



HAL
open science

Le rôle de l'ordre temporel des segments dans la reconnaissance des mots parlés

Élodie Bauduin, Étienne Personnaz

► **To cite this version:**

Élodie Bauduin, Étienne Personnaz. Le rôle de l'ordre temporel des segments dans la reconnaissance des mots parlés. Sciences du Vivant [q-bio]. 2020. dumas-02992046

HAL Id: dumas-02992046

<https://dumas.ccsd.cnrs.fr/dumas-02992046>

Submitted on 6 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mémoire de fin d'étude
en vue de l'obtention du certificat de capacité d'orthophonie



Le rôle de l'ordre temporel des segments dans la reconnaissance des mots parlés

Réalisé par Elodie Bauduin et Etienne Personnaz
sous la direction de Sophie Dufour

LABOATOIRE PAROLE ET LANGAGE



UMR 7309 • CNRS • AMU
**PAROLE ET
LANGAGE**

Table des matières

I)	INTRODUCTION.....	2
II)	CONTEXTE THEORIQUE	5
1.	Champ de la recherche en reconnaissance des mots parlés : modèles abstractionnistes et techniques comportementales	5
-	Modèles abstractionnistes	5
-	Techniques comportementales	6
2.	Théorie générale de la reconnaissance des mots parlés.....	7
-	Caractéristiques principales de la parole et conséquences pour son traitement cognitif	7
-	Postulats communs des modèles classiques.....	9
3.	Les différents mécanismes d'appariement et de sélection lexicale envisagés par les modèles classiques	11
-	Hypothèses lexicales considérées à l'écoute d'un mot.....	11
-	Régulation des hypothèses lexicales et sélection du meilleur candidat	13
-	Rôle des connaissances lexicales	14
4.	Traitement infra-lexical et ordre temporel des phonèmes.....	16
-	Enjeu de l'invariance temporelle.....	16
-	Codage de l'ordre temporel des phonèmes dans les modèles classiques	17
5.	Vers une conception nouvelle du traitement de la position des phonèmes ?.....	19
-	Apports du domaine de la reconnaissance des mots écrits.....	19
-	Un phonème mal placé empêche-t-il l'activation de la représentation du mot parlé ?	21
III)	HYPOTHESE DE RECHERCHE	23
IV)	PROTOCOLE EXPERIMENTAL.....	24
1.	Matériels et méthodes	24
2.	Procédure.....	26
3.	Résultats.....	26
V)	DISCUSSION	29
VI)	ANNEXES.....	34
VII)	BIBLIOGRAPHIE	39

I) INTRODUCTION

La phonologie étudie l'organisation des sons d'une langue et plus précisément des phonèmes, qui constituent la plus petite unité linguistique permettant de différencier les mots. Ainsi, l'ajout ou la suppression de phonèmes ([abɔʁ] - [bɔʁ] - [ɔʁ]), leur commutation ([bɔʁ] - [sɔʁ]) et leur permutation ([bɔʁ] - [ʁɔb]) peuvent former des mots différents (abord, bord, sort, robe). Si le phonème, par sa valeur distinctive, est un élément capital du langage, l'homme ne l'appréhende généralement qu'à l'âge de 5-6 ans. L'enfant commence alors à repérer des sons de différentes tailles et apprend à les manipuler : il exerce sa conscience phonologique. Cette compétence est essentielle dans notre société où l'accès à la lecture et à l'écriture passe par la maîtrise d'un code de conversion grapho-phonémique. Mais les faibles performances en manipulation de phonèmes des personnes illettrées ou maîtrisant un système d'écriture idéographique suggèrent qu'elle n'émerge pas spontanément (Morais et al., 1979), (Read et al., 1986).

S'il est possible de comprendre la parole sans avoir conscience de la phonologie, cette dernière est pourtant traitée de manière automatique par les individus. Des habiletés dans ce domaine, comme la discrimination entre phonèmes proches, sont même observées chez des nourrissons (Kuhl, 2000). Ce traitement précoce répond principalement à une quête de sens. La perception de contrastes sonores permet aux enfants d'isoler les unités signifiantes de leur langue, les mots, dans le flux continu de la parole. Ils lient progressivement des traits sémantiques à des formes auditives isolées et stockent ces associations signifiant - signifié en mémoire. Le stock phonologique est essentiel à la compréhension du mot, qui dépend de l'accès à sa forme sonore avant sa signification. En effet, des mesures électrophysiologiques permettent d'estimer que le traitement phonologique débute environ 200 ms avant le traitement sémantique (Connolly and Phillips, 1994).

Ce processus inconscient de sélection d'une forme mentale abstraite à partir du signal physique de parole est un domaine d'étude de la psycholinguistique. Cette opération, effectuée en quelques centaines de millisecondes, débute avant que le mot n'ait été entendu en entier. De ce fait, la plupart des psycholinguistes considère que la comparaison du signal acoustique aux représentations mentales se fait par le biais d'unités plus petites qu'elles. La discrimination phonémique est donc également impliquée dans la constitution d'un code linguistique suffisamment abstrait et stable pour permettre d'accéder rapidement aux représentations sonores des mots en de multiples conditions d'écoute.

Cependant, si les compétences précocement mesurables chez l'humain révèlent des habiletés phonétiques, il est moins certain qu'elles témoignent d'un véritable traitement phonologique. En effet, ce dernier suppose non seulement une sensibilité à des différences acoustiques mais également le repérage de celles qui sont pertinentes dans la langue. Seule l'appréhension des sons en tant que parties d'un système

cohérent semble pouvoir expliquer cela. La compétence phonologique repose donc sur des capacités perceptives mais aussi probablement sur des compétences plus élaborées.

Le traitement temporel est sans doute l'une d'entre elles et les orthophonistes le considèrent raisonnablement comme un prérequis à la rééducation d'un trouble du développement phonologique. D'après certains auteurs (Maillart et al., 2005), le premier niveau de traitement phonologique correspond à l'élaboration d'un code à partir d'informations perceptives. Le deuxième se rapporte au lexique phonologique, comprenant les exemplaires des signifiants de la langue. Le traitement temporel intervient à tous ces niveaux. Le code phonologique ne peut être conçu sans prise en compte de la distribution temporelle de l'information acoustique, qui permet le repérage de co-occurrences phonétiques et phonémiques fréquentes. Les premières sont indispensables à l'émergence des catégories phonémiques et les secondes à celle des règles phonotactiques, c'est-à-dire les combinaisons de phonèmes autorisées dans la langue. Quant à l'accès au lexique, il dépend de l'intégrité du premier niveau mais également du traitement de la position des phonèmes lors de l'écoute d'un mot.

La reconnaissance d'un mot est tributaire de l'association correcte entre des portions du signal et des catégories de phonèmes. Une erreur de catégorisation peut être à l'origine d'une confusion entre « bord » et « sort » par exemple. Mais ce processus implique également une analyse de la séquence phonémique, sans quoi nous confondrions les mots « robe » et « bord ». Parvenir à distinguer ces mots paraît élémentaire au premier abord car l'ordre des phonèmes est inhérent au signal de parole, qui s'écoule toujours dans la même direction. Mais le système de reconnaissance des mots parvient à convertir le signal en un langage abstrait tout en conservant des informations relatives à l'enchaînement de ses segments. La nature et la quantité des données temporelles codées, tout comme les niveaux de traitement impliqués dépendent de mécanismes complexes difficilement mesurables qui n'ont sans doute pas fini d'éveiller la curiosité des psycholinguistes.

Quel est le rôle de l'ordre temporel des segments dans la reconnaissance des mots parlés ?

Les psycholinguistes recueillent des données sur le traitement cognitif des mots oraux à l'aide de techniques comportementales et essaient de les organiser au sein d'un modèle cohérent. Avant d'opter pour la bonne entrée lexicale, le cerveau envisagerait plusieurs hypothèses sur la base d'un code infra-lexical ajusté aux exigences du signal. Les mécanismes d'appariement et de sélection standards suggèrent que le jeu de candidats, issu de différents points de l'input, est trié en fonction de l'actualisation des données perceptives et de l'influence de connaissances lexicales intégrées. Tous les modèles classiques s'accordent sur le principe d'un appariement séquentiel en totale conformité avec l'ordre du signal mais aucun n'explique de façon précise et réaliste comment notre cerveau parvient à cette invariance temporelle. Des chercheurs (Toscano et al., 2013), inspirés d'expérimentations menées à l'écrit, ont récemment

mesuré une activation de candidats différant du mot entendu par l'ordre de leurs phonèmes, suggérant un codage de la position des segments plus grossier que ce qui est actuellement envisagé par les modèles de référence.

Dans la lignée d'un précédent mémoire d'orthophonie, notre travail vise à collecter des preuves en français de l'effet découvert par l'équipe de Toscano et à montrer, comme elle, qu'il ne témoigne pas d'un simple effet pré-lexical, mais bien de la considération d'hypothèses non conformes à l'ordre du mot traité. Avec un paradigme d'amorçage phonologique, nous avons mesuré l'activation de mots aux phonèmes permutés tout en appréciant l'influence de la fréquence relative entre amorces et cibles. Les données recueillies avec ce protocole confirment notre hypothèse de recherche : l'effet d'amorçage est uniquement détectable lorsque l'amorce est moins fréquente que la cible, ce qui lui confère un locus lexical. Cette interprétation des résultats encourage la formulation de nouvelles hypothèses concernant le codage de l'ordre temporel des segments dans la reconnaissance des mots parlés.

II) CONTEXTE THEORIQUE

1. Champ de la recherche en reconnaissance des mots parlés : modèles abstractionnistes et techniques comportementales

Pour comprendre l'influence de l'ordre temporel des segments dans la reconnaissance des mots parlés, nous disposons, à la fois, de modèles qui tentent de simuler l'appariement entre le signal de parole et les représentations mentales, et de données expérimentales recueillies par les chercheurs à l'aide de techniques spécifiques.

- Modèles abstractionnistes

Organisation modulaire

Les modèles tentent d'établir une interface plausible entre des informations du monde extérieur et des connaissances et croyances de la vie mentale. La plupart repose sur l'idée d'une organisation modulaire des processus cognitifs, et du langage en particulier. Un domaine cognitif se compose de différents modules, eux-mêmes divisés en sous-modules, et ainsi de suite. Chaque module effectue un traitement spécifique de l'information et est relié à d'autres tout en gardant une certaine indépendance. Ainsi, la reconnaissance d'un mot dans la parole est notamment influencée par le contexte syntaxique et sémantique dans lequel il apparaît, mais la récupération de sa forme phonologique est obligatoire et reste possible sans contexte. La reconnaissance des mots parlés est donc étudiée en dehors de la parole et du langage, pour faciliter la pratique expérimentale, mais peut permettre de mieux la comprendre. Ce mémoire présente des études sur des mots isolés et il faudra se garder de généralisations hâtives à l'ensemble de la parole, processus évidemment plus complexe.

Interaction avec la recherche expérimentale

L'élaboration de modèles est exigeante car elle nécessite de préciser le nombre et la nature des modules ainsi que les connexions établies entre eux. Cela présente l'avantage d'aider le chercheur, et éventuellement le clinicien, à mieux saisir les paramètres en jeu dans un domaine cognitif donné et à affiner leurs investigations. Les modèles tentent de rendre compte des comportements observés chez les individus en recherche expérimentale. Il existe une interaction constante entre les expérimentations,

menées et interprétées en utilisant les modèles comme grilles d'analyse, et les modèles eux-mêmes, qui évoluent en intégrant un maximum de données comportementales. Mis à part les modèles verbaux, beaucoup sont des programmes informatiques qui peuvent simuler leurs propres résultats, ce qui permet de les confronter aux mesures relevées chez les individus. Les modèles de la reconnaissance des mots parlés les plus fréquemment cités dans la littérature et sur lesquels nous baserons nos propos sont le modèle de la Cohorte (Marslen-Wilson and Welsh, 1978), (Marslen-Wilson, 1987) pour sa version révisée, TRACE (McClelland and Elman, 1986), Shortlist (Norris, 1994) et NAM (Neighborhood Activation Model), (Luce and Pisoni, 1998).

