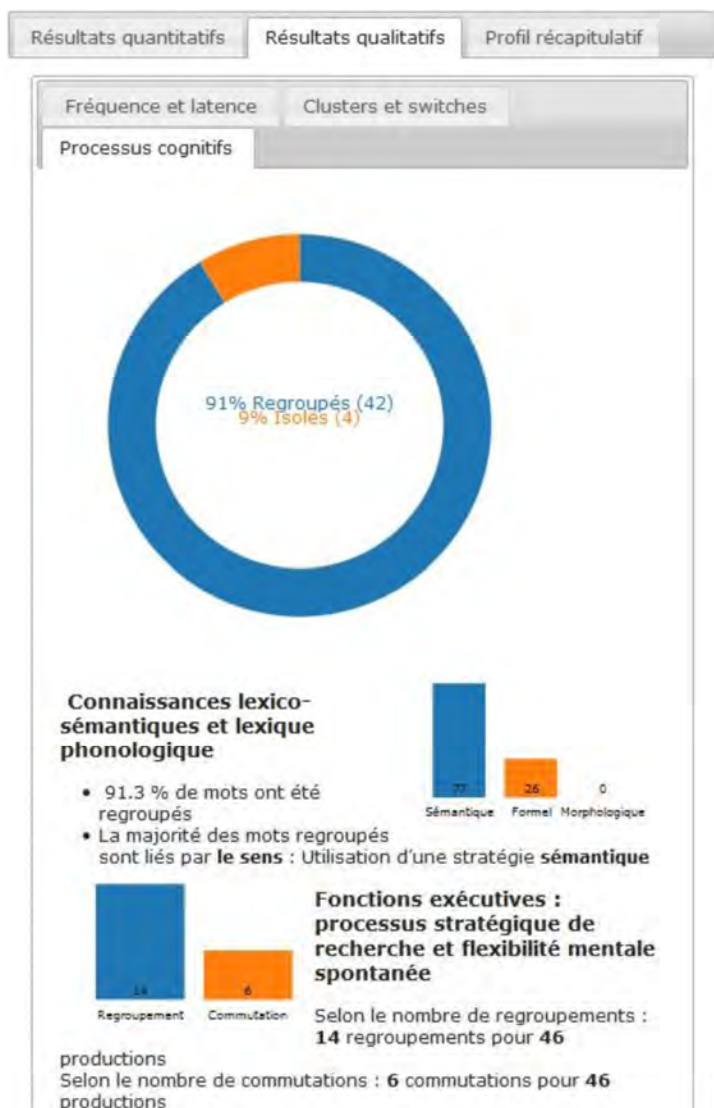


Figure 21 Vue détaillée des processus cognitifs effectués par le patient (regroupements, types de regroupements, flexibilité mentale)

➤ Profil récapitulatif

Le dernier écran du logiciel permet de montrer le **profil récapitulatif du patient**, présenté sous forme de cible. Celle-ci reprend les principaux indicateurs vus précédemment. Le point orange représente les performances du patient.

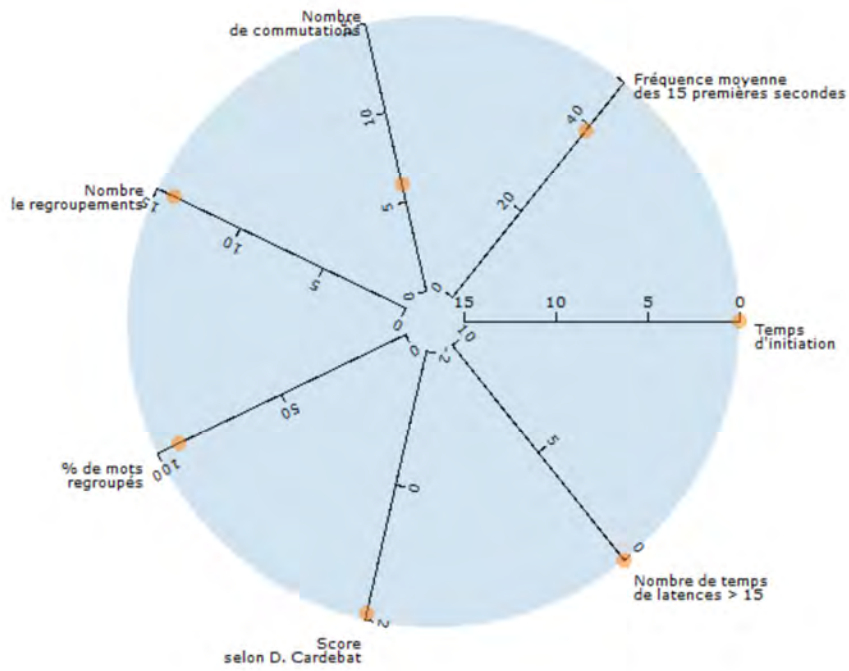


Figure 22 Profil récapitulatif du patient

1.3 Arborecence de la « version 2015 »

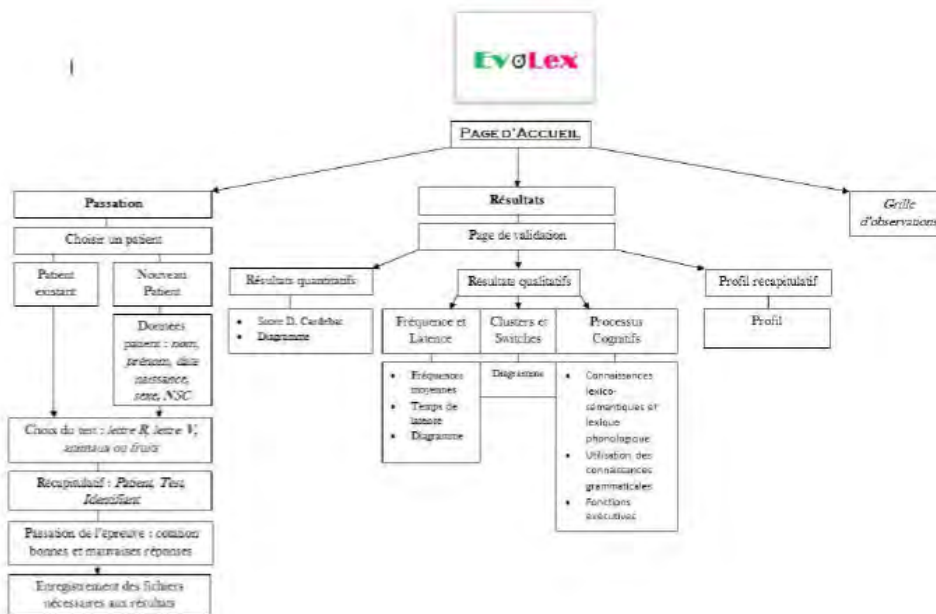


Figure 23 Architecture EvoLex – Version 2015

2. EvoLex « version 2016 »

En septembre 2015, l'équipe des informaticiens-enseignants-chercheurs de l'IRIT a remis une nouvelle version du logiciel. Des améliorations ont été apportées par rapport à la version précédente.

2.1 Modifications du logiciel

Un des grands changements de cette nouvelle version a été le stockage des informations (fichiers audio et codage temporels) sur un **serveur de l'IRIT**. Cela a permis d'améliorer grandement le traitement des informations. D'une part, toutes les informations sont désormais centralisées et conservées au même endroit, créant de ce fait une base de données globale. Un réapprentissage de la reconnaissance vocale peut ainsi s'envisager à partir de matériel plus conséquent et sur un plus long terme. D'autre part, la puissance de calcul est bien plus élevée que sur un ordinateur individuel et on peut espérer gagner en vitesse de traitement et en performance.

La communication entre le serveur et les ordinateurs individuels se fait par une **connexion Internet**. Ainsi, les informations recueillies sur nos ordinateurs respectifs (fichiers audio et codage temporel) sont envoyées au fur et à mesure sur le serveur, et les analyses quantitatives et qualitatives sont ensuite disponibles en ligne. Lors des enregistrements des tâches de fluence, en l'absence de réseau wifi, on peut envoyer les fichiers en différé par une synchronisation ultérieure.

De plus, cette dernière version permet de différencier les utilisateurs du logiciel : chaque utilisateur se voit attribuer un identifiant et un mot de passe personnels. Il dispose ainsi uniquement des fichiers dont il a procédé lui-même à l'enregistrement, dans le but indispensable de garantir l'anonymat des patients.

Un champ libre a été ajouté au moment de valider un fichier, afin d'ajouter des commentaires (qualité de la voix, problèmes rencontrés par la personne enregistrée, problèmes techniques apparus pendant l'enregistrement, etc.).

2.2 Changement de système

Début 2016, l'équipe de l'IRIT a proposé de changer de système d'exploitation du logiciel : une **version « Pocket Sphinx »** a été installée, version plus légère que la précédente, mais plus performante pour le type de fichiers que nous utilisons.

Cette version nous a été présentée et mise en place sur nos ordinateurs courant février 2016. C'est à partir de celle-ci que nous avons effectué les corrections de fichiers.

2.3 Arborescence du logiciel « version 2016 »

En fonction de toutes ces modifications, l'arborescence du logiciel a changé et devient la suivante :

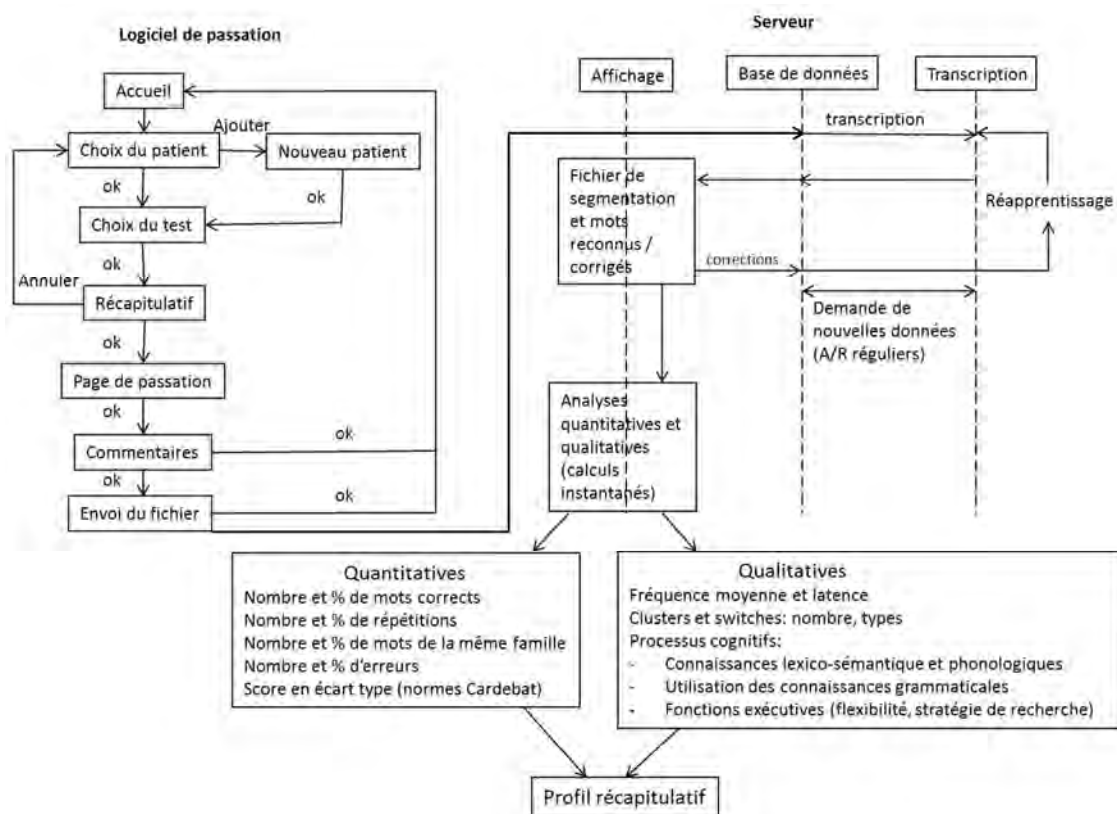


Figure 24 Architecture EvoLex – Version 2016

II. Méthodologie utilisée pour l'amélioration de la reconnaissance vocale

La condition sine qua non au bon fonctionnement et à la pertinence des analyses du logiciel est l'amélioration du taux de reconnaissance vocale. Comme cela a été détaillé dans la partie théorique, la reconnaissance vocale fonctionne avec l'aide d'un modèle acoustique et d'un modèle de langage fournis au logiciel, et peut être améliorée par des apprentissages.

Ainsi, pour que cette reconnaissance soit plus efficace, nous avons procédé à des **enregistrements de volontaires sur des tâches de fluence** (sémantique et phonémique) d'une part, et sur l'enregistrement de la lecture de 2 textes, d'autre part.

Pour la réalisation des enregistrements, nous sommes passées par différentes étapes : définition de notre échantillon, définition du matériel adéquat, tests du matériel et du logiciel, rédaction d'un protocole précis (procédure, conditions d'enregistrements), et enfin phase d'enregistrements proprement dite.

1. Enregistrements de tâches de fluence et de textes lus

1.1 Notre échantillon

1.1.1 Critères de sélection

Lorsque nous avons réfléchi à la constitution de notre échantillon, les premiers critères de sélection devaient mettre tous les atouts du côté du logiciel pour améliorer la reconnaissance vocale. Ainsi, nous avons défini les critères suivants :

- Avoir plus de 18 ans. En effet, le logiciel sera utilisé, dans un premier temps, auprès de personnes adultes atteintes de maladies neurologiques ou neurodégénératives.
- Avoir le français pour langue maternelle. Cela permet d'avoir des participants sans accent étranger, ce qui pourrait perturber la reconnaissance vocale, et dont l'exposition à la langue depuis la naissance et la maîtrise de la langue permettent d'augmenter les chances d'avoir des performances satisfaisantes aux tâches de fluence.

- Ne pas avoir eu d'atteintes neurologiques ou psychiatriques entraînant un trouble du langage. Bien entendu, dans notre objectif expérimental, nous devons avoir des sujets les plus sains possibles, et éviter les troubles de mémoire, les troubles arthriques, le bégaiement, ou une dysphonie, par exemple.

De plus, afin de répondre aux contraintes acoustiques, nous avons soin d'enregistrer autant d'hommes que de femmes.

Dans un second temps, nous nous sommes interrogées sur la faisabilité d'une normalisation des analyses de clusters et de switches à partir de notre échantillon. Pour qu'une normalisation soit valide, l'échantillon doit être le plus représentatif possible de la population générale, et doit donc contenir un nombre équitable d'au moins 30 personnes dans chaque tranche d'âge, sexe et niveau socio-culturel. Ce qui équivaldrait à 540 personnes dans notre cas.

Les conditions matérielles de réalisation de ce mémoire ne nous ont pas permis d'envisager cette option, nous avons donc recruté les participants parmi les personnes de notre entourage, en leur envoyant notre « lettre de recrutement » (Annexe 2) exposant notre étude, nos besoins et les critères d'inclusion et d'exclusion.

1.1.2 Vers une normalisation ?

L'objectif serait de pouvoir étalonner les résultats des analyses concernant les clusters et les switches, les temps d'initiation et de latence, et la fréquence des mots grâce à l'enregistrement de sujets sains sur les tâches de fluence. A ce jour, aucune normalisation à grande échelle n'a été réalisée.

Le serveur permet de stocker les fichiers audio et les codages temporels enregistrés. Dans une perspective de normalisation future, nous avons réalisé les enregistrements dans des conditions de test « authentiques », avec le respect strict des consignes et dans des conditions d'enregistrements optimales (silence, matériel approprié). D'autre part, nous avons créé un fichier qui permet de recenser les participants de façon anonyme, dont voici un extrait :

Code	Age	Etudes	NSC	Date Enreg
mdlf01	66	17	3	21/09/15
mdlf02	39	14	3	24/09/15
mdlf03	43	9	1	27/09/15
mdlf04	64	15	3	02/10/15
mdlf05	46	16	3	03/10/15

Mdl = initiales de l'orthophoniste

F ou h = femme ou homme

01 = numéro de passage

Figure 25 Anonymisation des participants

Les futurs enregistrements pourront s'ajouter aux nôtres et permettre, à terme, lorsque les critères quantitatifs et qualitatifs de l'échantillon seront atteints, de proposer une normalisation pertinente.

1.2 Matériel requis

Nous avons utilisé nos ordinateurs portables personnels (PC), avec un système d'exploitation Windows.

Sur les conseils des enseignants-chercheurs-informaticiens, nous avons acquis un micro-casque Sennheiser (TM), qui possède les caractéristiques suivantes :

- la réponse de fréquence est comprise entre 42 Hz et 17000 Hz, ce qui couvre largement le champ de fréquences de la voix humaine
- le micro contient un réducteur de bruit
- le casque permet d'aider au positionnement du micro et à sa stabilité
- le micro est toujours positionné au niveau du menton (en dehors du flux d'air).



1.3 Premiers tests et protocole

1.3.1 Premiers tests

Avant de faire les enregistrements « à grande échelle », nous avons fait des tests. Cela a permis de nous familiariser avec l'outil, de relever des bugs, des problèmes de manipulation, des difficultés dues aux ordinateurs eux-mêmes (mémoire vive insuffisante). Nous avons ainsi pu faire part de toutes ces remarques à l'équipe de l'IRIT qui a tout fait pour que le logiciel soit exploitable de façon simple et efficace.

1.3.2 Protocole de passation

Cette prise de contact avec le logiciel nous a permis de rédiger le protocole de passation (Annexe 3). Nous nous sommes largement inspirées de celui écrit par nos prédécesseurs, tout en tenant compte des modifications qui ont été apportées depuis.

Ce protocole explique toutes les étapes à respecter pour que les enregistrements se fassent dans les conditions techniques et orthophoniques imposées par le logiciel.

1.4 Tâches de fluence et lecture de textes

Afin de pouvoir contribuer à l'amélioration de la reconnaissance vocale, nous avons enregistré des sujets sains pendant qu'ils réalisaient des tâches de fluence, puis des tâches de lecture de texte.

Nous avons décidé d'enregistrer les participants sur les tâches de fluence en situation d'évaluation, pour plusieurs raisons :

- faire passer de nombreux tests de fluence et **aiguiser ainsi notre regard clinique sur les performances des sujets**. En effet, nous avons pu nous rendre compte de la complexité de ces tâches pour certains sujets, nous avons pu les questionner sur leur manière de fonctionner pour trouver des mots et les difficultés rencontrées.
- dans un souci de **normalisation future**, il fallait que les tests soient passés dans des conditions optimales et « réelles », comme pour des patients.
- faire lire des listes de mots aux participants aurait peut-être été plus performant pour l'amélioration de la reconnaissance vocale, mais on aurait alors perdu le côté « authentique » de la situation de test, d'une part, et d'autre part, la diction dans le cas de la lecture est moins naturelle qu'en production spontanée. Il aurait donc **manqué des éléments suprasegmentaux** essentiels pour que le logiciel dispose de données les plus proches possibles de situations écologiques, avec des patients.

De plus, nous avons demandé aux participants de lire deux petits textes : « La bise et le soleil » et un extrait du « Petit Prince » (Annexe 4). En effet, ces deux textes présentent la particularité d'être **phonétiquement équilibrés**, c'est-à-dire qu'ils contiennent quasiment tous les phonèmes du français, et les différentes façons de les prononcer suivant la coarticulation (la prononciation d'un phonème peut changer en fonction des phonèmes précédents et suivants). Ainsi, la lecture de ces 2 textes devrait permettre de contribuer à l'apprentissage du logiciel et d'aider à l'amélioration de la reconnaissance vocale.

1.5 Enregistrements des sujets sains

1.5.1 Période d'enregistrements

Nous avons donc procédé à ces enregistrements, après avoir fait des tests et apporté quelques modifications à l'interface (possibilité de faire une pause puis de reprendre

l'enregistrement, possibilité d'ajouter des commentaires après avoir validé le fichier, etc.). Presque tous les enregistrements ont été réalisés entre fin septembre et décembre 2015.

1.5.2 Conditions d'enregistrements

Nous avons prêté une **attention particulière aux conditions d'enregistrement** : environnement calme (pièce silencieuse à l'écart du bruit de la maison ou de la rue), dans le but de recueillir un signal acoustique de très bonne qualité. Nous avons ainsi pu expérimenter l'outil de nombreuses fois, et le maîtrisons bien.

1.5.3 Un apprentissage technique et « orthophonique »

Outre la maîtrise du logiciel sur un plan technique, nous avons retiré des **éléments humains et cliniques** lors des passations de ces tests. En effet, très souvent, les participants nous ont expliqué comment ils s'y prenaient pour trouver les mots. Certains, pour la catégorie des fruits, expliquaient qu'ils s'imaginaient chez le primeur, ou en train de faire des confitures. D'autres revivaient leurs voyages en Afrique, en Asie ou leurs randonnées dans les Pyrénées ou les Alpes, ou passaient en revue continent après continent pour donner des noms d'animaux. Certaines observations – toutes subjectives qu'elles fussent – nous ont surprises : par exemple, nous avons constaté que lorsque nous avons enregistré certains couples, ce sont souvent les mêmes mots qui revenaient en début de tâche.

La plupart des participants a fait état de la difficulté de ces tâches, en particulier pour la fluence formelle. En effet, ils disaient pouvoir s'organiser facilement pour trouver les noms d'animaux ou de fruits, mais pour les mots commençant par R ou V, ils avaient parfois des difficultés à mettre en place une stratégie, et à s'y tenir. Nous avons constaté que les personnes utilisant quotidiennement l'écrit (journaliste, éditeur, professeurs, personnes faisant des mots croisés ou jouant au scrabble, par exemple) avaient cependant plus de facilité que les autres en fluence formelle.

