

# MEMOIRE

En vue de l'obtention du  
Certificat de Capacité d'Orthophoniste  
présenté par

**Laëtitia ZMUDA**

qui sera présenté au jury en juin 2020

**Apprentissage statistique  
et segmentation de la parole en aphasie  
Études de cas uniques**

MEMOIRE dirigé par

**Anahita BASIRAT**, Maître de conférences, Département d'orthophonie, Faculté de Médecine,  
Lille

**Yves MARTIN**, Orthophoniste et coordonnateur de l'espace recherche, innovation et  
développement, Centre Espoir, Hellemmes

Mémoire réalisé dans le cadre du Parcours Recherche

Lille – 2020

---

## Remerciements

Je remercie chaleureusement mes directeurs de mémoire, Mme. Basirat et Mr. Martin, pour la richesse des échanges que nous avons pu avoir au cours de ce projet. Votre disponibilité et votre soutien ont été majeurs dans la réalisation de cette étude.

Je remercie Paolo Mairano, expert de la synthèse vocale, pour nous avoir apporté son aide lors de la création des stimuli de l'étude.

Je remercie l'ensemble des participants à ce protocole pour avoir accordé de leur temps à la recherche. Je remercie également l'ensemble des professionnels exerçant dans la structure où nous avons recruté les participants, et plus particulièrement les orthophonistes pour leur intérêt porté à ce projet. J'en profite ici pour remercier ma maîtresse de stage Ambre Paperman, pour l'aide précieuse qu'elle m'a apportée tout au long de cette cinquième année, et notamment lors du recrutement des participants.

Je souhaitais aussi remercier plus généralement l'ensemble de mes maîtres de stages pour leur enseignement et leur soutien qui m'ont permis d'évoluer et de grandir professionnellement.

Enfin, un grand merci à mes proches, famille et amis, qui m'ont toujours soutenue durant ces cinq années d'étude. Je remercie notamment mes parents sans qui le parcours que j'ai suivi n'aurait pas été possible. Merci également à mes relecteurs et en particulier à Tamara Massoni, ma binôme de promotion qui a toujours été présente pour moi. Je tenais aussi à remercier la personne qui partage ma vie pour son soutien sans faille et pour l'aide précieuse qu'elle m'a apportée lors de la rédaction de ce mémoire.

---

## **Résumé :**

Peu d'études se sont intéressées à la segmentation de la parole en aphasie ; et encore moins à la répercussion d'un éventuel déficit de ce processus sur la compréhension orale. Ainsi, ce travail vise à apporter de nouvelles données sur les processus sous-jacents à la segmentation, via l'étude du mécanisme d'apprentissage statistique chez les individus avec aphasie. Notre hypothèse est qu'en l'absence d'un défaut d'accès lexical, ces personnes peuvent bénéficier de l'aide d'un indice lexical pour segmenter une parole artificielle. L'étude a été menée auprès de neuf individus cérébrolésés en phase subaiguë après un Accident Vasculaire Cérébral (AVC) et de langue maternelle française. La tâche administrée consistait à écouter une langue artificielle pendant environ trois minutes avant de répondre à un test vérifiant dans quelle mesure les mots artificiels de cette langue ont été appris. Chaque participant a été exposé à deux langues. L'une d'elles contenait un indice lexical (vrai mot français). Nos résultats ont montré qu'un seul participant obtenait de meilleures performances en présence de cet indice, bien qu'il présentait, comme tous les autres, une atteinte lexicale. De plus, il semblerait qu'un déficit de segmentation puisse occasionner des troubles de compréhension orale. Cependant, en interprétant ces résultats, nous avons relevé toute la complexité du processus, qu'il semble difficile à étudier indépendamment d'autres fonctions cognitives. Nous avons aussi supposé l'existence d'un mécanisme cognitif coordonnant le recrutement des différents indices de segmentation. Néanmoins, ce postulat nécessite davantage d'investigations, d'autant que plusieurs biais méthodologiques ont été mis en évidence dans ce mémoire.

## **Mots-clés :**

Segmentation de la parole – apprentissage statistique – aphasie – compréhension orale

## **Abstract :**

Few studies have focused on speech segmentation in aphasia, and even fewer on the impact of a possible deficit in this process on listening comprehension. Thus, this work aims to provide new data on the processes underlying speech segmentation by studying statistical learning in people with aphasia. We hypothesize that, in the absence of lexical access defect, a lexical cue should be helpful to segment artificial speech. The study was carried out on nine french individuals with post-stroke sub-acute phase aphasia. The task consisted in listening to an artificial language for about three minutes, and then to respond to a test to verify the extent to which words in that language were learned. Each participant was exposed to two artificial languages. One language contained a lexical cue (real French word). Our results showed that only one participant performed better in the presence of this cue, even though he had a lexical impairment like all the others. In addition, it would appear that a segmentation deficit may lead to difficulties in listening comprehension. However, in interpreting these results, we noted how complex the process was, and that it was difficult to study it apart from other cognitive functions. We also assumed the existence of a cognitive mechanism, which function would be to coordinate the recruitment of the various segmentation cues. Nevertheless, this postulate requires further investigations, since several methodological biases have been highlighted in this paper.

## **Keywords :**

Speech segmentation – statistical learning – aphasia – listening comprehension

---

# Table des matières

<b>Introduction.....</b>	<b>1</b>
<b>Contexte théorique, buts et hypothèses.....</b>	<b>2</b>
1.La segmentation du signal de la parole.....	2
1.1.Les indices lexicaux de segmentation : rôle des connaissances lexicales.....	2
1.2.Les indices pré-lexicaux de segmentation.....	3
1.3.Interactions entre les différents indices de segmentation.....	3
1.4.Etudier la segmentation de la parole.....	4
1.5.Substrats neuronaux de la segmentation.....	5
2.Segmentation de la parole et aphasie.....	5
2.1.Les troubles de la compréhension orale en aphasie.....	6
2.2.Apprentissage statistique et aphasie.....	7
3.Objectifs de l'étude.....	7
<b>Méthodologie.....</b>	<b>8</b>
1.Participants.....	8
2.Evaluations langagières.....	8
3.Matériel : stimuli du test de segmentation.....	9
3.1.Flux de langue artificielle.....	9
3.2.Stimuli du test de reconnaissance.....	10
4.Procédure.....	10
4.1. Procédure suivie auprès des participants avec aphasie.....	10
4.2.Validation du test de segmentation auprès des jeunes adultes.....	11
5.Analyse des résultats.....	11
5.1.Validation du test de segmentation.....	12
5.2.Participants avec aphasie.....	12
<b>Résultats.....</b>	<b>13</b>
1.Validation du test de segmentation.....	13
2.Résultats des participants avec aphasie .....	14
2.1.Participant 46.....	14
2.2.Participant 23.....	15
2.3.Participant 51.....	17
2.4.Participant 99.....	18
2.5.Participant 85.....	19
2.6.Participant 66.....	20
2.7.Participant 39.....	22
2.8.Participant 27.....	22
2.9.Participant 59.....	23
<b>Discussion.....</b>	<b>24</b>
1.Validation du test de segmentation.....	24
2.Apprentissage statistique et segmentation en aphasie.....	25
2.1.Connaissances lexicales et segmentation d'une parole artificielle.....	25
2.2.Apprentissage statistique, segmentation et compréhension orale.....	27
3.Segmentation et orthophonie.....	30
<b>Conclusion.....</b>	<b>31</b>
<b>Bibliographie.....</b>	<b>32</b>
<b>Liste des annexes.....</b>	<b>35</b>
Annexe n°1 : Stimuli du test de segmentation.....	35
Annexe n°2 : Résumé des analyses statistiques effectuées dans le cadre de la validation du test de segmentation.....	35
Annexe n°3 : Résultats des participants à l'étude soumis à une seule condition expérimentale.....	35

# Introduction

Le signal de la parole est un flux continu de sons au sein duquel les mots ne sont pas clairement délimités par un silence, comme ils peuvent l'être par un espace à l'écrit. C'est en partie ce qui rend difficile la compréhension orale des langues étrangères. En effet, pour accéder à une compréhension satisfaisante, il est essentiel de pouvoir repérer les frontières entre les mots afin d'identifier les principales entités significatives de l'énoncé. Le processus permettant le découpage de la parole en mots est appelé « processus de segmentation lexicale ». Il constitue une étape indispensable à la compréhension de la parole.

De nombreuses études ont porté sur la segmentation dans le but spécifique de déterminer les indices sur lesquels nous nous appuyons pour segmenter la parole et identifier les mots de l'énoncé. Ainsi, il a été démontré que la segmentation reposait sur toute une série d'indices lexicaux et non-lexicaux (Bagou & Frauenfelder, 2018). Les indices lexicaux correspondent à nos connaissances lexicales. Lorsqu'ils sont disponibles, ils participent à la segmentation dans le cadre d'un mécanisme appelé « reconnaissance de mots ». Cependant, dès que l'accès à ces indices est insuffisant (lors de l'acquisition du langage par exemple), les indices non-lexicaux prennent de l'importance : ils participent à la segmentation dans le cadre d'un autre mécanisme appelé « apprentissage de mots ». C'est ce mécanisme en particulier que nous étudions dans ce mémoire.

Le processus de segmentation est très peu étudié dans le cadre de l'aphasie, bien que la présence de troubles de la compréhension orale chez les personnes avec aphasie soit fréquemment rapportée dans la littérature. Divers paradigmes peuvent être utilisés pour l'étudier. Dans ce mémoire, nous nous concentrons sur le paradigme d'apprentissage statistique. Celui-ci repose sur un mécanisme consistant à extraire des mots inconnus d'une langue artificielle en se basant sur les probabilités transitionnelles entre les syllabes, alors considérées comme l'un des indices non-lexicaux de segmentation (Saffran, Newport, & Aslin, 1996). Ce type d'apprentissage serait facilité chez l'enfant (Mersad & Nazzi, 2012) et chez l'adulte (Palmer, Hutson, White, & Mattys, 2019), lorsqu'un mot connu est présent dans le flux de parole de mots inconnus.

Peu d'études dans la littérature se sont intéressées au mécanisme d'apprentissage statistique dans le domaine de l'aphasie (Peñaloza et al., 2014), ou plus généralement après un AVC (Shaqiri, Danckert, Burnett, & Anderson, 2018). Les résultats de ces études suggèrent que l'apprentissage statistique pourrait être atteint chez certains patients. Plusieurs questions restent cependant sans réponses : (1) cet apprentissage est-il facilité chez les personnes avec aphasie lorsqu'un mot connu est présent dans le flux de mots inconnus ? (2) un déficit de l'apprentissage statistique peut-il perturber la compréhension orale d'énoncés chez ces personnes ?

L'enjeu de ce travail est d'affiner la compréhension des mécanismes mis en œuvre dans la segmentation en répondant à ces deux questions. Nous cherchons aussi à mieux comprendre le lien entre segmentation et troubles de la compréhension orale chez les personnes avec aphasie.

Plus précisément, nos objectifs sont les suivants : (1) évaluer les capacités de segmentation des personnes avec aphasie et l'effet de leurs connaissances lexicales sur leur performance, en utilisant le paradigme d'apprentissage statistique ; (2) mettre en lien les résultats obtenus avec leur profil langagier dans le but d'apporter de nouvelles données sur les mécanismes liés à la segmentation ; (3) à long terme, mener une réflexion quant à l'élaboration d'outils orthophoniques testant les capacités de segmentation de ces patients, ce type d'outils étant inexistant à ce jour.

# Contexte théorique, buts et hypothèses

## 1. La segmentation du signal de la parole

La parole est un signal acoustique composé d'une succession rarement interrompue de sons (Wauquier-Gravelines, 1999). Le caractère continu de ce flux sonore, associé au phénomène de la coarticulation, aux différences de prononciation interindividuelles et à des conditions d'écoute parfois dégradées, nous amène à poser le problème de la décomposition de la chaîne parlée en unités signifiantes, permettant à n'importe quel auditeur de comprendre un message oral.

Ce problème de segmentation est abordé dans de nombreux modèles traitant de l'accès au lexique. Bagou et Frauenfelder (2002) décrivent par exemple un modèle simplifié de la reconnaissance des mots, dans lequel le processus de segmentation intervient à plusieurs niveaux. Dans l'ensemble, il est défini comme un processus visant à déterminer le début et la fin de chaque unité de parole, au sein d'un signal dépourvu de pauses marquées acoustiquement de façon claire et systématique. Au niveau infra-lexical, il aurait pour objectif d'associer chaque partie du signal acoustique à une unité linguistique stockée en mémoire (phonème, syllabe ou mot), pour aboutir à une représentation infra-lexicale du signal. Au niveau lexical, cette représentation serait alignée et appariée aux unités stockées dans le lexique mental, permettant ainsi l'identification des mots de l'énoncé. Au niveau supra-lexical, la segmentation contribuerait à identifier des unités linguistiques plus larges, telles que les expressions idiomatiques (locutions figées dont le sens ne peut être déduit littéralement à partir du sens des mots qui les composent). Dans ce travail, nous entendons par « segmentation », l'opération permettant de localiser les frontières entre les mots.

Le processus de segmentation, indispensable à la reconnaissance des mots, joue aussi un rôle majeur dans l'acquisition du lexique, que ce soit chez les enfants ou chez les nouveaux apprenants d'une langue étrangère (Nazzi, 2008). Les mots n'étant que rarement perçus de façon isolée, il semble effectivement évident que la construction du lexique s'appuie sur un tel processus. Dans cette situation, il permet d'extraire les formes sonores des mots de la parole continue afin de les représenter mentalement et avant de pouvoir les apparier aux concepts.

Au final, la segmentation de la parole est un processus complexe dont de nombreux aspects sont encore étudiés aujourd'hui dans la littérature. La majorité des études portant sur ce processus visent notamment à mieux comprendre son fonctionnement en s'intéressant aux différents indices sur lesquels il repose. En effet, comme le rappellent Bagou et Frauenfelder (2018), la segmentation de la parole s'appuie sur toute une série d'indices lexicaux et non lexicaux.

### 1.1. Les indices lexicaux de segmentation : rôle des connaissances lexicales

La segmentation lexicale peut être considérée comme une conséquence de l'accès au lexique : l'identification d'éléments lexicaux individuels contribuerait à la localisation des frontières entre les mots (Bagou & Frauenfelder, 2002 ; Crouzet, 2000 ; Davis, Marslen-Wilson, & Gaskell, 2002). Cela suppose l'existence d'un lexique mental, composé d'unités lexicales connues aux informations multidimensionnelles (phonologiques, sémantiques, syntaxiques, orthographiques).

Les modèles d'accès au lexique décrits dans la littérature s'inscrivent dans deux types d'approche (Bagou & Frauenfelder, 2002 ; Davis et al., 2002). Dans l'approche séquentielle, issue du modèle Cohort (Marslen-Wilson & Welsh, 1978), de multiples candidats lexicaux sont activés

dès la perception de la partie initiale du mot. Ils sont progressivement éliminés par l'absence de correspondance avec l'entrée sensorielle, jusqu'à ce qu'il ne reste plus qu'un seul élément lexical. Ainsi, lorsqu'un mot est reconnu avant son articulation complète, les connaissances lexicales peuvent être utilisées pour prédire la frontière finale de ce mot et inférer le début du mot suivant. Cependant, il est parfois impossible d'identifier le mot avant la fin de son articulation. C'est pourquoi une approche alternative, issue du modèle TRACE (McClelland & Elman, 1986), a été proposée. Cette dernière considère la segmentation comme étant le résultat d'une compétition entre éléments lexicaux pouvant traverser les frontières entre les mots. L'activation des unités lexicales en compétition est modulée par des processus d'activation et d'inhibition. L'unité recevant le plus d'activation est ainsi identifiée comme celle correspondant au signal d'entrée. Les modèles Cohort révisé (Marslen-Wilson, 1987) et Shortlist (Norris, 1994) fonctionnent également sur ce principe.

## **1.2. Les indices pré-lexicaux de segmentation**

Dans le cas de l'acquisition du lexique, les connaissances lexicales sont indisponibles et seulement construites progressivement. Il est donc évident que d'autres indices de segmentation interviennent : ils sont qualifiés d'indices pré-lexicaux. Bagou et Frauenfelder (2002, 2018) en identifient plusieurs qu'ils classent dans deux catégories distinctes.

Les indices acoustico-phonétiques fourniraient des indications quant à la position des phonèmes dans les mots, signalant ainsi le début ou la fin d'un mot. Ils peuvent se baser sur les variations allophoniques liées à la coarticulation ou sur les variations de durée des phonèmes.