- Techniques comportementales

Principes généraux

En recherche expérimentale sur la reconnaissance des mots parlés, les sujets sont exposés à un stimulus choisi puis effectuent une tâche. Leur rapidité et leur taux de réussite varient en fonction du stimulus présenté, ce qui permet de déduire des informations sur la nature du traitement qui a été mis en jeu. Par exemple, en tâche de répétition, les temps de réaction et le nombre d'erreurs sont généralement moins importants pour des mots par rapport à des non-mots. Ceci suggère que le système de traitement est sensible à la lexicalité du stimulus. C'est toujours la comparaison entre plusieurs catégories d'items qui est informative et non les valeurs recueillies en soit. Ainsi, dans les études présentées dans ce mémoire, les données étudiées ont été comparées à des mesures contrôles.

Les techniques employées peuvent faire varier plusieurs paramètres comme la modalité de présentation du stimulus, le nombre de stimulus et l'intervalle de temps qui les sépare, la tâche ou encore la modalité de réponse. Les langues étudiées peuvent également changer et les conclusions des expérimentations ne sont pas nécessairement généralisables. Les citations qui suivent concernent pour l'essentiel des travaux en anglais et quelques-uns en français.

Toute la difficulté est de s'assurer que le résultat recueilli correspond à la mesure souhaitée. Pour cela, le psycholinguiste veille à réaliser des catégories d'items pertinentes et homogènes à l'aide de bases de données lexicales. Pour s'assurer que les réponses du sujet reflètent le produit d'un traitement automatique non conscient, l'intervalle de temps entre la présentation du stimulus et la réponse du sujet est minimisée : les tâches utilisées sont simples comme la répétition de mots, la détection de phonèmes ou encore la décision lexicale et il est demandé au sujet de répondre le plus rapidement possible.

L'amorçage

L'amorçage est une technique fréquemment employée. En paradigme d'amorçage, le sujet réalise une tâche sur un stimulus cible mais un stimulus amorce est préalablement présenté. Le traitement automatique et irrépessible de cette amorce influence le traitement subséquent de la cible et a fortiori la tâche à accomplir. Plus particulièrement, lorsque l'amorce est reliée à la cible, par des phonèmes ou des traits sémantiques communs, les chercheurs observent souvent des effets de facilitation ou d'inhibition. Dans ce cas, les temps de réponse et/ ou le nombre d'erreur varient significativement en fonction de la condition étudiée par rapport à la condition contrôle. La mise en évidence d'un effet stable renseigne sur le traitement cognitif de l'amorce, notamment sur les hypothèses lexicales générées. Il existe au moins deux types de liens entre amorces et cibles susceptibles de produire ces effets. Ils correspondent aux deux variantes principales de cette technique, l'amorçage phonologique et l'amorçage sémantique.

Les données recueillies en recherche expérimentale constituent une privilégiée pour la formulation d'hypothèses concernant le traitement cognitif du mot parlé. Comme certaines d'entre elles sont imposées par les propriétés de la parole, les modèles abstractionnistes standards se rejoignent sur des grands principes.

2. Théorie générale de la reconnaissance des mots parlés

-Caractéristiques principales de la parole et conséquences pour son traitement cognitif

Flux continu de parole et segmentation

Le débit de parole s'écoule en un flux continu : il n'y a pas de silence entre les mots. Pourtant, l'auditeur sain parvient à accéder à l'identité sonore de chaque mot, indispensable à la compréhension du discours. En effet, une erreur de segmentation peut générer des ambiguïtés sémantiques. De ce fait, les psycholinguistes recherchent quelle stratégie de segmentation est mise en œuvre dans le traitement de la parole.

La segmentation peut être envisagée comme un préalable au processus de reconnaissance des mots parlés. Dans cette hypothèse, le début des mots doit être connu avant que le processus nécessaire à sa reconnaissance soit enclenché. Les chercheurs ont ainsi exploré les différents indices acoustiques qui pourraient être

exploités par le traitement cognitif pour parvenir à segmenter. L'influence de divers paramètres a été mise en évidence, comme la structure prosodique, syllabique, métrique ou encore les contraintes phonotactiques. Mais aucun ne permet de segmenter à coup sûr et leur influence varie en fonction de la langue étudiée. En effet, si une sensibilité aux variations métriques permettrait de segmenter environ 85% des mots en anglais (Cutler and Carter, 1987), une stratégie de segmentation syllabique serait plus efficace en français (Content et al., 2001).

Une autre possibilité est que l'identification des mots permette de segmenter la parole. L'élimination ou l'inhibition progressive des hypothèses non compatibles avec le signal sur la base d'informations perceptives et/ou sur la mise en compétition des candidats lexicaux offre des pistes de segmentation dans le modèle de la Cohorte I, TRACE et Shortlist. Les mécanismes bottom-up et d'inhibition latérale à l'origine de cette identification des mots et des frontières lexicales sont développés dans le chapitre suivant.

Ces deux hypothèses n'étant pas incompatibles, il semble raisonnable de considérer que le processus de segmentation sollicite le concours de stratégies variées. Quoi qu'il en soit, la continuité de la parole permet de déduire deux caractéristiques du traitement cognitif des mots : une sensibilité aux régularités acoustiques de la langue et des mécanismes de sélection lexicale très performants.

Variabilité de la parole et catégorisation

En plus d'être continue, la parole est extrêmement variable. En effet, la voix du locuteur, la qualité de l'articulation ou encore le bruit ambiant sont autant de paramètres qui rendent unique chaque réalisation acoustique d'un mot. La variabilité des formes acoustiques rencontrées par un auditeur est augmentée par une caractéristique propre à la production de la parole : la coarticulation. Tout d'abord, l'enchaînement rapide de gestes articulatoires entraîne leur chevauchement partiel sur l'axe temporel. De plus, l'exécution d'un mouvement est influencée par les mouvements voisins. Comme la parole est programmée en avance, la configuration du conduit vocal est optimisée pour un enchaînement articulatoire spécifique. Par exemple, le phonème /ɔ/ dans « bord » est plus postérieur que dans « bol » car le /ʁ/ final, produit au niveau de la luvette, attire l'articulation vers le fond du tractus vocal. Aussi, un même phonème a des caractéristiques acoustiques différentes selon la séquence sonore dans laquelle il se trouve.

Une des propriétés étonnantes du traitement de la parole est de parvenir à la sélection d'une entrée lexicale unique à partir d'une infinité de réalisations possibles. Les mécanismes conduisant à cette invariance perceptive ne sont pas entièrement connus mais la plupart des psycholinguistes suppose qu'elle opère via la catégorisation en unités abstraites et stables. Ainsi, la variabilité de la parole confère

au traitement cognitif la capacité de regrouper des informations acoustiques pertinentes en fonction d'un code élaboré. Reste à déterminer la nature de ce code.

- Postulats communs des modèles classiques

Unité de représentation phonémique

Le choix d'unité de catégorisation dans la plupart des modèles de la reconnaissance des mots parlés est le phonème.

En linguistique, le phonème correspond à une combinaison spécifique de traits phonétiques. Les traits sont associés aux caractéristiques des mouvements articulatoires comme le mode d'articulation, le lieu d'articulation ou le voisement. Par exemple, le phonème /b/ est défini comme une consonne occlusive bilabiale voisée. Une variation d'un seul trait change la catégorie phonémique. Dans notre exemple, si au niveau du lieu d'articulation le trait bilabial est remplacé par un trait apico-dental, le phonème devient un /d/.

L'avantage du phonème comme unité de catégorisation est qu'il permet de distinguer les mots à partir d'un répertoire restreint de segments (36 en français). De plus, les psycholinguistes ont mis en évidence que cette unité était utilisée par le traitement cognitif pour réduire la perception de la variabilité acoustique à travers le phénomène de perception catégorielle. Lorsque des individus entendent un enchaînement de sons, créés artificiellement de manière à varier progressivement sur un continuum entre deux phonèmes, leur système perceptif ne perçoit pas les variations linéaires mais classe les stimuli entendus en deux catégories phonémiques distinctes. Ce changement franc de catégorie intervient au milieu du continuum.

Ainsi, dans les modèles classiques, une étape préalable à la reconnaissance des mots est l'identification des phonèmes associés au signal de parole. Le modèle TRACE explicite davantage cette catégorisation en incluant des détecteurs de traits en amont des détecteurs de phonèmes. Notons également que le modèle de la Cohorte II utilise uniquement le trait phonétique et que le choix d'unité ne fait pas l'unanimité dans la communauté scientifique. Dans ce travail, nous nous centrons sur le phonème car c'est l'unité la plus utilisée. Pour autant, d'autres unités comme le trait ou la syllabe jouent probablement un rôle dans les représentations infra-lexicale et lexicale.

L'unité de catégorisation est à mettre en relation avec l'unité de représentation des mots en mémoire. Le modèle TRACE détaille les images mentales en traits et en phonèmes, équivalents à ses détecteurs infra-lexicaux et la Cohorte II les imagine en gabarits de traits. Pour les autres modèles de référence, ce sont des séquences de phonèmes qui sont inscrites dans le lexique phonologique.

Déroulement global de la reconnaissance d'un mot : les différents niveaux de traitement

Les modèles classiques proposent les étapes suivantes pour décrire le processus cognitif de reconnaissance du mot.

Ce processus démarre dès que les premiers sons de parole sont entendus. Ils sont codés dans le cerveau, avant tout traitement cognitif, sous la forme d'une représentation d'entrée. Celle-ci n'est pas spécifiée et représente simplement l'homologue cérébral du signal acoustique. Au contact de cette représentation d'entrée, les unités abstraites sont détectées. La chaîne des segments identifiés constitue la représentation infra-lexicale.

La conversion du signal acoustique en une représentation interprétable par le cerveau définit le rôle du traitement pré-lexical.

Les unités infra-lexicales activent à leur tour de multiples entrées lexicales qui constituent autant de candidats à la reconnaissance du mot. Tant que l'information acoustique n'est pas disponible dans son intégralité, plusieurs mots sont susceptibles d'être reconnus. Par exemple, à l'écoute des segments « bo », les auditeurs envisagent « bord », « borne » ou encore « bol ». La reconnaissance d'un mot advient lorsqu'une forme lexicale est sélectionnée parmi un ensemble de candidats potentiels. L'existence de ces candidats lexicaux est prouvée expérimentalement (Zwitserslood, 1989).

Le niveau de traitement lexical est donc constitué de deux étapes : l'appariement, qui désigne l'activation de candidats à partir d'une représentation infra-lexicale, et la sélection lexicale, ou procédure de tri de ces candidats.