Leurs remarques nous ont permis de mieux appréhender les processus cognitifs et exécutifs et de saisir les difficultés en jeu lors de ces tâches.

2. **Reconnaissance vocale et correction des fichiers**

2.1 Reconnaissance vocale des fichiers enregistrés

Dans notre partie théorique, nous avons vu que les conditions optimales pour une bonne reconnaissance vocale de mots isolés passaient par plusieurs éléments :

- Un dictionnaire de référence contenant les mots du lexique concerné (animaux, fruits, mots commençant par R et par V) et leurs formes acoustiques (prononciations tenant compte des accents, des coarticulations, etc.).
- Un modèle acoustique et un modèle de langage décrivant les entités à reconnaître, et capables à partir du signal acoustique de donner la probabilité que le signal corresponde à chacun des phonèmes possibles de la langue cible.

Les enregistrements réalisés sont envoyés simultanément ou en différé sur le serveur de l'IRIT. Une phase de transcription est lancée, et par des rapprochements statistiques (via des algorithmes) entre l'onde acoustique du signal, les dictionnaires et les modèles acoustiques / linguistiques, les formes acoustiques des fichiers sont transcrites en mots. Nous avons donc ensuite pour chaque fichier une liste de mots reconnus par le logiciel, et des segments apparaissant sur l'onde acoustique, permettant de visualiser le début du mot et sa longueur.

2.2 Correction des fichiers

2.2.1 Première phase de correction

Les fichiers transcrits sont ensuite disponibles sur le serveur et la phase de correction peut alors commencer. Il s'agit en effet de **vérifier la correspondance entre ce qui a été dit par le sujet et ce qui a été reconnu par le logiciel.**

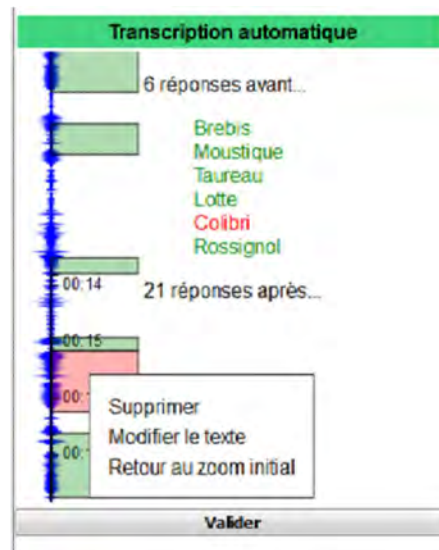


Figure 26 Ecran de transcription automatique avec onde acoustique, segments et liste de mots en cours de correction

Nous avons d'abord effectué une phase de tests pour aider à la mise au point de cette fonction de correction. Au fur et à mesure de nos retours, des discussions avec l'IRIT, et des modifications apportées, les « bugs » ont pu être corrigés et l'ergonomie améliorée. Il était important que ces corrections soient les plus faciles et les plus rapides possibles.

Ainsi il est possible d'écouter l'enregistrement, de faire une pause à tout moment, de corriger un mot, de modifier la taille ou la position des segments, d'ajouter un mot qui n'a pas été transcrit ou d'en supprimer un qui a été transcrit alors que rien n'a été dit. Il est aussi possible de réentendre uniquement le segment modifié pour s'assurer de la justesse de sa position sur l'onde acoustique.

Chaque correction a été chronométrée et nous avons pu disposer de données pour chaque fichier corrigé.

Le nombre de fichiers à corriger (fichiers A) a été déterminé avec l'IRIT en fonction de ce qu'ils estimaient avoir besoin pour lancer un apprentissage de la reconnaissance vocale à partir de ces fichiers « propres ».

2.2.2 Deuxième phase de correction

Suite aux premières corrections, un réapprentissage a été programmé sur tous les fichiers du serveur. Puis, nous avons recorrige les fichiers A, et en avons corrigé d'autres, que nous appellerons « fichiers B ». L'idée était de savoir si les premières corrections et le réapprentissage avaient permis d'améliorer la reconnaissance vocale.

2.2.3 Données fournies suite aux corrections

Après chaque phase de corrections, nous disposons de statistiques reprenant les éléments suivants : numéro de fichier, date de correction, type de fluence (test), temps de correction (en secondes), nombre de mots reconnus automatiquement, modifications de position de segment, modifications du mot, ajouts, suppressions et nombre de mots correctement transcrits.

- Le **nombre de mots automatiquement reconnus** correspond aux mots que le logiciel a transcrits, avant correction.
- Les **modifications de position** correspondent au nombre de fois où le segment d'un mot a été déplacé sur l'onde acoustique, afin d'obtenir des statistiques pertinentes notamment sur les temps de latence. Ces modifications comprennent des déplacements de segment, des agrandissements ou réductions de segment.

- Les **modifications de mots** reprennent le nombre de mots qui étaient mal reconnus et qu'il a fallu corriger. Nous devons faire un clic droit sur le segment et choisir l'option « modifier le mot » puis écrire le mot effectivement dit par le sujet.
- Les **ajouts** concernent les mots que le logiciel a oubliés ; il a alors fallu créer un segment, inscrire le mot et placer le segment de façon précise sur l'onde acoustique. Cette opération est longue, puisqu'il faut faire un clic droit sur l'onde acoustique et sélectionner « ajouter », puis faire un clic droit sur le nouveau segment pour sélectionner « modifier le mot », écrire le mot et ensuite placer correctement ce nouveau segment sur l'onde.
- Les **suppressions** concernent, à l'inverse, des mots que le logiciel a transcrits alors que rien n'était dit par le sujet. Il suffisait de faire un clic droit sur le segment concerné et choisir « supprimer » dans le menu.
- Les **mots correctement transcrits** reprennent les mots qui n'ont pas été modifiés du tout, et qui étaient donc bien reconnus d'emblée par le logiciel.

III. Tests en situation écologique

Avant de tester le logiciel en situation écologique, il nous a paru important de connaître les habitudes des orthophonistes en matière d'usage informatique et d'utilisation de tests de fluence dans leur pratique professionnelle. Nous avons donc opté pour l'élaboration d'un questionnaire (Annexe 5).

1. Choix d'un questionnaire préalable

1.1 Objectifs du questionnaire

Les objectifs de ce questionnaire n'étaient pas tant de dresser un portrait exhaustif de l'usage de l'informatique et des tests de fluence par les orthophonistes, que d'avoir une idée d'ensemble de leur pratique et de recueillir un avis a priori sur l'intérêt que présenterait pour eux un logiciel tel qu'EvoLex. Nous avons aussi établi ce questionnaire dans l'idée de disposer d'une base de comparaison avec la réalité du terrain.

1.2 Méthodologie du questionnaire

1.2.1 Savoir poser des questions

Nous nous sommes donc à la fois intéressées à des questions de fait, de type « Utilisez-vous... » et à des questions d'opinion, qui demandent un jugement, de type « Que pensez-vous de... ».

Suivant les conseils de De Singly (2014), nous avons essayé de construire un questionnaire avec une progression logique. Nous avons posé une seule question à la fois, évité les questions négatives ou doublement négatives, introduit la possibilité de répondre qu'on ne savait pas, gardé les questions personnelles pour la fin du questionnaire. Nous nous sommes attachées à poser des questions les plus objectives possibles afin de ne pas influencer les réponses dans un sens ou dans un autre.

Nous avons principalement utilisé des questions fermées pour pouvoir exploiter ensuite les réponses de manière quantitative et qualitative, mais nous avons aussi ajouté à plusieurs reprises la possibilité de préciser une réponse ou d'ajouter des commentaires. Tout en sachant que ceux-ci ne seraient pas exploitables de manière quantitative, les données qualitatives qu'ils apportaient nous ont semblé importantes.

Hormis la question 4, tributaire de la question 3, toutes les questions étaient obligatoires, c'est-à-dire que pour valider le questionnaire en ligne il fallait y avoir répondu, ce qui nous permettait d'avoir des questionnaires complets.

1.2.2 Tests préalables

L'élaboration du questionnaire s'est faite en plusieurs étapes. Nous avons d'abord testé une première « maquette » auprès de quelques orthophonistes. Leurs remarques, hésitations, demandes de précision nous ont orientées pour affiner nos questions et lever certaines ambiguïtés, et nous donner une idée du temps nécessaire pour y répondre.

1.2.3 Questionnaire

Au final nous avons conservé 20 questions, que nous détaillons ici.

Les **questions de 1 à 6** portent sur l'usage de l'informatique.

Question 1

Utilisez-vous l'informatique (bureautique) dans votre pratique professionnelle ?

Toujours, souvent, occasionnellement, jamais

Question 2

Utilisez-vous Internet dans votre pratique professionnelle ?

Toujours, souvent, occasionnellement, jamais

Question 3

Utilisez-vous des logiciels dans votre pratique professionnelle ?

Oui, non

Question 4

Si oui, quels types de logiciels utilisez-vous ?

Vous pouvez préciser lesquels dans la zone de commentaires.

Logiciels de rééducation et/ou logiciels de bilan

Question 5

Utilisez-vous des logiciels de reconnaissance vocale tels que Dragon, Siri, ou autres ?

Oui (précisez dans la zone de commentaires), non

Question 6

Quel(s) support(s) informatique (s) utilisez-vous dans votre pratique professionnelle ?

Cette première série de questions nous permet d'appréhender les rapports de l'orthophoniste avec l'informatique, de connaître sa fréquence d'utilisation et les types d'outils utilisés de façon privilégiée.

Les **questions 7 à 13** concernent l'utilisation de tests d'évocation lexicale et l'opinion sur les informations apportées par ces évaluations.

Question 7

Utilisez-vous des tests de fluence ?

Toujours, souvent, occasionnellement, jamais

Question 8

Utilisez-vous des tests de fluence...

Isolés (type Cardebat ou Isaacs) et/ou dans des batteries complètes

Question 9

Utilisez-vous des tests de fluence...

Sur papier et/ou informatisés

Question 10

Pour quels types de pathologies utilisez-vous les tests de fluence ?

L.O., LE, Neurologie, Autre

Question 11

A quel(s) moment(s) utilisez-vous les tests de fluence ?

En bilan initial et/ou en bilan d'évolution et/ou de renouvellement

Question 12

Ces tests apportent des informations quantitatives étalonnées (nombre de mots produits, nombre d'erreurs, etc.). Pensez-vous que ces analyses soient :

Satisfaisantes, moyennement satisfaisantes, insatisfaisantes, je ne sais pas

Question 13

Pensez-vous que des analyses qualitatives sur les types de mots produits (fréquence, concrétude, etc.) ou sur les stratégies utilisées par le patient constitueraient un apport pour un bilan ou un suivi de prise en charge ?

Oui, peut-être, non, je ne sais pas

Avec ces questions, nous souhaitons savoir de façon plus précise si des tests de fluence sont utilisés par l'orthophoniste, si oui lesquels, sur quels supports (papier, informatisés). De plus, nous abordons les éléments d'analyses quantitatives données par les tests existants, et nous introduisons les analyses qualitatives apportées par le logiciel.

Les **questions 14 à 17** portent sur l'opinion des orthophonistes sur l'intérêt a priori d'un logiciel comme EvoLex. Ces questions font suite à la présentation des analyses proposées par le logiciel.

Question 14

Pensez-vous que ce type d'analyses qualitatives soit utile ?

Oui, non, je ne sais pas

Question 15

Selon vous, quels seraient les atouts de ce logiciel ?

Gain de temps, disponibilité pour l'observation clinique, intérêt des analyses qualitatives, simplicité d'utilisation, autre

Question 16

Selon vous, quels seraient les inconvénients de ce logiciel ?

Fluence habituellement testées dans des batteries et non isolément, distance patient/orthophoniste induite par l'ordinateur, distance patient/orthophonique induite par le micro-casque, autre

Question 17

Auriez-vous d'autres attentes de ce type de logiciel ?

Oui (précisez dans zone de commentaire), non, je ne sais pas

Et enfin les dernières **questions de 18 à 20** donnent des informations sur la personne qui répond, à savoir le sexe, la tranche d'âge et le mode d'exercice de la profession d'orthophoniste (libéral, milieu hospitalier/centre de rééducation/soins de suite, institution ou autre).

1.2.4 Mode d'administration du questionnaire

Nous avons choisi d'utiliser un site spécialisé dans les créations de questionnaire afin de proposer un mode de réponse attrayant et rapide à traiter pour les orthophonistes. D'autre part, ce mode d'administration permet de toucher un grand nombre de personnes facilement et gratuitement. Sur un plan pratique, les résultats obtenus sont exportables sur un tableur, et peuvent ensuite être exploités, à travers des tableaux croisés dynamiques par exemple, pour essayer de dégager une tendance ou déterminer s'il existe des corrélations entre les différents facteurs.

Le questionnaire a été mis en ligne pendant 8 semaines et les orthophonistes ont été invités par e-mail à y répondre. Nous avons ensuite clos le questionnaire et exporté sur un tableur les résultats que nous exposerons plus loin.

2. Test et démonstration du logiciel auprès d'orthophonistes

Après avoir expérimenté le logiciel sur des sujets « sains », nous avons voulu savoir comment il était exploitable et perçu en situation écologique. Deux aspects nous intéressaient lors de cette présentation du logiciel : d'une part la comparaison du test de fluence manuel et informatisé auprès d'un patient, d'autre part l'avis de l'orthophoniste sur le logiciel et les analyses qu'il apporte.

2.1 Recrutement des orthophonistes et des patients

Nous avons demandé à nos maîtres de stage si elles étaient intéressées pour ce test, et si elles avaient un patient pouvant entrer dans nos critères :

- Avoir plus de 18 ans
- Etre de langue maternelle française
- Etre atteint d'une pathologie neurologique acquise ou neurodégénérative ayant des répercussions sur le langage
- Ne pas être dysarthrique, ni dysphonique
- Ne pas avoir passé de tests de fluence depuis au moins 6 mois (pour éviter l'effet re-test)

2.2 Test avec les patients et les orthophonistes

Nous avons choisi de faire passer aux patients une épreuve de fluence sémantique (animaux) et une épreuve de fluence formelle (mot en V). Nous avons rencontré le patient avec son orthophoniste sur son lieu habituel de rééducation ; au cabinet, à domicile, en EHPAD ou en centre de rééducation fonctionnelle.

Nous lui avons exposé le projet de mise au point d'un logiciel de test de fluence. Nous lui avons expliqué la teneur du test, en prévenant qu'il aurait un micro-casque pour que le logiciel l'enregistre et puisse reconnaître ce qu'il disait.

2.2.1 Ouverture du logiciel

Nous avons fait nos tests écologiques avec la « version 2016 » du logiciel. Dans cette version, chaque utilisateur a son propre identifiant, que l'on saisit pour ouvrir le logiciel.

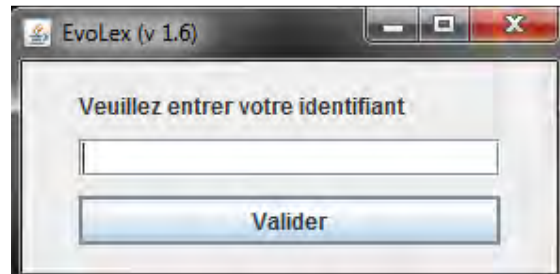


Figure 27 Ecran pour identification de l'utilisateur

Nous nous retrouvons alors sur la page suivante, qui offre le choix entre 3 possibilités :

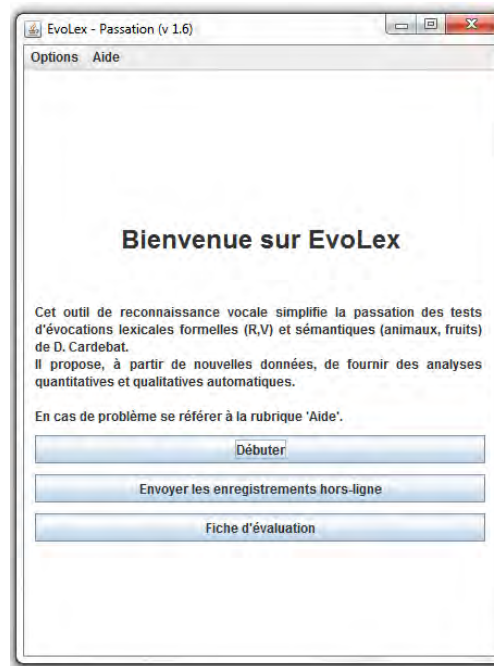


Figure 28 Menu principal du logiciel de passation

- « Débuter » : permet de créer un patient ou d'en retrouver un déjà créé
- « Envoyer les enregistrements hors-ligne » permet d'envoyer les enregistrements effectués sans connexion internet, en différé
- « Fiche d'évaluation » permet d'ouvrir un fichier au format pdf contenant la grille d'évaluation clinique des erreurs et des stratégies utilisées par le patient

Nous avons rentré quelques données personnelles dans le logiciel (Nom, prénom, Age, Sexe, NSC) et avons noté (sur papier) la pathologie.

2.2.2 Passation des épreuves

Nous plaçons le micro-casque sur la tête du patient de façon à ce que le microphone se situe au niveau du menton, pour éviter d'être dans le flux d'air.

Puis nous donnons la consigne et cliquons sur « Démarrer ». Le chronomètre s'enclenche automatiquement pour 2 minutes.



Figure 29 Chronomètre, validation/invalidation

Pendant que le patient dit les mots, nous cliquons soit sur A si le mot est jugé correct, soit sur P, s'il ne correspond pas à la consigne. Cela permet au logiciel de repérer que la forme acoustique précédant le clic est un mot à reconnaître et non un bruit de l'environnement par exemple.

A la fin du temps imparti, l'enregistrement s'arrête automatiquement. Nous avons alors la possibilité d'ajouter un commentaire libre, dans le cas où un événement serait arrivé pendant l'enregistrement par exemple.

Pendant l'épreuve chronométrée automatiquement par le logiciel, l'orthophoniste a noté manuellement les mots dits par le patient.

2.2.3 Anonymat assuré par le logiciel

Un des aspects très important du logiciel est qu'il garantit l'anonymat de la personne enregistrée, que ce soit dans un cadre « technique » d'amélioration de la reconnaissance vocale, ou auprès de patients en situation réelle de prise en charge orthophonique.

En effet, lorsque nous « créons » la fiche d'un nouveau participant, un code à 32 caractères (lettres et chiffres) lui est attribué de façon arbitraire. Ce code sera toujours le même pour le même participant. Par exemple, on peut trouver un code tel que celui-ci : FEF61EAB6955C97A722AF76B28E822EF. Il n'y a aucun moyen de savoir s'il s'agit d'un homme ou d'une femme, son âge ou son niveau d'études.

D'autre part, chaque fichier audio porte également un code à 32 caractères, qui lui est propre, comme par exemple : 771A607EF086552555749E967064755D.

Ces codes sont créés à partir du moment où les fichiers sont transmis au serveur, ce qui permet que les données soient traitées de façon complètement anonyme. Seul l'utilisateur du logiciel (l'orthophoniste) connaît l'identité des patients, car les noms apparaissent sur la page permettant la sélection du patient (ou la création d'un nouveau).

2.2.4 Envoi des fichiers sur le serveur

Lorsque les enregistrements sont terminés, si nous sommes connectées au wifi, ils sont envoyés automatiquement sur le serveur. Si nous n'avons pas de connexion, nous pouvons envoyer les fichiers sur le serveur par synchronisation, en différé. Dans ce cas, il faut cliquer sur « Envoyer les enregistrements hors-ligne ».

2.3 Démonstration des analyses et 2^{ème} questionnaire

Dans un second temps, nous avons revu l'orthophoniste pour lui montrer en détail le fonctionnement du logiciel (récupération des fichiers, corrections des fichiers), et les analyses qu'il apportait. La partie analyse du logiciel se présente ainsi :

2.3.1 Partie analyses

Il faut d'abord se connecter au serveur de l'IRIT avec identifiant et mot de passe. On obtient alors l'écran suivant qui présente la liste des patients (à la place des points d'interrogation) et le type d'épreuve réalisé :

Bienvenue Florence Déconnexion

Parcourir... Aucun fichier sélectionné.

Filtres

Afficher les passations corrigées

Afficher les passations non corrigées

Afficher les passations non transcrites

Dernières passations

17/05/16 ???		(Fluence - Mots en V)
17/05/16 ???	En attente de transcription	(Fluence - Animaux)
12/05/16 ???		(Fluence - Mots en V)
12/05/16 ???		(Fluence - Mots en V)
23/12/15 ???		(Fluence - Animaux)
16/12/15 ???		(Fluence - Mots en R)
16/12/15 ???		(Fluence - Fruits)
16/12/15 ???		(Fluence - Mots en V)
16/12/15 ???		(Fluence - Animaux)
16/12/15 ???		(Fluence - Fruits)
16/12/15 ???		(Fluence - Fruits)
15/12/15 ???		(Fluence - Animaux)
15/12/15 ???		(Fluence - Mots en V)

Figure 30 Identification de l'utilisateur et menu principal du serveur

On sélectionne ensuite le patient et l'épreuve qui nous intéressent :

Autres épreuves du patient

Mot transcrit

Barre d'affichage des analyses

Bienvenue Florence Déconnexion

Passation 98ED4F833ABDB82FC55BA168F96E3244

Date de passation
16 décembre 2015 00:00

Catégorie de test
Fluence

Test
Mots en R

Patient

Identifiant
CE2CE533AB1584E2156B95F247691F15

Autres tests de ce patient
16/12/15 - Mots en V
16/12/15 - Lecture Libre
16/12/15 - Fruits
16/12/15 - Lecture Libre
16/12/15 - Animaux
16/12/15 - Mots en R

14 en 2016 15:21

Afficher la légende

5 réponses exist.

Réponse
Réaliser
Rie
Rois
Rie.