Les indices phonologiques reposeraient quant à eux sur les variations segmentales et suprasegmentales de la langue. L'information statistique peut par exemple être utilisée pour isoler des mots sur la base des probabilités transitionnelles entre les syllabes (fréquences de co-occurrence des syllabes adjacentes). Il est également possible de se référer aux régularités phonotactiques (règles régissant l'agencement des phonèmes à l'intérieur des mots ou des syllabes) ainsi qu'aux informations prosodiques qui reposent notamment sur l'alternance de syllabes accentuées et atones.

Pris individuellement, les indices de segmentation ne rendent pas compte du processus dans sa globalité : aucun n'est disponible à chacune des frontières entre les mots et plusieurs sont généralement présents à une même frontière. Etudier séparément leur influence sur la segmentation peut alors limiter la compréhension que nous avons du processus. Il devient plus pertinent d'analyser leur contribution relative, et ce dans différentes langues puisque certains indices sont spécifiques à la langue étudiée (Bagou & Frauenfelder, 2018).

## **1.3. Interactions entre les différents indices de segmentation**

Mattys, White, et Melhorn (2005) ont ainsi établi une organisation hiérarchique de ces indices chez l'adulte en étudiant leur primauté relative. La prédominance des indices lexicaux est clairement mise en évidence lorsque les conditions d'écoute sont optimales. Lorsque ces dernières se dégradent (ex. dans le bruit), la quantité d'information lexicale disponible diminue, rendant les indices acoustico-phonétiques et phonotactiques plus prégnants. En revanche, lorsque les conditions d'écoute sont les plus dégradées, ce sont les indices prosodiques qui priment sur les autres. Les auteurs ont aussi montré que le recrutement des indices ne fonctionnait pas sur le principe du tout ou rien : il se fait plutôt de façon graduelle, selon la position des indices dans la hiérarchie.

Bagou et Frauenfelder (2018) ont décrit plus précisément la nature des interactions entre les différents indices en étudiant la façon dont les diverses combinaisons possibles d'indices influencent l'apprentissage des mots d'une langue artificielle. Ces combinaisons ont été construites selon quatre dimensions : convergence ou divergence des indices, redondance ou complémentarité, nombre d'indices présents, position de ces indices dans le mot (initiale/finale). Les résultats obtenus suggèrent que les indices complémentaires en position initiale et finale de mots (ex. en français : les indices phonotactiques en position initiale et prosodiques en position finale) sont intégrés de façon synergique : l'effet obtenu est supérieur à la somme des effets de chacun des indices pris séparément. À l'inverse, les indices prosodiques redondants (indices de durée et de hauteur situés sur la fin des mots) ne produisent pas d'effets synergiques : ils sont organisés hiérarchiquement, par ordre d'importance. Les auteurs émettent alors l'hypothèse suivante : les indices finaux ont un pouvoir prédictif de la segmentation tandis que les indices initiaux permettent de confirmer les décisions de segmentation de façon complémentaire et synergique.

En résumé, les différents indices convergent vers la même segmentation des mots. La façon dont ils interagissent varie selon leur nature et leur position, elles-mêmes dépendantes de la langue étudiée. Ainsi, les effets de certains indices priment sur d'autres ou peuvent se combiner.

#### **1.4. Etudier la segmentation de la parole**

Différents paradigmes peuvent être utilisés pour étudier la segmentation. Ils se distinguent selon le type de mécanisme impliqué : apprentissage versus reconnaissance de mots. Le poids des indices de segmentation est différent dans ces deux types de tâches (Bagou et Frauenfelder, 2018).

Dans le cadre de l'apprentissage de mots, les chercheurs utilisent généralement le paradigme d'apprentissage statistique (ex. Bagou & Frauenfelder, 2018 ; Palmer et al., 2019 ; Peñaloza et al., 2014 ; Saffran et al., 1996). Celui-ci consiste en deux phases : (1) une phase d'exposition à une langue artificielle et (2) une phase de test évaluant dans quelle mesure les participants sont parvenus à extraire des mots de cette langue artificielle. Dans cette deuxième phase, le test peut être une tâche de décision lexicale auditive (indiquer si oui ou non le mot présenté auditivement est un mot de la langue artificielle) ou une tâche en choix forcé à deux alternatives (sélectionner parmi deux mots présentés auditivement celui qui appartient à la langue artificielle). La tâche de décision lexicale est moins exigeante en terme de mémorisation, mais il est possible d'alléger le coût de la mémoire dans la tâche en choix forcé en présentant les paires de stimuli visuellement et auditivement (Palmer, Hutson, & Mattys, 2018). L'utilisation d'un tel paradigme permet d'étudier l'apprentissage des mots chez les adultes puisque les indices lexicaux ne sont pas disponibles (Bagou et Frauenfelder, 2018). Il est toutefois possible d'analyser l'effet des connaissances lexicales sur la segmentation avec ce type de paradigme en incorporant des vrais mots dans la langue artificielle (Palmer et al., 2019).

Dans le cadre de la reconnaissance des mots, les chercheurs peuvent par exemple utiliser une tâche de « word monitoring » (ex. Mattys et al., 2005) : un stimulus est présenté visuellement ou auditivement au participant, qui entend ensuite une phrase. Il doit indiquer si oui ou non le mot était présent dans la phrase. Il s'agit d'une tâche qui a été utilisée dans deux mémoires s'intéressant également à la segmentation de la parole en aphasie (Dupont, 2018 ; Inghilleri, 2018). Les auteurs avaient surtout mis en évidence le fait qu'il était difficile d'étudier la segmentation en la distinguant d'autres fonctions cognitives, telles que les capacités langagières ou mnésiques. Par ailleurs, il est aussi possible d'utiliser le paradigme d'amorçage (ex. Davis et al., 2002 ; Mattys et al., 2005) qui évalue l'effet de la présentation préalable d'un stimulus (l'amorce) sur le traitement d'un autre stimulus (la cible). Il y a effet d'amorçage lorsque la cible est traitée plus rapidement en présence de

l'amorce. Dans le cadre de la segmentation, on évalue généralement l'effet d'amorçage d'une séquence portant différents types d'indices (favorables ou non à la segmentation) sur les stimuli d'une tâche de décision lexicale.

### **1.5. Substrats neuronaux de la segmentation**

À ce jour, les chercheurs qui se sont intéressés aux substrats neuronaux de la segmentation ont utilisé la tâche d'apprentissage statistique et l'imagerie par résonance magnétique fonctionnelle (IRMf) chez des adultes sains. McNealy, Mazziotta, et Dapretto (2006) ont relevé l'importance du gyrus frontal inférieur gauche, tout comme Karuza et al. (2013), ainsi que de régions temporo-pariétales. Ces résultats ont été corroborés par Cunillera et al. (2009) qui ont mis en évidence l'implication du gyrus temporal supérieur postérieur et du cortex prémoteur, soit d'une interface auditivo-motrice. López-Barroso et al. (2015) ont quant à eux apporté des précisions sur les différents réseaux pouvant être impliqués. Ils en ont identifié quatre, spatialement indépendants et d'implication variable au cours de la période d'apprentissage : les réseaux dorsaux auditivo-prémoteur, sensitivo-moteur et fronto-pariétal, et le réseau ventral fronto-temporal. Ces réseaux auraient des rôles différents et sous-tendraient également d'autres processus tels que la mémoire de travail, la production de la parole ou le développement du lexique. Le réseau dorsal auditivo-prémoteur serait le plus fortement impliqué puisqu'il est le seul à avoir été continuellement activé.

Ces résultats suggèrent donc l'implication d'un réseau engageant les régions fronto-temporales de la voie dorsale. Le rôle de cette dernière est défini dans le modèle à double voie d'Hickok et Poeppel (2007) qui postule l'existence de deux voies de traitement à la perception auditive de la parole et du langage. La voie ventrale implique des régions temporales moyennes et inférieures et participe à la compréhension en appariant les informations auditives et phonologiques aux représentations conceptuelles lexico-sémantiques. La voie dorsale mobilise plutôt les régions pré-motrices et frontales inférieures gauches ainsi que la jonction temporo-pariétale qui sert d'interface sensorimotrice. En reliant les représentations auditives et articulatoires, cette voie participe notamment à l'acquisition du langage. Elle intervient aussi dans les tâches qui nécessitent un maintien actif des représentations auditives, telles que les tâches de mémoire de travail verbale.

Au cours de ce chapitre, nous avons montré que le processus de segmentation participait non seulement à la reconnaissance des mots mais aussi à l'acquisition du langage ; et qu'il se basait pour cela sur toute une série d'indices. Son fonctionnement est certes complexe, mais il semble évident qu'il est lié aux mécanismes qui sous-tendent la parole et le langage. Dans ce cadre, l'intérêt porté à ce processus dans le domaine de l'aphasie, un trouble acquis du langage, semble tout à fait justifié.

## **2. Segmentation de la parole et aphasie**

Dans le Dictionnaire d'Orthophonie, le terme d'aphasie est explicité selon la définition donnée par Trousseau en 1864 : l'aphasie est une « perturbation du code linguistique, affectant l'encodage (versant expression) et/ou le décodage (versant compréhension), et qui peut concerner le langage oral et/ou écrit. Ce trouble n'est ni lié à un état démentiel, ni à une atteinte sensorielle, pas plus qu'à un dysfonctionnement périphérique de la musculature pharyngolaryngée, mais à une atteinte cérébrale localisée ou diffuse, généralement dans la zone frontale, pariétale et/ou temporale de l'hémisphère gauche, d'origine essentiellement vasculaire, traumatique ou tumorale.» (« Aphasie », 2011, p. 20). Il s'agit donc d'un trouble acquis du langage, survenant chez un locuteur ayant été jusque-là en pleine possession de ses capacités langagières.

Les manifestations linguistiques de ce trouble et ses répercussions sur la communication sont multiples, de sévérité variable, et dépendent fortement de la nature et de la localisation de la lésion. Pour répondre à cette diversité, différents tableaux cliniques de l'aphasie ont été décrits et de nombreuses classifications ont été proposées. La plus utilisée est celle établie par l'équipe de Boston (Goodglass & Kaplan, 1972, cité dans De Partz & Pillon, 2014) qui distingue les aphasies de Broca, de Wernicke, de conduction, amnésique, transcorticale motrice, transcorticale sensorielle, et globale. La classification se base sur des caractéristiques de la production verbale spontanée (fluence vs. non-fluence), de la compréhension auditive et de la répétition. Dans la littérature, les aphasies de Broca (forme non fluente) et de Wernicke (forme fluente) restent les plus étudiées.

La segmentation étant un processus qui participe davantage à la réception de la parole qu'à son expression, nous nous concentrons plus spécifiquement dans la partie suivante aux troubles de la compréhension orale que peuvent présenter les personnes avec aphasie.

## **2.1. Les troubles de la compréhension orale en aphasie**

L'atteinte de la compréhension orale est fréquente dans le domaine de l'aphasie. Elle peut toucher différents niveaux linguistiques de façon plus ou moins sévère, d'où la diversité des troubles rencontrés. Pour comprendre la parole, il faut effectivement pouvoir identifier les sons issus de son signal (niveau phonologique), puis comprendre les mots qu'ils forment (niveau lexical), les énoncés formés par ces mots (niveau syntaxique), et enfin le discours (niveau discursif).

La compréhension des mots dépend principalement de leur reconnaissance dont nous avons décrit les mécanismes ci-dessus. Ainsi, les difficultés de compréhension lexicale présentées par certaines personnes avec aphasie sont à relier à d'éventuelles difficultés d'accès au lexique mental.

Toutefois, il est établi que les personnes aphasiques peuvent présenter des difficultés de compréhension d'origine syntaxique. Caramazza et Zurif (1976, cité dans Patil, Hanne, Burchert, De Bleser, & Vasisht, 2015, p. 6) ont par exemple montré que certains patients présentaient des troubles de compréhension dès lors que les phrases devenaient complexes (ex. phrases non canoniques réversibles), autrement dit dès lors qu'il était nécessaire de faire appel à la structure syntaxique des phrases pour en comprendre le sens. Les mécanismes qui sous-tendent la compréhension syntaxique sont complexes, comme en témoigne le nombre important de modèles décrits dans la littérature (pour une revue, voir Caplan, Michaud, & Hufford, 2015). Ces derniers sont traditionnellement classés selon deux groupes (Patil et al., 2015). Certains modèles attribuent les difficultés de compréhension à un trouble spécifique des représentations syntaxiques : les connaissances concernant la structure grammaticale des énoncés seraient elles-mêmes touchées. D'autres modèles considèrent que les représentations structurales sont opérationnelles mais pas les processus permettant l'analyse syntaxique : les difficultés de compréhension seraient alors la conséquence d'un ralentissement de la vitesse de traitement, d'une réduction des ressources en mémoire de travail, ou encore d'un déficit d'intégration lexicale. Patil et al. (2015) classent aussi le déficit d'accès au lexique dans ce deuxième groupe de modèles. Plus précisément, Ferril, Love, Walenski, et Shapiro (2012) ont montré que des personnes avec une aphasie de type Broca présentaient un défaut lors de l'accès au lexique nommé « retard d'activation lexicale » : le traitement syntaxique deviendrait alors trop lent pour permettre une bonne compréhension.

De nombreux mécanismes sont donc impliqués dans la compréhension orale et c'est ce qui rend difficile l'interprétation des troubles de compréhension chez les personnes atteintes d'aphasie.

## 2.2. Apprentissage statistique et aphasie

Peu d'études se sont intéressées à la segmentation de la parole chez les personnes avec aphasie dans le cadre de l'apprentissage statistique. Peñaloza et al. (2014) ont par exemple montré que des individus en phase chronique (plus de 6 mois après un AVC) étaient capables de segmenter les mots d'une parole artificielle, quelle que soit la sévérité de leur aphasie. Toutefois, ils ont aussi montré que la majorité présentaient des difficultés et que cela semblait être lié à la localisation des lésions : les patients avec une lésion antérieure (région frontale inférieure gauche) étaient effectivement plus en difficulté que ceux avec une lésion postérieure. Certains présentant une lésion dans cette région semblaient néanmoins capables de segmenter la parole artificielle, évoquant de ce fait une possible réorganisation neuronale du cerveau suite à l'AVC. Shaqiri et al. (2018) ont aussi étudié l'apprentissage statistique chez des patients cérébrolésés suite à un AVC. Ils ont également mis en évidence une atteinte du processus, sans que cela ne puisse être expliqué par l'aphasie ou par d'éventuels déficits cognitifs. Par ailleurs, ils n'ont pas trouvé de corrélations entre le déficit d'apprentissage statistique, l'étendue des lésions et leur localisation.

Les résultats de Peñaloza et al. (2014) ont aussi mis en évidence un lien entre la capacité en mémoire de travail auditivo-verbale des participants et leur capacité à segmenter la parole ; deux processus sous-tendus par les mêmes régions cérébrales. Pour rappel, les mémoires à court terme et de travail permettent de stocker temporairement de l'information, le temps de la manipuler dans le cadre d'opérations mentales diverses. Leur atteinte chez les personnes avec aphasie est souvent rapportée dans la littérature, tout comme l'influence qu'elle peut avoir sur leurs performances linguistiques (Christensen, Wright, & Ratiu, 2018 ; Ivanova, Dragoy, Kuptsova, Ulicheva, & Laurinavichyute, 2015 ; Salis, Murray, & Bakas, 2018). La mémoire de travail jouerait notamment un rôle dans la compréhension (Ivanova et al., 2015) : une corrélation entre troubles de la mémoire de travail et de la compréhension a été observée chez des individus avec une aphasie de type Broca. Par conséquent, il semble important de prendre en compte ces deux systèmes lorsqu'on s'intéresse à une population d'individus atteints d'aphasie.

## 3. Objectifs de l'étude

Actuellement, le processus de segmentation est très peu étudié dans le domaine de l'aphasie. Bien que certaines personnes avec aphasie en phase chronique semblent pouvoir segmenter et apprendre une langue artificielle sur la base d'indices statistiques, la majorité présentent des difficultés (Peñaloza et al., 2014). À notre connaissance, aucune étude ne s'est intéressée à l'influence des connaissances lexicales de ces personnes sur leurs capacités de segmentation ; ni sur les répercussions du déficit d'apprentissage statistique sur leurs capacités à comprendre des énoncés.

Ainsi, ce travail vise à apporter de nouvelles données sur les mécanismes mis en œuvre dans la segmentation lexicale de la parole, en passant par l'étude du mécanisme d'apprentissage statistique de mots chez les personnes présentant une aphasie en phase subaiguë post-AVC (un à six mois après l'AVC). Nous nous intéressons aux capacités de segmentation et d'apprentissage statistique de ces personnes, ainsi qu'à l'influence de leurs connaissances lexicales sur ces capacités.

Un autre enjeu de ce mémoire est d'affiner la compréhension du lien entre segmentation et troubles de la compréhension orale en aphasie. Aussi, nous nous intéressons aux capacités de segmentation des personnes avec aphasie en regard de leur profil langagier.