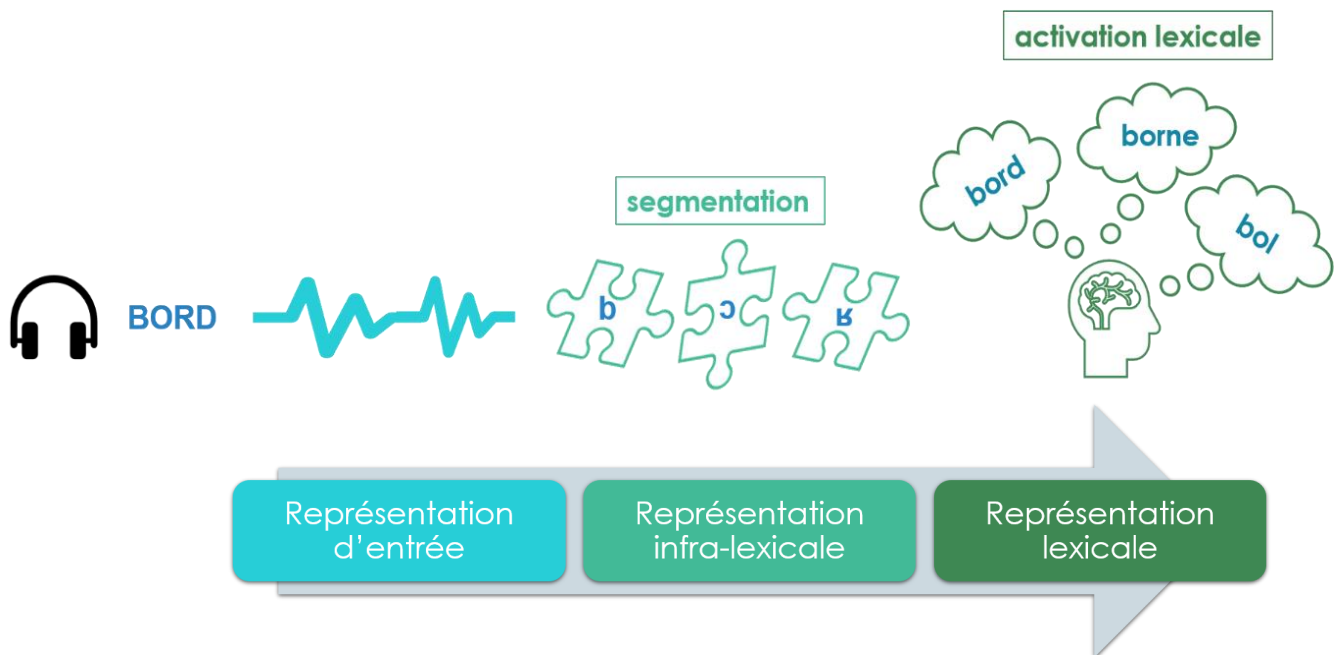


Schéma de la reconnaissance des mots parlés selon les modèles classiques

3. Les différents mécanismes d'appariement et de sélection lexicale envisagés par les modèles classiques

Nous avons vu que le traitement pré-lexical décrit dans les modèles classiques était globalement similaire et nous reviendrons sur certaines subtilités dans le chapitre concernant l'ordre temporel des phonèmes. En revanche, différents mécanismes sont envisagés pour le traitement lexical. La compréhension de ces mécanismes est essentielle à l'interprétation des données relevées à l'aide des techniques comportementales.

- Hypothèses lexicales considérées à l'écoute d'un mot

Les candidats appariés au niveau du premier phonème

Selon le modèle de la Cohorte I, un ensemble de candidats est activé dès les premiers phonèmes entendus soit 100 à 150 millisecondes après le début du signal de parole. Cette cohorte comprend tous les mots du lexique mental commençant par ces mêmes phonèmes. Par exemple, lorsqu'on écoute le mot « bord », les candidats « borne » ou « bol » sont activés mais pas « or », « sort » ou « sport ». C'est ce que suggère une étude avec un protocole d'amorçage sémantique (Marslen-Wilson and Zwitserlood, 1989). Ainsi, le modèle de la Cohorte I accorde aux premiers phonèmes un rôle central dans la reconnaissance d'un mot parlé puisqu'ils sont les seuls à générer des hypothèses lexicales.

Cependant, l'hypothèse de ce modèle est difficile à soutenir dès lors que l'on sort d'une situation expérimentale. Dans un environnement courant, les premiers phonèmes ne sont pas toujours intacts en raison de bruit ambiant ou d'un défaut de prononciation et le système de traitement les reconnaît quand même.

Si l'étude citée précédemment (Marslen-Wilson and Zwitserlood, 1989) soutient que l'accès lexical n'est pas préservé lorsque le début du mot entendu n'est pas intact, les travaux d'autres travaux (Connine et al., 1993), (Connine et al., 1997) suggèrent le contraire. En altérant artificiellement, dans leur protocole de recherche, le premier phonème des mots, ils mettent en évidence que les auditeurs activent la représentation du mot d'origine tant que l'écart avec celui-ci ne dépasse pas un ou deux traits. Selon cette interprétation, le mot « bord » pourrait être envisagé à l'écoute de la séquence « gor » mais pas de la séquence « chor », qui en est trop éloignée. De plus, leurs résultats suggèrent qu'un défaut d'appariement initial n'est pas plus invalidant qu'un défaut d'appariement médian.

A la suite de ces travaux, le modèle de la Cohorte I a été modifié. Pour qu'un défaut de prononciation impacte moins la reconnaissance du mot, la représentation infra-lexicale a été spécifiée en termes de traits phonétiques plutôt que de phonèmes.

Une autre étude va plus loin en montrant que l'accès lexical est préservé même si la modification du premier phonème d'un mot forme un nouveau mot (Allopenna et al., 1998). Les auteurs proposent un paradigme dans lequel les sujets entendent des instructions concernant un mot cible puis doivent réaliser la consigne sur un ordinateur en modalité visuelle, c'est-à-dire à partir d'une image qui lui correspond. L'enregistrement des mouvements oculaires leur permet d'établir que les candidats appariés à l'initiale sont considérés en première intention, et plus fréquemment, mais que ceux ne partageant pas le même début peuvent aussi être activés. D'après ces conclusions, l'écoute de « bord » activerait préférentiellement « bol » mais également « sport ».

Les mots partageant les premiers phonèmes constituent donc des candidats forts mais une procédure d'appariement lexical trop exigeante en termes de correspondance avec le signal ne semble pas suffisante pour décrire la diversité des hypothèses envisagées par les auditeurs.

Les candidats appariés avec d'autres portions du signal

Une autre hypothèse, développée par TRACE et Shortlist, suppose que la cohorte est renouvelée au fur et à mesure de l'écoute du signal, après chaque phonème entrant. Ainsi, lors de l'écoute de « bord », les mots « bol » et « borne » sont activés dans un premier temps puis par exemple « or » et « robe ». Il s'agit d'un alignement exhaustif où tous les mots partageant une partie commune avec l'entrée sont susceptibles d'être activés. Dans un tel modèle, les débuts de mots n'ont pas de statut privilégié et des candidats appariés en fin de signal peuvent être activés. Pour examiner cette éventualité, les psycholinguistes ont sélectionné des mots eux-mêmes constitués de mots au niveau de leurs derniers segments, comme « microbe » qui contient « robe ». Les expérimentations menées en amorçage sémantique ne vont pas toutes dans le sens de l'existence d'hypothèses alignées tardivement avec le signal (Gow and Gordon, 1995) mais une légère activation, mesurée par plusieurs chercheurs, laisse penser qu'elles sont effectivement envisagées par les auditeurs (Shillcock, 1990), (Luce and Cluff, 1998).

Enfin, le modèle NAM va dans le sens d'une conception encore plus large des candidats puisqu'il ne prend pas en compte la directionnalité de la parole. Selon lui, tous les mots pouvant être générés par addition, délétion ou substitution d'un phonème sont activés. Ainsi, à l'écoute du mot « sort », les mots « sorte » et « bord » sont envisagés ainsi que « sport » ou « or ». Cependant, ce modèle est spécialisé dans la reconnaissance mots monosyllabiques.

L'activation de candidats alignés par la fin est sérieusement envisagée par de nombreux chercheurs mais elle est plus difficile à mettre en évidence expérimentalement.

- Régulation des hypothèses lexicales et sélection du meilleur candidat

Nous avons vu que plusieurs candidats étaient activés à l'écoute d'un mot, a minima tous ceux commençant par le même phonème, et potentiellement d'autres. Cette activation multiple a lieu en cours d'écoute, alors que des informations acoustiques continuent d'affluer et rendent incompatibles certaines hypothèses. Quels mécanismes permettent de trier tous ces candidats afin de reconnaître le bon mot ?

Mécanisme bottom-up

Dans la plupart des modèles, les niveaux de traitement agissent en parallèle : ils opèrent dès l'arrivée des premiers sons et continuent de manière concomitante tout au long du processus d'écoute. Ainsi, même après avoir déclenché des hypothèses lexicales, le processus de reconnaissance continue de prendre en compte les informations provenant de l'input et celles-ci impactent le jeu de candidats constitué. Les modèles abstractionnistes ont donc tous développé un mécanisme « bottom-up » pour expliquer que nos interprétations du signal de parole s'ajustent au fur et à mesure de l'écoute, en temps réel.

De cette manière, les candidats ne correspondant plus au signal sont désactivés passivement après chaque phonème entrant selon le modèle de la Cohorte I. Le mot est reconnu lorsqu'il ne reste plus qu'une hypothèse compatible.

Ce mécanisme de sélection lexicale permet de mettre en évidence la notion de point d'unicité. Il s'agit du moment à partir duquel une entrée lexicale ne peut plus être confondue avec une autre. Par exemple, la séquence « spa » peut activer plusieurs mots comme « spatule », « spatial » ou « spacieux » mais l'occurrence du phonème /g/ dans cette séquence réduit la cohorte de candidats à une seule possibilité : le mot « spaghetti ». Cette caractéristique du lexique est également mise à profit dans ce modèle pour expliquer la segmentation de la parole puisque l'identification précoce d'un mot permet d'anticiper la prochaine frontière lexicale.

Le modèle de la Cohorte II, TRACE, NAM et Shortlist développent la notion de qualité d'ajustement, selon laquelle l'information acoustique nouvelle ne désactive pas instantanément les candidats inappropriés mais met à jour leur niveau d'activation. Ce sont des connexions ascendantes excitatrices qui permettent d'augmenter le niveau d'activation des entrées lexicales à chaque fois qu'elles coïncident avec le niveau infra-lexical.

Le modèle Shortlist est plus spécifique puisque le degré de correspondance avec le signal permet de limiter le nombre d'hypothèses générées. A chaque nouveau phonème reconnu, les hypothèses lexicales satisfaisantes reçoivent un score d'ajustement, + 1 pour un phonème bien apparié et - 3 pour un phonème mal apparié, et seules les plus conformes au signal constituent la Shortlist.

Mécanisme d'inhibition latérale

TRACE et Shortlist ajoutent à ce mécanisme « bottom-up » un mécanisme d'inhibition latérale. Il permet aux unités les mieux appariées au signal, et disposant d'un certain niveau d'activation, de diminuer celui des autres unités. Cette interaction est possible au sein d'un même niveau, entre mots, mais également entre phonèmes, à travers des connexions inhibitrices. L'inhibition exercée par un candidat lexical est alors proportionnelle à son niveau d'activation et à la puissance des connexions qui le relie aux autres. Ce mécanisme assure l'élimination progressive des hypothèses les moins probables et instaure entre les candidats une compétition lexicale pour la reconnaissance.