7 réponses après.

16/12/15
16/12/15
16/12/15
16/12/15
16/12/15
16/12/15
16/12/15
16/12/15
16/12/15
16/12/15

Valider

Lecteur du fichier audio

Echelle temporelle avec onde sonore



Figure 31 Page de présentation d'un fichier audio et de sa transcription

Pour corriger les erreurs de transcription, le clic droit de la souris permet d'ajouter un segment, de le modifier ou de le supprimer.

En cliquant sur « Voir les résultats de la segmentation sélectionnée », on obtient l'affichage des résultats des analyses instantanément. Les résultats quantitatifs font apparaître le pourcentage de mots valides, d'erreurs, de doublons et de mots de la même famille, ainsi que l'écart type lorsque les données (âge, sexe, NSC) du patient sont rentrées.

Résultats quantitatifs

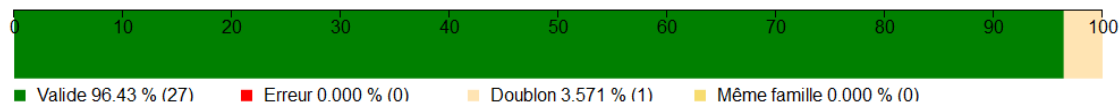


Figure 32 Présentation des résultats quantitatifs (normes Cardebat)

Un visuel présente la fréquence des mots et les temps de latence. Plus les barres sont hautes plus les mots sont fréquents, plus les barres sont espacées plus il y a un temps de latence entre les mots. Sur le schéma ci-dessous, on voit par exemple que les premiers mots prononcés sont beaucoup plus fréquents que ceux qui suivent. De même, il y a beaucoup plus de mots dits dans les 15 premières secondes que par la suite.

Résultats qualitatifs

Fréquence et latence

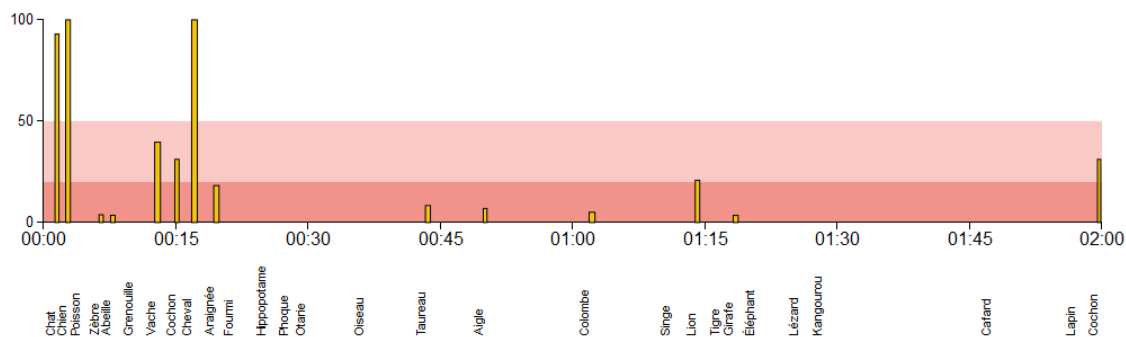


Figure 33 Vue détaillée des temps de latence et fréquence de mots

Les regroupements sont analysés afin de mettre en valeur les stratégies utilisées par le patient. Ainsi, des regroupements peuvent être de nature sémantique, phonologique, alphabétique, grammaticale (pour les fluences formelles) ou morphologique. Ils sont matérialisés en couleur, comme dans l'exemple ci-dessous : en vert, les lieux d'habitation, en orange les catégories d'animaux, etc.

Clusters et switches

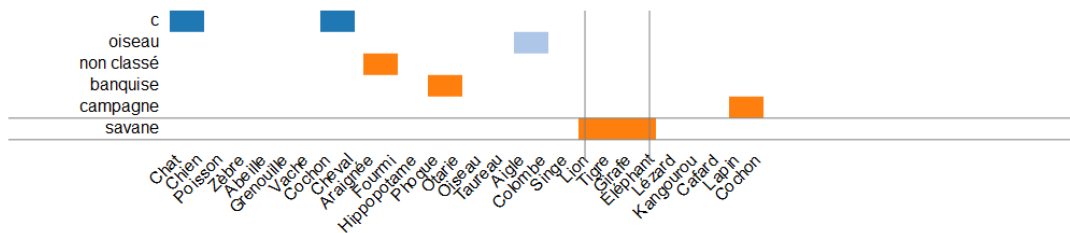


Figure 34 Vue détaillée des regroupements et commutations

Puis, nous trouvons des récapitulatifs, qui permettent de chiffrer les regroupements et commutations. On trouve également des données chiffrées sur les types de clusters, et enfin des données qui permettent de mettre en évidence le processus de recherche (% de mots regroupés) et de flexibilité mentale (mots isolés).

Processus cognitif

Regroupements

- 7 regroupements
- 1 commutations
- 28 productions
- 12 mots isolés
- 57,14 % des mots appartient à un regroupement

Connaissances lexico-sémantiques et lexique phonologique

La majorité des liaisons sont de nature **lieu**.

- 10 liens de nature lieu
- 4 liens de nature lettre_initiale
- 2 liens de nature famille

Fonctions exécutives : processus stratégique de recherche et flexibilité mentale spontanée

7 regroupements pour 28 productions (25.00 %)
12 mots isolés pour 28 productions (42.86 %)

Figure 35 Vue détaillée du profil cognitif du patient

Le profil récapitulatif présenté sous forme de cible permet de visualiser d'un coup d'œil les performances du patient, en reprenant les éléments quantitatifs (écart type selon les normes de Cardebat), et qualitatifs (fréquence, latence, initiation, regroupements et commutations).

Profil Récapitulatif

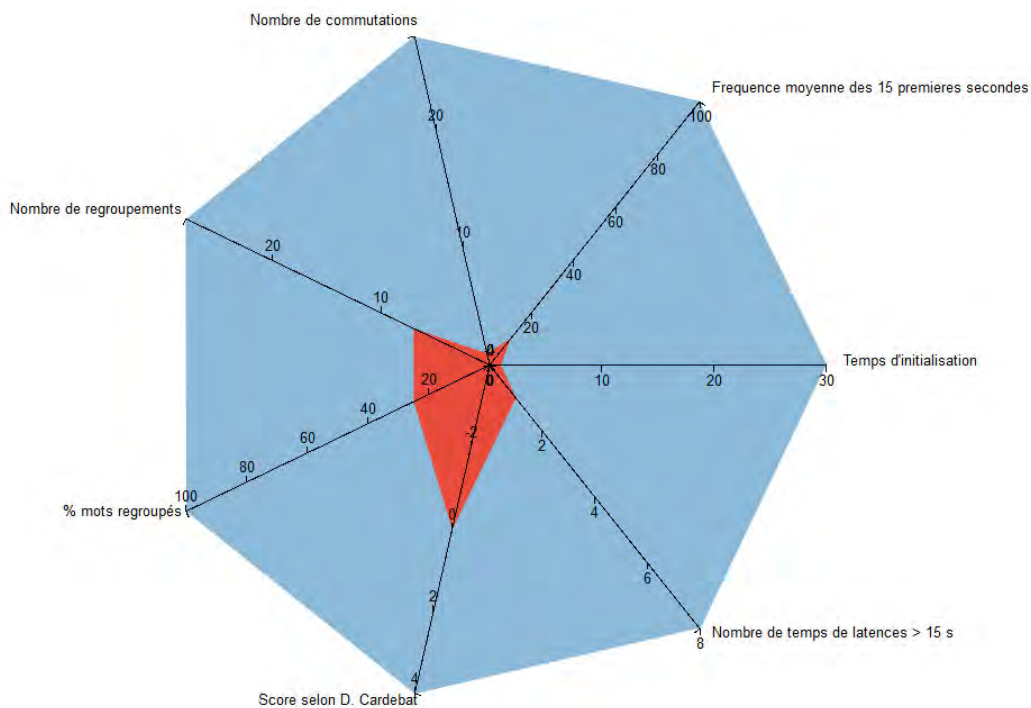


Figure 36 Présentation du profil récapitulatif des performances du patient sur une tâche de fluence

2.3.2 Questionnaire

Après avoir montré et expliqué le logiciel et ce qu'il propose, nous avons demandé aux orthophonistes de répondre à notre enquête. Pour collecter leurs impressions et critiques, nous nous sommes appuyées sur un deuxième questionnaire qui reprenait certaines questions abordées dans le premier (Annexe 5 Annexe 6). Cette fois, il s'agissait de recueillir un avis non plus a priori mais après démonstration - fut-ce avec un prototype. En utilisant le logiciel en situation écologique, nous avons pour but d'essayer de cerner les besoins, les appréhensions, les attentes du thérapeute, les atouts et les points faibles du logiciel.

Aux questions suivantes, il fallait se prononcer selon la gradation suivante : tout à fait d'accord, plutôt d'accord, moyennement d'accord ou pas du tout d'accord.

<p>Selon vous, le logiciel...</p> <ul style="list-style-type: none"> ▪ permet de gagner du temps par rapport aux tests de fluence "manuels ? ▪ est simple d'utilisation ? ▪ est ergonomique ? ▪ propose un large éventail d'analyses ? ▪ propose des analyses quantitatives intéressantes et pertinentes ? ▪ propose des analyses qualitatives intéressantes et pertinentes en matière de : temps de latence, fréquence des mots, temps d'initiation, nombre de clusters et de switches ? ▪ propose un profil récapitulatif intéressant et pertinent ? ▪ constitue une possibilité intéressante de stocker les données de patient et d'avoir ainsi une base pour un suivi ? ▪ permet d'être plus disponible pour l'observation clinique <p>Pensez-vous que le principe de reconnaissance vocale est intéressant ?</p>	<p>Pensez-vous qu'un ordinateur introduit une distance entre thérapeute et patient ?</p> <p>Pensez-vous qu'un micro-casque introduit une distance entre thérapeute et patient ?</p> <p>Pensez-vous qu'il soit intéressant d'avoir un test de fluence informatisée indépendamment d'une batterie ?</p> <p>Ce logiciel pourra permettre à terme de faire passer des épreuves de dénomination et de génération de mots, pensez-vous que ce soit intéressant ?</p> <p>Selon vous, quelles seraient les améliorations à apporter au logiciel ?</p> <p>Avez-vous des commentaires, remarques ou suggestions supplémentaires ?</p> <p>Sur une échelle de 1 à 10, comment vous jugeriez ce logiciel ?</p>
--	---

. **Présentation et analyses des résultats**

I. Présentation des résultats

1. Résultats de l'amélioration de la reconnaissance vocale

1.1 Enregistrements de sujets sains

1.1.1 Présentation de notre échantillon

Nous avons enregistré 86 sujets sains effectuant les tâches de fluence verbale sémantique (animaux et fruits) et littérale (mots commençant par R et par V). Lorsque nous avons créé la fiche « patient » au moment de la passation, nous avons renseigné le sexe, la date de naissance et le niveau d'études.

Afin que la reconnaissance vocale puisse être significativement améliorée, l'équipe de l'IRIT nous a demandé d'équilibrer le nombre d'hommes et de femmes, car les modèles acoustiques utilisés jusqu'à présent étaient basés sur des enregistrements de radio et majoritairement des voix d'hommes. Nous avons pu respecter cette demande, puisque parmi nos 86 participants, 42 sont des femmes et 44 sont des hommes.

L'histogramme ci-dessous permet de constater que les participants font principalement partie de la catégorie socio-culturelle 3 (\geq Baccalauréat), et sont majoritairement âgés de moins de 45 ans. Cela s'explique par le fait que nous avons ciblé des personnes de notre entourage, dont la majorité a des profils proches des nôtres.

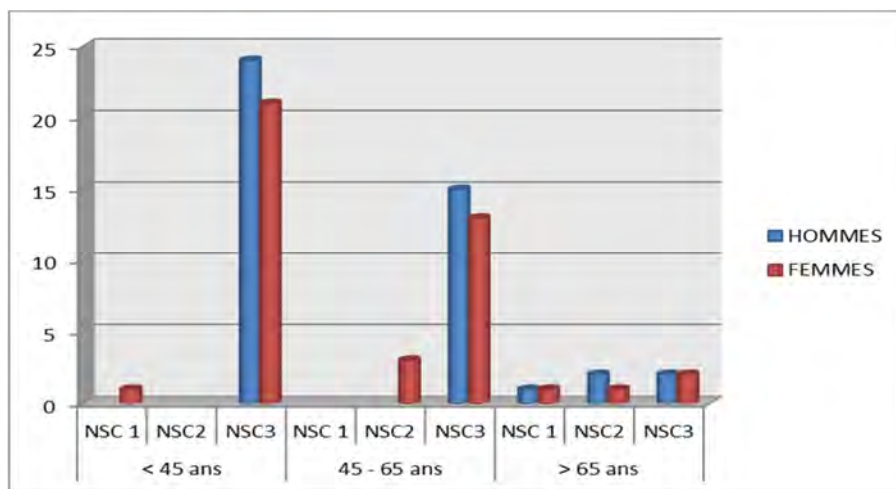


Figure 37 Répartition de l'échantillon par sexe, âge et catégorie socio-culturelle

1.1.2 Enregistrements

Pour chaque participant, nous avons pu enregistrer 4 tâches de fluence et 2 lectures de texte. Donc à ce jour, l'IRIT possède 344 enregistrements de tâches de fluence, et 172 enregistrements de textes lus.

Nous rappelons que ces enregistrements ont été réalisés dans des conditions strictes de tests (respect du protocole, lecture des consignes, etc.) et acoustiques, pour lesquelles nous avons été particulièrement vigilantes.

1.2 Résultats de l'amélioration de la reconnaissance vocale

Après chaque phase de corrections, nous avons accès via le serveur aux données suivantes, que nous avons exportées dans un tableau Excel pour faciliter les calculs et les analyses.

N° Fichier	Date correction	Test	Temps de correction (en secondes)	Nombre de mots reconnus automatiquement	Modifications de la position	Modifications du mot	Ajouts	Suppressions	Mots correctement transcrits
A9FB336ABC7E0642F67EEF5E7	16/04/16	Ani maux	369,346	54	0	40	0	8	5
D11030302B86DFE40CBA7C896	16/04/16	Ani maux	656,458	85	1	31	0	43	8
A88559231D6768084C02163005	17/04/16	Ani maux	442,380	50	2	23	0	16	11

Figure 38 Données des corrections

1.2.1 Taux de reconnaissance vocale

Ce ratio était très attendu, tant par nous que par l'équipe de l'IRIT, car il permet de voir où en est le logiciel.

Nous avons calculé ce ratio de 2 façons différentes. L'équipe des informaticiens-chercheurs souhaitait savoir combien de mots avaient été correctement transcrits par rapport à ceux reconnus automatiquement (ce ratio est calculé sur ce qui a été trouvé = mots automatiquement reconnus, appelé « précision » en langage mathématique). **Le ratio moyen est de 46,48%**. En ce qui nous concerne, nous voulions savoir ce que le logiciel avait bien reconnu par rapport à ce qu'il aurait dû bien reconnaître (le total des mots corrects après corrections). Ce mode de calcul est appelé « rappel » en langage mathématique. Avec ce mode de calcul, **le ratio moyen est de 49,83%**.

- Résultats obtenus sur les fichiers A, corrigés avant (A1) et après réapprentissage (A2)

		Avant réapprentissage		Après réapprentissage	
Test		Taux de reconnaissance vocale (Précision)	Taux de reconnaissance vocale (Rappel)	Taux de reconnaissance vocale (Précision)	Taux de reconnaissance vocale (Rappel)
Moyenne	Animaux	43,93%	51,87%	58,92%	70,27%
Moyenne	Fruits	51,89%	54,90%	67,12%	71,25%
Moyenne	Mots en R	37,59%	39,01%	50,31%	52,32%
Moyenne	Mots en V	52,51%	53,54%	54,53%	59,80%
Moyenne générale		46,48%	49,83%	57,72%	63,41%

Figure 39 Taux de reconnaissance vocale en précision et rappel avant et après réapprentissage

La comparaison de la reconnaissance vocale avant et après réapprentissage montre un gain de 24% en précision, et de 27% en rappel.

On constate des différences notoires entre les types de fluence. La reconnaissance vocale est meilleure sur la fluence sémantique, et plus faible en fluence formelle, en particulier pour les mots commençant par R dont le taux de reconnaissance vocale est de 52,32%. Le graphique en bâton suivant permet de mieux visualiser ces différences.

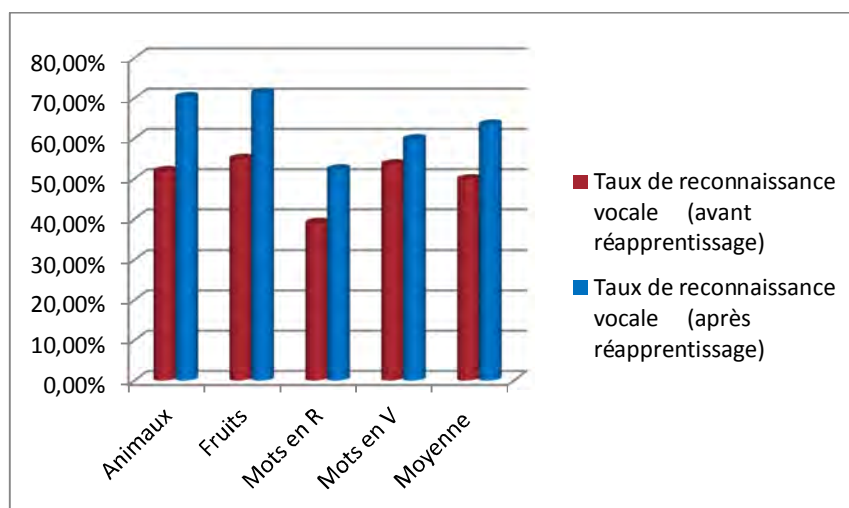


Figure 40 Taux de reconnaissance vocale par fluence

Les différences entre les fluences globalement sont significatives d'un point de vue statistique ; avec une analyse de variance ANOVA, nous obtenons en effet un degré de signification $p < 0,05$. Dans le détail, il existe des différences significatives (par exemple entre les Animaux et les mots en R, ou entre les Fruits et les mots en R), avec un degré de signification $p < 0,05$, et d'autres non, comme par exemple entre les Animaux et les Fruits, qui ont un degré de signification $p > 0,05$.

- Résultats obtenus sur les fichiers B, uniquement corrigés après réapprentissage

Test	Taux de reconnaissance vocale (Précision)	Taux de reconnaissance vocale (Rappel)
Moyenne Animaux	65,48%	72,70%
Moyenne Fruits	63,43%	64,71%
Moyenne Mots en R	40,60%	46,02%
Moyenne Mots en V	62,87%	67,58%
Moyenne générale	58,10%	62,75%

Figure 41 Taux de reconnaissance vocale en précision et en rappel

Les taux de reconnaissance vocale des fichiers B et des fichiers A2 sont assez proches : 63,41% pour les fichiers A2 et 62,75% pour les fichiers B.

Les différences entre les fluences, globalement, sont également significatives d'un point de vue statistique, avec un degré de signification $p < 0,05$. Nous retrouvons d'un autre côté les mêmes différences significatives et non significatives que sur les fichiers A.

1.2.2 Résultats sur les temps de corrections

L'année dernière, le gain de temps n'avait pas pu être mis en évidence car les temps de correction étaient importants (5:51 min).

Avec le prototype de cette année, le temps passé à corriger les fichiers a été de 6:12 minutes en moyenne avant le réapprentissage. Après le réapprentissage, il est passé à 4:44 minutes, ce qui représente une baisse de 19,8%. Concernant les fichiers B, corrigés uniquement après réapprentissage, le temps moyen est de 4:34 minutes. Le graphique ci-dessous montre les 3 courbes, ce qui permet de visualiser rapidement les écarts entre chacune d'elle.

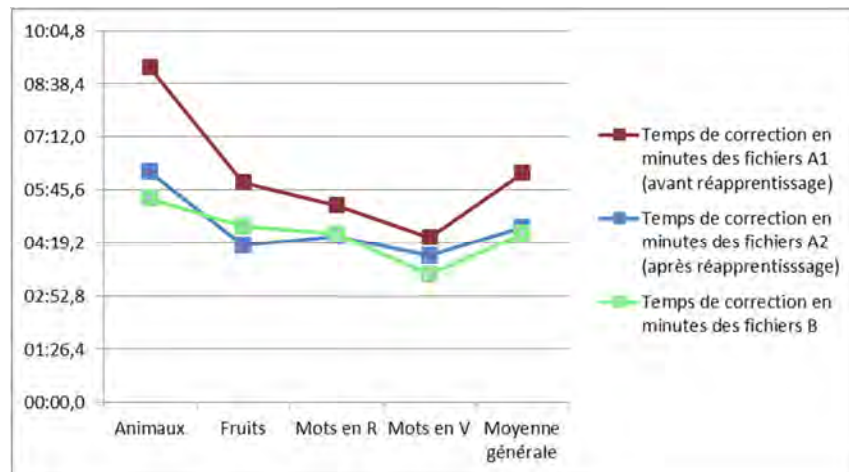


Figure 42 Ecarts des temps de correction avant et après réapprentissage par fluence

Grâce aux calculs de régression linéaire, nous avons pu mettre en évidence des corrélations entre le taux de reconnaissance vocale et le temps de correction. Le tableau ci-dessous résume les valeurs de coefficient de corrélation et de degré de signification.

	Coefficient de corrélation r	Degré de signification p
Animaux	-0,63	< 0,05
Fruits	-0,79	< 0,05
Mots en R	-0,71	< 0,05
Mots en V	-0,65	< 0,05

Figure 43 Résumé des valeurs de coefficients de corrélation et de degré de signification

Nous pouvons visualiser ces corrélations sur les nuages de points suivants, établis pour les fichiers A2, avec l'exemple des fichiers Animaux et des mots en V.