Sur la base des résultats obtenus par Peñaloza et al. (2014) et Shaqiri et al. (2018), nous nous attendons à observer des difficultés chez certains participants pour segmenter une langue artificielle. Par ailleurs, nous supposons qu'en l'absence d'un défaut d'accès lexical, les personnes avec aphasie pourraient bénéficier de l'aide d'un indice lexical pour segmenter une parole artificielle, tout comme les individus tout-venants de l'étude de Palmer et al. (2019). Enfin, nous pensons mettre en évidence des corrélations entre de potentiels déficits de segmentation et des troubles de compréhension orale.

## **Méthodologie**

### **1. Participants**

L'étude a inclus neuf individus de langue maternelle française présentant une aphasie post-AVC en phase subaiguë. Les critères d'inclusion étaient les suivants : (1) présenter une lésion cérébrale acquise suite à un AVC conduisant à une aphasie, (2) avoir un niveau de compréhension permettant au minimum de comprendre les consignes données lors de la passation de l'expérience, (3) présenter des capacités motrices satisfaisantes sur au moins un membre supérieur afin de pouvoir appuyer sur les touches du clavier de l'ordinateur. Les participants ont tous été recrutés au sein d'un centre de rééducation et de réadaptation fonctionnelles. Leur AVC datait de 3.22 mois en moyenne (ET = 1.79). Aucun des participants n'a rapporté de troubles auditifs non corrigés. Les informations démographiques et cliniques les concernant ont été précisées dans la partie « Résultats ».

Un groupe composé de 23 jeunes adultes a également été inclus afin de valider le test de segmentation créé pour l'étude. L'objectif était de vérifier la possibilité de segmenter les deux langues artificielles créées et de tester leur équivalence en terme de difficulté. Les jeunes adultes étaient uniquement des femmes de langue maternelle française. La moyenne d'âge était de 23.74 ans (ET = 2,88 ; étendue = 19-34). Les participantes avaient toutes une vue correcte ou corrigée et aucun déficit auditif connu. Aucune n'avait eu d'expérience antérieure avec une langue artificielle.

Le protocole suivi a bien obtenu l'avis favorable du Comité d'éthique en sciences comportementales de Lille. Avant de participer à l'étude, les participants ont été informés de son déroulement via une lettre d'information et nous ont donné leur consentement par écrit.

### **2. Evaluations langagières**

Une évaluation des capacités de compréhension et d'expression orales des participants avec aphasie a été menée la même semaine que la passation du test de segmentation. L'objectif était de déterminer leur profil langagier afin de le mettre en lien avec leurs capacités de segmentation.

La sévérité de l'aphasie de chaque participant a été déterminée à partir des propos spontanés échangés avec l'examineur au cours de l'ensemble des passations ; selon l'échelle de gravité de l'aphasie du BDAE (Echelle d'Évaluation de l'Aphasie, Goodglass & Kaplan, 1972).

Plusieurs épreuves de la BECLA (Batterie d'Évaluation Cognitive du LAngage, Macoir, Gauthier, Jean, & Potvin, 2016) ont permis d'évaluer le système lexical des personnes avec aphasie. Elles ont été administrées dans l'ordre suivant : (1) tâche de décision lexicale orale pour évaluer la reconnaissance de mots entendus et le lexique phonologique d'entrée, (2) tâche de dénomination d'images pour observer l'accès lexical, (3) tâche d'appariement sémantique d'images pour vérifier l'intégrité du système sémantique, (4) tâche d'appariement mot entendu-image pour tester la

compréhension orale lexicale, (5) tâche de jugement de rimes sur images pour observer l'accès au lexique phonologique de sortie, (6) tâches de répétition, différée et non différée, de mots et de non-mots, afin d'évaluer les voies lexicales et non lexicales de production orale ainsi que le fonctionnement de la mémoire tampon phonologique.

Le MT86 (Protocole Montréal-Toulouse d'examen linguistique de l'aphasie, Nespoulous et al., 1992) a quant à lui été utilisé pour évaluer la compréhension orale syntaxique, avec des tâches de manipulation d'objets sur consignes données à l'oral et d'appariement énoncé entendu-image.

Enfin, les capacités de discrimination phonologique des participants ont été testées avec l'épreuve informatisée de discrimination de consonnes créée par Van der Kaa et Majerus (2007). Dans cette épreuve, le participant devait déterminer si les deux pseudo-mots monosyllabiques qui lui étaient présentés auditivement étaient identiques ou non. Lorsque ce n'était pas le cas, ils formaient des paires minimales. Les stimuli étaient présentés à vitesse normale puis accélérée.

### **3. Matériel : stimuli du test de segmentation**

Les stimuli de notre étude ont été extraits de l'étude de Buiatti, Peña, et Dehaene-Lambertz (2009), ces derniers ayant été créés pour des participants français. Ils ont été adaptés à notre étude sur la base des méthodologies utilisées par Palmer et al. (2019) et par Peñaloza et al. (2014).

#### **3.1. Flux de langue artificielle**

Le test de segmentation a été administré dans deux conditions : une condition de base et une condition mot connu. Deux flux de langue artificielle ont été construits dans chaque condition.

Dans la condition de base, chaque flux était le résultat de la concaténation de six séquences trisyllabiques de structure consonne-voyelle (CV)-CV-CV, respectant les règles phonotactiques du français et ne contenant pas de vrais mots français de plus d'une syllabe (flux A : *milora*, *minoura*, *folobé*, *fonoubé*, *poulouki*, *pounouki* ; flux B : *bagavo*, *bapivo*, *togadou*, *topidou*, *ligafé*, *lipifé*). Ces séquences (nommées par la suite « mots artificiels » ou « MA ») ont été créées de sorte qu'elles puissent être réparties en trois « familles » : les MA d'une même famille possédaient les mêmes première et troisième syllabes tandis que la deuxième syllabe était différente (structure AxC). Par exemple, la famille « *mi\_ra* » était composée des MA « *milora* » et « *minoura* ». Chaque flux était ainsi composé de 45 occurrences de chacun des 6 MA, soit 270 mots au total concaténés dans un ordre semi-aléatoire de façon à ce que deux mots identiques ou appartenant à une même famille ne se suivent jamais. Le même ordre de présentation a été utilisé pour les deux flux.

Après que les flux ont été générés sous format texte, ils ont été synthétisés avec MBROLA (22050Hz, mono, 16 bits), un logiciel de synthèse vocale basé sur la concaténation de diphtonges, produits ici par une voix de femme française. Tous les phonèmes avaient la même durée (116 ms) et étaient de même fréquence (200 Hz). La plage d'intensité était similaire pour chaque syllabe.

Au total, chaque flux formait une séquence de parole continue monotone et dépourvue de sens, d'une durée de trois minutes environ. Comme dans l'étude de Peñaloza et al. (2014), les syllabes ne portaient pas d'indices acoustico-phonétiques ni prosodiques puisqu'elles avaient été concaténées et étaient toutes de mêmes durée et hauteur. Aucune pause ni aucun indice audible ne signalait la présence d'une frontière entre les mots. Ainsi, seuls les indices relevant des informations statistiques étaient disponibles : les probabilités transitionnelles étaient égales à 1 entre les syllabes formant les MA et variaient entre 0.1 et 0.4 entre les MA. Les probabilités transitionnelles au sein de chaque flux étaient identiques ; l'ordre de présentation des MA ayant été apparié.

Les flux de la condition mot connu (flux A+ et B+) correspondaient aux flux A et B de la condition de base, auxquels ont été ajoutés 45 occurrences d'un vrai mot français tétrasyllabique fréquent de structure CV-CV-CV-CV : le mot *médicament*. Ce mot a été inséré aléatoirement dans les deux flux. L'ordre de présentation des MA et du vrai mot était identique dans les flux A+ et B+. La synthèse vocale a été effectuée avec MBROLA de sorte que les flux des deux conditions aient les mêmes caractéristiques acoustiques. Cela signifie que le mot *médicament* a été synthétisé comme un MA, c'est-à-dire en synthétisant puis concaténant chacune des syllabes le composant.

Sur la base des résultats obtenus par Palmer et al. (2019), nous avons fait le choix d'ajouter le vrai mot à la langue artificielle plutôt que de venir le substituer à l'un des MA afin que le nombre de MA à apprendre dans chaque condition soit identique. Par ailleurs, nous avons préféré utiliser un mot tétrasyllabique et non trisyllabique pour ne pas faciliter la segmentation des MA en raison d'une possible attente métrique induite par le mot connu (Palmer et al., 2019).

### **3.2. Stimuli du test de reconnaissance**

Le test suivant la phase d'exposition à l'un des flux de langue artificielle se présentait comme une tâche de décision lexicale comprenant trois types d'items (cf. Tableau 1, Annexe A1) : des mots artificiels (MA), des mots-partiels (MP) et des mots-règles (MR). Les MA étaient les six mots du flux auquel le participant avait été exposé. Les MP correspondaient à six séquences trisyllabiques construites par concaténation des deux dernières syllabes d'un MA et de la première syllabe d'un autre MA adjacent à celui-ci. Les MP étaient donc présents dans la langue artificielle mais franchissaient une frontière de mots. Enfin, les MR correspondaient à six séquences trisyllabiques possédant la structure AxC. Ils ont été construits en remplaçant la deuxième syllabe de chaque MA par la première ou la troisième syllabe d'un MA appartenant à une famille différente. Contrairement aux MP, les MR n'apparaissaient jamais dans la langue artificielle.

Les stimuli du test ont été synthétisés avec MBROLA en utilisant les mêmes caractéristiques acoustiques que les stimuli des flux de langue artificielle. La tâche comptait en tout 54 items, répartis en 3 blocs. Chaque bloc présentait les mêmes dix-huit items (six MA, six MP et six MR) dans un ordre aléatoire. Il n'y avait pas de pause entre chaque bloc.

## **4. Procédure**

L'étude s'est déroulée en trois parties : (1) passations de pré-tests auprès de trois des participants avec aphasie, (2) validation du test de segmentation auprès des jeunes adultes, (3) passations de tests auprès de cinq des participants avec aphasie.

### **4.1. Procédure suivie auprès des participants avec aphasie**

Chaque participant a été exposé à un flux de langue artificielle de la condition de base (A ou B) et à un flux de la condition mot connu (A+ ou B+). Lorsqu'il s'agissait de pré-tests, les participants étaient soumis aux deux langues artificielles (flux A et B+ ; ou B et A+). Lors des tests, ils n'étaient soumis qu'à la langue artificielle A sous ses deux conditions (A et A+). Deux des participants aux tests n'ont pas pu passer l'expérience dans son intégralité en raison de la crise sanitaire provoquée par l'épidémie de Covid-19 : ils n'ont été soumis qu'à une seule condition (A ou A+) et leur capacité de discrimination phonologique n'a pas été testée. Les flux auxquels ont été exposés les participants sont précisés dans la partie « Résultats ».

La passation du test de segmentation a été divisée en deux sessions, chacune présentant une condition de l'expérience. Les personnes avec aphasie étant fatigables, nous avons fait le choix d'administrer chaque session à une semaine d'intervalle. Durant la semaine, les évaluations langagières décrites ci-dessus étaient administrées dans un ordre préétabli.

Les passations ont eu lieu dans un environnement calme afin de permettre des conditions d'écoute optimales et éviter toute distraction. Les participants étaient installés devant un ordinateur. Les stimuli leur étaient présentés via des haut-parleurs de qualité à une intensité confortable (les participants pouvaient régler l'intensité avant de commencer l'expérience à partir d'une bande sonore quelconque étalonnée sur la même intensité que les stimuli du test de segmentation). Les réponses au test devaient être sélectionnées via les touches du clavier de l'ordinateur, mises en évidence par des pastilles de couleur. Les consignes apparaissaient à l'écran. Elles étaient lues et expliquées par l'expérimentateur. Afin de s'assurer de la compréhension de la consigne, ce dernier accompagnait les participants lors des premiers items du test et rappelait la consigne si besoin.

Chaque session de l'expérience se déroulait de la façon suivante. Le principe de la langue artificielle était d'abord expliqué. La seule consigne donnée était d'écouter attentivement la langue afin d'essayer de découvrir les mots qui la composent. Aucune information relative à la structure ou à la longueur des MA n'était présentée. De même, il n'était pas indiqué le nombre de MA composant la langue. Le test de reconnaissance suivait immédiatement la phase d'exposition à la langue artificielle. Les participants devaient décider si les items présentés auditivement correspondaient oui ou non à un mot de la langue artificielle. Une fois la réponse « oui » ou « non » sélectionnée, ils devaient juger de leur certitude sur une échelle à trois niveaux (« peu sûr », « moyennement sûr », « très sûr » ; Palmer et al., 2019). Dans la condition mot connu, il était demandé aux participants d'énoncer ou de désigner parmi six images (selon les capacités d'expression orale des participants), le vrai mot inséré dans la langue, afin de vérifier qu'ils l'avaient bien perçu comme tel.

L'ordre de passation des deux conditions a été contrebalancé entre les participants, tout comme la disposition des touches « oui » et « non » du clavier de l'ordinateur.

## **4.2. Validation du test de segmentation auprès des jeunes adultes**

Les jeunes adultes ont été exposés aux deux flux de langue artificielle de la condition de base, soit aux flux A et B. La passation de la tâche a été divisée en deux sessions, chacune présentant un flux. Une pause de cinq à dix minutes a eu lieu entre les deux sessions afin de permettre aux participants d'oublier la première langue entendue. Pendant cette pause, le participant visionnait un court-métrage.

Chaque session se déroulait de la même façon que chez les participants avec aphasie à la seule différence que l'expérimentateur n'accompagnait pas les jeunes adultes lors des premiers items du test : il se contentait juste de lire et d'expliquer les consignes lorsqu'elles apparaissaient à l'écran. Les conditions expérimentales étaient également les mêmes, excepté le mode de présentation des stimuli. En effet, ces derniers ont été présentés au casque et non via des haut-parleurs.

L'ordre d'administration des flux A et B a été contrebalancé entre les participants, tout comme la disposition des touches « oui » et « non » du clavier de l'ordinateur.

## **5. Analyse des résultats**

Pour analyser les résultats obtenus au test de segmentation, plusieurs types de scores ont été calculés. Quel que soit le score, les réponses étaient considérées comme correctes si le participant répondait « oui » aux items correspondant aux MA, et « non » aux items correspondant aux MP et

aux MR. Elles étaient incorrectes s'il répondait « non » aux MA, et « oui » aux MP et aux MR. Dans ce dernier cas, les erreurs étaient qualifiées de « fausses reconnaissances ».

Ainsi, le *score B* correspond au pourcentage de réponses correctes obtenues ; une réponse correcte étant cotée « 1 » et une réponse incorrecte « 0 ». Il s'agit d'un score binaire. Le *score C* correspond à la moyenne des réponses de chaque participant pondérées par le degré de certitude : chaque réponse correcte est cotée « +1 », « +2 » ou « +3 » selon le degré de certitude (1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ») et chaque réponse incorrecte « -1 », « -2 » ou « -3 » sur le même principe. Ce score n'a été calculé que pour les jeunes adultes dans le cadre de la validation du test. Pour analyser de façon plus précise la certitude des participants avec aphasie, nous avons calculé la *moyenne de leur certitude* dans leurs réponses correctes et incorrectes.

## 5.1. Validation du test de segmentation

Les analyses statistiques se rapportant à la validation du test ont été effectuées avec le logiciel JASP (JASP Team, 2018). Le seuil de significativité considéré dans ce mémoire correspond à une valeur  $p$  ( $p$ ) inférieure à 0.05. Les tailles d'effet ont été précisées lorsque cela était possible.

Tout d'abord, nous avons calculé les moyennes et écart-types des *scores B* obtenus par les jeunes adultes dans chacune des langues. Pour vérifier si ces moyennes s'éloignaient significativement du hasard, nous avons utilisé un test de student pour échantillon unique avec un niveau de chance défini à 50% de bonnes réponses.

Puis, nous avons réalisé des analyses de la variance (ANOVA) à plusieurs facteurs explicatifs, afin d'évaluer l'influence de différentes variables catégorielles sur les capacités de segmentation des jeunes adultes, alors définies par les *scores B* et *C*. Plus précisément, nous avons effectué une première ANOVA pour évaluer l'influence de la *langue écoutée* (langue A vs. langue B), du *type d'items* (MA, MP, MR) et de l'*ordre d'administration des flux* (première session vs. deuxième session) sur les *scores B* obtenus par les jeunes adultes. Une deuxième ANOVA a été réalisée afin d'évaluer l'influence de ces mêmes facteurs sur les *scores C*.

Enfin, nous avons effectué deux analyses de type ANOVA pour tester l'influence du bloc d'items du test (bloc 1, bloc 2, bloc 3) sur les *scores B* et *C* des participants.