C'est cette compétition lexicale qui explique pourquoi il est difficile de mesurer l'activation d'un mot enchâssé dans un mot plus long. En effet, le mot qui correspond au signal depuis le début a un niveau d'activation élevé qui lui permet d'inhiber le mot qui ne correspond que tardivement à l'input. Selon ce principe, lorsqu'on entend le mot « microbe », le candidat « robe » est activé mais inhibé par la séquence porteuse. L'inhibition latérale a été mise en évidence avec différents paradigmes comme le « word spotting » (McQueen et al., 1994) ou l'amorçage phonologique (Dufour and Peereman, 2003).

Dans TRACE, la compétition lexicale permet également de déduire le début du mot suivant. Contrairement au modèle de la Cohorte I, des candidats sont activés en tout point du signal, ce qui signifie que des frontières lexicales sont postulées à tout moment. C'est l'issue de la compétition lexicale qui détermine quel mot et donc quelle frontière est retenue.

- Rôle des connaissances lexicales

Le processus cognitif de reconnaissance des mots consiste à proposer des candidats potentiels et à conserver celui qui correspond le mieux au signal acoustique entrant. Mais ce processus étant le fruit d'un apprentissage, il doit probablement évoluer et s'enrichir des expositions répétées de l'auditeur à la parole. L'écoute dépend à la fois de l'analyse du signal et des connaissances lexicales. Pourtant, si les modèles ne peuvent fonctionner sans l'existence préalable d'un stock lexical, l'impact des

informations provenant de ce lexique sur la reconnaissance des mots varie en fonction des mécanismes retenus.

La fréquence d'occurrence

Une propriété du lexique, dont l'influence sur la reconnaissance des mots fait consensus au sein de la communauté scientifique, est la fréquence d'occurrence. Plus un mot est fréquemment employé dans une langue, plus vite et mieux il est reconnu. Par exemple, le mot « bord », plus fréquent que « robe », sera reconnu plus rapidement et avec moins d'erreur. Cet effet a été mis en évidence dans plusieurs tâches (Connine et al., 1990), (Dupoux and Mehler, 1990). L'effet de fréquence est admis comme ayant un locus lexical et sa fiabilité constitue un marqueur efficace pour déterminer le niveau de traitement mis en jeu lors de la réalisation d'une tâche en recherche expérimentale (Radeau et al., 1995).

Alors que le modèle de la Cohorte I ne prend pas du tout en compte la fréquence d'occurrence, sa version révisée l'intègre avec la notion de seuil d'activation. Un mot fréquent est reconnu plus vite qu'un mot peu fréquent car le niveau d'activation nécessaire à sa reconnaissance est plus bas.

D'autres modèles intègrent cette caractéristique en attribuant aux entrées lexicales un niveau d'activation de base proportionnel à leur fréquence. C'est le cas dans TRACE et Shortlist où les mots fréquents ont d'emblée un avantage dans la compétition lexicale. Enfin, la fréquence peut aussi être prise en compte plus tardivement, au moment de la reconnaissance du mot. C'est le cas dans le modèle NAM où un mot est reconnu après un calcul prenant en compte la fréquence des candidats en lice.

L'avantage de ces trois modèles par rapport au modèle de la Cohorte II est qu'ils prennent en compte la fréquence relative entre le mot cible et ses compétiteurs. Or, des études montrent qu'un mot cible est reconnu plus lentement s'il a des voisins lexicaux plus fréquents que lui (Luce et al., 1990), (Luce and Pisoni, 1998), (Dufour and Frauenfelder, 2010, en français). Cet effet est difficilement détectable lorsqu'un mot a beaucoup de voisins lexicaux.

Mécanisme top-down

Le modèle TRACE accorde une grande importance aux connaissances lexicales : non seulement il prend en compte la fréquence relative entre un mot et ses compétiteurs mais il intègre également un mécanisme de rétroaction du lexique sur la représentation infra-lexicale, appelé mécanisme « top-down ». En cela, il se distingue

des modèles autonomes qui font l'hypothèse de l'indépendance des différents niveaux de traitement.

A partir d'un certain niveau d'activation, les candidats lexicaux renforcent celui des phonèmes qu'ils contiennent grâce à des connexions excitatrices descendant vers le niveau inférieur. Ensuite, le candidat est activé en retour via les connexions « bottom-up » entre phonèmes et entrées lexicales appariées. Ces mécanismes, combinés à l'inhibition latérale, assurent la reconnaissance du candidat correspondant le mieux au signal malgré la présence de nombreux compétiteurs. Cette reconnaissance intervient lorsque le niveau d'activation du mot cible domine celui des autres hypothèses.

La rétroaction du lexique sur la représentation infra-lexicale permet également au modèle TRACE d'expliquer la restauration phonémique, bien que d'autres interprétations ont été proposées. Il s'agit du phénomène perceptif selon lequel l'identité d'un phonème bruité, ou mal prononcé, est reconstruite lorsqu'il se situe dans un mot (Warren, 1970). Il semble alors que l'information lexicale pallie son absence, voire modifie notre perception du signal, en nous donnant l'illusion que le phonème est bien présent.

4. Traitement infra-lexical et ordre temporel des phonèmes

Nous venons de décrire les mécanismes d'appariement et de sélection lexicale en supposant que l'enchaînement des phonèmes était inhérent à la représentation infra-lexicale. Pourtant, expliquer comment le traitement cognitif parvient à cette équivalence temporelle n'est pas acquis.

- Enjeu de l'invariance temporelle

Indépendance temporelle du flux de traitement

Comme nous venons de l'aborder, l'écoute du signal stimule des détecteurs de bas niveau qui relaient l'activation vers des niveaux d'abstraction croissante. Selon la hiérarchie fonctionnelle classique, la décharge d'activation progresse des traits acoustiques aux phonèmes puis jusqu'aux représentations lexicales. L'activation se propage également de façon horizontale au sein d'un même niveau, dans les modèles incluant une compétition lexicale, et de façon descendante vers les niveaux inférieurs dans les modèles interactifs. Cette propagation d'activation correspond au flux de traitement du mot. Ce flux, concomitant à celui du signal, n'est pourtant pas toujours synchronisé avec lui. Par exemple, lorsqu'un mot est reconnu avant la fin de sa réalisation, le traitement est en avance sur le signal de parole.

Pour ne pas confondre l'ordre des segments infra-lexicaux, le processus de reconnaissance du mot doit donc mettre en œuvre une stratégie pour mémoriser l'information temporelle jusqu'à la sélection de la bonne image mentale. En effet, la parole étant éphémère, cette donnée n'est plus disponible dès que les détecteurs de bas niveaux ont été stimulés.

Indépendance temporelle des unités infra-lexicales

Contrairement aux portions du signal auxquelles elles sont associées, les unités infra-lexicales envisagées dans les modèles de référence sont en nombre limité et ne sont pas spécifiques à une position dans le mot. Pourtant, les images mentales qu'elles activent sont détaillées en séquences fixes d'unités.

Ainsi, la catégorisation ne suffit pas à expliquer la constitution d'une représentation infra-lexicale en une suite ordonnée de segments. Doté de ce seul mécanisme, le système de traitement confondrait par exemple les mots « bonbon » et « bond ». Dans cet exemple, la deuxième portion du signal, c'est-à-dire les segments /b/ et /ɔ̃/, respectivement en troisième et quatrième position, stimulent les mêmes unités que la première partie du mot. Si la position des phonèmes n'est pas codée, la deuxième portion du mot n'est pas traitée comme un nouvel événement acoustique mais comme une activation supplémentaire des phonèmes /b/ et /ɔ̃/.

- Codage de l'ordre temporel des phonèmes dans les modèles classiques

Modèles spécifiques au traitement lexical

Certains modèles, comme celui de la Cohorte (I et II), ne décrivent que le traitement lexical du mot, sans préciser comment la séquence des segments est représentée cognitivement. Ce modèle, proposant une théorie purement verbale, suggère que les unités reconnues par l'auditeur activent un ensemble de candidats, en se basant sur l'hypothèse que l'ordre des segments dans la représentation infra-lexicale est équivalent à celui du signal. Il n'explique pas comment cette équivalence est obtenue par les processus cognitifs. Cette question prend tout son sens pour les modèles computationnels, cherchant à simuler la reconnaissance d'un mot à l'aide d'un programme informatique, car ils doivent spécifier toutes les étapes qui interviennent.

TRACE propose une stratégie complète de codage de la position des unités. Dans ce modèle, le temps est codé à l'aide d'une matrice divisée en différents intervalles temporels. Chaque intervalle correspond à l'analyse d'une portion du signal. En fonction de la taille du grain d'analyse, la durée de l'intervalle augmente : 5 ms pour les traits acoustiques et 15 ms pour les phonèmes. Des détecteurs de traits et de phonèmes sont déployés à chaque intervalle de temps dédié : ils sont donc spécifiques à une position dans le mot. Ainsi, selon notre précédent exemple, à l'écoute de « bonbon », le premier intervalle de 15 ms détecte la présence du phonème /b/ et c'est un autre détecteur qui est stimulé 30 ms plus tard, lors de l'occurrence du deuxième /b/, en troisième position du mot.

Dans Shortlist, le codage de la position des segments s'opère via le score d'ajustement attribué, par le système de traitement, aux hypothèses lexicales. Ce score assure que les candidats qui constituent la « shortlist » respectent l'ordre temporel des phonèmes ou soient désavantagés dans la compétition lexicale.

L'inconvénient des stratégies élaborées par TRACE et Shortlist est qu'elles prennent en compte la position absolue des unités. Le processus de traitement est alors très sensible à une délétion ou une adjonction de phonème et ne rend pas bien compte des similarités entre mots proches comme « sort » et « sport ». Dans ces deux modèles, le code infra-lexical de ces mots diverge complètement dès l'occurrence du deuxième segment.

TRACE rend compte, dans une certaine mesure, des similarités entre mots phonologiquement proches car il mémorise la trace de l'activation provenant du signal. Les détecteurs infra-lexicaux sont renouvelés après chaque unité de temps mais leur stimulation est mémorisée durant 11 intervalles temporels. Le poids de cette activation résiduelle diminue progressivement après chaque intervalle jusqu'à disparaître. Ainsi, l'activation générée par la reconnaissance des segments /s/, /ɔ/ et /ʁ/, à l'écoute de « sort », est renvoyée aux mots composés de ces phonèmes, dont « sport », durant un certain temps. Cependant, l'inhibition latérale minimise ce phénomène et il génère un coût de traitement considérable. D'après une étude (Hannagan et al., 2013), si TRACE utilisait un répertoire de phonème et un lexique de taille réelle, son fonctionnement nécessiterait approximativement 4 millions d'unités et 80 billions de connexions.

En revanche, le modèle NAM code de manière plus relative la position des segments puisqu'il admet des candidats formés par addition, délétion ou substitution d'un phonème. Pour garantir la reconnaissance du mot strictement apparié, il effectue un calcul de probabilité qui prend en compte l'ajustement des candidats au signal, en plus de leur fréquence et de leur voisinage. Mais l'extension de ce modèle aux mots

plurisyllabiques mettrait probablement en évidence un phénomène de traitement très coûteux de l'information à cause de l'augmentation exponentielle des calculs à effectuer.

Les modèles abstractionnistes standards se basent tous sur une série ordonnée d'unités pour décrire le traitement lexical. Néanmoins, aucun ne propose une solution satisfaisante pour expliquer l'invariance temporelle entre le signal et la représentation infra-lexicale. Le champ d'étude dans ce domaine reste donc largement ouvert.

5. Vers une conception nouvelle du traitement de la position des phonèmes ?

Si les modèles classiques conçoivent le traitement infra-lexical comme l'assemblage d'une série ordonnée de segments, une vision moins stricte de l'agencement des composants est envisageable.