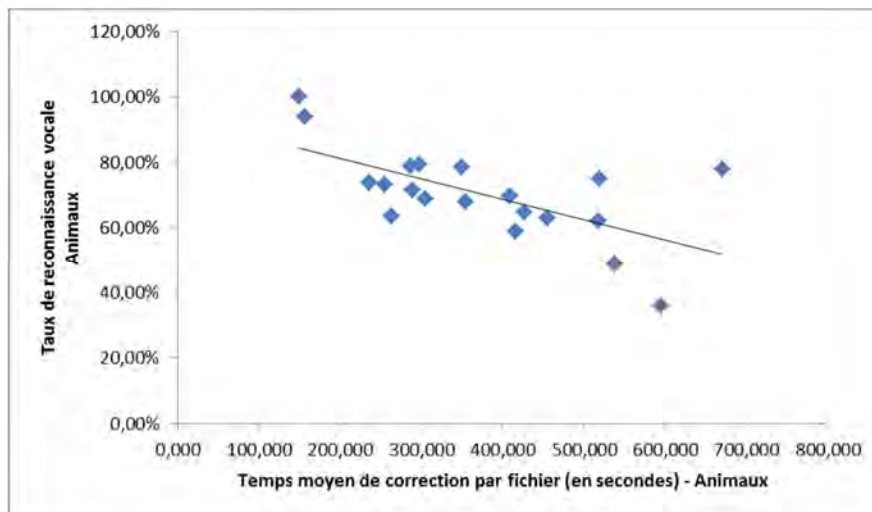


Figure 44 Fichiers A2 Animaux

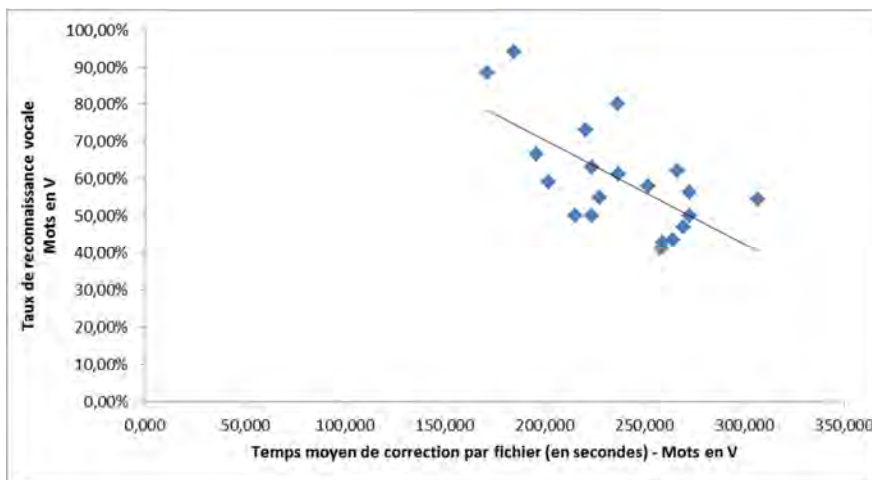


Figure 45 Fichiers A2 mots en V

Nous constatons qu'il y a une corrélation forte (car les coefficients de corrélation sont proches de $|1|$) entre le taux de reconnaissance vocale et le temps passé à corriger les fichiers : meilleure est la reconnaissance vocale, plus court est le temps de correction.

2. Résultats des questionnaires et des tests en situation écologique

2.1 Résultats du questionnaire préalable

2.1.1 Echantillon et taux de réponse

L'échantillon a été déterminé par les adresses e-mails que nous avons pu collecter, majoritairement celles d'orthophonistes exerçant dans la région Midi-Pyrénées-Languedoc-Roussillon. Nous avons envoyé le questionnaire à 545 orthophonistes et avons reçu 181 réponses, et obtenons ainsi un taux de réponse de 33,2%.

2.1.2 Profil des participants

Nous avons comparé le profil des orthophonistes qui ont répondu avec les chiffres décrivant la profession publiés dans la revue l'*Orthophoniste* en décembre 2014.

	FNO	Notre Questionnaire
Sexe	Femmes : 96 % Hommes : 4%	Femmes : 96 % Hommes : 4%
Age	Moyen : 43,3 ans	entre 21 et 30 ans : 40% entre 31 et 40 ans : 24% entre 41 et 50 ans : 17% entre 51 et 60 ans : 14% plus de 60 ans : 5%
Mode d'exercice	Libéral : 80,86% Mixte : 7,48 % Salariat : 11,66 %	Libéral : 69% Mixte : 27 % Salariat : 4 %

Figure 46 Comparaison de notre échantillon et des données de la FNO

La répartition par sexe est identique dans notre échantillon et dans le sondage effectué par la FNO. La répartition par âge n'est pas vérifiable. Concernant le lieu d'exercice, on constate une large proportion de libéraux, mais un pourcentage plus élevé d'orthophonistes ayant un exercice mixte salariat-libéral dans notre échantillon que dans l'ensemble de la profession.

2.1.3 Usage de l'informatique

Plus de la moitié des orthophonistes utilisent toujours ou souvent l'informatique et Internet dans leur pratique professionnelle, et seuls 3,3 % n'ont recours ni à l'un ni à l'autre.

➤ Reconnaissance vocale

88,4 % ont répondu ne pas se servir de la reconnaissance vocale. Ceux qui l'utilisent citent principalement « Dragon » comme logiciel, mais ne s'en disent pas très satisfaits.

➤ Sur quels supports ?

L'usage de l'ordinateur fixe ou celui du portable sont premiers à égalité, puis vient celui de la tablette et enfin celui du smartphone en dernière position.

➤ Quels logiciels ?

80% des orthophonistes qui ont répondu utilisent des logiciels. Seuls 11% d'entre eux n'utilisent que des logiciels de bilan, 43% utilisent uniquement des logiciels de rééducation et 46% utilisent les 2. D'un point de vue statistique, il y a des différences significatives entre ces pourcentages ($p < 0,05$). Les commentaires nous ont permis de nous faire une idée des logiciels utilisés et de dresser la liste des plus cités :

- Logiciels de rééducation
 - Lecture flash (logicieleducatif.fr)
 - Logatron (OrthoEdition)
 - Audiolog (Happyneuronpro)
 - CréaLangage (Happyneuronpro)
 - L'imagier phonétique (OrthoEdition)
- Logiciels de bilan
 - Exalang (Orthomotus)
 - Vocalab (Gérip)

2.1.4 Usage en matière de tests de fluence

68,5% ont répondu utiliser les tests de fluence toujours ou souvent.

On constate également que 44% des participants utilisent les tests de fluence uniquement inclus dans une batterie d'évaluation, contre 21% qui ne l'utilisent que de façon isolée.

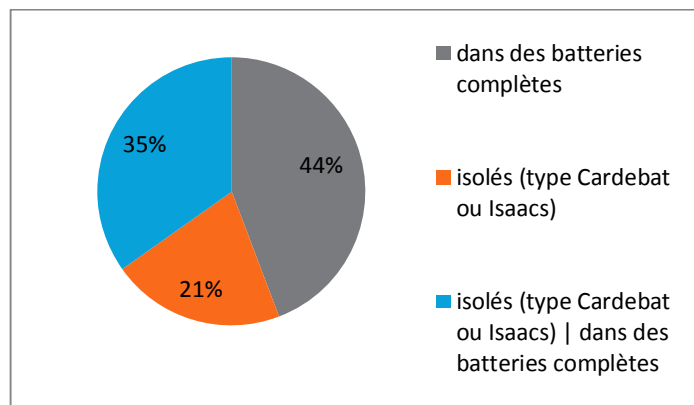


Figure 47 Mode d'usage des tests de fluence

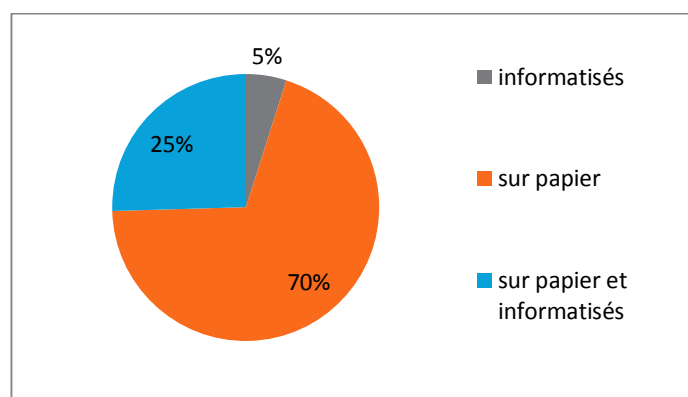


Figure 48 Support du test de fluence privilégié par les utilisateurs de l'informatique

La majorité des tests de fluence sont passés de façon manuelle (70%), contre seulement 5% d'orthophonistes qui utilisent exclusivement des tests informatisés.

Voici les domaines dans lesquels sont souvent proposés des tests de fluence :

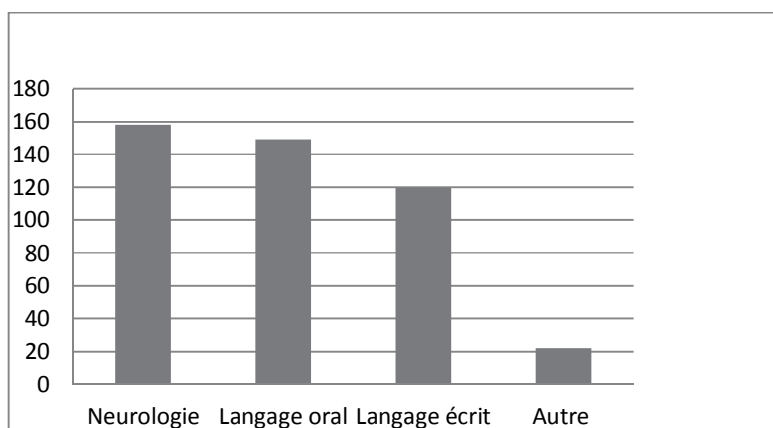


Figure 49 Domaines d'utilisation du test de fluence

2.1.5 Avis a priori sur un logiciel de fluence

Analyses quantitatives et qualitatives

Les graphiques ci-dessous présentent la perception des analyses quantitatives (fig. 49) puis qualitatives (fig. 50) :

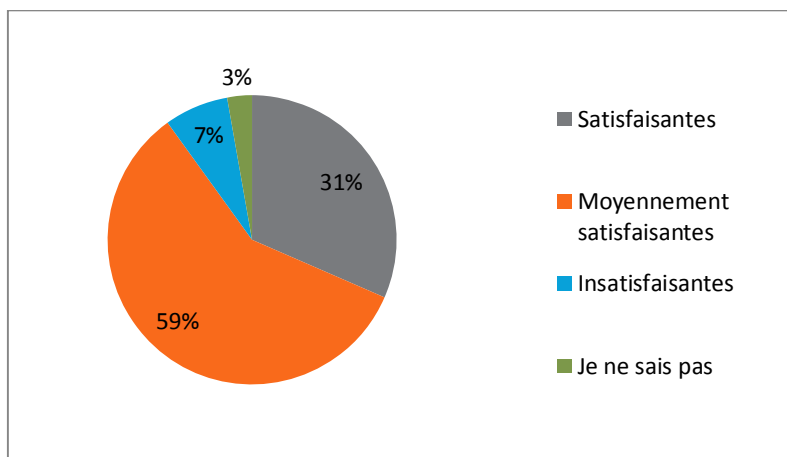


Figure 50 Satisfaction concernant les analyses quantitatives

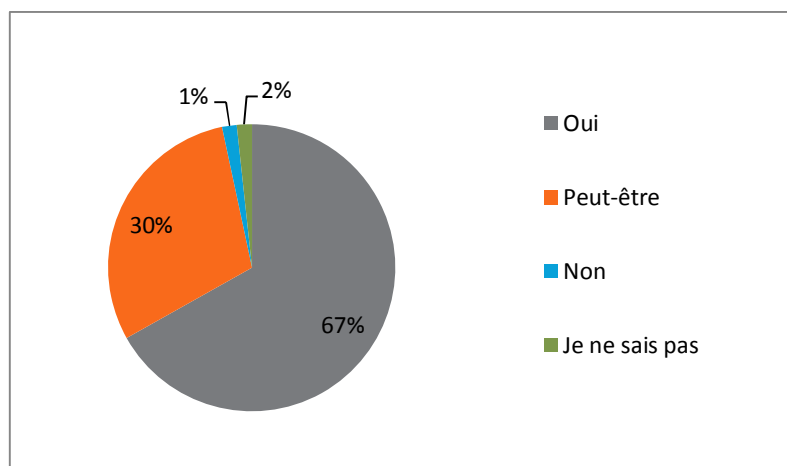


Figure 51 Des analyses qualitatives constitueraient un apport pour un bilan ou un suivi

Inconvénients a priori du logiciel

C'est d'abord la distance patient/orthophoniste induite par le micro-casque qui est citée, puis celle induite par l'ordinateur. Vient ensuite le fait que les fluences ne sont pas habituellement testées isolément.

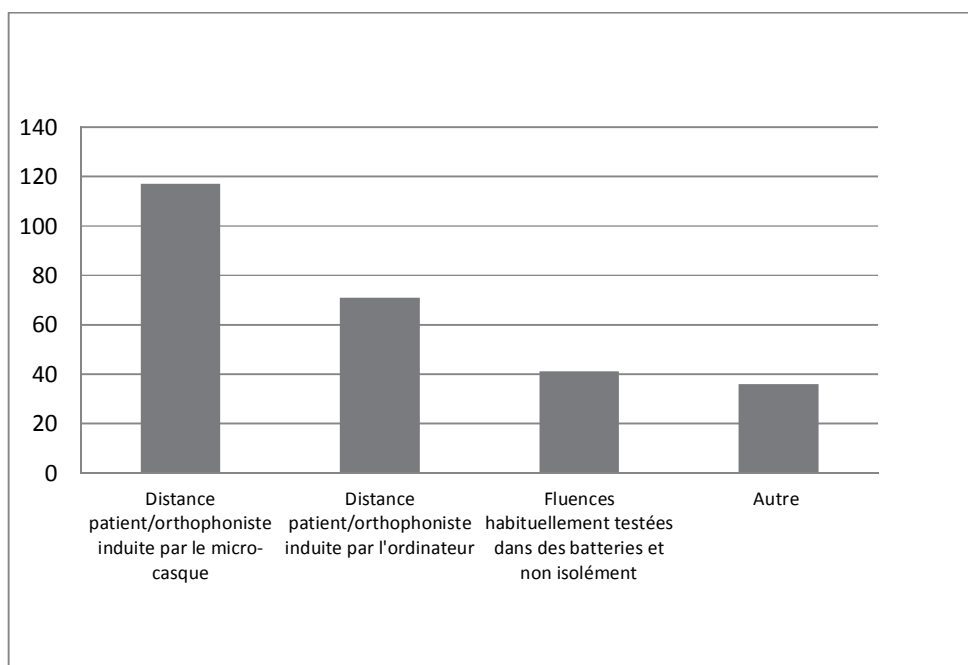


Figure 52 Inconvénients a priori du logiciel

La lecture des commentaires à cette question nous a permis de lister un certain nombre de réserves :

- Réserves par rapport à la fiabilité technologique et aux contraintes techniques
 - Problème de la fiabilité de la reconnaissance vocale, notamment en cas de trouble phonologique
 - Nécessité d'un ordinateur lors des visites à domicile
 - Question du temps de mise en place
 - Complexité de l'utilisation (nécessité d'une grande simplicité d'utilisation)
 - Risque de problème technique
- Réserves par rapport au coût supplémentaire avec le microphone
- Réserves par rapport à la relation avec le patient :
 - Rupture de l'empathie et de l'interlocution avec le patient due au microphone, situation peu écologique du patient qui s'adresse à un ordinateur, risque d'inhibition du patient
 - Possible perte d'attention liée au casque
 - Inconfort du patient dans cette situation, voire stress
 - Particularité des patients ayant un handicap auditif

➤ Réserves par rapport aux analyses

- Non prise en compte du contexte de la séance ou de la particularité du patient avec le risque d'une analyse trop théorique
- Questionnement sur la façon d'exploiter concrètement ces analyses en rééducation
- Statistiques induites par un ordinateur et non un humain

Atouts a priori du logiciel

Les principaux atouts du logiciel selon les personnes interrogées résideraient dans l'apport d'analyses qualitatives. Viennent ensuite la disponibilité pour l'observation clinique, le gain de temps, et la simplicité d'utilisation.

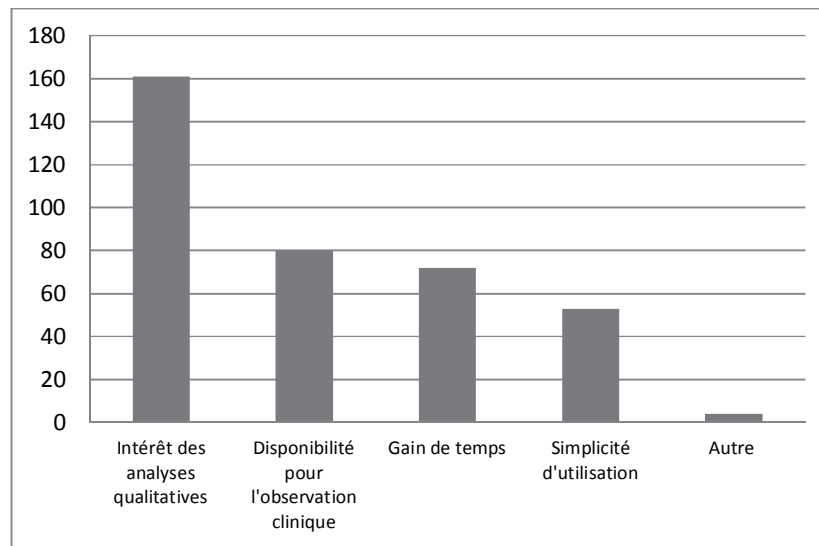


Figure 53 Atouts a priori du logiciel

Cinq personnes seulement ont répondu avoir des attentes par rapport au logiciel, mais les commentaires sont plus nombreux et suggèrent :

- De pouvoir utiliser le smartphone pour résoudre la distance induite par l'ordinateur ou le microphone et préserver ainsi la situation naturelle
- D'avoir la possibilité d'élargir l'analyse à plusieurs fluences, plus variées que celles déjà connues.
- Une comparaison avec les productions d'une population témoin pour les différentes variables de l'analyse qualitative serait très intéressante (normalisation)

- De disposer d'une grande étendue d'âges (enfants et adultes) pour les étalonnages et de différentes durées de passation (une minute par exemple)
- De pouvoir repérer la fréquence des mots
- Un intérêt pour le temps de latence entre les réponses

2.2 Résultats des tests en situation écologique

2.2.1 Reconnaissance vocale (étude sur les fichiers « patients »)

Taux de reconnaissance vocale

Nous avons pu enregistrer 8 patients avec la version 2 d'EvoLex, ce qui représente 16 fichiers ; 8 sur une tâche de fluence sémantique (animaux) et 8 sur une tâche de fluence formelle (mots commençant par V).

	Avant réapprentissage		Après réapprentissage	
	Taux de reconnaissance vocale (Précision)	Taux de reconnaissance vocale (Rappel)	Taux de reconnaissance vocale (Précision)	Taux de reconnaissance vocale (Rappel)
Moyenne Animaux	22,07%	34,93%	42,64%	59,66%
Moyenne Mots en V	19,93%	26,91%	20,35%	28,02%
Moyenne générale	21,00%	30,92%	31,49%	43,84%

Figure 54 Taux de reconnaissance vocale sur les fichiers « patients »

Rappelons que le taux de reconnaissance vocale en précision est calculé sur la base des mots reconnus automatiquement (ce que cherchent les informaticiens-chercheurs), alors que le taux de reconnaissance vocale en rappel est calculé sur la base des mots corrects après correction, ce qui nous intéresse en tant qu'orthophonistes.

On constate que le taux de reconnaissance vocale, que ce soit en précision ou en rappel, est meilleur après réapprentissage (43,84% contre 30,92% en rappel). Cependant, le taux de reconnaissance vocale est plus faible que sur les sujets sains (63,41%).

Temps de correction

	Temps de correction en secondes (avant réapprentissage)	Temps de correction en secondes (après réapprentissage)
Moyenne Animaux	421,123	298,141
Moyenne Mots en V	269,128	228,992
Moyenne générale	345,126	263,566

Figure 55 Comparaison des temps de correction avant et après apprentissage sur les fichiers « patients »

Nous constatons que le temps passé aux corrections a nettement baissé : -23,63%.

2.2.2 Questionnaire aux orthophonistes

Profil des orthophonistes interrogées

		2 ^{ème} questionnaire	1 ^{er} questionnaire
Sexe		100% femmes	96% femmes - 4% hommes
Age	21 - 30 ans	10%	40%
	31 - 40 ans	60%	24%
	41 - 50 ans	20%	17%
	51 -60 ans		14%
	> 60 ans	10%	5%
Lieu d'exercice	Libéral	70%	73%
	Milieu hospitalier	20%	11%
	Institutions		14%
	Autres	10% (SSR)	2%

Figure 56 Comparaison des profils des orthophonistes entre le 1^{er} et le 2nd questionnaire

Nous pouvons voir que l'âge diffère légèrement entre les participants aux 1^{er} et 2nd questionnaires en comparaison de tranches d'âge, même si on peut montrer que le pourcentage de participants dont l'âge est inférieur ou égal à 40 ans est proche (70% pour le 2^{ème} questionnaire et 64% pour le 1^{er}). Enfin, les lieux d'exercices sont proches.