## 5.2. Participants avec aphasie

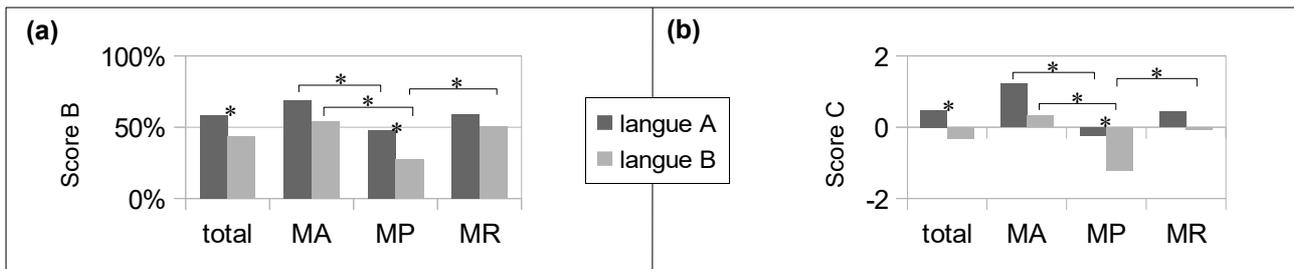
Compte tenu du faible effectif de la population, nous avons préféré étudier les participants avec aphasie comme des cas uniques. Pour chacun d'eux, plusieurs scores ont été calculés et présentés sous la forme de graphiques : (1) les *scores B* obtenus au total et pour chaque type d'items (MA, MP, MR) dans les deux conditions ; (2) les *moyennes de leur certitude* dans leurs réponses, correctes et incorrectes, au total et pour chaque type d'items. Ces scores ont été calculés sur l'ensemble des items du test (items 1 à 54,  $N = 54$ ) mais aussi pour le bloc 1 (items 1 à 18,  $N = 18$ ). Nous avons utilisé le test exact de la loi binomiale afin de vérifier si les *scores B* totaux supérieurs au niveau de chance (soit à 50%) dépassaient significativement le hasard.

À partir de ces scores, nous avons réalisé une analyse descriptive détaillée des performances de chaque participant au test de segmentation. Tout d'abord, nous avons comparé leur performance dans les deux conditions sur les *scores B* totaux. Puis, nous avons analysé leur capacité à discriminer les différents types d'items dans chaque condition, en nous reportant aux *scores B* obtenus pour les MA, les MP et les MR. Enfin, nous avons comparé les *moyennes de leur certitude* dans leurs réponses dans chaque condition. Lorsque cela était nécessaire pour comprendre les résultats, des éléments d'analyse qualitative ont été ajoutés.

# Résultats

## 1. Validation du test de segmentation

Tout d'abord, les moyennes des *scores B* obtenus par les jeunes adultes ont été calculées dans chacune des langues. Dans la langue A, la moyenne était de 58.615 % (ET = 26.137) ; celle-ci était significativement meilleure que le hasard,  $t(68) = 2.738$ ,  $p = 0.008$ ,  $d = 0.330$ . Dans la langue B, la moyenne était de 43.881 % (ET = 22.909) ; celle-ci était significativement inférieure au niveau de chance,  $t(68) = -2.219$ ,  $p = 0.030$ ,  $d = 0.267$  (cf. Figure 1(a)).



**Figure 1.** Scores obtenus par les jeunes adultes dans chaque langue (A vs. B), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). L'astérisque (\*) indique la présence d'une différence significative entre les scores. (a) Pourcentages de réponses correctes ou *scores B*. (b) Moyennes des réponses pondérées par la certitude ou *scores C*.

Puis, des analyses statistiques de type ANOVA à plusieurs facteurs explicatifs ont permis d'évaluer l'influence de différentes variables sur les capacités de segmentation des jeunes adultes : *langue écoutée* (langue A vs. langue B), *type d'items* (MA, MP, MR), et *ordre d'administration des flux* (première session vs. deuxième session). Les résultats de ces analyses ont été résumés dans le Tableau 2.1 (cf. Annexe A2). Nous avons mis en évidence une influence significative des variables *langue écoutée* et *type d'items* sur les moyennes des *scores B* et *C* obtenus par les jeunes adultes. En revanche, nous n'avons pas relevé d'effets de l'*ordre d'administration des flux*.

Des comparaisons post-hoc ont été réalisées afin d'identifier les moyennes contribuant aux effets mis en évidence (cf. Tableau 2.2, Annexe A2). Nous avons montré que les performances des jeunes adultes étaient significativement meilleures pour la langue A que pour la langue B (cf. Figure 1). En ce qui concerne la variable *type d'items*, nous avons montré que les *scores B* et *C* obtenus pour les MA et pour les MR étaient significativement meilleurs que ceux obtenus pour les MP (cf. Tableau 2.2, Annexe A2). Nous avons également noté qu'il existait une différence significative entre les MA et les MR, mais uniquement pour le *score C* (cf. tableau 2.2, Annexe A2).

En outre, nous nous sommes intéressée aux interactions entre les variables *langue écoutée* et *type d'items* afin d'évaluer l'effet du type d'items dans chacune des langues (cf. tableau 2.2, Annexe A2). Nous avons remarqué que la différence significative relevée entre les MA et les MP était observée dans les deux langues (cf. Figure 1) : les scores concernant les MP étaient significativement inférieurs à ceux concernant les MA dans les langues A et B (cf. Figure 1). Par ailleurs, dans la langue B, les scores obtenus pour les MP étaient aussi significativement inférieurs à ceux obtenus pour les MR (cf. Figure 1). Nous avons alors soulevé le fait que les scores obtenus dans les langues A et B ne différaient significativement que pour les MP (cf. Figure 1) : ce type d'items occasionnerait significativement plus d'erreurs dans la langue B.

Enfin, pour tester l'influence de la variable *bloc*, des ANOVA à plusieurs facteurs ont été réalisées. Le modèle incluait les variables explicatives *langue écoutée*, *type d'items* et *bloc*. Les résultats de ces analyses ont montré qu'il n'y avait pas d'influence significative de la variable *bloc* sur le *score B*,  $F(2,396) = 0.231$ ,  $p = 0.794$ , ni sur le *score C*,  $F(2,396) = 0.373$ ,  $p = 0.689$ .

## 2. Résultats des participants avec aphasie

### 2.1. Participant 46

Le participant 46 (nommé « P46 ») était une femme droitrière âgée de 49 ans et de langue maternelle française. Elle présentait une aphasie de Broca sévère, en phase subaiguë après un AVC ischémique survenu sur dissection de la carotide interne gauche. Les lésions étaient localisées dans le territoire des artères cérébrales moyenne, antérieure et jonctionnelle postérieure gauche.

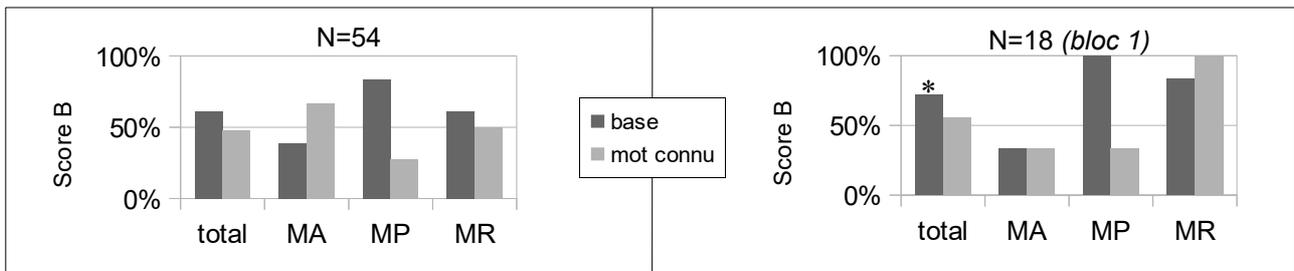
Les troubles phasiques de P46 se caractérisaient principalement par une réduction majeure de la fluence liée à une apraxie bucco-linguo-faciale ainsi qu'à une apraxie de la parole. Les productions orales étaient parasitées par des persévérations et une stéréotypie verbale encore incontrôlable. Sur le plan réceptif, seul le lexique phonologique d'entrée était préservé. P46 présentait quelques difficultés de discrimination phonologique, une légère atteinte du système sémantique et de la compréhension orale lexicale, ainsi que de sévères difficultés de compréhension orale syntaxique. La compréhension était meilleure en contexte.

P46 a participé à l'étude dans le cadre d'un pré-test. Elle a été exposée à la condition de base (flux A) en première session et à la condition mot connu (flux B+) en deuxième session.

Dans la condition mot connu, P46 semblait affirmer avoir perçu la présence d'un vrai mot dans la langue artificielle. Cependant, elle n'a pas su l'identifier précisément : la participante a désigné les images correspondant aux mots « médicament » et « médecin ».

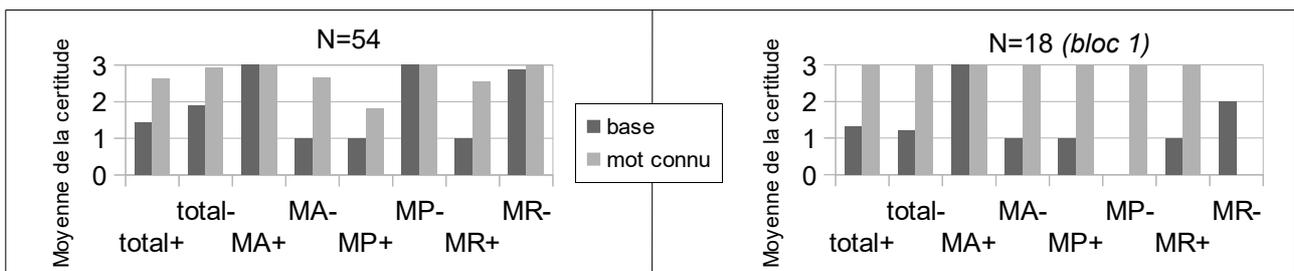
La Figure 2.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P46 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* total est mis en évidence dans la condition de base pour  $N=54$  et  $N=18$ . Seul le *score B* total obtenu dans la condition de base pour  $N=18$  est significativement meilleur que le hasard (Test binomial, unilatéral,  $p = 0.048$ ).

L'analyse des *scores B* en fonction du type d'items met en évidence certaines tendances (cf. Figure 2.1). Dans la condition de base, les *scores B* obtenus pour les MP et les MR sont supérieurs au niveau de chance mais pas celui obtenu pour les MA. Cela suggère que P46 présentait des difficultés pour reconnaître les MA de la langue écoutée. Dans la condition mot connu et à l'échelle du test ( $N=54$ ), la tendance s'inverse : seul le *score B* obtenu pour les MA dépasse le niveau de chance. P46 a donc reconnu davantage de MA dans cette condition mais a aussi fait plus de fausses reconnaissances. En début de test ( $N=18$ ), la participante avait identifié moins de MA et plus de MR. Ainsi, quelle que soit la condition, la répartition des réponses entre les différents types d'items suggère que P46 n'a pas été en capacité de discriminer ceux appartenant à la langue écoutée (les MA) de ceux n'en faisant pas partie (les MP et les MR).



**Figure 2.1.** Pourcentages de réponses correctes (scores *B*) obtenus par P46 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. L'astérisque (\*) indique les scores totaux significativement supérieurs au niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 2.2 représente la *moyenne de la certitude* de P46 dans ses réponses, correctes et incorrectes, et dans chaque condition. Globalement, nous observons une certitude moyenne plus élevée en présence du mot connu, que les réponses de la participante se soient révélées correctes ou incorrectes. Cela n'est pas corrélé avec l'exactitude des réponses, le *score B* total étant inférieur dans la condition mot connu (cf. Figure 2.1). De plus, dans la condition de base, nous pouvons remarquer que P46 était peu sûre de ses réponses lorsqu'elle répondait « non » (cf. MA-, MP+, MR+, Figure 2.2) et très sûre lorsqu'elle répondait « oui » (cf. MA+, MP-, MR-, Figure 2.2), indépendamment du fait que ses réponses soient correctes ou non.



**Figure 2.2.** Moyennes de la certitude de P46 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.2. Participant 23

Le participant 23 (nommé « P23 ») était une femme droitrière âgée de 56 ans et de langue maternelle française. Elle présentait une aphasie de Broca modérée, en tout début de phase chronique après un AVC ischémique d'étiologie indéterminée. Les lésions occasionnées touchaient les régions cortico-sous-corticales pariéto-temporo-insulaires gauches et frontales gauches. Le tableau clinique évoquait en phase aiguë une aphasie globale.

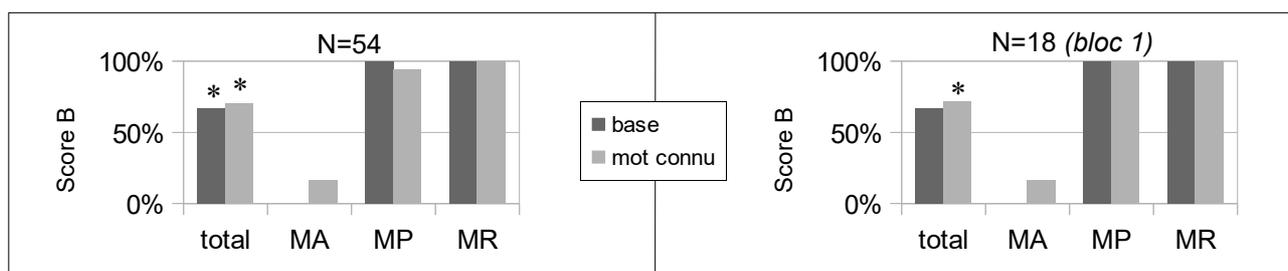
A son inclusion dans l'étude, P23 présentait sur le plan expressif un manque du mot modéré et un léger agrammatisme, à l'origine d'une réduction de la fluence de la parole. La répétition de mots et de non-mots, qu'elle soit différée ou non, était altérée. Le système sémantique était en revanche préservé, suggérant de ce fait une atteinte du lexique phonologique de sortie. Sur le plan réceptif, la participante présentait quelques difficultés de discrimination phonologique, accentuées avec la rapidité de la parole. Le lexique phonologique d'entrée était atteint. La compréhension orale était préservée en contexte et au niveau lexical mais pas au niveau syntaxique : P23 présentait un trouble sévère de la compréhension d'énoncés.

P23 a participé à l'étude dans le cadre d'un pré-test. Elle a été exposée à la condition mot connu (flux B+) en première session et à la condition de base (flux A) en deuxième session.

Dans la condition mot connu, P23 a bien identifié le vrai mot inséré dans la langue artificielle. Elle n'a pas su l'énoncer à l'oral mais a pu désigner l'image correspondante.

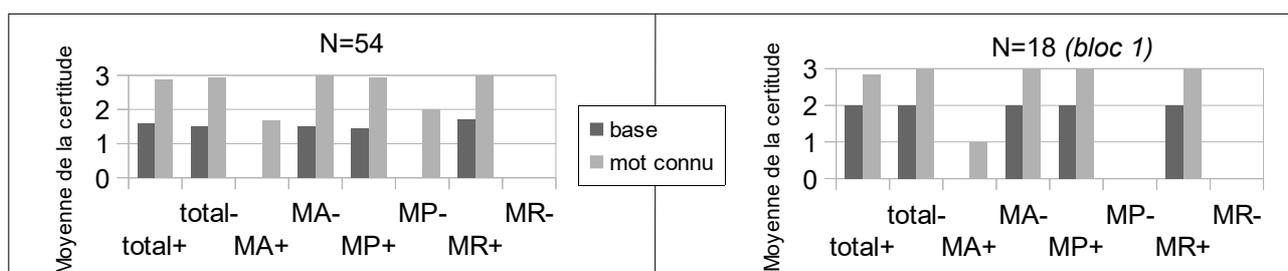
La Figure 3.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P23 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* total est mis en évidence dans la condition mot connu pour N=54 et N=18. Trois des *scores B* totaux sont significativement meilleurs que le hasard : ceux de la condition mot connu pour N=18 (Test binomial, unilatéral,  $p = 0.048$ ) et N=54 (Test binomial, unilatéral,  $p = 0.010$ ) ; et celui de la condition de base pour N=54 (Test binomial, unilatéral,  $p = 0.002$ ).

L'analyse des *scores B* en fonction du type d'items met clairement en évidence des difficultés de discrimination dans les deux conditions (cf. Figure 3.1). En effet, les *scores B* se rapportant aux MA sont très faibles, voire nuls dans la condition de base, tandis que ceux concernant les MP et les MR atteignent quasiment tous la perfection. Cela suggère que P23 n'a donné que des réponses négatives lors du test dans la condition de base et une grande majorité de réponses négatives dans la condition mot connu. Elle n'a pas été capable d'identifier les items correspondant à des MA de la langue écoutée, tout en rejetant ceux correspondant à des non-mots (MP et MR). Toutefois, nous pouvons souligner que les quelques réponses positives données dans la condition mot connu étaient pour la plupart correctes et démontrent que P23 a identifié l'un des MA de la langue : elle a effectivement reconnu le mot *togadou* dans les trois blocs. Elle n'a fait qu'une seule fausse reconnaissance sur un MP formé en partie des deux dernières syllabes du mot reconnu (*gadouba*).