- Apports du domaine de la reconnaissance des mots écrits

Les premières expérimentations à propos du codage de la position des unités viennent du domaine de la reconnaissance des mots écrits.

Enjeux communs à l'oral et à l'écrit

S'intéresser à la reconnaissance des mots écrits présente un intérêt multiple. Tout d'abord, la recherche expérimentale est facilitée en langage écrit car l'unité de représentation des mots est unanimement reconnue et aisément manipulable. De plus, ce domaine partage des enjeux communs avec la parole concernant la représentation cognitive de l'ordre des segments.

Comme à l'oral, un des défis de la recherche en reconnaissance des mots écrits est de comprendre comment les mots sont reconnus aussi rapidement et efficacement par les adultes sains malgré une infinité de formes de surface. La variabilité des mots écrits est moindre qu'à l'oral car l'intégrité des lettres est souvent préservée et elles ne changent pas de forme en fonction de leur place dans le mot. Mais le système de reconnaissance des mots écrits réalise tout de même une invariance perceptive puisque les mots rencontrés par le lecteur peuvent changer de taille, de casse ou encore de police.

De plus, si la modalité visuelle offre davantage de possibilités de retours en arrière qu'à l'oral, son système de reconnaissance des mots n'est pas pour autant

insensible à l'agencement des composants. En effet, un mot n'est pas toujours fixé exactement au même endroit sur la rétine et a fortiori ses lettres n'apparaissent pas toujours à la même position. Le traitement des mots écrits doit donc coder l'ordre des segments pour parvenir à cette invariance spatiale.

Comme à l'oral, le système de reconnaissance des mots écrits doit représenter l'identité des lettres et leur agencement. La reconnaissance des mots parlés a donc tout intérêt à s'inspirer des expériences menées autour du traitement cognitif des mots écrits.

Expériences d'amorçage en reconnaissance des mots écrits

Comme à l'oral, le codage de la position des segments à l'écrit a d'abord été envisagé comme concomitant à leur identification. Dans le modèle à activation interactive (McClelland and Rumelhart, 1981), qui a inspiré le modèle TRACE, ce sont des détecteurs de lettres spécifiques à une position qui permettent de reconstituer la chaîne de segments infra-lexicaux. Mais des études en paradigme d'amorçage invalident cette conception.

Une étude en français montre que le traitement cognitif des mots écrits prend en compte la position relative des segments (Peressotti and Grainger, 1999). Plus précisément, les résultats mettent en évidence qu'une amorce respectant la position absolue des lettres du mot cible n'est pas plus facilitante qu'une amorce conservant leur position relative. Ainsi, les amorces « bdlcdn » ou « b-lc-n », qui conservent la position absolue des unités, favorisent autant le traitement du mot « balcon » que l'amorce « blcn ». En revanche, une amorce ne préservant pas la position relative des segments, comme « bcln », ne génère aucun effet de facilitation.

D'autres travaux en anglais suggèrent qu'un codage de la position relative des lettres est encore trop strict. En effet, une étude (Perea and Lupker, 2003) met en évidence que l'effet de facilitation mesuré classiquement avec la technique d'amorçage sémantique résiste à l'inversion de deux lettres internes de l'amorce. Selon ce principe, l'amorce « balcon » deviendrait « baclon ». Les mêmes auteurs répliquent l'effet en transposant des lettres internes non adjacentes (Perea and Lupker, 2004). Ils ne relèvent toutefois pas d'effet significatif avec une permutation de voyelles.

L'ensemble de ces données suggère que le codage de la position des lettres dans un mot écrit n'est pas strictement séquentiel.

- Un phonème mal placé empêche-t-il l'activation de la représentation du mot parlé ?

En s'inspirant des travaux réalisés à l'écrit, des chercheurs ont effectué en 2013 une expérience inédite qui montre que l'activation d'un mot-cible résiste à une inversion de phonèmes (Toscano et al., 2013). Ces résultats ont ensuite été confirmés en français, avec un protocole différent, dans le cadre d'un mémoire d'orthophonie.

Expérimentation en anglais

L'équipe de Toscano (Toscano et al., 2013) utilise le paradigme du « visual world » décrit précédemment (Alloppenna et al., 1998). Le matériel est constitué d'amorces de 3 phonèmes, identiques à ceux des mots cibles associés mais dans un ordre différent, à la manière d'anagrammes à l'oral. En suivant cette procédure, le mot « bord » serait relié à l'amorce « robe ». Ils montrent que dans cette condition, l'amorce facilite la cible et concluent que les auditeurs envisagent des hypothèses qui ne respectent pas l'ordre temporel des phonèmes.

Les chercheurs excluent la possibilité d'un effet généré par le recouvrement phonétique des premiers phonèmes. Ce pourrait être le cas, par exemple, avec une paire comme « patte – tape », où les consonnes initiales ne s'éloignent que d'un trait phonétique. Mais cela reviendrait à confirmer les conclusions des travaux sur les altérations en début de mots (Connine et al., 1993), (Connine et al., 1997). Cette éventualité est rejetée car un net effet de facilitation est mesuré lorsque les paires d'items ne partagent aucun trait phonétique au niveau de leur consonne initiale, comme c'est le cas des mots « robe » et « bord ».

L'hypothèse d'un effet dû au recouvrement vocalique est également écartée, sans quoi leurs résultats s'apparenteraient à ceux obtenus avec des amorces ne partageant pas le même début que les cibles (Alloppenna et al., 1998). En suivant notre exemple, l'amorce « robe » pourrait faciliter « bord » en raison du deuxième phonème commun (/ɔ/) et d'une tolérance du système de reconnaissance à un mésappariement au niveau du premier phonème. Mais l'analyse des fixations oculaires sur les items reliés par rapport aux contrôles avec partage vocalique, révèle un effet facilitateur plus important lorsque les mots ont des phonèmes transposés. D'après ces résultats, l'amorce « robe » faciliterait davantage « bord » que l'item contrôle « bol ».

A la suite de ces analyses, les auteurs concluent que c'est bien un phonème mal placé qui active le mot-cible et initie l'effet facilitateur mesuré.

Expérimentation en français

En 2019, un mémoire d'orthophonie réplique ces résultats en français avec un paradigme d'amorçage phonologique et une tâche de décision lexicale (Colin et al., n.d.). Une facilitation de traitement du mot-cible est mesurée lorsque l'intervalle inter-stimuli est court, c'est-à-dire lorsque la cible est présentée 20 ms après l'amorce. En revanche, en paradigme d'amorçage à long terme, lorsque toutes les amorces sont présentées d'un seul bloc et suivies d'une tâche de décision lexicales sur les cibles, les résultats ne sont pas répliqués. L'interprétation donnée est en faveur d'une dissipation rapide de l'activation du mot-cible.

Ces nouvelles données rejoignent les résultats collectés à l'écrit. Elles suggèrent que le codage de la position des segments n'est pas strictement séquentiel et ne respecte pas complètement la position relative des unités.

III) HYPOTHESE DE RECHERCHE

Le précédent mémoire sus-cité met en évidence l'existence d'un effet dû à l'inversion de phonèmes en français, tel que dans l'exemple « robe – bord » (Colin et al., n.d.). Celui-ci est mesurable en amorçage phonologique à condition d'utiliser un intervalle inter-stimuli court. Pour augmenter l'échantillon de preuves de cet effet, nous utilisons donc ce même paradigme. Or, l'amorçage phonologique à court terme met en jeu des effets pré-lexicaux et lexicaux (Radeau et al., 1995).

Un effet pré-lexical de facilitation est mis en évidence par ces auteurs lorsque l'amorce rime avec la cible. Il se produit en tâche de répétition et de décision lexicale. D'après eux, l'effet est généré par l'activation résiduelle des détecteurs de phonèmes qui établissent le code infra-lexical à l'écoute d'un mot.

De la même manière, la présentation 20 ms au préalable du mot « robe » pourrait faciliter le traitement de « bord » parce que les détecteurs de phonèmes /ʁ/, /ɔ/ et /b/ viennent tout juste d'être stimulés.

L'effet lexical mesuré au cours de cette étude est, quant à lui, inhibiteur et se produit uniquement en tâche de répétition, lorsque l'amorce et la cible se recouvrent sur leur partie initiale. Comme il se combine à un effet de fréquence, c'est-à-dire qu'il varie en fonction de la fréquence relative entre l'amorce et la cible, les chercheurs concluent qu'il ne peut avoir lieu qu'au niveau lexical. Dans ce cas, ils supposent que l'inhibition de traitement résulte d'une compétition entre candidats partageant les mêmes premiers phonèmes.

Nous supposons que l'effet de transposition mesuré en amorçage phonologique à court terme intervient aussi au niveau lexical. Cela signifierait que la facilitation de traitement serait bien due à la considération d'hypothèses lexicales non conformes à la position des segments de l'entrée et non à une activation résiduelle des détecteurs infra-lexicaux. L'étude réalisée par l'équipe de Toscano (Toscano et al., 2013) ainsi que l'analyse additionnelle menée à la suite de l'expérimentation des étudiantes en orthophonie, sous la direction de Sophie Dufour, (Colin et al., n.d.) vont dans ce sens. Pour en apporter la preuve expérimentale, nous avons conçu notre protocole de façon à pouvoir faire varier la fréquence relative entre les amorces et les cibles, à la manière de l'étude décrite ci-dessus (Radeau et al., 1995).

IV) PROTOCOLE EXPERIMENTAL

1. Matériels et méthodes

Notre protocole expérimental se base sur une tâche de décision lexicale et un paradigme d'amorçage phonologique à court terme.

Il a été délivré à 80 sujets âgés de 18 à 40 ans de langue maternelle française. En raison de la nature de la tâche et de sa modalité essentiellement auditive, les sujets sélectionnés ne présentent aucun trouble du langage ou ORL.

Pour la constitution de notre échantillon d'items, nous avons sélectionné 26 mots qui, après mélange de leurs phonèmes, forment 26 nouveaux mots de la langue française, comme les items « robe – bord ». Les 26 paires de mots ainsi obtenues seront qualifiées de paires d'items reliés ou transposés. Dans notre expérimentation, chacun de ces mots est utilisé une fois en amorce et une fois en mot-cible, ce qui nous permet d'obtenir 52 séquences différentes.

Les mots choisis sont composés d'1 syllabe et de 3 phonèmes ; ils constituent ainsi des stimuli brefs, captant au mieux les effets d'amorçage. L'utilisation de mots courts permet par ailleurs de réduire les difficultés méthodologiques et l'insertion de biais. Les mots sont de structure CVC et les paires ont donc toujours une voyelle commune et 2 consonnes interverties. Les mots sont issus de la base de données VOCOLEX. Leur point d'unicité se situe après le dernier phonème de sorte que le sujet soit contraint de les écouter en entier afin de décider s'ils existent, ou non, dans la langue française.

Afin de pouvoir contrôler la fréquence relative entre les amorces et les mots-cibles, les items que nous avons sélectionnés peuvent être répartis en 2 groupes homogènes de 2 types de fréquence lexicale. Ainsi, nos 26 paires sont chacune composées d'un mot de haute fréquence lexicale et d'un mot de basse fréquence lexicale. L'écart entre ces 2 catégories de fréquence est de 50 occurrences par million.