Résultats des réponses

Afin de présenter les résultats de notre enquête suite aux démonstrations d'EvoLex aux orthophonistes, nous reprenons chaque question avec le détail des réponses.

Selon vous, le logiciel permet de gagner du temps par rapport aux tests « manuels »

6 des 10 orthophonistes que nous avons rencontrées ont un avis positif par rapport au gain de temps, en particulier pour les cotations qui se calculent automatiquement. Elles émettent toutefois des réserves par rapport à la reconnaissance vocale trop faible pour le moment, ce qui fait perdre trop de temps pour les corrections des fichiers. Certaines pensent également qu'il faut prendre du temps pour la vérification du fonctionnement du matériel (son du micro, démarrage du logiciel).

D'un autre côté, 4 d'entre elles ne pensent pas que le logiciel garantisse un gain de temps du fait des temps de correction encore trop élevés. Elles disent avoir l'habitude de faire des analyses rapidement et d'y consacrer très peu de temps.

Selon vous, le logiciel est simple d'utilisation

8 professionnelles pensent qu'en effet, le logiciel est simple à utiliser. 2 orthophonistes ne l'ont pas testé (du fait de l'absence de réseau wifi) et n'ont donc pas d'avis.

Selon vous, le logiciel est ergonomique

Sur les 10 avis, 6 sont positifs, en particulier pour les orthophonistes qui ont l'habitude d'utiliser EVALO. 2 ont un avis négatif du fait de l'aspect « expérimental » de la maquette et du fait de devoir valider/invalider chaque réponse du patient (en tapant sur A ou P). 2 n'ont pas testé le logiciel du fait du manque de réseau wifi, et n'ont donc pas d'avis.

Selon vous, le logiciel propose un large éventail d'analyses

Sur les 10 professionnelles interrogées, 9 le pensent, en effet. Seule une est moyennement d'accord.

Selon vous, le logiciel propose des analyses quantitatives intéressantes et pertinentes

8 orthophonistes ont un avis positif sur la pertinence des analyses quantitatives. 2 ont un avis mitigé. Elles pensent en effet que les analyses quantitatives peuvent être intéressantes si le patient est fluent, mais moins pertinentes si le patient est non fluent, auquel cas, selon elles, le logiciel n'apporterait pas plus que les tests manuels.

Selon vous, le logiciel propose des analyses qualitatives intéressantes et pertinentes

9 des 10 orthophonistes ont un avis favorable sur les analyses qualitatives, que ce soit sur le temps de latence, le temps d'initiation, la fréquence des mots ou les clusters et switches.

Seule une est mitigée sur le sujet, car elle dit faire ces analyses (clusters et switches) de façon manuelle.

Selon vous, le logiciel propose un profil récapitulatif intéressant et pertinent

On compte 6 avis positifs et 3 négatifs. Les principales remarques concernent la présentation en cible difficile à lire, et des questionnements sur la façon de lire les résultats (la partie en rouge montre-t-elle des mesures pathologiques ?). Une orthophoniste n'a pas donné de réponse.

Selon vous, le logiciel constitue une possibilité intéressante de stocker les données du patient et d'avoir ainsi une base pour un suivi

Toutes les orthophonistes ont une bonne opinion sur ces possibilités de suivi des données du patient. Pouvoir disposer d'éléments de comparaison quantifiés est intéressant, de même que le fait de pouvoir réécouter l'épreuve entière et notamment avec le patient. Il y a deux remarques qui nuancent ce plébiscite : ce serait un vrai atout de pouvoir extraire les données, et d'un autre côté, le respect de la confidentialité doit être garanti.

Selon vous, le logiciel permet d'être plus disponible pour l'observation clinique

8 des 10 orthophonistes interrogées pensent en effet que le logiciel permet d'être plus disponible pour l'observation clinique (mimiques, signes non verbaux), notamment grâce aux enregistrements que l'on peut réécouter. Elles disent toutefois que cette disponibilité est diminuée par le fait de valider chaque réponse. 2 d'entre elles pensent que la situation est la même que pour un passage de test en manuel.

Pensez-vous que le principe de reconnaissance vocale est intéressant ?

Les avis sont plutôt positifs (9/10), à condition que cela fonctionne bien. Elles pensent que la reconnaissance vocale est encore trop fragile. Lorsqu'elle sera fonctionnelle, cela pourrait permettre d'analyser les déformations, en ayant une transcription phonétique ou de pouvoir être un indicateur de dysarthrie (en fonction de la déformation du mot), ou encore d'analyser les modalisations (conscientes, non conscientes).

Pensez-vous qu'un ordinateur introduit une distance entre thérapeute et patient ?

Les avis sont partagés. 6 sur 10 pensent que l'ordinateur ne crée pas de distance. Certaines pensent que la distance est générée par la situation de bilan qui en elle-même est jugée artificielle, plus que par l'ordinateur. D'autres soulignent que la manière dont on présente le test peut changer les choses et que l'ordinateur est un outil comme les autres. Tout dépend de la manière dont il est utilisé, de son emplacement physique ; le placer entre le thérapeute et le patient ou sur le côté n'aura pas le même impact. Le contexte et la fréquence d'utilisation avec le patient peut aussi créer une familiarité avec l'outil.

4 sur 10 pensent que cela crée une distance. Une personne souligne que ça rend la situation de test plus apparente.

Pensez-vous qu'un micro-casque introduit une distance entre thérapeute et patient ?

Une majorité ne le pense pas (7/10). Certaines y voient même une aide à la concentration, d'autres expliquent que le casque est oublié rapidement.

3 orthophonistes sur 10 pensent en revanche que le micro-casque génère une distance. Certaines expliquent que c'est surtout gênant pour les personnes malentendantes ou les personnes âgées. D'autres regrettent qu'avec un micro-casque on ne s'adresse plus au thérapeute mais on produit pour l'ordinateur, ce qui peut notamment entraîner une modification de la voix (intensité, prosodie ?...).

Pensez-vous qu'il soit intéressant d'avoir un test de fluence informatisée indépendamment d'une batterie ?

Les avis sont largement favorables à cette idée (8/10). Divers avantages du test isolé sont mis en exergue :

- Dans une batterie il y a des effets planchers ou des effets plafonds possibles, un test isolé permet d'affiner une donnée.
- Disposer d'un test de fluence différent des batteries permet d'éviter un effet re-test.
- Les tests de fluence sont souvent utilisés seuls pour compléter un bilan ou confirmer/infirmier des hypothèses de diagnostic.

2 avis sont négatifs. Le coût notamment pour un test de fluence seul ne serait pas justifié.

Ce logiciel pourra permettre à terme de faire passer des épreuves de dénomination et de génération de mots, pensez-vous que ce soit intéressant ?

Presque toutes les personnes interrogées y sont favorables. Certaines à la condition qu'ils apportent des analyses qualitatives comme pour le test de fluence. D'autres soulignent que cela serait intéressant d'avoir des liens ou des corrélations entre les 3 tests pour un même patient.

Une personne ne se prononce pas car l'intérêt de ces épreuves dépendrait du contenu précis proposé et de leur plus-value par rapport aux tests existants.

Améliorations suggérées

- Techniquement :
 - Améliorer la reconnaissance vocale
 - Revoir le fait de valider/invalider les réponses avec le clavier
 - Par rapport au micro-casque : ne pas être contraint de porter un casque.
- Analyses quantitatives :
 - Faire un étalonnage des latences, clusters/switches, fréquence, initiation...
 - Elargir l'étalonnage à toutes les tranches d'âge : y compris les enfants (cf. dans le 1^{er} questionnaire, la question sur les types de pathologies concernés par les tests de fluence).
 - Avoir un lien ou une corrélation entre les épreuves de fluence sémantique et les épreuves de fluence formelle
- Analyses qualitatives
 - Ajouter des analyses concernant les registres de langue (familier, courant, soutenu)
 - Ajouter des analyses des erreurs : types d'erreurs, paraphasies...
 - Consolider, élargir, affiner les analyses des clusters et switches
 - Avoir le choix du temps de fluence : 1 min, 2 min (ce qui implique 2 étalonnages)

Commentaires

- Sur le logiciel même :
 - Le top départ est trop long (il faut en effet attendre 3 secondes avant que l'enregistrement ne commence)
 - Le visuel des clusters ne parle pas
 - Il manque des diagrammes pour les processus cognitifs
 - Le profil récapitulatif est une bonne base pour un bilan d'évolution
- En général :
 - Un logiciel permet de calculer les temps de latence, ce qu'on ne peut pas faire avec papier/crayon.
 - 2 minutes peuvent paraître longues mais peuvent faire émerger un décrochage attentionnel, des temps de latence (certains trouvent des mots juste avant la fin du test), et mettre vraiment en exergue des stratégies de recherche.

- Que faire concrètement des analyses en rééducation ?
 - L'outil informatique est intéressant et propose des calculs rapides et automatiques avec toutefois un biais : les critères de l'ordinateur sont renseignés avec la subjectivité de celui qui a rentré les données.
- Jusqu'où reconnaît-on un mot ? Dans les consignes du test manuel, le mot, s'il est reconnu par l'orthophoniste, même s'il est mal prononcé, est compté juste, alors que le logiciel peut ne pas le reconnaître.

Jugement du logiciel (prototype) sur une échelle de 1 à 10 :

La moyenne est de 6,2/10. Cinq personnes ont souhaité ajouter une note au concept du produit (à la condition d'une reconnaissance vocale fonctionnelle) et ont donné une note de 9/10.

II. Analyses des résultats

1. Analyse des résultats sur la reconnaissance vocale

1.1 Remarques préalables

Nous émettons quelques réserves sur les résultats présentés précédemment. Tout d'abord, d'un point de vue purement statistique, nous avons privilégié le hasard dans la sélection des fichiers à corriger, pour que le **plus grand nombre de voix différentes** soient représentées et utiles à un réapprentissage. Toutefois, il aurait été intéressant d'apparier les fichiers à corriger, c'est-à-dire, corriger les fichiers de chaque fluence pour un même sujet sain.

Ensuite, concernant les corrections en elles-mêmes, les conditions dans lesquelles nous avons réalisé la 2^{ème} phase de correction (le feu vert pour les dernières corrections a été donné très tardivement) ne nous ont pas permis de corriger autant de fichiers B que souhaité. Nous disposons de données sur 40 fichiers B, ce qui est moins que sur les fichiers A (80 au total). Nous notons cependant qu'au total **178 fichiers ont été corrigés** au minimum 1 fois : 42 fichiers ont servi de tests, mais sont « propres », 16 fichiers patients, 80 fichiers ont été corrigés 2 fois et 40 fichiers ont été corrigés 1 fois après le réapprentissage.

De plus, afin de **limiter l'effet intra-individuel**, nous aurions dû échanger nos fichiers. Toutefois, cela n'a été envisagé que tard dans l'avancement des travaux et des contraintes

techniques auraient généré des difficultés (bugs) pour lesquelles nous n'avions plus le temps de refaire des tests.

Nous avons cependant relevé des **écarts dans le nombre de mots corrects** après correction, entre les 2 phases de correction. Nous sommes conscientes que nos critères de jugement de la validité d'un mot ont pu être changeants. Nous nous sommes demandées si nous devons conserver les répétitions conscientes (le sujet redit un mot qu'il vient de dire, tout en précisant qu'il l'a déjà dit, mais cela l'aide à se relancer pour la suite, par exemple) ? D'autre part, les mots n'entrant pas dans la consigne (par exemple les noms propres pour la fluence formelle) ont été parfois marqués comme des erreurs, et n'ont finalement pas été comptabilisés dans le nombre total de mots corrects, alors qu'ils pourraient être traités comme erreur dans les analyses, sans que cela n'impacte le nombre de mots reconnus.

1.2 Comparaison de la reconnaissance vocale entre les versions 2015 et 2016 du logiciel

Nous constatons ainsi que le taux de reconnaissance vocale a été amélioré depuis la version 2015 du logiciel. En effet, il était de 30% en juin 2015, et la moyenne est passée à 49,83% avec la version 2 du logiciel (calculé lors de la première phase de corrections).

La version 2016 du logiciel a permis une nette amélioration de la reconnaissance vocale. L'amélioration technique du logiciel, la qualité des 344 fichiers enregistrés ont permis d'avoir une meilleure transcription.

Il est indéniable que le réapprentissage lancé après la première phase de corrections des 80 fichiers A1 a permis une nouvelle **augmentation de la reconnaissance vocale**. En passant à 63,41% ; elle a été doublée par rapport à la version initiale du logiciel, ce qui est significatif.

Les enseignants-chercheurs de l'IRIT ont mis au point trois modèles de réapprentissage :

- Le 1^{er} considérant les enregistrements de 2 minutes comme une grande phrase en continu
- Le 2^{ème} extrayant uniquement les mots (supprimant les silences...) et les mettant bout à bout
- Le 3^{ème} utilisant à la fois les textes lus (soit 1h02 au total) et les mots extraits

Le 1^{er} modèle a été écarté d'emblée car était totalement insatisfaisant. Le 3^{ème} bien que permettant un meilleur alignement dans le temps du début et de la fin des mots engendrait une moins bonne détection des mots et n'a donc pas été retenu non plus. C'est donc le 2^{ème} qui a été choisi pour la phase de réapprentissage.

1.3 Analyses des résultats par fluence

Lorsque nous tentons d'analyser les résultats pour chaque fluence, et d'autant plus si nous entrons dans le détail de chaque fichier, nous relevons des **écarts très importants** qui nous empêchent de tirer des conclusions pertinentes sur l'explication des phénomènes et des liens entre le nombre de mots reconnus automatiquement et le nombre de corrections effectuées, par exemple. Nous pouvons parler d'une tendance générale, mais pas d'une certitude. Il faudrait un nombre beaucoup plus important de fichiers corrigés pour prétendre à faire des corrélations entre les différents éléments analysés.

1.4 Temps de correction

En revanche, les analyses de temps de correction sont intéressantes, et le fait d'avoir **gagné 1 minute 30 par fichier** en moyenne est un réel atout. Nous nous sommes questionnées sur les raisons de cette diminution. Nous évoquons plusieurs hypothèses :

- Nous sommes devenues plus rapides dans les tâches de corrections à effectuer, car nous en avons pris l'habitude
- La reconnaissance vocale étant plus performante, il y a beaucoup moins de corrections à apporter
- Les fichiers patients nous ont semblé plus faciles à corriger. Nous pensons que c'est principalement dû au nombre de mots inférieur à celui des sujets sains (263 secondes pour les fichiers patients contre 284 secondes sur les fichiers recorrectés)
- La combinaison de ces 3 éléments peut également expliquer le gain de temps

1.5 Nombre et types de corrections effectuées

Nous avons volontairement écarté les analyses sur les nombres et types de corrections effectuées car nous pensons que les chiffres fournis par le serveur ne reflètent pas les corrections que nous avons effectivement réalisées.

En effet, nous avons appris que les modifications de position de segment (augmenter / diminuer la taille d'un segment, déplacer un segment) sont calculées de la façon suivante. Si la modification concerne un déplacement inférieur à 0,2 seconde, alors elle est comptabilisée comme « modification de position ». Si la modification concerne un déplacement supérieur à 0,2 seconde, alors, en fonction de l'action effectuée (réduction de taille ou augmentation de taille ou déplacement), la modification peut être comptabilisée comme un ajout ou une suppression. Or, de notre point de vue, cela **modifie le comptage des manipulations** ; pour nous, un ajout concerne un mot qui n'a pas été transcrit car non reconnu mais qui a été dit par le patient. A l'inverse, une suppression concerne un mot qui a été transcrit donc reconnu, alors que rien n'a été dit par le patient.

Dans ce cas précis, le point de vue retenu pour les statistiques était un point de vue « informatique ».

D'autre part, nous nous sommes rendu compte que nos **interprétations des erreurs** a pu varier : faut-il supprimer cette répétition, car elle est consciente, et laisser les répétitions inconscientes ? D'autre part, concernant les homophones, comme par exemple pour [VɛR], le mot transcrit automatiquement était systématiquement « vers ». Or, dans certaines situations, le sujet a dit « vert », car il était dans les couleurs et avait dit « vermillon » juste avant. Ou encore, le sujet venait de donner un nom d'animal (« verrat »), donc nous avons déduit qu'il parlait du « ver de terre ».

2. Analyse des résultats du questionnaire sur les avis a priori des orthophonistes sur le logiciel

Ce questionnaire a été très riche d'enseignements, et les avis a priori favorables au logiciel, notamment pour l'intérêt des analyses proposées nous ont confortées dans la plus-value que représente EvoLex.

2.1 Sur l'usage de l'informatique

2.1.1 La reconnaissance vocale :

L'usage de la reconnaissance vocale est peu répandu : 88,4% des personnes interrogées ont répondu ne pas s'en servir. Et les commentaires montrent que ceux qui y ont recours le font la plupart du temps dans le cadre de la rédaction (de compte-rendu par exemple). Il s'agit donc plus de dictée vocale (qui nécessite un apprentissage de la voix de l'utilisateur), ce qui diffère de la reconnaissance vocale dont nous parlons dans ce mémoire.

2.1.2 Le support :

Sachant que la mise au point du logiciel sur tablette est envisagée, nous avons posé la question du type de support utilisé. Comme on pouvait s'y attendre, la tablette est moins utilisée que l'ordinateur fixe ou portable, mais l'est plus que le smartphone. Avoir une application sur ce type de support vraisemblablement appelé à se développer semble donc intéressant.

2.1.3 Les logiciels :

Une très grande majorité des orthophonistes qui ont répondu (80%) utilisent des logiciels dans leur pratique professionnelle. C'est donc un fait plutôt encourageant pour le développement d'un nouveau logiciel.

2.2 Sur l'usage des tests d'évocation lexicale

2.2.1 Le test de fluence :

Les réponses des orthophonistes nous ont confirmé que les tests de fluence sont largement utilisés dans cette profession (68,5% ont répondu les utiliser toujours ou souvent), et la plupart les font passer manuellement avec papier/crayon et chronomètre. Rappelons que même informatisés, les tests de fluence existants ne font pas appel à la reconnaissance vocale. EvoLex constitue donc un apport innovant.

2.2.2 Les types de pathologies concernées :

Cette question concernant les types de pathologies pour lesquelles sont utilisés des tests de fluence nous paraît apporter une information intéressante dans le cadre de notre projet. On peut voir que ces tests sont souvent utilisés dans d'autres pathologies que celles relevant de la neurologie, même si elle demeure la plus importante. Cela rejoint ce que nous avons vu dans la partie théorique où il est expliqué que la fluence est souvent faible dans les troubles spécifiques du langage oral ou langage écrit, et donc utile à explorer dans ces domaines

également. La catégorie « Autre » est précisée dans les commentaires et concerne le bégaiement, le handicap mental, la surdité, les troubles envahissants du développement.

2.3 Sur l'intérêt a priori d'un logiciel comme EvoLex

2.3.1 Analyses quantitatives et qualitatives.

Les analyses quantitatives apportées sont majoritairement perçues comme moyennement satisfaisantes ou insatisfaisantes. Et les analyses qualitatives (types de mots produits ou stratégies utilisées par le patient) sont largement plébiscitées. On peut donc penser qu'un outil permettant d'obtenir des résultats quantitatifs immédiats et précis ainsi que des analyses qualitatives brossant un profil cognitif est susceptible d'intéresser ces orthophonistes.

2.3.2 Inconvénient a priori du logiciel.

La distance induite par le port d'un micro-casque est ce qui semble le plus gêner les orthophonistes. Parmi les points faibles du logiciel, les contraintes techniques (micro-casque et ordinateur) sont citées majoritairement, alors que le fait que la fluence soit testée habituellement dans une batterie ou un autre logiciel ne vient qu'en 3^{ème} position et ne semble donc pas constituer un obstacle en soi.

2.3.3 Atouts a priori du logiciel.

Ce sont principalement les analyses qualitatives qui constitueraient un atout.

3. Analyse des résultats des tests en situation écologique

3.1 Reconnaissance vocale des fichiers « patients »

3.1.1 Démonstration avec une « version locale » du logiciel

Présenter la version 2016 d'EvoLex suppose d'avoir un accès à Internet, notamment pour la communication avec le serveur et pour obtenir les analyses. Or, nous nous sommes trouvées dans des situations (hôpital, centre de rééducation) où cela n'était pas possible. L'équipe de l'IRIT a donc installé une version embarquée sur nos PC respectifs. Cela signifie que tout le processus, depuis la passation des tests jusqu'aux analyses pouvait se faire en autonomie sur notre ordinateur, sans passer par le serveur.

Cette version a été présentée à deux orthophonistes, mais il s'est avéré qu'elle était très lente et ne fonctionnait pas bien (notamment, le fichier audio ne pouvait pas être réécouté). Elle n'a donc pas permis de faire une démonstration représentative des possibilités du logiciel à partir des tests des patients. Mais nous avons pu a posteriori montrer les analyses obtenues avec des sujets sains.

3.1.2 Pathologies des patients

Les tests ont été effectués sur des patients aux pathologies diverses, à la fois des maladies neurodégénératives et des maladies acquises :

- AVC sylvien gauche (3 patients)
- Maladie d'Alzheimer (2 patients)
- Maladie de Parkinson (1 patient)
- Aphasie primaire progressive (1 patient)
- Hématome frontal gauche (1 patient)

3.1.3 Résultats

Les résultats obtenus pour le taux de reconnaissance vocale et les détails des corrections effectuées sont à nuancer du fait du petit nombre de fichiers.

3.2 Avis des orthophonistes ayant vu la démonstration du logiciel

Nous n'avons pas pu faire la démonstration prévue initialement dans la mesure où la phase de réapprentissage n'a pas été réalisée à temps. Les résultats de la reconnaissance vocale annoncés aux orthophonistes lors des démonstrations était alors encore faible (49,83%) et le temps de correction moyen de 6:12 minutes. Ces éléments ont pu engendrer une réticence ou un a priori négatif.