**Figure 3.1.** Pourcentages de réponses correctes (*scores B*) obtenus par P23 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. L'astérisque (\*) indique les scores totaux significativement supérieurs au niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 3.2 représente la *moyenne de la certitude* de P23 dans ses réponses, correctes et incorrectes, et dans chaque condition. Nous pouvons relever que la certitude moyenne est plus élevée en présence du mot connu, que les réponses se soient révélées correctes ou incorrectes. De plus, nous observons dans cette condition une variation du degré de certitude corrélée à la nature de la réponse : P23 était plus sûre de ses réponses lorsqu'elle répondait « non » (cf. MA-, MP+, MR+, Figure 3.2) que lorsqu'elle répondait « oui » (cf. MA+, MP-, Figure 3.2).



**Figure 3.2.** Moyennes de la certitude de P23 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.3. Participant 51

Le participant 51 (nommé « P51 ») était une femme droitrière âgée de 44 ans et de langue maternelle française. Elle présentait une aphasie non fluente modérée, en phase subaiguë après un AVC hémorragique causé par une rupture d'anévrisme. Les lésions occasionnées touchaient les régions fronto-temporo-insulaires gauches.

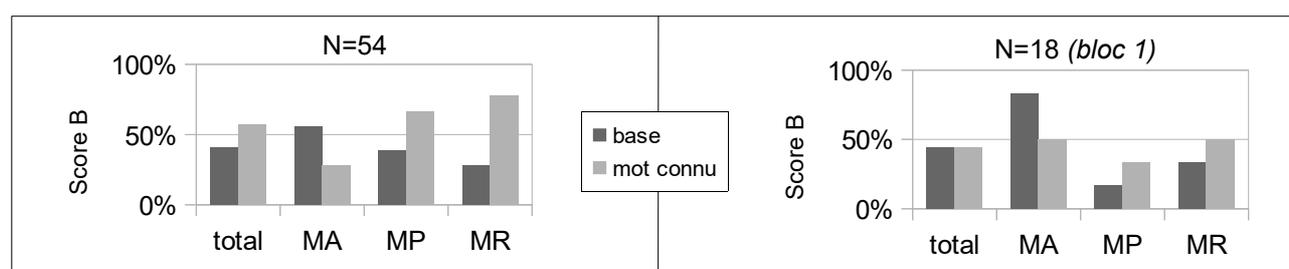
Les troubles phasiques de P51 évoquaient un tableau d'aphasie transcorticale mixte. Sur le plan expressif, elle présentait un manque du mot important et une légère dyssyntaxie. La répétition de mots et de non-mots était préservée malgré une altération de la mémoire tampon phonologique (difficultés en répétition différée). Le système sémantique était aussi légèrement atteint. Sur le plan réceptif, seul le lexique phonologique d'entrée était préservé. P51 présentait quelques difficultés de discrimination phonologique, un léger trouble de compréhension orale lexicale et des difficultés plus sévères de compréhension orale syntaxique. La compréhension était meilleure en contexte.

P51 a participé à l'étude dans le cadre d'un pré-test. Elle a été exposée à la condition mot connu (flux A+) en première session et à la condition de base (flux B) en deuxième session.

Dans la condition mot connu, P51 semblait affirmer avoir perçu la présence d'un vrai mot dans la langue artificielle. Cependant, elle n'a pas su l'identifier : la participante a désigné l'image correspondant au mot « médecin » et non à « médicament ».

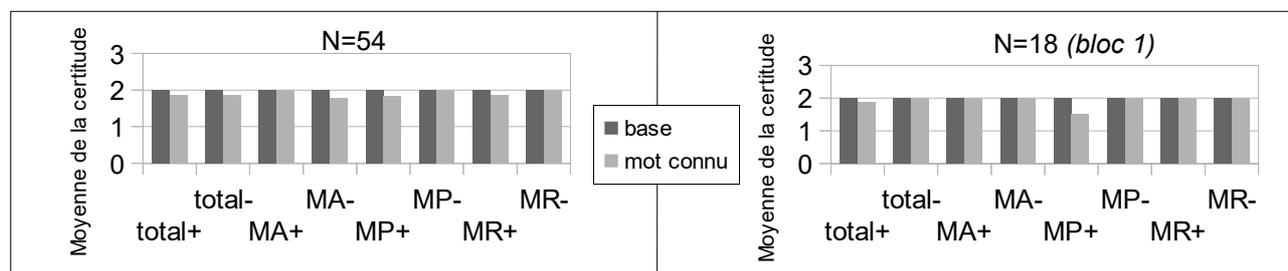
La Figure 4.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P51 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* a été mis en évidence dans la condition mot connu pour N=54. Il s'agit du seul *score B* total supérieur au niveau de chance ; la différence avec le hasard n'est pas significative (Test binomial, unilatéral,  $p = 0.170$ ).

L'analyse des *scores B* en fonction du type d'items met en évidence des différences entre les deux conditions (cf. Figure 4.1). Dans la condition de base, seul le *score B* obtenu pour les MA est supérieur au niveau de chance : P51 a donc reconnu plus de la moitié des MA mais a aussi fait de nombreuses fausses reconnaissances. Dans la condition mot connu et à l'échelle du test (N=54), nous pouvons observer le contraire : seuls les *scores B* obtenus pour les MP et les MR dépassent le niveau de chance, suggérant ainsi une difficulté pour reconnaître les MA de la langue écoutée. En début de test dans cette même condition (N=18), aucun *score B* ne dépassait le niveau de chance. La comparaison avec les résultats obtenus sur l'ensemble du test (N=54) souligne le fait que les quelques MA identifiés dans la condition mot connu étaient surtout ceux présentés en début de test, et que P51 donnait davantage de réponses négatives après le bloc 1. Cela explique l'amélioration des *scores B* pour les MP et les MR à l'échelle du test. Ainsi, quelle que soit la condition administrée, la répartition des réponses entre les différents types d'items suggère que P51 n'a pas été en capacité de différencier ceux appartenant à la langue écoutée et ceux n'en faisant pas partie.



**Figure 4.1.** Pourcentages de réponses correctes (*scores B*) obtenus par P51 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 4.2 représente la *moyenne de la certitude* de P51 dans ses réponses, correctes et incorrectes, et dans chaque condition. Globalement, nous observons une certitude moyenne similaire dans les deux conditions. La certitude est légèrement inférieure en présence du mot connu quand P51 donnait une réponse négative lors du test (cf. MA-, MP+, MR+, figure 4.2).



**Figure 4.2.** Moyennes de la certitude de P51 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.4. Participant 99

Le participant 99 (nommé « P99 ») était un homme droitier âgé de 66 ans, présentant a priori un bilinguisme précoce français/créole. Il était atteint d'une aphasie légère de type anomique, en phase subaiguë après un AVC ischémique sous-cortical. Les lésions occasionnées étaient localisées au niveau des centres semi-ovales gauche et droit. Le participant présentait également une surdité pour laquelle il était appareillé. L'examineur a veillé au port de ses prothèses lors de l'expérience.

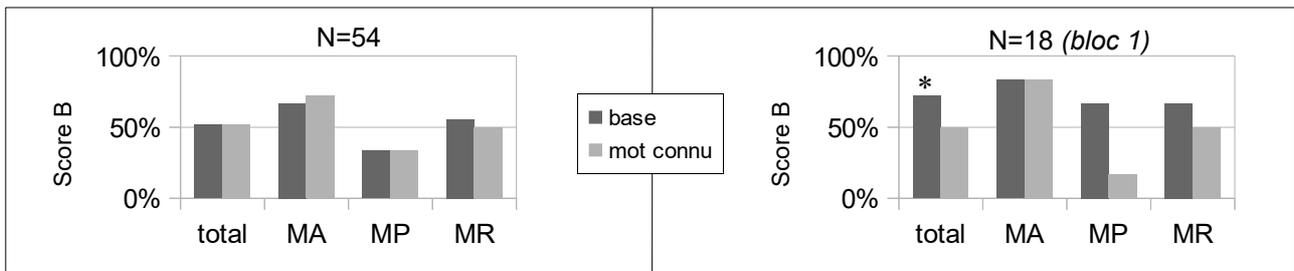
Les difficultés phasiques présentées étaient à peine perceptibles et surtout ressenties par le participant. Elles se caractérisaient notamment par un léger manque du mot. Il n'y avait pas d'atteinte du système sémantique, ni du lexique phonologique d'entrée. La compréhension orale, lexicale comme syntaxique, était préservée. Le participant présentait quelques difficultés de discrimination auditive, probablement liées à son déficit auditif.

P99 a participé à l'étude dans le cadre du test. Il a été exposé à la condition de base (flux A) en première session et à la condition mot connu (flux A+) en deuxième session.

Dans la condition mot connu, le participant a bien identifié le vrai mot inséré dans la langue artificielle : il a su l'énoncer à l'oral.

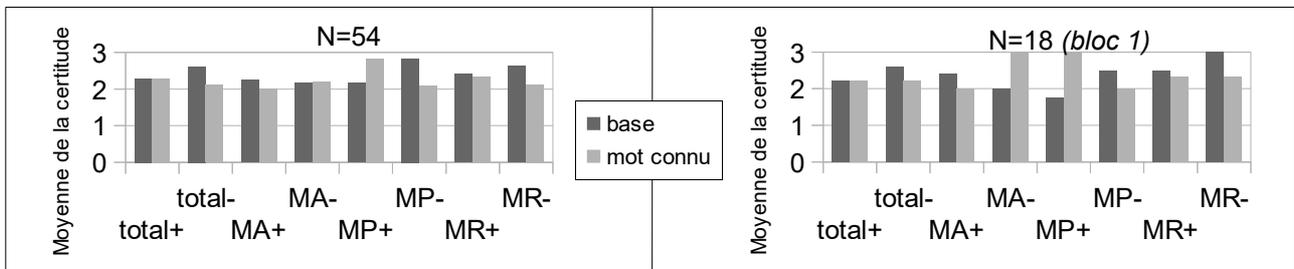
La Figure 5.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P99 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* total a été mis en évidence dans la condition de base pour N=18. Ce score est significativement supérieur au niveau de chance (Test binomial, unilatéral,  $p = 0.048$ ). Aucune différence n'est notée entre les *scores B* totaux pour N=54.

L'analyse des *scores B* en fonction du type d'items met en évidence une répartition des réponses similaire dans les deux conditions pour N=54 (cf. Figure 5.1) : les *scores B* se rapportant aux MA et MR sont supérieurs au niveau de chance mais pas ceux se rapportant aux MP. P99 aurait réussi à identifier la majorité des MA de la langue écoutée tout en rejetant la présence de la majorité des MR. Il n'aurait en revanche pas suffisamment bien discriminé les MA des MP, ces derniers faisant l'objet de nombreuses fausses reconnaissances. Cette répartition des réponses était déjà présente dans la condition mot connu dès le début du test (N=18). Dans la condition de base et en début de test (N=18), P99 aurait réussi à discriminer les MA des non-mots : les *scores B* des trois types d'items étaient supérieurs au niveau de chance ; il n'y avait pas de différence entre les *scores B* obtenus pour les MP et les MR.



**Figure 5.1.** Pourcentages de réponses correctes (*scores B*) obtenus par P99 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. L'astérisque (\*) indique les scores totaux significativement supérieurs au niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 5.2 représente la *moyenne de la certitude* de P99 dans ses réponses, correctes et incorrectes, et dans chaque condition. La certitude au total sur les réponses correctes est identique dans les deux conditions. En revanche, celle sur les réponses incorrectes est plus élevée dans la condition de base uniquement. Dans la condition mot connu, la certitude moyenne de P99 variait pour les MP en fonction de l'exactitude de la réponse : P99 était plus sûr de lui lorsqu'il rejetait leur présence de la langue artificielle que lorsqu'il les reconnaissait à tort comme des MA.



**Figure 5.2.** Moyennes de la certitude de P99 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.5. Participant 85

Le participant 85 (nommé « P85 ») était une femme droitrière âgée de 85 ans et de langue maternelle française. Elle présentait une aphasie légère de type anomique, en phase subaiguë après un AVC ischémique sylvien profond gauche d'étiologie cardio-embolique. Le tableau clinique évoquait en phase aiguë une aphasie de Broca.

Lors de son inclusion dans l'étude, les troubles avaient bien régressé. Ils se caractérisaient principalement par un léger manque du mot. Le système sémantique était préservé. Sur le plan réceptif, la discrimination auditive était correcte. Il n'y avait pas d'atteinte du lexique phonologique d'entrée, ni de la compréhension orale lexicale. En revanche, P85 présentait des difficultés modérées de compréhension syntaxique, touchant notamment les phrases passives.

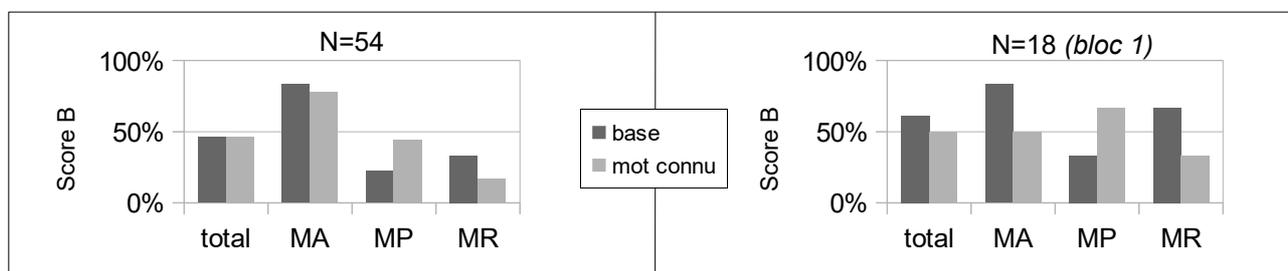
P85 a participé à l'étude dans le cadre du test. Elle a été exposée à la condition mot connu (flux A+) en première session et à la condition de base (flux A) en deuxième session.

Dans la condition mot connu, P85 a bien identifié le vrai mot inséré dans la langue artificielle : elle a su l'énoncer à l'oral.

La Figure 6.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P85 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* a été mis en évidence dans la condition de base pour N=18. Il s'agit du seul score supérieur au niveau de chance mais la

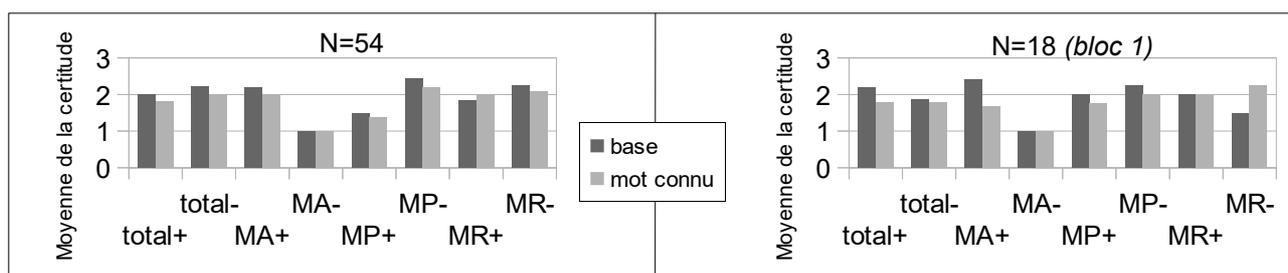
différence avec le hasard n'est pas significative (Test binomial, unilatéral,  $p = 0.240$ ). Aucune différence n'est notée entre les *scores B* pour  $N=54$ .

L'analyse des *scores B* en fonction du type d'items met en évidence la même tendance dans les deux conditions pour  $N=54$  (cf. Figure 6.1) : seul le *score B* se rapportant aux MA est supérieur au niveau de chance. Cela suggère que P85 a reconnu plus de la moitié des MA mais qu'elle a aussi fait de nombreuses fausses reconnaissances. Ainsi, elle aurait été en difficulté pour différencier les items correspondant à des mots de la langue écoutée et ceux correspondant à des non-mots. En début de test ( $N=18$ ), la répartition des réponses est différente. Dans la condition de base, les *scores B* se rapportant aux MA et aux MR sont supérieurs au niveau de chance mais pas celui se rapportant aux MP : P85 aurait réussi à identifier la majorité des MA de la langue et à rejeter la présence de la majorité des MR. Elle n'aurait en revanche pas suffisamment bien discriminé les MA des MP, ces derniers occasionnant de nombreuses fausses reconnaissances. Dans la condition mot connu, seul le *score B* se rapportant aux MP dépasse le niveau de chance, suggérant davantage de difficultés pour discriminer les différents types d'items.