Nous avons également conçu 52 amorces que nous avons associé aux mots cibles précédents afin de constituer 52 séquences contrôles. C'est la comparaison entre cette condition contrôle et la condition où les items sont reliés qui nous permettra de déterminer si notre expérimentation génère un effet d'amorçage transposé. Ces amorces contrôles n'ont pas les mêmes consonnes que les mots-cibles qui leur sont appariés mais partagent la même voyelle. Ce partage vocalique nous permet d'éviter que les sujets déduisent la relation existant au sein des 52 séquences transposées et ainsi qu'ils élaborent des stratégies de réponse, conscientes ou non. Comme précédemment, 26 amorces de fréquences d'occurrence élevée sont appariées aux mots-cibles de fréquence d'occurrence faible et inversement.

Toutes ces données sont réparties en 4 listes afin que chaque liste contienne 13 séquences transposées et 13 séquences contrôles.

Pour les besoins de la tâche de décision lexicale, nous avons ajouté des séquences contenant un non-mot en position de cible, de sorte que les sujets aient à répondre autant de fois « mot » que « non-mot ». Nous en avons créé, comme précédemment, 13 transposées et 13 contrôles avec partage vocalique. Notre étude portant uniquement sur les mots, ces données ne seront pas analysées et nous utilisons donc les mêmes pour chacune des 4 listes. Les non-mots ont été formés en substituant le 3^{ème} phonème (consonne) d'un mot existant. Par exemple, nous avons généré le non-mot « mide » à partir du mot « mise ».

Enfin, nous avons constitué des séquences de remplissage dans lesquelles les amorces et les mots-cibles n'ont aucun lien consonantique ou vocalique. Il s'agit d'un contrôle de biais de réponse. En effet, si les voyelles se chevauchent systématiquement dans les séquences entendues par les sujets, les consonnes et les liens qui les unissent risquent d'être d'autant plus perceptibles. Nous avons donc intégré à notre matériel 39 séquences de remplissage où la réponse attendue est « mot » et 39 où la réponse attendue est « non-mot ». Ces 78 séquences ne seront pas analysées et nous utilisons donc les mêmes pour chacune des 4 listes.

	Amorce _ Cible	Amorce _ Cible
Condition transposée	mot _ mot ex : robe _ bord	mot _ non-mot ex : jupe _ puje
Condition contrôle avec partage vocalique	mot _ mot ex : vol _ bord	mot _ non-mot ex : pire _ mide
Remplisseurs sans partage vocalique	mot _ mot ex : calque _ grotte	mot _ non-mot ex : grippe _ stabe

Tableau récapitulatif des conditions d'amorçage phonologique

L'expérience est programmée via le logiciel d'expérimentation E-prime qui permet de lancer aléatoirement les séquences au sein de chaque liste.

2. Procédure

L'expérimentation a lieu dans un environnement calme et isolé. Le participant est équipé d'un casque audio et dispose devant lui d'un boîtier de réponses. Nous lui demandons de garder ses deux index sur les boutons tout au long de l'expérience, l'index de sa main dominante étant positionné au-dessus de la réponse « mot » et l'autre au-dessus de la réponse « non-mot ».

La consigne est administrée en modalité écrite via un écran d'ordinateur. Pour chaque séquence de 2 items, le sujet doit décider si le deuxième item est un mot ou un non-mot puis appuyer le plus rapidement possible sur le bouton correspondant. Une phase d'entraînement composée de 12 séquences précède la phase de test.

Le délai entre la fin de présentation de l'amorce et le début de présentation du mot-cible est de 20 ms. La pression d'un bouton réponse déclenche automatiquement la présentation d'une nouvelle séquence amorce - cible. La passation dure environ 15 minutes.

3. Résultats

Les données expérimentales ont été analysées à l'aide d'un modèle mixte à 2 facteurs aléatoires et à 8 facteurs fixes. Les facteurs aléatoires sont les participants et les items retenus pour le protocole. Les facteurs fixes correspondent, quant à eux, aux 2 variables de fréquence lexicale (haute et basse), aux 2 variables d'amorces (reliée et contrôle) ainsi qu'à leurs interactions.

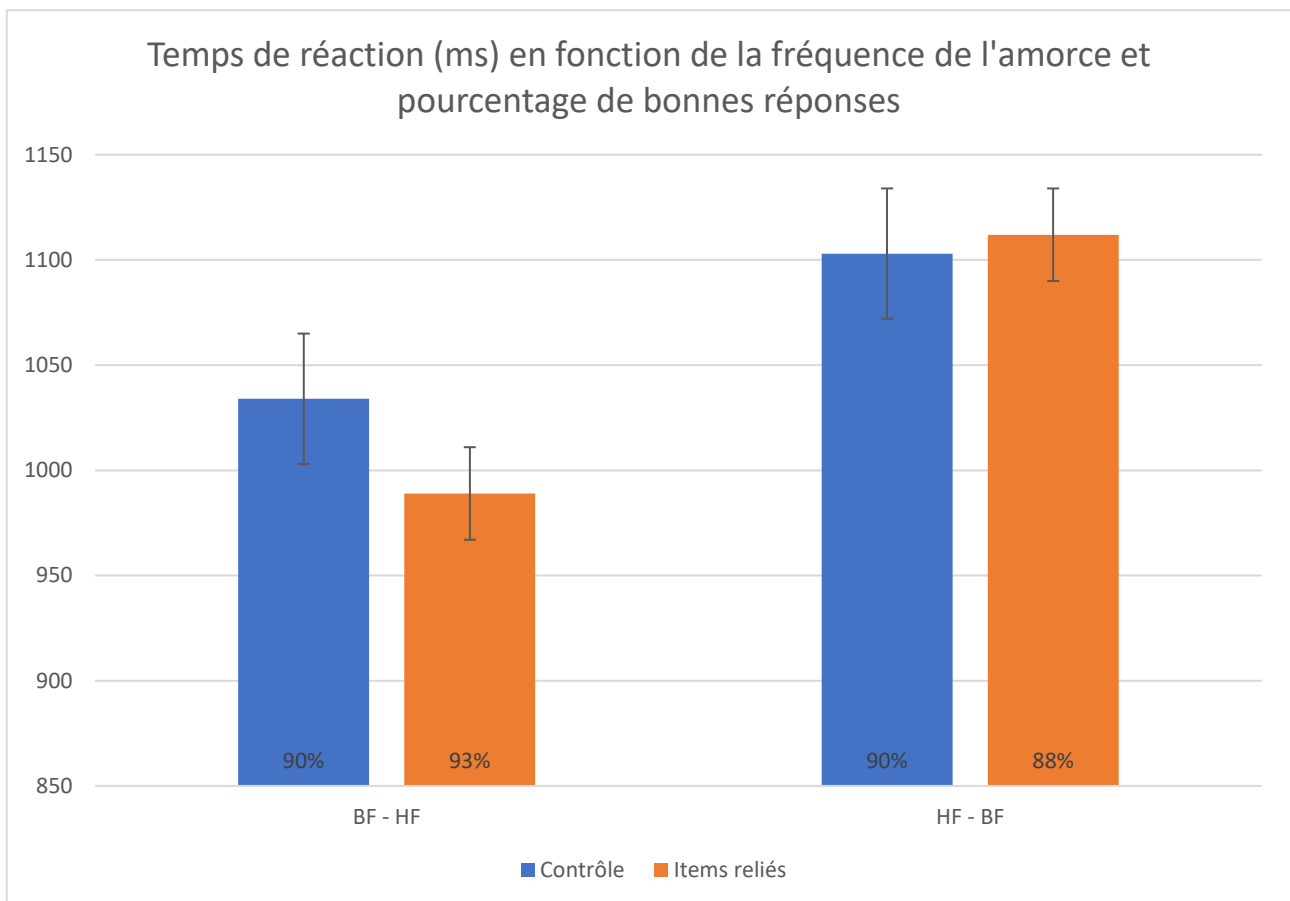
Nous avons analysé le temps de réaction (TR) et le taux d'erreur pour chaque sujet et pour chaque item.

Aucun participant n'a été écarté de l'étude.

En revanche, nous avons été contraints de supprimer 3 items de l'analyse car ils présentaient un taux d'erreur aberrant. Il s'agit des mots « chope », « rab » et « rosse », qui ont été considérés à tort comme des non-mots par plus de 60% des sujets. Il s'agissait des items ayant la fréquence d'occurrence la plus faible de notre sélection. Ainsi, sur les 26 paires d'items transposés constituées, 23 ont été conservées.

Parmi les données restantes, nous avons retiré 190 TR correspondant à des réponses erronées, leur taux s'élevant à 10,3%, ainsi que 37 TR que nous avons considéré comme non significatifs car supérieurs à 2200 ms.

Au total, les données analysées correspondent aux réponses justes et suffisamment rapides pour être exploitables de 80 sujets confrontés chacun à 23 items transposés, soient 1613 données. Le taux de données rejetées s'élève à 2,24%.



Analyse des TR :

Le modèle révèle un effet proche de la significativité pour le type d'amorce ($F(1,354) = 3,58 ; P = .06$), ce qui signifie que l'avantage des amorces transposées par rapport aux amorces contrôles pour le traitement du mot cible n'est pas suffisamment robuste sur l'ensemble de l'expérimentation.

Le type de fréquence lexicale montre quant à lui un effet significatif ($F(1,108) = 5,84 ; P < .05$). Il s'agit d'un effet de relation déjà observé dans d'autres expérimentations, selon lequel les mots fréquents sont reconnus plus rapidement que les mots moins fréquents.

Enfin, l'interaction entre le type d'amorce et le type de fréquence lexicale est significative ($F(1,354) ; P < .05$ ($P = 4,13$)). Plus précisément, le modèle met en évidence un effet d'amorçage significatif avec les amorces de basse fréquence lexicale ($Z = -2,79 ; P < .05$) et aucun effet significatif avec celles de haute fréquence lexicale ($Z = 0,21 ; P > .20$).

Ainsi, nous ne retrouvons pas d'effet d'amorçage transposé sur l'ensemble de notre expérimentation car celui-ci n'intervient que lorsque l'amorce reliée au mot-cible est de basse fréquence d'occurrence. Les amorces transposées de haute fréquence lexicale ne constituent pas un avantage par rapport aux contrôles pour la reconnaissance du mot cible, sans générer pour autant d'effet inhibiteur significatif.

Analyse du taux d'erreur :

L'analyse des erreurs n'a révélé aucun effet d'amorçage significatif.

V) DISCUSSION

Notre travail confirme l'existence d'un effet d'amorçage de transposition en français et ajoute la preuve qu'il intervient au niveau lexical.

Nos résultats indiquent que lorsqu'un mot de basse fréquence lexicale (BF), comme « robe », est présenté en amorce, le mot de haute fréquence lexicale (HF) associé, ici « bord », est facilité. Mais l'inverse ne fonctionne pas. Cette dissociation de résultats observée entre les deux conditions de fréquences relatives (HF - BF et BF - HF) montre que l'effet intervient au niveau lexical (Radeau et al., 1995).

Nous savons que les mots de fréquence lexicale élevée sont reconnus plus vite et plus facilement que les mots de fréquence plus basse. Il est donc logique que l'effet mesuré soit plus important lorsque la décision lexicale porte sur un mot fréquent. Mais d'autres paramètres lexicaux entrent en jeu puisque, dans la condition HF - BF, l'effet de facilitation est annulé voire légèrement inhibiteur (inhibition non significative dans notre expérimentation).

Nous supposons que c'est la compétition lexicale qui génère ce différentiel, ce qui confirmerait l'existence d'un mécanisme d'inhibition latérale. Ainsi, dans la condition HF - BF, le mot-cible est bien envisagé mais la fréquence élevée de l'amorce, d'un niveau d'activation plus forte que la cible, induit une importante inhibition latérale sur elle. Ce phénomène est diminué dans la condition BF - HF.