D'autre part, nous signalons que seules 3 orthophonistes ont pu voir les résultats de la reconnaissance vocale sur les fichiers de leurs patients. Pour les autres, nous avons fait les démonstrations et donné les explications sur la base de fichiers d'autres patients ou de sujets sains. En effet, la transcription des enregistrements des patients n'étant pas lancée simultanément à la passation des tests, nous n'avons pas pu corriger ces fichiers à temps pour qu'ils soient prêts pour la démonstration.

Enfin, certaines orthophonistes ont eu des difficultés à imaginer une version moins « expérimentale » et sont restées trop focalisées sur les graphiques et visuels du prototype.

. Validation des hypothèses et discussion

I. Validation des hypothèses

1. Sur l'amélioration de la reconnaissance vocale depuis juin 2015

1.1 Hypothèses techniques

- H1-1 : Nous avons supposé que la reconnaissance vocale pouvait être significativement améliorée par les enregistrements de la lecture de « La bise et le soleil » et d'un extrait du « Petit Prince ».

Nous ne pouvons pas valider cette hypothèse car les textes lus n'ont pas été utilisés dans le projet pour l'instant. Les enregistrements des lectures ont été inclus par l'IRIT dans un modèle de réapprentissage avec les mots des tâches de fluence extraits de chaque enregistrement et mis bout à bout. Ce modèle générait de meilleurs alignements sur le début et la fin des mots dans le temps mais de moins bonnes détections de mots. Il n'a donc pas été retenu.

- H1-2 : Nous nous sommes demandé si l'amélioration du système d'exploitation du logiciel et les enregistrements permettraient une amélioration de la reconnaissance vocale

Nous pouvons valider cette hypothèse. En effet, la moyenne des résultats de la reconnaissance vocale des fichiers A1, c'est-à-dire avant la phase de réapprentissage, montre une amélioration par rapport au taux de reconnaissance vocale de l'année dernière évalué à 30% : il est passé à 49,83%.

- H1-3 : Nous avons espéré que la correction des fichiers permettrait l'amélioration de la reconnaissance vocale par réapprentissage

Nous pouvons valider cette hypothèse, car nous avons pu montrer une nette amélioration du taux de reconnaissance vocale entre la correction des fichiers A1 (sans réapprentissage) et

A2 (les mêmes fichiers recorrectés après le réapprentissage) : 49,83% pour les premiers à 63,41% pour les seconds.

De même, on note une amélioration de la reconnaissance vocale entre les fichiers A1 (sans réapprentissage) et les fichiers B (fichiers corrigés une seule fois, après réapprentissage) qui passe de 49,83% à 62,75%,

1.2 Hypothèse écologique

- H1-4 : Nous avons voulu vérifier que la reconnaissance vocale fonctionnerait aussi bien avec des patients et donc que le taux de reconnaissance vocale avec des patients serait le même qu'avec des sujets sains.

Cette hypothèse est invalidée. En effet, pour les animaux, le taux de reconnaissance vocale des enregistrements de patients est de 59,66% contre 70,27% pour les fichiers A2, et de 28,02% pour les mots en V, contre 59,80% pour les fichiers A2.

2. Sur l'avis a priori des orthophonistes sur le logiciel

- H2 : ce type de logiciel est susceptible d'intéresser les orthophonistes

Notre 1^{er} questionnaire a permis de mettre en exergue l'intérêt suscité par le logiciel et les analyses qu'il propose auprès des orthophonistes sondés. 66,85% d'entre eux pensent que les analyses proposées par le logiciel sont un apport pour le bilan et la prise en charge des patients.

Nous pouvons donc valider cette hypothèse.

3. Sur l'utilisation du logiciel en situation écologique

- H3-1 : le logiciel apporte un gain de temps aux orthophonistes

Cette hypothèse s'est avérée difficilement vérifiable. Nous n'avons pas pu chronométrer le temps passé à faire les tests et les analyses de façon manuelle et via le logiciel. Nous ne pouvons donc pas valider cette hypothèse, par manque de preuve matérielle.

Toutefois, les orthophonistes pensent qu'avoir tous les calculs réalisés immédiatement et automatiquement avec des résultats fiables compense le temps passé à corriger le fichier du

patient. De plus, nous pouvons noter que le temps moyen de correction avec la version 2015 était de 5:51 minutes par fichier, contre 4:44 avec la version 2016 après réapprentissage. Le gain de temps en matière de correction est incontestable.

- H3-2 : les orthophonistes « testeurs » pensent que le logiciel apporte des analyses quantitatives et qualitatives pertinentes pour la clinique (aide au diagnostic et à la définition des axes thérapeutiques)

Le 1^{er} questionnaire « a priori » nous a laissé penser que les analyses apportées par EvoLex étaient jugées tout à fait intéressantes et pertinentes. Après démonstration du logiciel, cela s'est confirmé avec des réponses favorables des orthophonistes à ce sujet avec une moyenne de 6,22/10 pour le prototype et de 9/10 pour le logiciel lorsqu'il sera performant. Nous pouvons donc valider cette hypothèse.

II. Discussion

1. Amélioration de la reconnaissance vocale

1.1 Discussion sur notre contribution

- Au niveau clinique

Rencontrer de nombreuses personnes pour nos enregistrements s'est révélé être une expérience enrichissante à la fois dans l'observation de l'évocation lexicale et dans la pratique de la situation de test. Nous avons aussi souvent pris le temps de discuter avec les participants pour avoir leurs réflexions sur les tâches de fluence qu'ils venaient de réaliser et acquis ainsi une certaine aisance qui nous sera utile dans notre future pratique.

- Au niveau technique

Nous avons expérimenté le travail en groupe avec les enseignants-chercheurs, avec des attentes, des besoins, un vocabulaire spécifiques, des échelles de temps différentes. Nous aurions aimé avoir des résultats plus significatifs, mais la dernière version du réapprentissage nous a été remise trop tard. Nous retiendrons de ce partenariat qu'un engagement de tous, une bonne communication et des synthèses régulières sont nécessaires au bon déroulement d'un projet. Cela pourra nous être utile pour notre pratique

future, dans le cas de travail en équipe (milieu hospitalier, institutions, ou encore équipe pluridisciplinaire, par exemple).

1.2 Discussion sur les aspects techniques

Avant tout, nous soulignerons que l'outil est beaucoup plus fonctionnel que l'an dernier. Le fichier audio s'écoute et se réécoute facilement, des pauses et des retours-arrière sont possibles, les analyses s'obtiennent en un clic...

D'un point de vue purement technique et pratique, les informaticiens-chercheurs avaient besoin d'enrichir une base de données avec des enregistrements de qualité puis une correction précise des fichiers : nous avons donc passé beaucoup de temps à relever des bugs, à faire des tests lors de la mise au point de la phase de correction, à faire les corrections de 178 fichiers. L'IRIT dispose désormais d'une base de fichiers « propres » qui sera amenée à être réutilisée à l'avenir dans le processus d'amélioration de la reconnaissance vocale de ce logiciel.

Dans cette perspective, nous suggérons quelques améliorations possibles ou points à analyser.

1.2.1 Concernant les corrections à effectuer

➤ Ajouts

Les ajouts de mots seraient à simplifier, par exemple au moyen d'une liste de propositions prédictives, établie d'après les mots précédents ou suivants, ou d'après les erreurs le plus souvent commises, ou d'autres options à définir

➤ Propositions de mots

Prévoir un menu déroulant avec des propositions de mots pour les modifications de mots, afin de ne pas perdre le temps de la saisie du mot correct

1.2.2 Concernant la base de données

➤ Mots dits fréquemment

A partir des fichiers enregistrés et corrigés, répertorier les mots les plus fréquemment donnés pour qu'ils soient reconnus à coup sûr.

➤ Utiliser la prédiction de mots (pour faciliter et perfectionner la reconnaissance vocale)

Si l'on dit « chien », il y a de très fortes chances qu'on dise « chat » en suivant. Utiliser les analyses de clusters de la base de données pour affiner cette prédiction.

- Enrichissement des dictionnaires

Parallèlement certains mots qui reviennent très fréquemment ne sont pour l'instant jamais reconnus. Par exemple, pour la catégorie des fruits : goyave, coing, litchi, clémentine, mandarine, fruit de la passion, noix de coco... Il serait intéressant de procéder à des vérifications sur leur appartenance au dictionnaire, leur transcription en phonétique informatique ou leur correspondance avec le signal acoustique, etc.

- Homophones

Les homophones constituent une réelle limite pour les analyses : un substantif à la place d'un autre n'induit pas le même lien sémantique, et donc modifie les clusters, de même qu'un substantif à la place d'un verbe, ou inversement. Par exemple, **virer** quelqu'un ou faire une **virée** dans tel endroit est très différent. Sans avoir de certitude à 100%, peut-être serait-il possible d'utiliser le contexte pour décider d'un mot ou d'un autre ?

- Mots composés

Nous avons constaté également des erreurs lorsque le sujet disait des mots composés ou pouvant être décomposés en plusieurs mots. Par exemple, « pivert » était segmenté en « pie » + « ver », ou encore « cochon d'Inde » en « cochon » + « dinde ». Il fallait donc supprimer le 2^{ème} segment et corriger le mot du premier.

1.2.3 Concernant la passation :

- Données du patient

Parmi les données du patient, il nous semble intéressant de pouvoir renseigner sa pathologie et la date d'apparition de la pathologie

- Apparition de la consigne

il nous paraît indispensable que la consigne de l'épreuve soit spécifiée sur l'écran de passation

- Discussion sur la consigne de Cardebat

Celle-ci nous paraît trop complexe. Nous avons en effet dû la réexpliquer quasiment systématiquement aux patients, mais aussi, souvent, aux sujets sains.

➤ Pause

Même si l'on est dans une situation de test, il nous paraît important dans la réalité écologique d'avoir la possibilité d'arrêter momentanément le chronomètre.

➤ Remarques sur la validation / invalidation des réponses

Un abandon ou une simplification de la tâche de validation/invalidation des réponses. Ce procédé permet d'avoir un codage temporel repéré par un point vert sur l'onde acoustique lorsqu'un mot a été validé par l'orthophoniste (donc prononcé par le sujet). Si aucun mot n'est reconnu, on sait qu'il y a une correction à apporter à cet endroit précis. Une variation du signal acoustique permet de savoir qu'il y a du son, mais ne permet pas de déterminer s'il s'agit d'un bruit de souffle, d'un « heu » ou d'un « mmm ».

L'inconvénient majeur de ce procédé est la perte de disponibilité clinique de l'orthophoniste, atout a priori important du logiciel. L'IRIT ne s'est pas servi du codage temporel pour les réapprentissages, car certaines validations arrivaient avant la fin des mots ou quelques millisecondes après, et on a constaté que les repères (points verts) sur certains fichiers ne correspondaient à rien ; il y avait donc eu des décalages entre le « clic » pendant la passation et sa matérialisation sur l'onde acoustique. Si les codages temporels ne sont pas utilisés, on pourrait tout simplement abandonner cette idée de validation/invalidation. Mais si l'on veut garder une trace visuelle qu'un mot a été dit, on pourrait la simplifier en demandant au thérapeute d'appuyer sur une même touche chaque fois qu'un mot est prononcé, qu'il soit correct ou non (ce qui ne devrait pas requérir trop de concentration de sa part).

1.2.4 Tentative d'explication sur les différences de chiffres entre les tâches de fluence

Nous avons vu dans la figure 40 que les différences de taux de reconnaissance vocale étaient significatives entre les animaux ou les fruits et les mots en R. Nous pouvons tenter d'expliquer ces différences. Pour les fluences sémantiques, les dictionnaires sont les plus restreints (1489 mots pour les animaux et 542 pour les fruits, contre 5430 pour les mots en R et 2819 pour les mots en V). Ainsi, il y a plus de risques d'erreurs pendant la transcription des productions en fluence formelle, ce qui peut expliquer un taux de reconnaissance vocale plus faible.

2. Enquête sur les habitudes des orthophonistes en matière d'informatique et de tests de fluence

➤ Utilisation des analyses

Nous avons été confrontées à des questions, notamment sur l'utilité des données apportées par les analyses des clusters : de quelle façon les intégrer en rééducation ? Nous pouvons ici proposer les pistes suivantes :

- Travail sur le réseau sémantique avec des classifications (mettre ensemble ce qui va ensemble, et en affinant au fur et à mesure les sous-catégories)
- Stratégies sans erreur, en multi modalité, en travaillant la désignation, la dénomination orale, écrite ou par tout autre moyen (gestes, dessin, etc.), des appariements image/mot écrit, des complétions de mots, puis de phrases, par exemple.
- Fluence en séance de temps en temps avec aide à la mise en place de stratégies. Par exemple, faire nommer les vêtements qui sont dans le placard du patient, commencer par ce qui est suspendu, puis ce qui est dans un tiroir, ou alors ce que l'on met en fonction de la saison, etc.

Les réponses à ce questionnaire nous ont beaucoup apporté. Nous avons recueilli l'avis de professionnels par rapport à leur pratique quotidienne, ce qui nous a permis d'avoir une vision objective de la réalité du terrain.

Ainsi, nous avons pu constater qu'une majorité des orthophonistes interrogés étaient plutôt favorables au fait que des analyses sur les stratégies utilisées par le patient dans les tâches de fluence soient réalisées de façon automatique. Cela nous a bien entendu confortées dans notre conviction sur le bien-fondé du logiciel.

➤ Sur un plan technico-pratique

Certains orthophonistes ont fait état de leurs craintes quant aux risques de problèmes techniques, au fait que l'on doive transporter un ordinateur à domicile, la nécessité d'avoir une connexion internet, même en différé, mais également sur une éventuelle complexité et/ou perte de temps de mise en place, de contrôle du fonctionnement du matériel. Pour avoir pratiqué le logiciel de nombreuses fois, nous pensons que la démonstration de sa simplicité d'utilisation pourrait finir de convaincre les plus réticents.

➤ Concernant les supports des tests de fluence

Nous avons noté qu'une large majorité des orthophonistes utilisent encore exclusivement des tests manuels (70%). Cela peut s'expliquer par le fait que peu de logiciels de bilan sont disponibles sous forme de logiciels (BIA, EXALANG pour les principaux) et que d'autre part, les batteries largement utilisées telles que le MT-86 ou le BDAE, ou encore les tests tels que Cardebat ou Grefex sont sous format « papier ». Certes, les habitudes sont parfois difficiles à changer, mais cela laisse penser qu'un logiciel tel qu'EvoLex a un bel avenir !

➤ Concernant la relation au patient

D'autres professionnels ont fait état d'une possible perte de contact avec le patient, notamment par le fait que le casque puisse trop l'isoler, qu'il soit trop intimidé ou stressé par le dispositif et donc limite sa participation. Nous avons pu constater que ce n'était pas le cas lorsque nous avons réalisé les tests auprès de patients ; seule une patiente a été perturbée par le casque, mais l'orthophoniste a expliqué qu'elle l'était de toute façon dès qu'elle faisait un bilan.

➤ Orthophoniste versus logiciel ?

Certains orthophonistes ont indiqué qu'on ne pouvait pas remplacer une analyse clinique humaine du patient et de son fonctionnement par une machine. Cela nous semble évident, et ce n'est pas l'objectif du logiciel. Nous travaillons ici sur un outil qui permettra, nous l'espérons, de dégager de la disponibilité à l'orthophoniste pendant la réalisation du test, afin de mieux observer le patient (ses gestes, ses signes de recherche, les paraphasies, etc.), sans avoir à écrire tous les mots produits et à surveiller le chronomètre. D'autre part, nous pouvons répondre que le fait que l'orthophoniste prenne des notes pendant que le patient dit les mots peut également le stresser ; il peut se demander quelles sont les annotations du thérapeute, si elles sont positives ou négatives, etc. ?

Ces remarques nous ont fait réfléchir sur le rôle et la place de l'orthophoniste en situation de test : empathie, neutralité, encouragements, confiance. La situation de test est par essence antinaturelle, puisque les résultats aux tests peuvent dépendre de la forme du patient à ce moment-là, de son état de stress, de son attitude face aux difficultés, de sa nosognosie ou non. Même si l'utilisation d'un micro-casque et d'un ordinateur peuvent introduire une distance entre le patient et l'orthophoniste, la manière d'aborder la situation de test et donc d'expliquer l'utilisation de ces outils devrait limiter leurs effets négatifs. L'attitude de l'orthophoniste peut elle aussi être bénéfique ou non au patient qui réalise les tests.

3. Tests en situation écologique

3.1 Un travail de terrain

Comme avec les sujets sains, ces démonstrations ont été très riches pour nous, notamment en matière d'observation clinique, de contact avec le patient, ou des discussions qui s'en sont suivies avec les orthophonistes. Nous avons ainsi pu mieux appréhender la réalité du terrain et percevoir les atouts et les limites du logiciel.

Cette phase de tests et de démonstration nous a permis d'être au plus près des pratiques des orthophonistes et des réactions des patients face au dispositif.

- Peu se sont finalement montrés gênés par le casque, contrairement aux réticences a priori de certaines orthophonistes ayant répondu au 1^{er} questionnaire
- Il apparaît clairement que la confiance entre l'orthophoniste et le patient permet de bonnes conditions de passation ; la situation de test étant de toute façon artificielle, cette confiance permet au patient d'être dans les meilleures conditions possibles, avec ou sans ordinateur (dans la mesure où celui-ci n'est pas disposé comme une « barrière » physique entre le professionnel et le patient, mais sur le côté du bureau par exemple).

3.2 Les analyses : la plus-value du logiciel

Disponibles juste après la passation de l'épreuve, elles permettent d'avoir immédiatement un profil. Elles permettront un gain de temps certain, mais aussi d'avoir des résultats chiffrés pouvant être utilisés dans les rédactions de bilan et dans l'appréciation du suivi du patient. D'ici-là, l'analyse des clusters aura dû être affinée.

3.3 Intérêt de la reconnaissance vocale

La reconnaissance vocale est un vrai plus, car cela décharge l'orthophoniste de l'écriture des mots, permet de conserver un enregistrement des productions du patient et ainsi d'y revenir éventuellement pour analyser de façon plus fine les modalisations, déformations de mots, etc.

4. Perspectives du logiciel

- La reconnaissance vocale pourrait être utilisée pour analyser les modalisations et déformations (paraphasies, etc.), pour une transcription phonétique des homophones
- Le logiciel est prévu pour des tâches de fluence, de dénomination et de génération de mots. Cela sera un atout dans la mesure où tous les résultats pourront être analysés et mis en lien. Il y aura ainsi un réel gain de temps cumulé.
- Des étalonnages des clusters, switches, temps de latence, d'initiation, fréquence de mots seraient un atout important, car cela n'existe pas encore. La réalisation de ces étalonnages est possible avec les fichiers déjà enregistrés, et sera à compléter pour une normalisation future.
- L'ergonomie et les visuels devront être perfectionnés, afin d'avoir une version moins « expérimentale ». De même, la cible devra être harmonisée ; la partie en rouge ne devra regrouper que les résultats faibles ou pathologiques.
- Il est prévu un changement majeur de système utilisant une technologie extrêmement complexe mais pleine de promesses ; à savoir l'utilisation d'un système DNN (Deep Neuronal Network), ou Intelligence Artificielle, dont le fonctionnement veut se rapprocher du fonctionnement neuronal humain, pour l'analyse de la parole. Une personne de l'IRIT va y travailler à l'avenir, ce qui devrait permettre le franchissement d'un grand palier dans la performance de la reconnaissance vocale

Conclusion

A travers ce mémoire, nous avons poursuivi le travail entrepris l'année dernière, en tentant d'améliorer EvoLex, un logiciel d'évaluation de la fluence verbale avec reconnaissance vocale.

Rappelons qu'une tâche de fluence verbale permet de mettre en évidence le fonctionnement cognitif et exécutif d'un patient, notamment par les stratégies mises en place, la flexibilité mentale, la planification, dans le but de trouver le plus de mots possibles dans une catégorie donnée ou commençant par une lettre donnée. Grâce au logiciel - après l'enregistrement du patient et la transcription de ses productions via la reconnaissance vocale - ces analyses peuvent être automatisées.