**Figure 6.1.** Pourcentages de réponses correctes (*scores B*) obtenus par P85 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. L'astérisque (\*) indique les scores totaux significativement supérieurs au niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 6.2 représente la *moyenne de la certitude* de P85 dans ses réponses, correctes et incorrectes, et dans chaque condition. La certitude moyenne au total est légèrement inférieure dans la condition mot connu, et ce indépendamment de l'exactitude des réponses. Dans les deux conditions, P85 était un peu moins sûre d'elle lorsqu'elle rejetait à tort un item de type MA.



**Figure 6.2.** Moyennes de la certitude de P85 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.6. Participant 66

Le participant 66 (nommé « P66 ») était une femme gauchère âgée de 86 ans et de langue maternelle française. Elle présentait une aphasie anomique légère en phase subaiguë après un AVC ischémique sylvien droit d'étiologie cardio-embolique. Les lésions retrouvées étaient peu étendues dans le territoire sylvien droit. La participante présentait également une surdité pour laquelle elle était appareillée. L'examineur a veillé au port de ses prothèses lors de l'expérience.

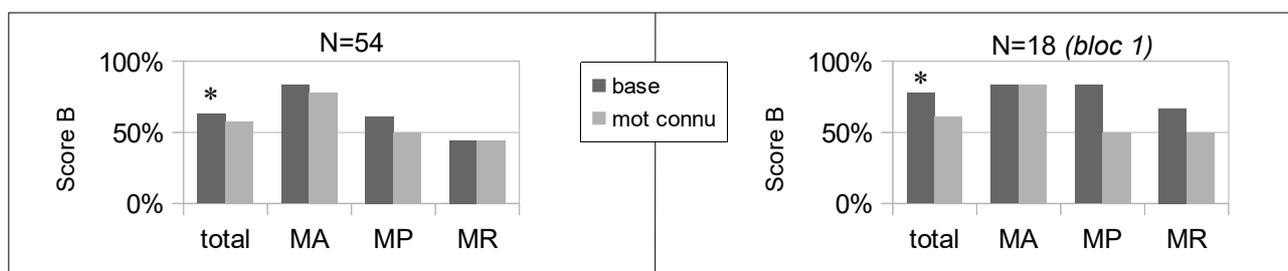
Les troubles phasiques présentés par P66 se caractérisaient notamment par un léger manque du mot, surtout ressenti par la participante. Le lexique phonologique d'entrée ainsi que le système sémantique étaient préservés. Les capacités de compréhension orale lexicale et syntaxique étaient également satisfaisantes. La discrimination auditive était en revanche perturbée par le déficit auditif.

P66 a participé à l'étude dans le cadre du test. Elle a été exposée à la condition de base (flux A) en première session et à la condition mot connu (flux A+) en deuxième session

Dans la condition mot connu, P66 a bien identifié le vrai mot inséré dans la langue artificielle. Initialement, elle disait ne plus se souvenir du mot en question. Elle a pu le retrouver à partir des images proposées par l'examineur.

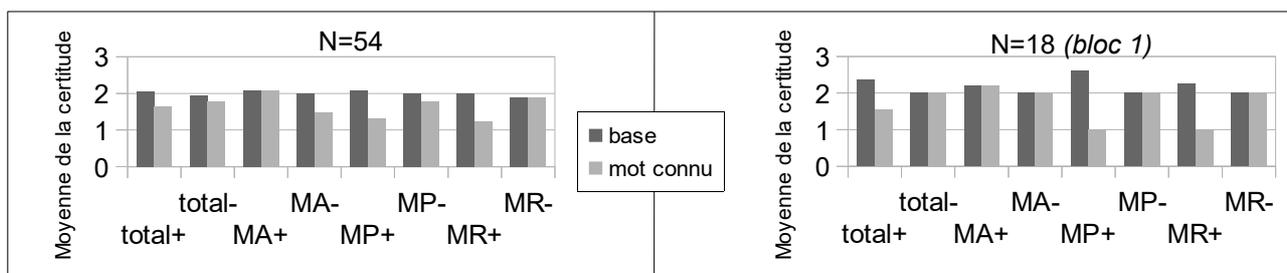
La figure 7.1 illustre les pourcentages de réponses correctes (*scores B*) obtenus par P66 dans chaque condition, au total et pour chaque type d'items. Un meilleur *score B* total est mis en évidence dans la condition de base pour N=54 et N=18. Tous les *scores B* totaux sont supérieurs au niveau de chance. Cependant, seuls ceux obtenus dans la condition de base dépassent significativement le hasard (N=54, Test binomial, unilatéral,  $p = 0.038$  ; N=18, Test binomial, unilatéral,  $p = 0.015$ ).

L'analyse des *scores B* en fonction du type d'items témoigne d'une meilleure discrimination de ces derniers dans la condition de base (cf. Figure 7.1). En effet, dans cette condition et à l'échelle du test (N=54), seuls les *scores B* se rapportant aux MA et aux MP sont supérieurs au niveau de chance : P66 aurait donc réussi à identifier la majorité des MA de la langue écoutée tout en rejetant la présence de la majorité des MP. Elle aurait moins bien discriminé les MA des MR, plus de la moitié de ces derniers faisant l'objet de fausses reconnaissances. Dans le bloc 1 (N=18), les *scores B* des trois types d'items sont tous supérieurs au niveau de chance, suggérant une meilleure discrimination entre les items en début de test. Dans la condition mot connu (N=54 et N=18), la répartition des réponses est similaire à celle décrite dans la condition de base pour N=54, à la seule différence que le *score B* se rapportant aux MP ne dépassait pas le niveau de chance.



**Figure 7.1.** Pourcentages de réponses correctes (*scores B*) obtenus par P66 dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). La ligne horizontale ( $y = 50\%$ ) représente le niveau de chance. L'astérisque (\*) indique les scores totaux significativement supérieurs au niveau de chance. « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

La Figure 7.2 représente la *moyenne de la certitude* de P66 dans ses réponses, correctes et incorrectes, et dans chaque condition. La certitude moyenne au total était un plus élevée dans la condition de base, notamment sur les réponses correctes. Dans la condition mot connu, nous relevons davantage de variations de la certitude en fonction des réponses, notamment en début de test (N=18) : P66 était moins sûre d'elle dès lors qu'elle rejetait, pourtant sans faire d'erreurs, la présence des MP et des MR dans la langue artificielle (cf. MP+, MR+, Figure 7.2).



**Figure 7.2.** Moyennes de la certitude de P66 dans ses réponses, correctes (+) et incorrectes (-), dans chaque condition (base vs. mot connu), au total et pour chaque type d'items (MA = mots artificiels, MP = mots-partiels, MR = mots-règles). Moyenne de la certitude : 1 = « peu sûr », 2 = « moyennement sûr », 3 = « très sûr ». « N » correspond au nombre d'items du test à partir duquel les scores ont été calculés.

## 2.7. Participant 39

Le participant 39 (nommé « P39 ») était un homme droitier de 68 ans de langue maternelle française. Il présentait une aphasie fluente légère, en phase subaiguë après un probable AVC ischémique secondairement hémorragique de localisation lenticulo-caudée. Le participant avait déjà été victime d'un AVC hémorragique temporal il y a 40 ans, et d'un AVC pariétal gauche il y a 4 ans.

Les troubles phasiques de P39 évoquaient un tableau d'aphasie transcorticale sensorielle, touchant principalement le versant réceptif. Le participant présentait une légère atteinte du système sémantique. Il possédait de bonnes capacités de compréhension orale au niveau lexical mais pas au niveau syntaxique : la compréhension orale de phrases était modérément atteinte. P39 n'avait pas de difficultés de discrimination auditive. Son lexique phonologique d'entrée et de sortie étaient préservés, tout comme ses capacités de répétition.

P39 a été inclus dans le cadre d'un pré-test. Il a été exposé à la condition de base (flux B) en première session et à la condition mot connu (flux A+) en deuxième session. Cependant, nous avons dû exclure les résultats de cette deuxième session en raison d'une mauvaise compréhension de la consigne, signalée par le participant lui-même à la fin de l'épreuve. Nous tenons toutefois à préciser que P39 avait bien identifié le vrai mot inséré dans la langue artificielle.

La figure 8(a) (cf. Annexe A3) illustre les pourcentages de réponses correctes (*scores B*) obtenus par P39 au total et pour chaque type d'items. Le *score B* total est supérieur au niveau de chance uniquement pour N=54, mais sans que cela ne soit significatif (Test binomial, unilatéral,  $p = 0.170$ ). L'analyse des *scores B* en fonction du type d'items témoigne de difficultés pour discriminer les items entre eux : les *scores B* obtenus pour les MP et les MR sont supérieurs au niveau de chance mais pas celui obtenu pour les MA.

La Figure 8(b) (cf. Annexe A3) représente la *moyenne de la certitude* de P39 dans ses réponses, correctes et incorrectes. Elle reste moyenne sur l'ensemble du test, quoique légèrement meilleure en début de test (N=18).

## 2.8. Participant 27

Le participant 27 (nommé « P27 ») était un homme droitier de 54 ans de langue maternelle française. Il présentait une aphasie de Broca sévère en phase subaiguë après un AVC ischémique sylvien gauche survenu sur une occlusion en tandem de la carotide et de l'artère cérébrale moyenne gauches, dans les suites d'une thrombo-endarterectomie d'une sténose athéromateuse du bulbe carotidien gauche. Le participant avait présenté un premier AVC ischémique sylvien gauche un mois avant celui-ci, suite auquel il n'avait présenté aucune séquelle hormis un léger manque du mot.

Les troubles phasiques de P27 étaient caractérisés sur le plan expressif par une réduction majeure de la fluence, liée à une apraxie de la parole. Les productions orales étaient parasitées par des persévérations et une stéréotypie verbale difficilement contrôlables, rendant impossible la passation des épreuves de dénomination et de répétition. Sur le plan réceptif, P27 présentait de sévères difficultés de compréhension orale lexicale et syntaxique. Son système sémantique était atteint. Son lexique phonologique d'entrée semblait l'être également. Nous n'avons pas pu évaluer ses capacités de discrimination auditive. La compréhension semblait meilleure en contexte.

P27 a été inclus dans le cadre d'un test. Il n'a été exposé qu'à la condition de base (flux A).

La Figure 9(a) (cf. Annexe A3) illustre les pourcentages de réponses correctes (*score B*) de P27 au total et pour chaque type d'items. Les *scores B* totaux ne dépassent pas le niveau de chance. L'analyse en fonction du type d'items témoigne de difficultés pour discriminer les items entre eux : les *scores B* pour les MP et MR sont égaux ou supérieurs au niveau de chance ; ils sont inférieurs à ce dernier, voire nuls en début de test (N=18) pour les MA.

La figure 9(b) (cf. Annexe A3) représente la *moyenne de la certitude* de P27 dans ses réponses, correctes et incorrectes. Globalement, la certitude était meilleure en début de test (N=18), et ce indépendamment de l'exactitude des réponses et du type d'items.

## 2.9. Participant 59

Le participant 59 (nommé « P59 ») était une femme droitrière âgée de 54 ans et de langue maternelle française. Elle présentait une aphasie globale évoluant progressivement vers une aphasie de Broca sévère, en phase subaiguë après un AVC ischémique sylvien gauche.

Les troubles phasiques de P59 étaient caractérisés par d'importantes difficultés sur les plans expressif et réceptif. P59 était non fluente en raison d'un trouble phonologique sévère et d'une apraxie de la parole. De plus, ses productions orales étaient limitées par la présence d'une stéréotypie verbale qu'elle commençait à savoir contrôler. Elles étaient aussi parasitées par des persévérations. Son système sémantique était quant à lui préservé. Le lexique phonologique d'entrée semblait atteint. Ses capacités de discrimination auditive n'ont pas pu être évaluées. En ce qui concerne la compréhension orale, P59 présentait de sévères difficultés au niveau lexical et syntaxique. Sa compréhension semblait néanmoins meilleure en contexte.

P59 a été inclus dans le cadre d'un test. Elle n'a pu être exposée qu'à la condition mot connu (flux A+). La participante n'aurait pas perçu la présence du vrai mot dans le flux.

La Figure 10(a) (cf. Annexe A3) illustre les pourcentages de réponses correctes (*score B*) obtenus par P59 au total et pour chaque type d'items. Les *scores B* totaux ne dépassent pas le niveau de chance. L'analyse en fonction du type d'items témoigne de difficultés pour discriminer les items entre eux. En début de test (N=18), P59 a reconnu tous les MA mais aucun MP et très peu de MR : elle n'était donc pas capable de distinguer les items appartenant à la langue écoutée de ceux n'en faisant pas partie. A l'échelle du test (N=54), le *score B* obtenu pour les MA chute en dessous du niveau de chance tandis que ceux se rapportant aux MP et aux MR le dépassent légèrement : cela suggère que P59 a donné davantage de réponses négatives après le bloc 1.

La Figure 10(b) (cf. Annexe A3) représente la *moyenne de la certitude* de P59 dans ses réponses, correctes et incorrectes. Globalement, la certitude était meilleure en début de test (N=18), et ce indépendamment de l'exactitude des réponses et du type d'items. Sur l'ensemble du test (N=54), la participante était moins sûre de ses réponses négatives (cf. MA-, MP+, MR+).

# Discussion

L'objectif principal de ce mémoire était d'apporter de nouvelles données sur les mécanismes mis en œuvre dans la segmentation lexicale de la parole chez les personnes présentant une aphasie. Pour cela, nous nous sommes intéressée au mécanisme d'apprentissage statistique de mots. Nous avons cherché à mieux comprendre ce mécanisme chez les personnes avec aphasie, en étudiant l'influence de leurs connaissances lexicales sur leurs capacités à segmenter une langue artificielle. Autrement dit, la présence d'un mot connu dans une langue artificielle facilite-t-elle la segmentation et l'apprentissage des mots inconnus de cette langue ? Un autre objectif était d'affiner la compréhension du lien entre segmentation et troubles de la compréhension orale de phrases chez les personnes avec aphasie.

Pour mener à bien ce projet, nous avons construit une tâche évaluant la segmentation et basée sur le paradigme d'apprentissage statistique en nous appuyant sur les méthodologies utilisées dans trois études (Buiatti et al., 2009 ; Palmer et al., 2019 ; Peñaloza et al., 2014). L'un des objectifs secondaires de ce mémoire était aussi de valider la tâche créée auprès d'individus tout-venants, en l'occurrence des jeunes adultes : nous souhaitions vérifier et comparer les possibilités d'apprentissage des mots inconnus dans les deux flux de langue artificielle créés.

## 1. Validation du test de segmentation

Les analyses réalisées dans le cadre de la validation de la tâche ont montré que les performances des jeunes adultes étaient significativement meilleures avec l'une des deux langues artificielles construites (langue A). Pour éviter cet effet, nous avons apparié les deux langues créées en terme de probabilités transitionnelles entre les syllabes. Toutefois, nous savions qu'il était fréquent d'observer ce type de différence lorsqu'on utilise plusieurs langues artificielles en raison d'inégalités dans la familiarité entre les syllabes et séquences de syllabes présentes dans les deux langues (Palmer et al., 2019). Le fait que nos deux flux ne contenaient pas les mêmes syllabes constitue une autre explication possible à la différence observée.

D'autres analyses ont démontré que les performances des participants entre les deux langues ne différaient significativement que pour l'un des trois types d'items utilisé dans le test : les mots-partiels étaient significativement plus souvent reconnus à tort comme des mots appartenant à la langue artificielle avec la langue B. Pour rappel, les mots-partiels chevauchaient une frontière de mots dans la langue. Or, nous n'avons pas contrôlé les probabilités d'apparition des séquences trisyllabiques formant les mots-partiels dans les deux flux. Si ces probabilités étaient inférieures dans le flux A, il est possible que les participants aient mieux distingué les mots-partiels des mots artificiels, d'où un nombre moins important de mots-partiels reconnus à tort comme des mots de la langue. Pour limiter cet effet, nous aurions pu construire les mots-partiels de notre étude à partir de nos flux de langue artificielle au lieu d'utiliser ceux créés par Buiatti et al. (2009). Ainsi, nous aurions pu apparié leurs probabilités d'apparition dans les deux langues.