Dans certains modèles, les mots de haute fréquence lexicale ont également des connexions avec les phonèmes qui le composent plus performantes. Une excitation top-down pourrait donc intervenir et creuser l'écart entre ces deux conditions. Selon ce mécanisme, à l'écoute de l'amorce « bord », l'effet de rétroaction du niveau lexical vers les phonèmes /b/, /ɔ/ et /ʁ/ renforce son propre niveau d'activation et donc son pouvoir inhibiteur sur le mot-cible de basse fréquence lexicale.

L'effet de transposition de phonèmes observé dans notre expérience ne peut être dû qu'à la présence d'un ou plusieurs phonèmes mal placés, ce qui remet en question le codage de la position absolue des segments postulé par les modèles classiques.

L'effet de facilitation mesuré n'est pas généré par le partage vocalique car la comparaison avec cette condition contrôle exclut une telle hypothèse. Il n'est pas non plus l'effet d'un recouvrement phonétique au niveau de la consonne initiale puisque l'équipe de Toscano (Toscano et al., 2013) a déjà écarté cette possibilité.

Parmi les modèles classiques, seuls TRACE et Shortlist prédisent l'activation du mot-cible dès la présentation de l'amorce, grâce au mécanisme d'alignement exhaustif. Selon ces modèles, à l'écoute de « robe », « bord » est activé à l'occurrence du phonème

/b/. Mais ce mécanisme ne suffit pas à expliquer sa facilitation de traitement lors de la tâche de décision lexicale.

En effet, nous avons vu que les mécanismes d'inhibition latérale et d'excitation top-down interféraient avec l'activation du mot-cible. Ils sont encore plus performants si ce dernier n'est activé qu'à la fin de la réalisation de l'amorce, comme le suggèrent TRACE et Shortlist. Aussitôt activé, le mot « robe » est inhibé par « bord » et, comme l'entrée auditive est interrompue, plus aucune activation ne vient le renforcer. Aussi, la stimulation des détecteurs infra-lexicaux /b/, /ɔ/ et /ʁ/, ainsi que l'éventuelle rétroaction top-down, profite à l'amorce « bord » mais pas au mot-cible « robe », qui entre trop tard dans la compétition lexicale.

La présence d'un effet facilitateur laisse donc penser que le mot-cible est activé plus tôt, dès le /b/ de « bord » pour l'item « robe ». Le mot-cible bénéficierait alors pleinement du partage des mêmes phonèmes que l'amorce, ce qui pourrait expliquer un effet de transposition. Mais cette hypothèse est incompatible avec les modèles qui utilisent des détecteurs infra-lexicaux spécifiques à une position dans le mot car le phonème /b/ ne se situe qu'en troisième position de « robe ». Une manière plus souple de coder la position des segments s'avère donc nécessaire.

Un modèle récent, inspiré de la reconnaissance des mots écrits, le modèle TICS (Hannagan et al., 2013), développe une nouvelle hypothèse compatible avec nos résultats.

Selon les auteurs, les détecteurs infra-lexicaux ne sont pas spécifiques à une position de l'entrée auditive et c'est le codage de la position relative des unités qui assure la reconnaissance du bon mot. Ce codage temporel intervient après la reconnaissance des segments, à un niveau infra-lexical supérieur à celui des phonèmes, correspondant aux biphones extraits du signal. Plus précisément, les auteurs décrivent 4 niveaux de traitement : la représentation d'entrée, le niveau des phonèmes, le niveau des n-phones (phonèmes + biphones) et le niveau lexical.

La grande nouveauté de ce modèle est que les niveaux infra-lexicaux et lexicaux ont une certaine indépendance temporelle vis-à-vis du signal. Les détecteurs de segments activent toutes les entrées lexicales qui contiennent les unités reconnues. Appliqué à notre expérimentation, les mêmes détecteurs de phonèmes /b/, /ɔ/ et /ʁ/ sont stimulés, que l'amorce soit « bord » ou « robe », et les deux candidats sont activés de manière concomitante. Le mot-cible étant envisagé dès la prononciation du premier phonème de l'amorce, son niveau d'activation augmente avec l'occurrence des phonèmes suivants.

C'est la présence des biphones qui assure la reconnaissance du mot-cible dans ce modèle. Ils suivent la direction du signal et diffèrent donc entre items transposés. Les biphones de « robe » sont /ʁɔ/, /ʁb/ et /ɔb/ alors que ceux de « bord » sont /bɔ/, /ɔʁ/ et /bʁ/.

TICS ressemble à TRACE pour le reste de son fonctionnement (alignement exhaustif, mécanisme d'excitation « bottom-up et compétition lexicale) à l'exception qu'il n'a pas encore développé de mécanisme d'excitation top-down. Mais grâce à ce codage de la position relative des segments, il prédit, pour sa part, un effet de transposition.

Dans le modèle TICS, le mot-cible « robe » est aussi inhibé par « bord » mais moins que dans le modèle TRACE. En effet, par rapport à ce dernier, la différence d'activation entre les deux items est bien plus faible et un effet de transposition semble pouvoir être mesuré. Mais aucune simulation de ce type n'a été réalisée à ce jour.

Le modèle TICS semble donc adapté pour prédire l'effet de transposition de phonèmes mais de nombreuses hypothèses sont envisageables et attendent d'être développées et intégrées au sein d'autres modèles cohérents.

La difficulté de cette nouvelle donnée pour les modèles est d'autoriser une large activation de candidats tout en assurant que le mot-cible soit bien reconnu. De plus, un codage plus grossier de la position des segments se fait au détriment du mot-cible et de son avantage dans la compétition lexicale. En effet, dans un modèle comme TICS, les mots-cibles ont un niveau d'activation proche des items transposés et le bénéfice de traitement obtenu par la répétition du même item en amorce et en cible est moindre. L'effet de répétition est donc diminué.

Dans leur article (Toscano et al., 2013), Toscano et ses collaborateurs émettent l'hypothèse que l'ordre temporel des segments n'est pas du tout codé. Ils supposent que ce sont des différences acoustiques fines qui garantissent le respect de la séquence des segments et donc la reconnaissance du bon mot. En effet, comme nous l'avons évoqué précédemment, la coarticulation génère des exemplaires variés d'un même phonème, appelés allophones, en fonction de leur position dans le mot. Selon cette hypothèse, la confusion entre items transposés est possible car les segments qui les composent ne sont pas complètement équivalents. Par exemple, le phonème /ʁ/ n'est pas prononcé tout à fait de la même manière dans « bord » ou dans « robe ».

Cette hypothèse suppose que des traits acoustiques fins influencent l'activation lexicale, ce qui est effectivement envisagé par certains auteurs (Marslen-Wilson and Warren, 1994). Nous pourrions alors imaginer un modèle où les détecteurs de phonèmes activent des entrées lexicales quelles que soient leur position dans le mot et qu'un niveau inférieur, composé de détecteurs de traits, envoie une activation supplémentaire vers le niveau lexical afin d'opter pour le bon mot.

Mais une telle architecture s'éloigne des modèles existants et de nombreuses questions restent en suspens. Notamment, combien de traits acoustiques seraient nécessaires à la distinction de tous les mots et quel est le prix à payer en termes de coût de traitement ?

Rappelons que notre travail présente certaines limites, notamment parce que les expérimentations sur des mots isolés ne reflètent pas complètement le traitement cognitif de la parole.

Tout d'abord, notre protocole n'est pas à l'abri d'un biais de sélection : les sujets ont été choisis au hasard, dans la limite de nos critères d'inclusion, et notre échantillon ne représente donc peut-être pas parfaitement la population cible.

De plus, nous n'avons pas abordé dans ce mémoire les effets du contexte sémantique et syntaxique à l'œuvre dans la parole. Or, le contexte pourrait influencer l'activation lexicale voire restreindre le nombre de candidats. Certaines études montrent que c'est le cas dans certaines conditions (Tabossi, 1988), (Spinelli and Alario, 2002), (Dahan et al., 2000).

L'avantage d'une telle hypothèse est qu'elle suggère que tous les candidats respectant les contraintes d'alignement soumises par les modèles ne sont pas effectivement envisagés en condition réelle. Ainsi, elle permettrait aux modèles autorisant l'activation d'entrées lexicales éloignées de l'input d'économiser une partie de leur coût de traitement. En effet, malgré une grande capacité de calcul, le processus de traitement des mots dans la parole est probablement optimisé par le cerveau des individus, ce que ne peut pas faire un modèle computationnel.

Une collaboration entre les psycholinguistes, qui estiment les calculs nécessaires à la reconnaissance du bon mot, et les chercheurs qui tentent de simuler des modèles à intelligence artificielle, capables d'apprentissages, permettrait peut-être de nous approcher un peu plus encore de notre fonctionnement cérébral à l'écoute de la parole.

Enfin, la question du rôle de l'ordre temporel des segments reste ouverte et ce champ d'étude promet d'affiner notre compréhension du développement phonologique, notamment dans le cadre de troubles du langage.

Contrairement à l'hypothèse des modèles classiques, notre interprétation de l'effet de transposition suggère que l'ordre temporel des segments n'est pas concomitant à leur identification. Il admet la possibilité d'un traitement spécifique de l'agencement temporel des segments et ainsi l'existence d'un nouveau module à intégrer au sein des modèles de la reconnaissance des mots parlés. De nombreuses questions cliniques se posent alors. Est-ce qu'un trouble isolé du traitement temporel des segments est envisageable ? Est-ce qu'un déficit de ce module pourrait expliquer certaines erreurs d'inversion de phonèmes, commises notamment par les enfants présentant un trouble développemental du langage ?

Bien que le traitement de l'agencement temporel des segments soit très spécifique et intriqué au sein de nombreuses compétences cognitives, la manière de l'appréhender peut avoir une certaine influence sur notre compréhension du développement phonologique et d'éventuelles pathologies. Par exemple, les deux

hypothèses que nous avons retenues, pour intégrer l'effet de transposition de phonèmes, n'admettent pas la même interprétation d'un déficit de traitement de l'ordre des segments. D'après le modèle TICS, un tel dysfonctionnement interviendrait au niveau du codage des biphones et serait dépendant de bonnes capacités de catégorisation phonémique. En revanche, si l'ordre temporel des segments n'était pas codé, comme le propose l'équipe de Toscano, des difficultés dans ce domaine seraient davantage liées à un défaut de perception acoustique fine. Ainsi, l'orthophonie s'enrichira peut-être des avancées de la recherche dans ce champ de la psycholinguistique.