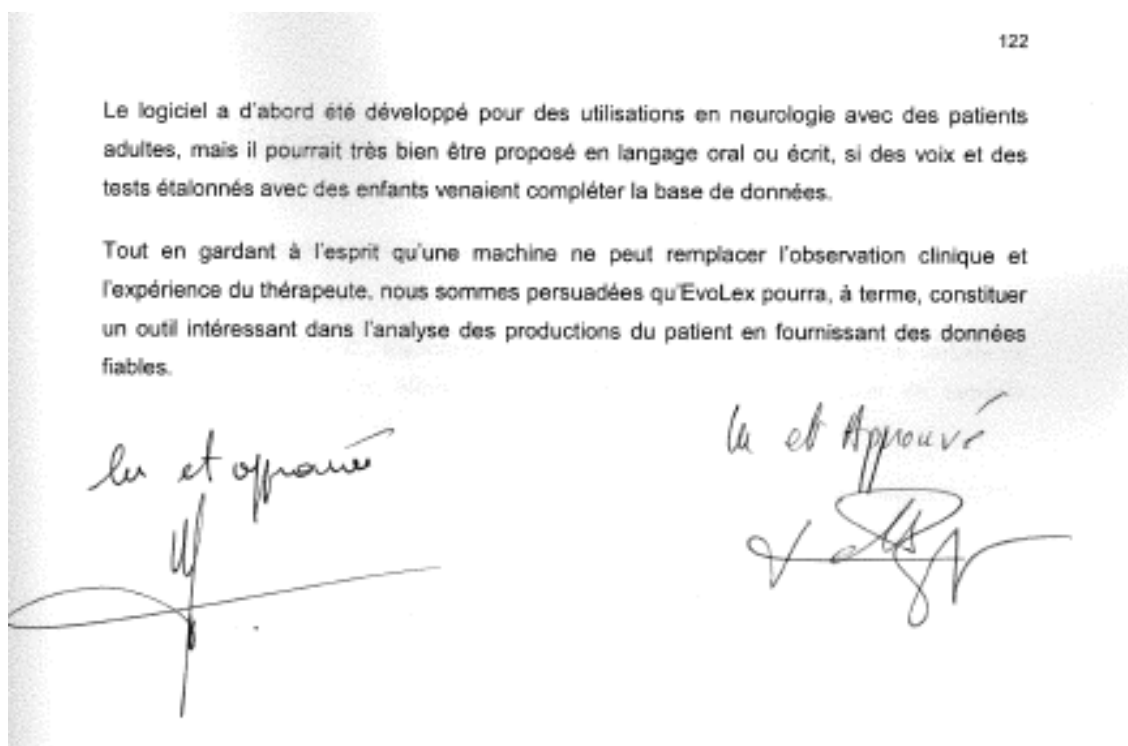
Ce mémoire est le fruit d'un travail conjoint avec une équipe d'enseignants-chercheurs en informatique de l'IRIT. L'objectif était double : il s'agissait en premier lieu de contribuer à l'amélioration de la reconnaissance vocale du logiciel EvoLex, créé l'année dernière, afin de permettre de mettre en valeur les analyses quantitatives et qualitatives apportées automatiquement par l'outil lors de la réalisation de tâches de fluence. Et en second lieu, de s'intéresser au point de vue écologique en recueillant les avis et impressions des orthophonistes sur le concept et le produit.

Nous nous sommes attachées à effectuer des enregistrements de sujets sains sur 4 tâches de fluence (sémantique et formelle), et à corriger ces fichiers transcrits via le logiciel. Ces enregistrements de qualité, les corrections rigoureuses, ainsi que des améliorations techniques ont permis d'améliorer la reconnaissance vocale du logiciel. EvoLex devient donc plus performant et pertinent, comme l'ont souligné les orthophonistes auprès desquelles nous avons pu faire des tests avec des patients, et ainsi recueillir leur avis suite à la démonstration du logiciel.

Bien entendu, des améliorations sont encore à apporter, sur un plan technique d'abord, car la reconnaissance vocale bien qu'indéniablement améliorée n'est pas encore suffisamment au point - surtout en situation réelle - pour permettre une exploitation immédiate du logiciel, mais aussi linguistique (dictionnaires à étoffer et à mettre à jour, par exemple). Il serait également intéressant d'envisager un affinage des analyses des stratégies et peut-être un étalonnage des clusters, des temps de latence, etc.

Le logiciel a d'abord été développé pour des utilisations en neurologie avec des patients adultes, mais il pourrait très bien être proposé en langage oral ou écrit, si des voix et des tests étalonnés avec des enfants venaient compléter la base de données.

Tout en gardant à l'esprit qu'une machine ne peut remplacer l'observation clinique et l'expérience du thérapeute, nous sommes persuadées qu'EvoLex pourra, à terme, constituer un outil intéressant dans l'analyse des productions du patient en fournissant des données fiables.



. Bibliographie

Bonin P., Le choix des mots, Cerveau et Psycho n°30, Paris : Pour la science, novembre-décembre 2008

Boulangé A., MARTIN C., Elaboration et normalisation d'épreuves de fluence verbale de verbes et de substantifs. Mémoire en vue de l'obtention du certificat de capacité d'orthophoniste, Lyon I, 2014

Brin-Henry F., Courrier C., Lederlé E., Masy V., Dictionnaire d'orthophonie, Isbergues: Ortho Edition, 2011

Caelen J., L'ordinateur au doigt et à l'œil, A l'écoute de la parole humaine, La Recherche mars 1996, n°285

Calliope (nom collectif) sous la responsabilité de Tubach J.-P., Boë L.-J., Caelen J., Descout R. et al., La parole et son traitement automatique, IV^{ème} partie, Paris : Masson, 1989

Chomel-Guillaume S., Leloup G., Bernard I. Les aphasies – évaluation et rééducation, Issy-les-Moulineaux : Masson, 2010

Courtade-Jouanicq A., Balland O., Etude de faisabilité d'un logiciel de reconnaissance vocale adapté à des tâches d'évocation lexicale. Mémoire en vue de l'obtention du certificat de capacité d'orthophoniste, Toulouse III, 2015

Coutaz J., L'ordinateur au doigt et à l'œil, L'art de communiquer à plusieurs voies, La Recherche mars 1996, n°285

Crowe S., Decrease in performance on the verbal fluency test as a function on time : evaluation in a young health sample. Journal of clinical and experimental neuropsychology, vol. 20, n°3, pp. 391-401, 1998

De Boysson-Bardies B., Comment la parole vient aux enfants ? Paris : Odile Jacob, 1996

Dehaene S., Les neurones de la lecture, Paris : Odile Jacob, 2007

De Singly F., L'enquête et ses méthodes – le questionnaire (3^{ème} édition), Paris : Armand Colin, 2014

De Weck G., Marro P., Les troubles du langage oral chez l'enfant, Issy-les-Moulineaux : Masson, 2010

Dubois B., Michon A. et al. Démences, Montrouge : Doin, 2015

Eustaches F., Desgranges B., Vers un modèle unifié de la mémoire, Cerveau et Psycho n°28. Paris : Pour la science, juillet-août 2008

Gierski F., Ergis A.-M., Les fluences verbales : aspects théoriques et nouvelles approches. L'année psychologique volume 104, 2004 : 331-353

Godefroy O. et le GREFEX, Fonctions exécutives et pathologies neurologiques et psychiatriques – évaluation en pratique clinique, Marseille : Solal, 2008

Haton J-P, Fohr D., Laprie Y., Cerisara C. Reconnaissance automatique de la parole - Du signal à son interprétation, Paris : Dunod, 2006

Hirshon E., Thompson-Schill S., Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. Neuropsychologia, 2006, vol.44 : 2547-2557

Lussier F., Flessas J., Neuropsychologie de l'enfant – Troubles développementaux et de l'apprentissage. Paris : Dunod 2009

Marchal A., Précis de physiologie de la production de la parole, Marseille : Solal, 2011

Mariani J. Reconnaissance automatique de la parole : progrès et tendances, Traitement du Signal, Vol. 7 n°4, 1990

Mariani J., Reconnaissance de la parole, Traitement automatique du langage parlé, Hermes Science Publications, Paris, 2002

Marmion J.-F., Le langage sur le bout de la langue, Sciences humaines n°274, septembre-octobre 2015

Marquier-Morin E., Lahille P., L'évocation lexicale formelle et sémantique : établissement d'une norme dans une population française. Mémoire en vue de l'obtention du certificat de capacité d'orthophoniste, Toulouse III, 1989

Martin G., Comment la parole est venue aux machines, La Recherche, septembre 2003 n°367

Mayr U., On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss, *Neuropsychologia*, 2002, vol. 40.562-566

Mazaux J.-M., Pradat-Diehl P., Brun V., *Aphasies et aphasiques*, Issy-les-Moulineaux : Masson, 2007

Mazaux J.-M., de Boissezon X., Pradat-Diehl P., Brun V., *Communiquer malgré l'aphasie*, Montpellier : Sauramps Médical, 2014

Rousseau T., Gatignol P., Topouzkhianian S. et al. *Les approches thérapeutiques en orthophonie, Tome 1 - Prise en charge orthophonique des troubles du langage oral*, Isbergues : Ortho Edition, 2013

Rousseau T., Gatignol P., Topouzkhianian S. et al. *Les approches thérapeutiques en orthophonie, Tome 2 - Prise en charge orthophonique des troubles du langage écrit*, Isbergues : Ortho Edition, 2013

Seron X., Van Der Linden M., *Traité de neuropsychologie clinique de l'adulte, Tome 1 – Evaluation*, Louvain-La-Neuve (Belgique) : De Boek, 2014

Siciak-Tartaruga A., *Démographie professionnelle*, dossier publié dans la revue *L'Orthophoniste* n°344, Isbergues : Ortho Edition, décembre 2014

Thibault M.-P., Lenfant M., Helloin M.-C., *EXALANG 11-15 ans, Batterie informatisée d'examen du Langage Oral, Langage écrit, Compétences transversales*, Grenade : Orthomotus, 2009

Thibault M.-P., Helloin M.-C., Croteau B., *EXALANG 5-8 ans*, Grenade : Orthomotus, 2010

Thibault M.-P., Lenfant M., Helloin M.-C., *EXALANG 8-11 ans, Batterie informatisée d'examen du Langage Oral, Langage écrit, Mémoire, Attention, Compétences transversales*, Grenade : Orthomotus, 2012

Troyer A., Moscovitch M., Winocur G., Alexander M., Stuss D., Clustering and switching on verbal fluency: the effects of focal frontal-and temporal-lobe lesions, *Neuropsychologia*, vol. 36, n°6, pp 499-504, 1998

Vaissière J., *La phonétique*, Paris : Puf, 2e édition, 2011

Zarino B., Crespi M., Launi M., A new standardization of semantic verbal fluency test. *Neurol Sci*, vol. 35, pp 1405-1411, 2014

Sites

Cardebat D et al, Fluence verbale catégorielle et lexicale, animaux, fruits, meubles, lettres P, R et V, 1990, consulté via <http://www.notices-pdf.com/fluence-categorielle-de-cardebat-pdf.html#a0>

Lebel H., Paquette R., Critères diagnostiques démence / trouble neurocognitif majeur, site Psychomedia (Canada), article publié le 17/08/2015. <http://www.psychomedia.qc.ca/dsm-5/2015-08-17/criteres-diagnostiques-demence-trouble-neurocognitif-majeur>, consulté le 07/01/2016

Lussier F., test de fluidité verbale (1995), http://www.orthophonie.fr/fluidite_verbale.pdf, consulté le 13/01/2016

<http://www.cochlea.eu/oreille-generalites/oreille-externe> site consulté le 18/01/2016

https://fr.wikipedia.org/wiki/Reconnaissance_automatique_de_la_parole Site de Wikipédia consulté le 11/01/2016

<https://synthesevoix.wordpress.com/reconnaissance-vocale/definition-et-principe/> Site consulté le 6/01/2016

http://rdorat.free.fr/Enseignement/ELN/Seance7/Reconnaissance_vocale.pdf consulté le 6/01/2016

<http://lexique.org/>, Université de Savoie (Chambéry – Annecy) et CNRS, site consulté le 23/05/2016

<http://pontt.net/> - Partage Orthophonie Neuropsychologie Théories Thérapies, consulté le 15/01/2016

<http://flenet.unileon.es/> Site espagnol de Français Langue Etrangère, consulté le 18/01/2016

. Table des annexes

Annexe 1	Epreuves et étalonnage de Cardebat	128
Annexe 2	Lettre de recrutement des participants	131
Annexe 3	Protocole de passation	132
Annexe 4	Textes lus « La bise et le soleil » et extrait du « Petit Prince ».....	139
Annexe 5	Questionnaire (enquête sur les pratiques professionnelles en matière d'informatique et de tests de fluence)	140
Annexe 6	Questionnaire visant à recueillir l'avis des orthophonistes après démonstration du logiciel.....	144

FLUENCE VERBALE CATÉGORIELLE ET LEXICALE
ANIMAUX, FRUITS, MEUBLES
LETTRE P, R, V
(Cardebat D, et al 1990)

MOTS CLÉS :

Test de langage, Test des fonctions exécutives, Fluence verbale

FAMILLE :

Test d'évaluation de plusieurs fonctions cognitives : le langage/ Fluence verbale/ mémoire de travail / mémoire sémantique

PRINCIPES:

La fluence verbale est un test simple de production de langage. Il est bien sûr très lié au NSC.

CONSIGNES:

L'ordre de passation est aléatoire, trois critères sémantiques ont été retenus : « Animaux » considéré comme riche en items, « Meubles » considéré comme pauvre en items et « Fruits » considéré comme intermédiaire. Une fluence lexicale et une fluence catégorielle suffisent en clinique courante pour se faire une idée.

Pour la fluence Lexicale, dire au patient :

« Vous allez devoir me dire le plus de mots français possibles, soit des noms, soit des verbes, soit des adjectifs, etc... commençant par la lettre que je vais vous donner et ceci en deux minutes. Ne dites pas des mots de la même famille, des noms propres et ne vous répétez pas.... Par exemple avec la lettre L, vous pouvez dire Lune, Laver, Laborieuse, ...Avez-vous bien compris ?».

Pour la fluence Sémantique ou catégorielle, dire au patient :

« Vous allez devoir me dire le plus de noms que vous connaissez, dans la catégorie que je vais vous donner et ceci en deux minutes ; sans noms de la même famille, et sans répétitions.... Par exemple pour la catégorie des Fleurs, vous pouvez me dire Rose, Dahlia, Violette ...Avez-vous bien compris?»

Déclencher le chronomètre et noter les réponses à la 1^{ère} et à la 2^{ème} minute (ambigu... « noter les réponses correspondant à la 1^{ère} minute et à la 2^{ème} minute » (. A des fins de recherche, il est proposé de noter par un tiré le passage de toutes les 15 secondes)

1. Noter le total de mots produits.**2. Noter le nombre de répétitions.**

Une répétition (ou persévération) est un même mot reproduit une ou plusieurs fois, par exemple un mot produit trois fois donne deux répétitions).

3. Noter le nombre d'erreurs.

L'erreur est représentée par le non-respect de la consigne ci-dessus (mot commençant par une autre lettre que celle donnée ou ne correspondant pas à la catégorie), les logatomes, les noms propres ou mots étrangers.

COTATION :

Par exemple en fluence sémantique animaux, « Canard, cane, caneton » seront comptabilisés pour 3 mots produits et 2 erreurs. En revanche, « cheval, jument, poulain » comptent simplement pour trois mots produits.

Si un hyperonyme est décliné (nom de la catégorie sémantique), seules les déclinaisons comptent : « **animaux**, pie, corbeau, perroquet » comptent pour quatre mots produits et une erreur (« oiseau » doit être considéré comme une erreur, dès lors que le patient a été prévenu).

Le patient peut donner des noms d'animaux commençant par la lettre qui avait été demandée précédemment.

Ex : le patient peut donner des mots commençant par P qui soient des animaux.

En cas de mot produit à la 1^{ère} minute et répété à la 2^{ème} minute, le rapporter dans le nombre de répétitions de la 2^{ème} colonne.

N.B. : Si le patient produit plus de 22 mots à la 1^{ère} minute, continuer de noter les réponses dans la 2^{ème} colonne et noter dans la 2^{ème} colonne le passage entre la 1^{ère} minute et la 2^{ème} minute.

Le score total est le nombre total de mots produits moins les erreurs ou répétitions.

INTERPRÉTATION :

la fluence verbale est précocement touchée dans la maladie d'Alzheimer mais aussi lors de toute atteinte des fonctions exécutives, (ou encore chez des patients aphasiques après AVC). il faut cependant noter les autres composantes influençant la performance à ce test : influence de la mémoire de travail et de la mémoire sémantique pour la fluence sémantique

Il faut rapporter le nombre de mots corrects aux normes publiées dans l'article de référence selon le sexe, l'âge et le NSC. (tableau reproduit ci-dessous). Cet article détaille l'étalonnage réalisé chez des sujets normaux (cf : tableaux ci-dessous).

RÉFÉRENCES ET COPYRIGTS**REFERENCE :**

Cardebat D, Doyon B, Puel M, Goulet P, Joannette Y. Formal and semantic lexical evocation in normal subjects. Performance and dynamics of production as a function of sex, age and educational level. Acta Neurol Belg. 1990;90(4):207-17. (in French).

COPYRIGHT PROTÉGÉ POUR TOUTE REPRODUCTION

© : équipe de Toulouse

Contact : Dr Jean-François DEMONET <demonet@toulouse.inserm.fr>

OU Dr Michèle PUEL <puel.m@chu-toulouse.fr>

Scores chez des Hommes sans pathologie cérébrale (n= 84) selon Cardebat D. et al 1990

Niveau I	P	R	V	Animaux	Fruits	Meubles
30-45 ans	17,28 ± 5,49	13,07 ± 5,49	12,71 ± 4,92	29,35 ± 7,79	15,78 ± 3,48	12,21 ± 2,32
50-65 ans	19,21 ± 5,56	17,64 ± 6,77	14,64 ± 8,82	33,64 ± 6,59	18,71 ± 4,23	16,28 ± 3,02
70-85 ans	19,28 ± 7,05	16,78 ± 6,04	14,50 ± 6,46	27,14 ± 8,53	15,42 ± 3,85	11,71 ± 3,53
Niveau II	P	R	V	Animaux	Fruits	Meubles
30-45 ans	20,64 ± 6,83	18,14 ± 8,39	15,00 ± 5,9	35,78 ± 10,67	17,50 ± 3,91	15,00 ± 3,82
50-65 ans	22,71 ± 7,24	19,42 ± 8,37	16,64 ± 5,04	31,85 ± 6,93	18,64 ± 4,43	14,85 ± 3,97
70-85 ans	20,78 ± 7,32	18,21 ± 7,08	15,71 ± 7,05	29,71 ± 11,97	16,00 ± 7,86	14,35 ± 4,55

Niveau d'étude I : ≤ 9 ans d'études

Niveau II : > 9 ans d'études

Scores chez des Femmes sans pathologie cérébrale (n= 84) selon Cardebat D. et al 1990

Niveau I	P	R	V	Animaux	Fruits	Meubles
30-45 ans	19,07 ± 10,01	14,35 ± 5,91	13,07 ± 5,85	26,00 ± 7,46	17,64 ± 4,19	15,42 ± 4,12
50-65 ans	14,86 ± 5,36	13,07 ± 5,79	11,21 ± 3,62	27,35 ± 7,72	19,35 ± 3,34	14,50 ± 4,03
70-85 ans	14,71 ± 7,56	13,00 ± 6,97	11,42 ± 4,98	29,71 ± 5,32	17,31 ± 4,97	13,92 ± 3,58
Niveau II	P	R	V	Animaux	Fruits	Meubles
30-45 ans	20,57 ± 5,99	19,00 ± 6,44	16,07 ± 5,66	30,14 ± 11,42	19,85 ± 3,32	17,42 ± 5,77
50-65 ans	25,57 ± 6,02	25,57 ± 8,57	20,78 ± 6,05	38,71 ± 9,78	22,42 ± 3,93	19,35 ± 4,06
70-85 ans	23,78 ± 8,35	21,14 ± 7,16	17,71 ± 5,91	32,64 ± 7,90	19,35 ± 3,24	17,00 ± 4,64

Niveau d'étude I : ≤ 9 ans d'études

Niveau II : > 9 ans d'études

Annexe 2 Lettre de recrutement des participants

18/09/2015

Création du logiciel EVOLEX pour les orthophonistes

Dans le cadre de notre **mémoire de recherche**, nous participons à la création d'un logiciel de reconnaissance vocale adapté à des tâches d'évocation lexicale. L'évocation lexicale se mesure au nombre de mots émis par minute. On distingue 2 types de fluences (ou évocation lexicale) : sémantique (noms de fleurs, d'aliments, etc.) ou formelle (mots qui commencent par telle lettre).

Notre objectif est **d'optimiser et de fiabiliser la reconnaissance vocale de ce logiciel**, par l'enregistrement audio de la lecture de 2 textes courts et de tâches de fluence. Pour cela, nous avons besoin de **volontaires pour réaliser ces tâches** ; cela ne prend qu'une vingtaine de minutes !

Pour faire partie de notre échantillon, nous recherchons des hommes et des femmes, à partir de 18 ans, dont le français est la langue maternelle.

En revanche, nous ne pouvons pas inclure dans notre recherche les personnes ayant des troubles psychiatriques, ou des troubles neurologiques ayant un impact sur le langage ou la mémoire.

Nous garantissons que l'anonymat des participants est respecté et qu'aucune identification ne peut être faite une fois l'enregistrement archivé sur le serveur.

Nous vous remercions chaleureusement de votre collaboration à ce projet.

Florence Fugier-Fresne et Magali Ségui-de Lapasse

Etudiantes en 4^{ème} année d'orthophonie – Faculté de Médecine de Toulouse

Annexe 3 Protocole de passation

Protocole de passation – EVOLEX**Prérequis informatiques :**

- Le logiciel Evolex : il est compatible avec Windows, Linux, OS X... Il doit être dézippé, et fonctionne avec Java qui peut être installé gratuitement sur votre ordinateur (<https://www.java.com/fr/>).
- Un micro-casque : **Sennheiser PC3 chat Micro-Casque PC VOIP binaural Connecteur 2 x Jack 3,5 mm Noir (suggéré en date du 15/09/2015)**. Pensez à vérifier que les enregistrements se font bien à partir de ce microphone et non à partir du microphone interne de l'ordinateur.

Conditions de passation

Il est très important que les enregistrements soient réalisés **dans un environnement calme, avec le minimum de bruits** parasites (éviter climatisation, clavier, paroles autres que le test, bruits extérieurs tels que tondeuse, etc.)

Le patient est installé face à l'orthophoniste, un peu à distance (1,50 m).

Le micro doit être placé sous le flux d'air buccal (et non en face), soit **au niveau du menton** environ.