Au final, seule la langue A semblait pouvoir être segmentée et apprise : le pourcentage de réponses correctes obtenu dépassait significativement le niveau de chance. Par ailleurs, les analyses effectuées ont mis en évidence une influence du type d'items sur les performances des jeunes adultes dans les deux langues : davantage de réponses correctes ont été relevées pour les mots artificiels comparativement aux mots-partiels et aux mots-règles. La différence observée entre les

mots artificiels et les mots-partiels était significative. Par conséquent, les participants auraient été capables de reconnaître la majorité des mots artificiels de la langue à laquelle ils ont été exposés tout en rejetant plus aisément la présence des mots-règles dans la langue par rapport aux mots-partiels. Nous nous attendions à observer un tel effet, étant donné la nature des mots-partiels qui, au contraire des mots-règles, apparaissent dans les flux de langue artificielle : ils sont donc plus enclins à être confondus avec des mots de la langue. Palmer et al. (2019) avaient également relevé ce type d'effet avec leurs stimuli mais en utilisant une tâche en choix forcé à deux alternatives dans la phase de test. En revanche, Buiatti et al. (2009) qui utilisaient, comme nous, une tâche de décision lexicale, n'observaient pas de différences entre les trois types d'items. Cependant, le nombre de stimuli dans leur flux de parole artificielle ainsi que le temps d'exposition des participants à la langue étaient plus élevés. Ainsi, les différences observées entre les trois types d'items témoignent de la sensibilité des participants à différents degrés de probabilités transitionnelles entre les syllabes (Palmer et al., 2019). Les jeunes adultes s'appuieraient de façon implicite sur les indices statistiques présents dans les flux afin de les segmenter et d'identifier les mots inconnus les composant.

Enfin, les analyses réalisées n'ont pas montré d'influence (1) du bloc d'items du test administré sur les performances des jeunes adultes, excluant de ce fait l'hypothèse d'un oubli de la langue au cours du test ; ni (2) de l'ordre d'administration des deux langues. La langue administrée en deuxième session aurait effectivement pu être plus facilement segmentée en raison de la connaissance de la tâche par les participants. Ces derniers auraient pu inférer la longueur des mots composant la langue en se basant sur celle des items proposés dans le test en première session.

En résumé, nous avons montré qu'il était possible de segmenter et d'apprendre la langue A et que cela était plus difficile avec la langue B. C'est d'ailleurs pour cette raison que le protocole expérimental utilisé pour les participants avec aphasie était différent selon qu'il s'agissait de tests ou de pré-tests (cf. méthodologie pour plus de détails). De plus, nous avons relevé dans les deux langues une sensibilité à l'information statistique présente dans les flux, se traduisant par davantage de fausses reconnaissances pour les mots-partiels.

## **2. Apprentissage statistique et segmentation en aphasie**

### **2.1. Connaissances lexicales et segmentation d'une parole artificielle**

La présence d'un vrai mot dans la langue artificielle de la condition mot connu a-t-elle aidée les participants avec aphasie à segmenter et apprendre les mots inconnus de cette langue ? En ne tenant compte que du pourcentage total de réussite au test, il est difficile de répondre avec certitude à cette question. Il existe effectivement un biais sur le score total lié à la présentation des résultats sous la forme d'un pourcentage de réponses correctes. Plus concrètement : le pourcentage total de réussite inclut le nombre de réponses correctes obtenues pour chaque type d'items. Or, l'exactitude des réponses binaires (« oui » / « non ») dépend du type d'items. Par exemple, répondre « oui » pour un mot artificiel est correct mais pas pour les non-mots (mots-partiels et mots-règles), et inversement. En outre, le test compte deux fois plus de non-mots que de mots artificiels. De ce fait, si un participant répond majoritairement « oui » dans une condition et « non » dans l'autre, il obtiendrait un pourcentage total de réussite plus élevé dans la deuxième condition, sans pour autant mieux réussir la tâche. Dans les deux cas, ce type de réponses suggère plutôt des difficultés de discrimination des différents types d'items présents dans le test. Par conséquent, il nous a semblé

plus judicieux de comparer les deux conditions en terme de capacités de discrimination entre mots artificiels et non-mots, en nous reportant aux résultats obtenus pour chaque type d'items.

En procédant ainsi, il semblerait qu'un seul participant ait présenté de meilleures performances en présence d'un vrai mot. Il s'agit du participant 23 (P23). En effet, ce dernier n'a reconnu aucun mot artificiel dans la condition de base (il répondait « non » à tous les items). Dans la condition mot connu, il présentait également des difficultés pour discriminer les différents types d'items. Cependant, il est tout de même parvenu à identifier l'un des six mots artificiels de la langue (reconnu trois fois lors du test), au prix d'une unique fausse reconnaissance sur un mot-partiel. Par ailleurs, il est possible que l'effet du vrai mot sur les performances de P23 soit ici sous-estimé, ce participant ayant été exposé à la langue la plus difficile (langue B) dans la condition mot connu.

Ces résultats ne nous permettent pas de vérifier notre hypothèse de départ, postulant qu'en l'absence d'un défaut d'accès au lexique, les personnes avec aphasie peuvent bénéficier de l'aide d'un indice lexical pour segmenter une parole artificielle, tout comme les individus tout-venants de Palmer et al. (2019). Les participants de notre étude présentant tous une atteinte lexicale, nous ne nous attendions pas à ce que l'un d'entre eux puisse bénéficier de cet indice, d'autant qu'il ne s'agissait pas du participant ayant l'atteinte la plus légère. En effet, deux des participants ne présentaient qu'un léger manque du mot, et pourtant, leur performance était moins bonne dans la condition mot connu (P99, P66). Dans l'ensemble, l'étude des profils langagiers ne nous a pas permis de mettre en évidence de liens entre un déficit particulier et l'absence de bénéfice évident d'un indice lexical. Par exemple, nous avons pu remarquer que seuls les participants présentant une compréhension lexicale déficitaire n'avaient pas identifié le vrai mot dans la condition mot connu (P46, P51). Cela aurait pu justifier l'absence d'amélioration de leur performance dans cette condition. Seulement, les participants à la compréhension orale de mots préservée n'ont pas tous bénéficié de l'indice lexical (P99, P66, P85). Cette hypothèse n'est donc pas recevable.

L'étude des profils langagiers ne nous permettant pas d'expliquer nos résultats, nous nous sommes intéressée de plus près aux capacités de segmentation des participants dans les deux conditions. Ainsi, seuls P99 et P66 semblaient en capacité de segmenter les langues artificielles dans les deux conditions. Le participant 85 aurait également présenté cette capacité mais uniquement en début de test dans la condition de base. Les scores obtenus par les autres participants (dont P23) indiquaient clairement des difficultés de discrimination des différents types d'items, quelle que soit la condition administrée, et bien que le pourcentage total de réussite soit parfois significativement meilleur que le hasard (cf. biais sur le score total). Sur la base de ces résultats, nous ne pouvons pas établir de liens entre la capacité à utiliser les indices statistiques de segmentation et la présence d'un indice lexical dans le flux de parole artificielle. Selon Palmer et al. (2019), il semblerait effectivement que des individus capables de segmenter une langue artificielle sur la base d'indices statistiques puissent améliorer leur performance en présence d'un indice lexical. Dès lors, il est difficile d'expliquer pourquoi P99 et P66, qui en étaient également capables, n'ont pas bénéficié de cet indice ; et pourquoi P23, qui n'en était pas capable, a pu en bénéficier.

En résumé, l'analyse des capacités langagières et de segmentation des participants ne nous a pas permis de comprendre pourquoi seul P23 semblait pouvoir bénéficier de la présence du vrai mot dans la condition mot connu. C'est pourquoi nous avons formulé d'autres hypothèses explicatives de nos résultats.

Tout d'abord, les divergences observées entre nos participants d'une part, et avec ceux de Palmer et al. (2019) d'autre part, pourraient être le fruit de différences dans les méthodologies employées. Nous tenons effectivement à rappeler que deux protocoles ont été utilisés dans ce mémoire (pré-tests vs. tests), ce qui limite non seulement les comparaisons entre participants, mais aussi entre conditions expérimentales chez un même participant. Par exemple, il est difficile de conclure à un effet de l'indice lexical chez P23, les langues auxquelles il a été exposé dans chacune des conditions étant différentes. En outre, il est possible que les différences observées entre ce participant et les participants 99 et 66 quant à la présence de cet indice, soient uniquement liées à la langue dans laquelle il a été inséré (respectivement, langues B et A). Enfin, nous pouvons dire que les différences méthodologiques observées entre notre étude et celle de Palmer et al. (2019) se rapportent non seulement aux stimuli utilisés, mais aussi à la tâche employée dans la phase de test (tâche de décision lexicale vs. tâche en choix forcée à deux alternatives). Ces deux tâches ne mobilisent pas les mêmes ressources cognitives (Palmer et al., 2018) et n'évaluent pas les capacités de segmentation avec la même sensibilité. Dans la tâche en choix forcé, par exemple, il est difficile de savoir si le participant sélectionne l'item correct parce qu'il l'a reconnu et identifié comme un mot ; ou parce qu'il rejette la présence du mot-partiel qui lui est associé (Buiatti et al., 2009). Ainsi, il serait intéressant de répliquer les résultats de Palmer et al. (2019) avec notre méthodologie et auprès d'individus tout-venants. L'objectif serait de vérifier si l'ajout du vrai mot *médicament* dans nos langues artificielles peut en faciliter la segmentation.

En faisant abstraction de ces différences méthodologiques et étant donné le rejet de notre hypothèse initiale, nous pouvons formuler une nouvelle hypothèse. Un déficit d'accès au lexique ne permettant pas d'expliquer à lui seul l'absence d'amélioration des performances en présence d'un indice lexical, nous avons imaginé l'existence d'un mécanisme cognitif particulier. Ce dernier serait potentiellement atteint chez P99 et P66, mais pas chez P23, ni chez les individus tout-venants de Palmer et al. (2019). Il aurait pour fonction de coordonner le recrutement des différents indices de segmentation qui, comme l'ont montré Mattys et al. (2005), ne fonctionnent pas sur le principe du tout ou rien. Une atteinte de ce mécanisme se traduirait par l'incapacité à bénéficier des effets synergiques ou de primauté de certains indices, mis par exemple en évidence par Bagou & Frauenfelder (2018). Cela expliquerait pourquoi P99 et P66 ne bénéficient pas de la synergie des indices lexicaux et statistiques alors qu'elle apparaîtrait a priori chez des individus tout-venants. En effet, Palmer et al. (2019) font l'hypothèse que les indices lexicaux peuvent améliorer la sensibilité aux indices statistiques. Chez P23, ce mécanisme serait fonctionnel et lui permettrait de coordonner les deux types d'indices ; néanmoins, leur effet synergique serait moindre du fait d'un accès déficitaire à l'un d'eux (a priori aux indices statistiques). La mise en évidence d'un tel mécanisme nécessite toutefois davantage de preuves. L'étude des interactions entre d'autres indices de segmentation chez des individus cérébrolésés est une première piste de recherche (cf. étude menée chez des individus sains par Bagou et Frauenfelder, 2018).

## **2.2. Apprentissage statistique, segmentation et compréhension orale**

Comme dit précédemment, la majorité des participants avec aphasie inclus dans ce mémoire présentaient des difficultés pour segmenter une langue artificielle et identifier les mots la composant. Ces observations viennent corroborer celles d'études antérieures s'intéressant également à l'apprentissage statistique chez les personnes cérébrolésées (Peñaloza et al., 2014 ; Shaqiri et al., 2018). L'hypothèse que nous avons formulée à ce sujet est donc validée : les personnes avec une cérébrolésion conduisant à une aphasie peuvent présenter une atteinte de l'apprentissage statistique.

Les deux participants réussissant le mieux le test de segmentation avaient globalement un profil similaire (P99 et P66) : ils présentaient tous deux une aphasie fluente de type anomique et de sévérité légère, se traduisant par un léger manque du mot. Les troubles étaient davantage ressentis par la personne elle-même que perceptibles par l'auditeur. Contrairement aux observations faites par Dupont et Inghilleri (2018) dans leurs mémoires et par Peñaloza et al. (2014), il semblerait que la sévérité de l'aphasie soit corrélée aux capacités de segmentation : les participants réussissant le mieux le test présentaient une aphasie légère ; ceux en difficulté, une aphasie modérée à sévère. Seul un participant présentant des troubles phasiques légers n'a pas réussi à segmenter la parole artificielle (P39) ; en notant tout de même que celle-ci était plus difficile à apprendre (langue B).

En nous intéressant de plus près aux capacités langagières des participants, nous pouvons relever certaines corrélations entre leurs performances au test et leur niveau de compréhension orale. Par exemple, ceux réussissant le mieux le test étaient aussi les seuls à ne pas présenter de difficultés de compréhension orale, que ce soit au niveau lexical ou syntaxique (P99, P66).

Tous les participants à la compréhension syntaxique déficitaire étaient en difficulté pour segmenter les langues artificielles et davantage lorsqu'ils présentaient des troubles sévères. D'où le fait que P85, dont le trouble de compréhension syntaxique était modéré, présentait moins de difficultés que les autres participants pour segmenter la parole. Ces résultats soutiennent l'idée qu'un défaut de segmentation puisse être à l'origine d'un trouble de la compréhension orale des énoncés, comme l'avait montré Dupont (2018) dans son mémoire avec une tâche de « word monitoring ». Toutefois, la corrélation inverse s'observe dans notre étude : des difficultés de compréhension syntaxique pourraient causer des difficultés de segmentation. Nos résultats suggèrent donc l'existence d'un lien entre difficultés de segmentation et trouble de la compréhension orale syntaxique. Cependant, il est difficile de savoir si les résultats déficitaires obtenus au test sont liés à (1) un déficit du processus de segmentation de la parole, qui précéderait donc les processus de compréhension syntaxique ou (2) à un déficit de la compréhension orale des énoncés. Cette dernière hypothèse, qui semble moins probable, nous permet de souligner l'une des limites de ce mémoire : la compréhension de la consigne par les participants. Il est effectivement possible que ceux ne réussissant pas la tâche n'aient tout simplement pas compris les instructions données. Shaqiri et al. (2018) soulèvent également ce point dans leur étude. Pour limiter ce biais, il serait bienvenue d'ajouter un entraînement à la tâche lors des passations, comme l'ont fait Peñaloza et al. (2014).

La nature du lien entre compréhension orale lexicale et segmentation est plus difficile à définir. Les participants présentant une bonne compréhension lexicale pouvaient réussir le test (P99, P66, P85) ou non (P23), tandis que ceux présentant un déficit rencontraient systématiquement des difficultés. Tout d'abord, il est difficile de postuler qu'un défaut de segmentation puisse être à l'origine d'un défaut de compréhension lexicale ; le cas de P23 démontrant qu'il n'est pas nécessaire de présenter de bonnes capacités de segmentation pour comprendre les mots. Nous tenons plutôt des arguments pour affirmer l'inverse : un déficit de la compréhension lexicale pourrait occasionner un trouble de la segmentation, soutenant ainsi les modèles selon lesquels la segmentation de la parole est le résultat de l'accès au lexique (ex. Bagou & Frauenfelder, 2002 ; Crouzet, 2000 ; Davis et al., 2002). Cependant, le paradigme d'apprentissage statistique utilisé ici ne nous permet pas d'expliquer les résultats obtenus en regard de cette hypothèse. En effet, il s'agissait de segmenter une parole artificielle ; l'accès au lexique mental n'était donc pas nécessaire (du moins dans la condition de base). Ainsi, une compréhension lexicale déficitaire ne permet pas d'expliquer l'échec de certains

participants au test de segmentation, bien que cela pourrait justifier l'absence d'amélioration des scores en présence d'un mot connu.

Dans leur étude, Peñaloza et al. (2014) n'ont pas analysé le lien entre compréhension et segmentation. Nous pouvons cependant noter que peu de difficultés de compréhension avaient été relevées chez leurs participants, et que certains étaient en difficulté pour segmenter une parole artificielle malgré une compréhension relativement préservée. Quant à Shaqiri et al. (2018), ils n'ont pas mis en évidence de corrélations entre la performance des participants lors de la tâche d'apprentissage statistique et leurs capacités de compréhension orale. Au final, l'ensemble de ces observations souligne le fait que d'autres mécanismes sous-jacents à la segmentation et plus particulièrement à l'apprentissage statistique puissent être atteints.

Les capacités de discrimination phonologique par exemple, pourraient être considérées comme un pré-requis pour segmenter la parole en mots, étant donné qu'il s'agit d'une des premières étapes au traitement de la parole avant d'arriver à la compréhension orale. Dans notre étude, tous les participants échouant au test de segmentation présentaient des difficultés de discrimination phonologique, suggérant ainsi qu'une atteinte de ce pré-requis pourrait être à l'origine d'un déficit d'apprentissage statistique. Toutefois, les participants 99 et 66 réussissent tous deux le test malgré des difficultés de discrimination auditive. Notons néanmoins que leurs difficultés de discrimination phonologique semblaient plutôt liées à une atteinte auditive et non cérébrale. L'hypothèse d'un lien entre les mécanismes de perception phonologique de la parole et la segmentation ne peut donc être rejetée. L'interprétation de ce lien nécessite toutefois de la prudence, les erreurs commises lors d'une tâche de discrimination auditive pouvant être multifactorielles, d'origine exécutives, mnésiques ou attentionnelles par exemple.