Critères d'inclusion des sujets sains pour les enregistrements visant à l'amélioration de la reconnaissance vocale :

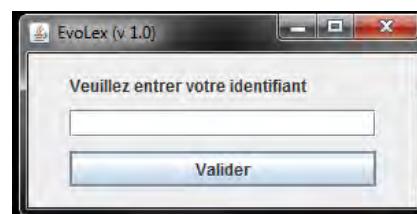
- Avoir plus de 18 ans
- La langue maternelle doit être le français
- Ne pas avoir ou avoir eu de troubles psychiatriques ou neurologiques ayant un impact sur le langage

Exécution du logiciel

Double-cliquez sur le fichier Evolex-p.jar pour ouvrir le logiciel

Identifiant

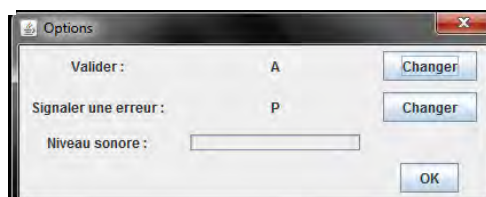
- **Saisissez un identifiant** choisi librement, mais prenez garde à le reproduire exactement à chaque ouverture d'Evolex (respect des majuscules / minuscules, signes, chiffres...) sous peine de ne pas retrouver les fichiers déjà enregistrés.
- Puis, cliquez sur **Valider**



Page d'accueil



L'onglet **Options** permet, si vous le souhaitez, de modifier les touches du clavier de validation et d'invalidation, et permet de vérifier que le microphone fonctionne et son niveau sonore. Cliquez sur **OK**, puis sur **Débuter**



Choix du patient

Vous devez saisir les données du patient ou du sujet à enregistrer :

- Cliquez sur **+** s'il s'agit d'un nouveau patient
- Si le patient existe déjà, ses nom et prénom sont consultables dans le menu déroulant à gauche ▾



S'il s'agit d'un nouveau patient ou sujet

- Saisir les nom et prénom, la date de naissance, le sexe et le niveau d'études :
 - avant le brevet = Brevet,
 - du brevet à la terminale = Bac,
 - avec le Bac au moins = bac+.
- Puis cliquez sur **Ajouter**

Choix de la tâche à accomplir



- Fluence sémantique (fruits)
- Fluence sémantique (animaux)
- Fluence formelle (R)
- Fluence formelle (V)
- Génération de mots nom-nom
- Génération de mots nom-verbe
- Génération de mots verbe-nom
- Génération de mots verbe-verbe
- Lecture de texte libre

Puis cliquez sur **Choisir**, une page **RECAPITULATIF** du patient s'affiche.

Tâches de fluence

Les tâches de fluence seront réalisées **avant les tâches de lecture** de préférence, afin de ne pas introduire de biais.

Lisez les consignes et assurez-vous qu'elles sont bien comprises.

Evocation lexicale sémantique :

Animaux :

« Vous allez devoir me dire le plus de noms français que vous connaissez dans la catégorie que je vais vous donner sans noms de la même famille et sans répétitions. Vous aurez 2 minutes pour le faire et je vous arrêterai à la fin du temps imparti. Par exemple, pour la catégorie des fleurs, vous pouvez me dire : rose, dahlia, violette. Avez-vous bien compris ? »
Lorsque vous êtes sûr que le sujet a bien compris, dites-lui : « Dites-moi le plus de noms d'animaux que vous connaissez. »

Fruits

« Vous allez devoir me dire le plus de noms français que vous connaissez dans la catégorie que je vais vous donner sans noms de la même famille et sans répétitions. Vous aurez 2 minutes pour le faire et je vous arrêterai à la fin du temps imparti. Par exemple, pour la catégorie des fleurs, vous pouvez me dire : rose, dahlia, violette. Avez-vous bien compris ? »
Lorsque vous êtes sûr que le sujet a bien compris, dites-lui : « Dites-moi le plus de noms de fruits que vous connaissez. »

Evocation lexicale formelle :

Lettre R :

« Vous allez devoir me dire le plus de noms français possible, soit des noms, soit des verbes, soit des adjectifs... commençant par la lettre que je vais vous donner et ceci en 2 minutes. Ne dites pas de noms de la même famille, de noms propres et ne vous répétez pas. Par exemple avec la lettre L, vous pouvez me dire : lune, laver, laborieuse. Avez-vous bien compris ? » Lorsque vous êtes sûr que le sujet a bien compris, dites-lui : « Dites-moi le plus de mots français possible commençant par la lettre R. »

Lettre V :

« Vous allez devoir me dire le plus de noms français possible, soit des noms, soit des verbes, soit des adjectifs... commençant par la lettre que je vais vous donner et ceci en 2 minutes. Ne dites pas de noms de la même famille, de noms propres et ne vous répétez pas. Par exemple avec la lettre L, vous pouvez me dire : lune, laver, laborieuse. Avez-vous bien compris ? » Lorsque vous êtes sûr que le sujet a bien compris, dites-lui : « Dites-moi le plus de mots français possible commençant par la lettre V ».

Puis, cliquez sur **Valider** pour démarrer l'enregistrement.

Après un décompte de 3 secondes visible sur l'écran, le chronomètre se met en route pendant 2 minutes et s'arrête automatiquement (défilement du temps au milieu de l'écran). Il est recommandé de faire le décompte au patient pour qu'il sache quand il doit démarrer.

Pause

Vous avez la possibilité de faire une pause pendant l'épreuve. Si par exemple, vous devez préciser la consigne, si le micro est mal positionné, ou autres.

Pour cela appuyez sur le bouton PAUSE, puis, pour continuer la tâche, appuyez sur REPRENDRE



Validation/invalidation des réponses

Le principe est que pour chaque mot dit, l'orthophoniste appuie sur la touche **A** si la réponse est considérée comme **valide** (couleur verte), ou sur **P** si la réponse est considérée comme une **erreur** (couleur rouge). Cela permet au logiciel de savoir que ce qui précède la validation ou l'invalidation est un mot.

Les réponses valides concernent toutes les réponses possibles correspondant à la consigne du test, sans tenir compte des mots répétés, mots de la même famille, ou autres. Ce sera le logiciel qui fait ensuite ce type d'analyses.

Par conséquent, les réponses invalides concernent tous les mots qui n'ont rien à voir avec la consigne (par exemple, banane pour la catégorie des animaux).

Ne rien faire en cas de modalisations (« je ne sais plus » « *chat*, je l'ai déjà dit »).

Commentaires



A la fin de la tâche, on peut ajouter des commentaires sur la passation (patient fatigué...)

Ecrivez EXCLUDE si vous jugez que l'enregistrement pour l'apprentissage de la reconnaissance vocale n'est pas de bonne qualité (trouble arthrique, bégaiement, trouble vocal, etc.).

Si vous considérez que l'enregistrement est valide, cliquez sur **Enregistrer**.

Changement d'épreuve

Pour enregistrer le patient sur une autre tâche, revenir au menu initial, sélectionner le nom du patient, puis sélectionner l'épreuve à faire passer. Reprendre le protocole au paragraphe « Choix de la tâche à accomplir »

Trois fichiers sont générés dans le dossier *db* ; l'enregistrement audio, le fichier renseignement et un fichier réponse (correspondant aux validations/invalidations effectuées pendant l'épreuve).

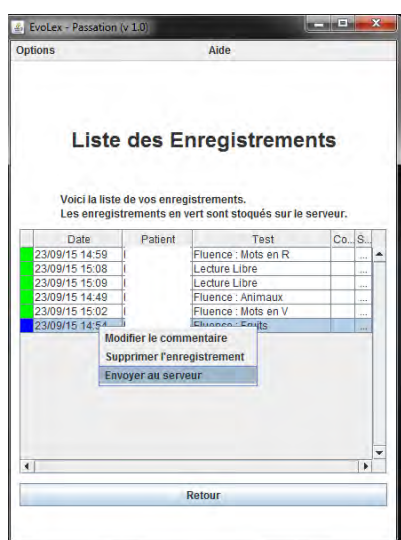
Lecture des textes.

Choisissez « Lecture libre » dans le menu déroulant, et proposez la lecture de *La Bise et le soleil* ou l'extrait du *Petit Prince*. Puis cliquer sur VALIDER

Appuyer sur **Stop** quand la lecture est terminée, puis sur **ENREGISTRER**.

Il est recommandé **d'enregistrer les 2 textes séparément** pour qu'il soit plus facile de rechercher le fichier ultérieurement, si besoin.

Envoi des fichiers sur le serveur à l'IRIT



Les enregistrements se trouvent désormais dans la **Liste des enregistrements**.

Si l'on est connecté à Internet lors de l'enregistrement, l'envoi des fichiers sur le serveur se fait immédiatement et automatiquement. Sinon, il faut se connecter à Internet, puis retourner dans la liste des enregistrements et faire un **clic droit** sur celui qu'on veut envoyer. Le carré de couleur verte signifie que le fichier est sur le serveur, en bleu qu'il n'a pas encore été expédié. N.B. les couleurs ne sont visibles que lorsqu'on est connecté sur Internet.

N.B. Sur certains ordinateurs, il est nécessaire d'être connecté à Internet avant d'ouvrir le logiciel EvoLex.

Il est aussi possible avec un clic droit de modifier le commentaire ou de supprimer l'enregistrement (tant qu'il n'a pas été envoyé sur le serveur).

Fichier Excel

En vue d'une éventuelle **normalisation des tâches de fluence**, il convient de remplir un fichier Excel annexe, reprenant des éléments plus précis sur les patients, notamment le nombre exact d'années d'études à compter à partir du CP (soit par exemple CM2 = 5 ans, 3^{ème} = 9 ans, Bac = 12 ans, etc.). Une copie de fichier Excel est jointe au protocole de passation.

Pour les **niveaux socio-culturels**, il existe 3 niveaux (normes du GREFEX) :

- 1 : études jusqu'en 3^{ème} (sans l'obtention du brevet)

Annexe 4 Textes lus « La bise et le soleil » et extrait du « Petit Prince »

La bise et le soleil

La bise et le soleil se disputaient, chacun assurant qu'il était le plus fort, quand ils ont vu un voyageur qui avançait, enveloppé dans son manteau. Ils sont tombés d'accord, que celui qui arriverait le premier à faire ôter son manteau au voyageur, serait regardé comme le plus fort. Alors, la bise s'est mise à souffler de toute sa force, mais plus elle soufflait, plus le voyageur serrait son manteau autour de lui ; et à la fin, la bise a renoncé à le lui faire ôter. Alors, le soleil a commencé à briller et au bout d'un moment, le voyageur, réchauffé, a ôté son manteau. Ainsi la bise a dû reconnaître que le soleil était le plus fort des deux.

*Tiré de http://www.er.uqam.ca/nobel/labophon/cours/LIN8206/La_bise_et_le_soleil.pdf,
Laboratoire de phonétique de l'Université du Québec à Montréal*

Le Petit Prince

C'est alors qu'apparut le renard :

- Bonjour, dit le renard.
- Bonjour, répondit poliment le petit prince, qui se retourna mais ne vit rien.
- Je suis là, dit la voix, sous le pommier...
- Qui es-tu ? dit le petit prince. Tu es bien joli...
- Je suis un renard, dit le renard.
- Viens jouer avec moi, lui proposa le petit prince. Je suis tellement triste...
- Je ne puis pas jouer avec toi, dit le renard. Je ne suis pas apprivoisé.
- Ah, pardon, fit le petit prince.

Mais après réflexion, il ajouta :

- Qu'est-ce que signifie « apprivoiser » ?
- Tu n'es pas d'ici, dit le renard. Que cherches-tu ?
- Je cherche les hommes, dit le petit prince. Qu'est-ce que signifie « apprivoiser » ?
- Les hommes, dit le renard, ils ont des fusils et ils chassent. C'est bien gênant ! Ils élèvent aussi des poules. C'est leur seul intérêt. Tu cherches aussi des poules ?
- Non, dit le petit prince. Je cherche des amis. Qu'est-ce que signifie « apprivoiser » ?

Extrait de « Le Petit Prince », d'après l'œuvre d'Antoine de Saint-Exupéry

Annexe 5 Questionnaire (enquête sur les pratiques professionnelles en matière d'informatique et de tests de fluence)

Etude des pratiques des orthophonistes en matière d'informatique et de tests de fluence verbale

Par ce questionnaire anonyme, nous cherchons à connaître vos habitudes dans l'utilisation de l'informatique et des tests de fluence verbale dans votre pratique professionnelle.
Nous vous remercions de bien vouloir répondre à ces quelques questions rapides !

Question 1

Réponse obligatoire

Utilisez-vous l'informatique (bureautique) dans votre pratique professionnelle ?

un seul choix possible

- Toujours
 Souvent
 Occasionnellement
 Jamais

Question 2

Réponse obligatoire

Utilisez-vous Internet dans votre pratique professionnelle ?

un seul choix possible

- Toujours
 Souvent
 Occasionnellement
 Jamais

Question 3

Réponse obligatoire

Utilisez-vous des logiciels dans votre pratique professionnelle ?

un seul choix possible

- Oui
 Non

Question 4

Si oui, quels types de logiciels utilisez-vous ? Vous pouvez préciser lesquels dans la zone de commentaires.

plusieurs choix possibles

- Logiciels de rééducation
 Logiciels de bilan

Commentaires

Question 5

Réponse obligatoire

Utilisez-vous des logiciels de reconnaissance vocale tels que Dragon, Siri, ou autres ?

un seul choix possible

- Oui (précisez dans la zone de commentaires)
 Non

Commentaires

Question 6

Réponse obligatoire

Quel(s) support(s) informatique(s) utilisez-vous dans votre pratique professionnelle ?

- 0 = jamais
 1 = occasionnellement
 2 = souvent
 3 = toujours

Saisissez une priorité allant de priorité 0 (minimum) à priorité 3 (maximum)

Choix	Priorité
Ordinateur fixe	<input type="text"/>
Ordinateur portable	<input type="text"/>
Tablette	<input type="text"/>
Smartphone	<input type="text"/>

Les questions suivantes concernent votre pratique pour les tests de fluence verbale (production du maximum de mots dans une catégorie ou commençant par une lettre donnée, en un temps limité).

Question 7

Réponse obligatoire

Utilisez-vous des tests de fluence ?

un seul choix possible

- Toujours
 Souvent
 Occasionnellement
 Jamais

Question 8

Réponse obligatoire

Utilisez-vous des tests de fluence...

plusieurs choix possibles

- isolés (type Cardebat ou Isaacs)
 dans des batteries complètes

Question 9

Réponse obligatoire

Utilisez-vous des tests de fluence...

plusieurs choix possibles

- sur papier
 informatisés

Question 10

Réponse obligatoire

Pour quels types de pathologies utilisez-vous les tests de fluence ?

plusieurs choix possibles

- Langage oral
 Langage écrit
 Neurologie
 Autre

Question 11

Réponse obligatoire

A quel(s) moment(s) utilisez-vous les tests de fluence ?

plusieurs choix possibles

- En bilan initial
- En bilan d'évolution ou de renouvellement

Question 12

Réponse obligatoire

Ces tests apportent des informations quantitatives étalonnées (nombre de mots produits, nombre d'erreurs, etc.). Pensez-vous que ces analyses soient :

un seul choix possible

- Satisfaisantes
- Moyennement satisfaisantes
- Insatisfaisantes
- Je ne sais pas

Question 13

Réponse obligatoire

Pensez-vous que des analyses qualitatives sur les types de mots produits (fréquence, concrétude, etc.) ou sur les stratégies utilisées par le patient constitueraient un apport pour un bilan ou un suivi de prise en charge ?

un seul choix possible

- Oui
- Peut-être
- Non
- Je ne sais pas

EVOLEX est un logiciel de reconnaissance vocale adapté aux tâches de fluences verbales. Il permettra, grâce à l'enregistrement du patient avec un micro-casque, d'obtenir automatiquement des analyses quantitatives et qualitatives de ses productions.

Les analyses quantitatives traditionnelles seront affinées par des données temporelles. Les analyses qualitatives permettront d'avoir des repères en matière de stratégies de recherche de mots (regroupements, mots fréquents, etc.) et sur les fonctions exécutives sous-tendues par ces tâches d'évocation lexicale.

Question 14

Réponse obligatoire

Pensez-vous que ce type d'analyses qualitatives soit utile ?

un seul choix possible

- Oui
- Non
- Je ne sais pas

Question 15

Réponse obligatoire

Selon vous, quels seraient les atouts de ce logiciel ?

plusieurs choix possibles

- Gain de temps
- Disponibilité pour l'observation clinique
- Intérêt des analyses qualitatives
- Simplicité d'utilisation
- Autre

Question 16

Réponse obligatoire

Selon vous, quels seraient les inconvénients de ce logiciel ?

plusieurs choix possibles

- Fluences habituellement testées dans des batteries et non isolément
- Distance patient/orthophoniste induite par l'ordinateur
- Distance patient/orthophoniste induite par le micro-casque
- Autre

Question 17

Réponse obligatoire

Auriez-vous d'autres attentes de ce type de logiciel ?

un seul choix possible

- Oui (précisez)
- Non
- Je ne sais pas

Commentaires

Question 18

Réponse obligatoire

Vous êtes...

un seul choix possible

- un homme
- une femme

Question 19

Réponse obligatoire

Vous avez...

un seul choix possible

- entre 21 et 30 ans
- entre 31 et 40 ans
- entre 41 et 50 ans
- entre 51 et 60 ans
- plus de 60 ans

Question 20

Réponse obligatoire

Vous exercez...

plusieurs choix possibles

- en libéral
- en milieu hospitalier / centre de rééducation / soins de suites
- en institution (IME, CMP, etc.)
- Autre

Merci de votre participation !

Annexe 6 Questionnaire visant à recueillir l'avis des orthophonistes après démonstration du logiciel

Questionnaire aux orthophonistes pour tests en situation écologique.				
	Tout à fait d'accord	Plutôt d'accord	Moyennement d'accord	Pas du tout d'accord
Selon vous, le logiciel...				
permet de gagner du temps par rapport aux tests de fluence "manuels"				
est simple d'utilisation				
est ergonomique				
propose un large éventail d'analyses				
propose des analyses quantitatives intéressantes et pertinentes				
propose des analyses qualitatives intéressantes et pertinentes				
temps de latence				
fréquence des mots				
temps d'initiation				
nombre de clusters et de switches				
propose un profil récapitulatif intéressant et pertinent				
constitue une possibilité intéressante de stocker les données de patient et d'avoir ainsi une base pour un suivi				
permet d'être plus disponible pour l'observation clinique				
Pensez-vous que le principe de reconnaissance vocale est intéressant ?				

Questionnaire aux orthophonistes pour tests en situation écologique.

	Tout à fait d'accord	Plutôt d'accord	Moyennement d'accord	Pas du tout d'accord							
Pensez-vous qu'un ordinateur introduit une distance entre thérapeute et patient ?											
Pensez-vous qu'un micro-casque introduit une distance entre thérapeute et patient ?											
Pensez-vous qu'il soit intéressant d'avoir un test de fluence informatisée indépendamment d'une batterie ?											
Ce logiciel pourra permettre à terme de faire passer des épreuves de dénomination et de génération de mots, pensez-vous que ce soit intéressant ?											
Selon vous, quelles seraient les améliorations à apporter au logiciel ?											
Avez-vous des commentaires, remarques ou suggestions supplémentaires ?											
Sur une échelle de 1 à 10, comment vous jugeriez ce logiciel ?	0	1	2	3	4	5	6	7	8	9	10

Données personnelles

Sexe	H	F				
Tranche d'âge	21 - 30	31 - 40	41 - 50	51 - 60	+ 60	
Lieu d'exercice	Libéral	Milieu hospitalier	Institutions	Autres		

RESUME

EvoLex est né d'une idée de disposer automatiquement, grâce à la transcription des productions du patient par reconnaissance vocale, des analyses quantitatives et qualitatives mettant en valeur les fonctions exécutives en jeu sur des tâches de fluence verbale. Cette reconnaissance vocale n'étant pas assez opérationnelle dans la première version du logiciel, nous avons enrichi la base de données avec l'enregistrement de 86 sujets sains sur 4 tâches de fluence. Ensuite, après correction des fichiers transcrits, le réapprentissage du logiciel a permis d'améliorer cette reconnaissance vocale, pour atteindre 63,4%. De plus, nous avons réalisé une enquête auprès de 181 orthophonistes qui a montré leurs avis a priori plutôt positifs sur ce type de logiciel et d'analyses, en particulier sur les analyses chiffrées montrant les stratégies exécutives. Enfin, nous avons procédé à des tests avec des patients et avons montré à 10 orthophonistes le fonctionnement d'EvoLex et les analyses proposées. Leurs avis après démonstration sont positifs quant à la pertinence et l'intérêt du logiciel, à la condition d'une reconnaissance vocale encore plus opérationnelle. EvoLex, a donc pu être enrichi, perfectionné, et même s'il y a encore des améliorations à apporter, a un bel avenir devant lui.

Mots-clés : Fluence verbale, Evocation lexicale, Test, Logiciel, Reconnaissance vocale, Regroupement, Fonctions exécutives, Pratique orthophonique

EvoLex was born from the idea of getting automatically quantitative and qualitative analyses that showcase executive functions while processing fluency tasks (using speech recognition for the patients' productions). In the first version, the speech recognition was not efficient enough, so we contributed to enlarge the database recording 4 fluency tasks in 86 healthy individuals. We corrected a number of transcribed files so that they can be used in a relearning process of speech recognition. This has been a success, increasing the speech recognition at a rate of 63.4%. Furthermore, we conducted a survey of 181 speech therapists that showed a rather positive opinion especially on statistical analyses showing executive processes used in fluency tasks. Finally, we tested the software with patients and demonstrated the operation and available analyses of EvoLex to 10 speech therapists. After the demonstration, positive views regarding relevance and benefit of the software were expressed, provided the speech recognition is effective. Thus, EvoLex, has been refined and developed, even though further improvements have still to be made.

Keywords : Verbal fluency, Test, Cluster, Switch, Speech recognition, Software, Executive functions, Speech therapy practice.