Ces mêmes facteurs sont aussi à prendre en compte dans l'interprétation des résultats au test de segmentation : ce dernier étant inhabituel, il aurait pu perturber certains participants. De plus, la tâche demandait beaucoup de concentration et de ressources sur les plans cognitif et mnésique. Rappelons que Peñaloza et al. (2014) avaient mis en évidence que des ressources limitées en mémoire de travail pouvaient réduire l'efficacité de l'apprentissage statistique. Dans notre étude, nous avons fait le choix de ne pas évaluer les mémoires à court-terme et de travail auditivo-verbales des participants, en gardant néanmoins à l'esprit leur atteinte fréquente chez les personnes avec aphasie. A posteriori, nous pouvons dire qu'il aurait été difficile de les évaluer chez l'ensemble de nos participants en raison de la sévérité de l'atteinte des fonctions langagières chez certains. Il est donc probable que les difficultés de segmentation observées chez certains participants soient sous-tendues par des difficultés mnésiques.

Par ailleurs, nos résultats semblent avoir mis en évidence une possible diminution des performances au cours du test de segmentation chez certains participants ; laquelle pourrait être liée à des difficultés attentionnelles mais aussi mnésiques. La tâche de décision lexicale utilisée incluant un grand nombre d'items, il existe effectivement un potentiel effet d'oubli des mots de la langue écoutée au fur et à mesure du test. Même si nous n'avons pas observé de tel effet chez les jeunes adultes, nous l'avons tout de même pris en compte dans l'analyse des résultats des participants avec aphasie en choisissant de présenter d'une part, les résultats obtenus sur l'ensemble du test, et d'autre part, ceux obtenus uniquement dans le bloc 1.

L'implication de la mémoire dans ce type de tâche constitue une autre limite de ce mémoire. C'est pourquoi il serait intéressant de réaliser des études EEG auprès de personnes avec aphasie afin d'étudier « on-line » les mécanismes neuronaux sous-jacents à la segmentation (Buiatti et al., 2009).

Enfin, ce mémoire ne nous a pas permis de mettre en évidence de liens entre la localisation des lésions et les déficits observés. Peñaloza et al. (2014) avaient montré que des lésions plutôt antérieures semblaient être prédictives d'un échec lors d'une tâche d'apprentissage statistique en comparaison avec des lésions du territoire postérieur. Les participants de notre étude présentant tous des lésions de topographie antérieure, il n'a pas été possible de relever de corrélations.

### **3. Segmentation et orthophonie**

Comme nous l'avons dit précédemment, le processus de segmentation est très peu étudié dans le cadre de l'aphasie. Or, comme le suggèrent nos résultats conjoints à ceux de Peñaloza et al. (2014) et Shaqiri et al. (2018), il semblerait que certains mécanismes sous-jacents à la segmentation puissent être atteints chez les personnes présentant une aphasie. C'est le cas du mécanisme d'apprentissage statistique. Comme le suggèrent Shaqiri et al. (2018), ces résultats sont particulièrement intéressants dans le cadre de la prise en charge post-AVC. En effet, de nombreux aspects de la rééducation reposent sur un apprentissage (ou ré-apprentissage) implicite de compétences. Or, nous venons ici de démontrer qu'un processus d'apprentissage implicite pouvait être atteint après une lésion cérébrale. L'efficacité des techniques de rééducation sous-tendues par un tel processus pourraient donc potentiellement être impactée. Ainsi, il serait pertinent de pouvoir évaluer les capacités d'apprentissage statistique des patients cérébrolésés afin de mettre en place des stratégies de prise en soin plus adaptées à leurs possibilités d'apprentissage.

Par ailleurs, il semblerait que le mécanisme d'apprentissage statistique présente des liens avec certaines capacités langagières pouvant être atteintes dans le cadre d'une aphasie, telles que la compréhension orale. De ce fait, il semble évident qu'une meilleure compréhension de son fonctionnement permettrait d'améliorer la pratique orthophonique à long terme.

De façon générale, il est important que toute intervention orthophonique soit spécifique aux difficultés présentées par le patient, afin d'optimiser la récupération des fonctions cognitives parfois sévèrement atteintes. Cela suggère la réalisation d'évaluations précises, permettant non seulement de décrire les déficits, mais également de rechercher les processus sous-jacents lésés. Toutefois, les outils d'évaluation dont disposent les orthophonistes n'aident pas toujours à localiser l'origine fonctionnelle probable du déficit. C'est par exemple le cas de la majorité des outils évaluant la compréhension orale de phrases ; un processus complexe impliquant différents niveaux de traitements, eux-mêmes sous-tendus par d'autres fonctions cognitives telles que l'attention ou la mémoire. Ainsi, l'épreuve de compréhension de phrases du MT-86 (Nespoulous et al., 1992), fréquemment utilisée en clinique, ne permet d'évaluer que la dernière étape du processus, sans localiser les mécanismes sous-jacents atteints. Ces dernières années, de nouveaux outils ont été créés dans le but de combler ce manque. C'est le cas de la Batterie d'évaluation de la compréhension syntaxique (BCS), développée par Caron, Le May, Bergeron, Bourgeois et Fossard (2015).

Sur le même principe, il serait intéressant d'imaginer un outil orthophonique évaluant spécifiquement la segmentation. La mise en évidence d'un trouble permettrait alors d'expliquer certaines difficultés de compréhension orale chez les patients ; une rééducation ciblée du processus atteint pourrait être proposée. Cependant, nous ne disposons pas à ce jour de suffisamment de données pour créer un tel outil. De nombreuses questions subsistent encore quant aux liens que la segmentation entretient avec les autres fonctions cognitives. Or, une meilleure compréhension de ces liens permettrait de l'évaluer plus précisément. Pour cela, il faudrait mener davantage d'études auprès de personnes cerébrolésées, atteintes d'aphasie ou d'autres troubles cognitifs.

## Conclusion

Ce mémoire avait pour objet l'étude du processus de segmentation de la parole chez les personnes avec aphasie, via le mécanisme d'apprentissage statistique. Il s'agit d'un processus peu étudié auprès de cette population, alors qu'il pourrait contribuer aux troubles de compréhension orale présents dans les différents tableaux d'aphasie. Ainsi, la recherche menée visait à mieux comprendre les mécanismes sous-jacents impliqués, en évaluant l'influence des connaissances lexicales des personnes avec aphasie sur leurs capacités à segmenter une langue artificielle. Nous avons supposé qu'en l'absence d'un défaut d'accès au lexique, les personnes avec aphasie pouvaient bénéficier, comme les individus tout-venants, de l'aide d'un indice lexical pour segmenter une parole artificielle. Nous souhaitons aussi affiner la compréhension du lien entre segmentation et troubles de la compréhension orale chez ces personnes.

Pour mener à bien ce projet, nous avons construit une tâche évaluant la segmentation et basée sur le paradigme d'apprentissage statistique. Les participants entendaient d'abord un flux de langue artificielle pendant environ trois minutes. Puis, ils participaient à un test de décision lexicale évaluant leur capacité à identifier les mots de la langue artificielle. Une partie de l'étude a consisté à valider les stimuli créés auprès de jeunes adultes ; tandis qu'une autre organisait la passation du test de segmentation auprès de participants avec aphasie dans deux conditions. L'une était la condition de base et l'autre présentait une langue artificielle contenant un vrai mot français. Nous avons ainsi comparé pour chaque participant les résultats obtenus à cette tâche dans les deux conditions, tout en faisant le lien avec leur profil langagier.

Les résultats obtenus ne nous ont pas permis de valider notre hypothèse principale : un seul participant a obtenu de meilleures performances en présence du vrai mot ; et pourtant, il présentait comme tous les autres, une atteinte lexicale. Nous avons alors confronté les profils langagiers des participants et leurs résultats au test de segmentation pour expliquer ce résultat. Ainsi, il semblerait qu'un déficit d'accès lexical ne soit pas suffisant pour empêcher l'accès aux indices lexicaux de segmentation. Cette observation nous a mené à formuler une hypothèse en faveur d'un mécanisme cognitif particulier, qui aurait pour fonction de coordonner le recrutement des différents indices de segmentation. Cependant, plusieurs biais méthodologiques ne nous permettent ni de la valider, ni d'infirmar celle de départ.

Les résultats obtenus nous ont en revanche permis de valider notre hypothèse secondaire : certaines personnes avec aphasie peuvent présenter des difficultés de segmentation dans ce type de tâche. En confrontant les scores du test de segmentation au profil langagier des participants, nous avons pu relever des liens entre segmentation et compréhension orale d'énoncés. Toutefois, les participants présentaient des profils hétérogènes limitant nos conclusions. Aussi, nous avons soulevé le fait que la segmentation impliquait divers mécanismes ; et qu'il était difficile de les étudier indépendamment d'autres fonctions cognitives, langagières ou mnésiques par exemple.

Notre étude apporte donc de nouvelles données sur la segmentation et plus particulièrement sur les processus qui sous-tendent l'utilisation des différents indices de segmentation. Cependant, cela mériterait davantage d'investigations. Il serait notamment intéressant de poursuivre l'étude tout en contrôlant les différents biais méthodologiques mis en évidence. Il serait aussi bienvenue d'étudier d'autres combinaisons d'indices auprès des personnes avec aphasie. La réalisation d'études EEG pourraient être envisagées afin d'étudier ce processus sous un autre angle. À long terme, l'objectif serait de pouvoir élaborer un outil d'évaluation orthophonique testant la segmentation de la parole, et ce afin de proposer une rééducation plus ciblée.

## Bibliographie

- Aphasie. (2011). Dans Brin-Henri, F., Courrier, C., Lederlé, E., & Masy, V. (dir.), *Dictionnaire d'Orthophonie*. Isbergues, France : OrthoEdition.
- Bagou, O. & Frauenfelder, U.H. (2002). Alignement lexical et segmentation de la parole. *Revue française de linguistique appliquée*, 7(1), 67-82.
- Bagou, O., & Frauenfelder, U. H. (2018). Lexical Segmentation in Artificial Word Learning: The Effects of Converging Sublexical Cues. *Language and Speech*, 61(1), 3-30.
- Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, 44(2), 509-519.
- Caplan, D., Michaud, J., & Hufford, R. (2015). Mechanisms underlying syntactic comprehension deficits in vascular aphasia: new evidence from self-paced listening. *Cognitive Neuropsychology*, 32(5), 283-313.
- Caron, S., Le May, M.-E., Bergeron, A., Bourgeois, M.-E. & Fossard, M. (2015). Batterie d'évaluation de la compréhension syntaxique. Québec, Canada : Centre intégré universitaire de santé et services sociaux de la Capitale-Nationale – Institut de réadaptation en déficience physique de Québec.
- Christensen, S.C., Wright, H.H., Ratiu, I. (2018). Working memory in aphasia : peeling the onion. *Journal of Neurolinguistics*, 48, 117-132.
- Crouzet, O. (2000). *Segmentation de la parole en mots et régularités phonotactiques: Effets phonologiques, probabilistes ou lexicaux ?* (Thèse en psychologie, Université Paris 5-René Descartes, Paris, France).
- Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, 48(3), 541-553.
- Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading Up the Lexical Garden Path: Segmentation and Ambiguity in Spoken Word Recognition. *Journal of Experimental Psychology Human Perception & Performance*, 28(1), 218-244.
- De Partz, M.P., & Pillon, A. (2014). Sémiologie, syndromes aphasiques et examen clinique des aphasies. Dans X. Seron, & M. Van Der Linden (dir.), *Traité de neuropsychologie clinique de l'adulte. Tome 1-Evaluation* (2<sup>e</sup> ed., p. 249-265). Paris, France : De Boeck-Solal.
- Dupont, M. (2018). *Segmentation lexicale de la parole en aphasie et capacités langagières : quels liens ?* (mémoire en vue de l'obtention du Certificat de Capacité d'Orthophonie, Département d'orthophonie de Lille, Lille, France).

- Ferrill, M., Love, T., Walenski, M. & Shapiro, L. P. (2012). The time-course of lexical activation during sentence comprehension in people with aphasia. *American Journal of Speech-Language Pathology*, 21(2), 179-189.
- Goodglass, H. & Kaplan, E. (1972). Boston Diagnostic Aphasia Examination. France : ECPA.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, 8(5), 393-402.
- Inguilleri, M. (2018). *Segmentation lexicale de la parole en aphasie et capacités mnésiques: quels liens ?* (mémoire en vue de l'obtention du Certificat de Capacité d'Orthophonie, Département d'Orthophonie de Lille, Lille, France).
- Ivanova, M. V., Dragoy, O. V., Kuptsova, S. V., Ulicheva, A. S. & Laurinavichyute, A. K. (2015). The contribution of working memory to language comprehension: differential effect of aphasia type. *Aphasiology*, 29(6), 645-664.
- JASP Team. (2018). *JASP (Version 0.11.1) [computer software]*. Consulté à l'adresse <https://jasp-stats.org/>
- Karuza, E.A., Newport, E.L., Aslin, R.N., Starling, S.J., Tivarus, M.E., Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, 127(1), 46–54.
- López-Barroso, D., Ripollés, P., Marco-Pallarés, J., Mohammadi, B., Münte, T.F., Bachoud-Lévi, A.C., ..., Diego-Balaguer, R. (2015). Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *NeuroImage*, 110, 182-193.
- Macoir, J., Gauthier, C., Jean, C., & Potvin, O. (2016). BECLA, a new assessment battery for acquired deficits of language : Normative data from Quebec-French healthy younger and older adults. *Journal of the Neurological Sciences*, 361, 220-228.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71-102.
- Marslen-Wilson, W. D. & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive psychology*, 10(1), 29–63.
- Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of Multiple Speech Segmentation Cues: A Hierarchical Framework. *Journal of Experimental Psychology: General*, 134(4), 477-500.
- McClelland, J. L. & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
- McNealy, K., Mazziotta, J. C. & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(29), 7629-7639.

- Mersad, K., & Nazzi, T. (2012). When Mommy comes to the rescue of statistics: Infants combine top-down and bottom-up cues to segment speech. *Language learning and Development*, 8(3), 303-315.
- Nazzi, T. (2008). Segmentation précoce de la parole continue en mots : évaluation inter-linguistique de l'hypothèse d'initialisation rythmique. *L'Année psychologique*, 108(02), 309-342.
- Nespoulous, J. L., Lecours, A. R., Lafond, D., Lemay, A., Puel, M., Joannette, Y., ... & Rascol, A. (1992). Protocole Montréal-Toulouse d'examen linguistique de l'aphasie (MT86). France : Ortho Edition.
- Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.
- Palmer, S.D., Hutson, J., Mattys, S.L. (2018). Statistical Learning for Speech Segmentation: Age-Related Changes and Underlying Mechanisms. *Psychology and aging*, 33(7), 1035-1044.
- Palmer, S. D., Hutson, J., White, L., & Mattys, S. L. (2019). Lexical knowledge boosts statistically-driven speech segmentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(1), 139-146.
- Patil, U., Hanne, S., Burchert, F., De Bleser, R. & Vasishth, S. (2015). A Computational Evaluation of Sentence Processing Deficits in Aphasia. *Cognitive Science*, 40(1), 5-50.
- Peñalosa, C., Benetello, A., Tuomiranta, L., Heikius, I., Järvinen, S., Majos, M. C., ... Rodriguez-Fornells, A. (2014). Speech segmentation in aphasia. *Aphasiology*, 29, 724-743.
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language*, 35(4), 606-621.
- Salis, C., Murray, L., & Bakas, K. (2018). An international survey of assessment practices for short-term and working memory deficits in aphasia. *American Journal of Speech-Language Pathology*, 27(2), 574-591.
- Shaqiri, A., Danckert, J., Burnett, L., & Anderson, B. (2018). Statistical learning impairments as a consequence of stroke. *Frontiers in human neuroscience*, 12, 339.
- Van der Kaa, M. A., & Majerus, S. (2007). Evaluation et rééducation du versant « phonologique » de la mémoire à court terme. Dans G. Aubin, F. Coyette, P. Pradat-Diehl, & C. Vallat-Azouvi (Eds.), *Neuropsychologie de la mémoire de travail*, 87-111. Marseille : Solal.
- Wauquier-Gravelines, S. (1999). Segmentation lexicale de la parole continue : la linéarité en question. *Recherches linguistiques de Vincennes*, (28), 133-156.

## **Liste des annexes**

**Annexe n°1 : Stimuli du test de segmentation.**

**Annexe n°2 : Résumé des analyses statistiques effectuées dans le cadre de la validation du test de segmentation.**

**Annexe n°3 : Résultats des participants à l'étude soumis à une seule condition expérimentale.**