VI) ANNEXES

	Mot hF	NEWF2	NPHONS	NSYLL	PUPHONS	Durée	Mot bF	NEWF2	NPHONS	NSYLL	PUPHONS	Durée
1	bord	142	3	1	4	641	robe	85	3	1	4	659
2	chaque	483	3	1	4	686	cache	28	3	1	4	537
3	chair	185	3	1	4	736	rèche	4	3	1	3	664
4	dire	789	3	1	4	624	ride	13	3	1	4	565
5	mal	393	3	1	4	574	lame	30	3	1	4	596
6	mille	129	3	1	4	483	lime	2	3	1	4	659
7	nulle	102	3	1	4	506	lune	54	3	1	4	605
8	mur	201	3	1	4	646	rhume	2	3	1	4	605
9	tard	247	3	1	4	596	rate	3	3	1	4	677
10	sac	92	3	1	4	639	casse	23	3	1	4	650
11	corps	393	3	1	4	596	roc	26	3	1	4	637
12	salle	199	3	1	4	650	lasse	9	3	1	4	718
13	mort	426	3	1	4	691	rhum	6	3	1	4	569
14	part	5217	3	1	4	614	rape	1	3	1	4	677
15	sors	101	3	1	4	732	rosse	0	3	1	4	596
16	sure	4458	3	1	4	695	russe	72	3	1	4	686
17	tous	964	3	1	4	628	soute	1	3	1	4	610
18	tache	122	3	1	4	560	chatte	16	3	1	4	714
19	barre	66	3	1	4	695	rabe	0	3	1	4	628
20	poche	81	3	1	4	506	chope	1	3	1	4	637
21	douce	55	3	1	4	668	soude	4	3	1	4	632
22	jour	912	3	1	4	592	rouge	218	3	1	4	590
23	sel	760	3	1	4	736	laisse	144	3	1	4	659
24	patte	87	3	1	4	560	tape	12	3	1	4	533
25	verre	961	3	1	4	700	rêve	114	3	1	4	714
26	tire	74	3	1	4	637	rite	25	3	1	4	655

Condition transposée

	Mot bF	NEWF2	NPHONS	NSYLL	PUPHONS	Durée	Mot hF	NEWF2	NPHONS	NSYLL	PUPHONS	Durée
1	vol	49	3	1	4	554	somme	259	3	1	4	593
2	vase	23	3	1	4	735	femme	599	3	1	4	615
3	pelle	8	3	1	4	515	belle	240	3	1	4	650
4	pile	25	3	1	4	519	ville	290	3	1	4	568
5	tasse	19	3	1	4	681	car	472	3	1	4	623
6	fiche	20	3	1	4	773	type	250	3	1	4	546
7	puce	9	3	1	4	494	dur	131	3	1	4	682
8	fugue	4	3	1	4	686	sud	79	3	1	4	592
9	bave	3	3	1	4	714	face	255	3	1	4	763
10	vache	36	3	1	4	621	dame	101	3	1	4	560
11	loge	19	3	1	4	664	bonne	281	3	1	4	655
12	matte	5	3	1	4	659	date	77	3	1	4	682
13	botte	25	3	1	4	610	sol	153	3	1	4	630
14	gaffe	11	3	1	4	686	bal	75	3	1	4	723
15	vote	27	3	1	4	605	comme	2929	3	1	4	470
16	bulle	13	3	1	4	558	lutte	64	3	1	4	664
17	bouc	5	3	1	4	580	pour	5332	3	1	4	601
18	dalle	17	3	1	4	601	basse	75	3	1	4	723
19	galle	1	3	1	4	641	passe	201	3	1	4	479
20	loque	5	3	1	4	619	nord	114	3	1	4	695
21	loupe	4	3	1	4	574	cours	479	3	1	4	592
22	bouche	160	3	1	4	702	doute	305	3	1	4	628
23	fête	248	3	1	4	714	mer	750	3	1	4	637
24	gaz	56	3	1	4	723	rare	88	3	1	4	750
25	neige	52	3	1	4	636	même	2653	3	1	4	664
26	quiche	0	3	1	4	555	vide	154	3	1	4	575

Condition contrôle avec partage vocalique

	Amorce	Cible (non-mot)	Condition
1	jambe	bamje	Transposé
2	chance	sanche	Transposé
3	jupe	puje	Transposé
4	couche	chouc	Transposé
5	bac	cabe	Transposé
6	mouche	choume	Transposé
7	pince	sainpe	Transposé
8	tombe	bonte	Transposé
9	vite	tive	Transposé
10	cure	ruque	Transposé
11	moule	loume	Transposé
12	latte	tale	Transposé
13	bête	têbe	Transposé
14	signe	liffe	Ctr avec Voyelle
15	pire	mide	Ctr avec Voyelle
16	duc	tuve	Ctr avec Voyelle
17	nonne	gotte	Ctr avec Voyelle
18	fente	lenbe	Ctr avec Voyelle
19	mythe	bire	Ctr avec Voyelle
20	pull	duve	Ctr avec Voyelle
21	case	chabe	Ctr avec Voyelle
22	beige	fède	Ctr avec Voyelle
23	four	boute	Ctr avec Voyelle
24	ronce	londe	Ctr avec Voyelle
25	nage	saffe	Ctr avec Voyelle
26	banque	genffe	Ctr avec Voyelle

Condition avec non-mots

	Amorce	Cible	Lexicalité
1	calque	grotte	Mot-Mot
2	carpe	danse	Mot-Mot
3	charme	poire	Mot-Mot
4	corse	sang	Mot-Mot
5	serpe	grave	Mot-Mot
6	vaste	frère	Mot-Mot
7	borne	champ	Mot-Mot
8	brun	chien	Mot-Mot
9	merle	blanche	Mot-Mot
10	cable	chiot	Mot-Mot
11	verte	deuil	Mot-Mot
12	prix	blond	Mot-Mot
13	brave	cloche	Mot-Mot
14	cing	froid	Mot-Mot
15	crainte	luge	Mot-Mot
16	grec	digne	Mot-Mot
17	boire	genre	Mot-Mot
18	braille	ciel	Mot-Mot
19	brie	dieu	Mot-Mot
20	cake	drap	Mot-Mot
21	brousse	pulpe	Mot-Mot
22	drain	chat	Mot-Mot
23	tri	kyste	Mot-Mot
24	bourse	fuite	Mot-Mot
25	foin	rue	Mot-Mot
26	larme	flou	Mot-Mot
27	fric	plaque	Mot-Mot
28	gland	fraise	Mot-Mot
29	flot	riz	Mot-Mot
30	suite	trait	Mot-Mot
31	crampe	fière	Mot-Mot
32	crème	gag	Mot-Mot
33	cruche	gain	Mot-Mot
34	brosse	file	Mot-Mot
35	guerre	poste	Mot-Mot
36	crâne	tigre	Mot-Mot
37	trappe	miel	Mot-Mot
38	fringue	star	Mot-Mot
39	stock	trousse	Mot-Mot
40	grippe	stabe	Mot-NonMot
41	glace	viope	Mot-NonMot
42	brute	plaibe	Mot-NonMot
43	groupe	flanne	Mot-NonMot
44	place	myobe	Mot-NonMot

45	plante	filn	Mot-NonMot
46	claque	poime	Mot-NonMot
47	prime	cleude	Mot-NonMot
48	bref	plinde	Mot-NonMot
49	fleur	miède	Mot-NonMot
50	brique	spode	Mot-NonMot
51	braise	volp	Mot-NonMot
52	crise	blade	Mot-NonMot
53	frite	blèque	Mot-NonMot
54	frappe	styre	Mot-NonMot
55	graine	mioffe	Mot-NonMot
56	grappe	floche	Mot-NonMot
57	plaine	stanbe	Mot-NonMot
58	liane	crouche	Mot-NonMot
59	lynx	grine	Mot-NonMot
60	grille	clote	Mot-NonMot
61	piège	brugne	Mot-NonMot
62	couette	gloppe	Mot-NonMot
63	store	plande	Mot-NonMot
64	duel	priffe	Mot-NonMot
65	belge	prove	Mot-NonMot
66	tresse	blope	Mot-NonMot
67	trêve	snope	Mot-NonMot
68	masque	tranffe	Mot-NonMot
69	corde	lieffe	Mot-NonMot
70	troupe	flaine	Mot-NonMot
71	chouette	trope	Mot-NonMot
72	parc	solpe	Mot-NonMot
73	match	frine	Mot-NonMot
74	score	plaine	Mot-NonMot
75	foire	culde	Mot-NonMot
76	buste	crède	Mot-NonMot
77	geste	bourpe	Mot-NonMot
78	carte	bli	Mot-NonMot

Matériel de remplissage

VII) BIBLIOGRAPHIE

Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K., 1998. Tracking the time course of spoken word recognition using eye movements : evidence for continuous mapping models. *J. Mem. Lang.* 38, 419-439.

Colin, M.-P., Cousseau, S., Dufour, S., Aix-Marseille Université Faculté de médecine, Aix-Marseille Université, n.d. L'ordre temporel des phonèmes dans la reconnaissance des mots parlés.

Connine, C.M., Blasko, D.G., Titone, D., 1993. Do the beginnings of spoken words have a special status in auditory word recognition ? *J. Mem. Lang.* 32, 193-210.

Connine, C.M., Mullennix, J., Shernoff, E., Yelen, J., 1990. Word familiarity and frequency in visual and auditory word recognition. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 1084-1096.

Connine, C.M., Titone, D., Deelman, T., Blasko, D., 1997. Similarity mapping in spoken word recognition. *J. Mem. Lang.* 37, 463-480.

Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *J. Cogn. Neurosci.* 6, 256-266.

Content, A., Kearns, R.K., Frauenfelder, U.H., 2001. Boundaries versus onsets in syllabic segmentation. *J. Mem. Lang.* 45, 177-199.

Cutler, A., Carter, D.M., 1987. The predominance of strong initial syllables in the English vocabulary. *Comput. Speech Lang.* 2, 133-142.

Dahan, D., Swingle, D., Tanenhaus, M.K., Magnuson, J.S., 2000. Linguistic gender and spoken-word recognition in French. *J. Mem*

Dufour, S., Frauenfelder, U.H., 2010. Phonological neighbourhood effects in French spoken-word recognition. *Q. J. Exp. Psychol.* 63, 226-238.

Dufour, S., Peereman, R., 2003. Inhibitory priming effects in auditory word recognition : when the target's competitors conflict with the prime word. *Cognition* 88, B33-B44.

- Dupoux, E., Mehler, J., 1990. Monitoring the lexicon with normal and compressed speech : frequency effects and the prelexical code. *J. Mem. Lang.* 29, 316–335.
- Gow, D.W., Gordon, P.C., 1995. Lexical and prelexical influences on word segmentation : evidence from priming. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 344–359.
- Hannagan, T., Magnuson, J.S., Grainger, J., 2013. Spoken word recognition without a TRACE. *Front. Psychol.* 4, 563.
- Kuhl, P.K., 2000. A new view of language acquisition. *Proc. Natl. Acad. Sci. U. S. A.* 97, 11850–11857.
- Luce, P.A., Cluff, M.S., 1998. Delayed commitment in spoken word recognition : evidence from cross-modal priming. *Percept. Psychophys.* 60, 484–490.
- Luce, P.A., Pisoni, D.B., 1998. Recognizing spoken words : the neighborhood activation model. *Ear Hear.* 19, 1–36.
- Luce, P.A., Pisoni, D.B., Goldinger, S.D., 1990. Similarity neighborhoods of spoken words, in : *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. MIT Press, p. 122.
- Maillart, C., Van Reybroeck, M., Alegria, J., 2005. Représentations phonologiques et troubles du développement linguistique : théorie et évaluation.
- Marslen-Wilson, W., Warren, P., 1994. Levels of perceptual representation and process in lexical access : words, phonemes, and features. *Psychol. Rev.* 101, 653–675.
- Marslen-Wilson, W., Zwitserlood, P., 1989. Accessing spoken words: the importance of word onsets. *J. Exp. Psychol. Hum. Percept. Perform.* 15, 576–585.
- 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. *Cognit. Psychol.* 10, 29–63.

McClelland, J.L., Elman, J.L., 1986. The TRACE model of speech perception. *Cognit. Psychol.* 18, 1-86.

McClelland, J.L., Rumelhart, D.E., 1981. An interactive activation model of context effects in letter perception : Part I. An account of basic findings. *Psychol. Rev.* 88, 375-407.

McQueen, J.M., Norris, D., Cutler, A., 1994. Competition in spoken word recognition : spotting words in other words. *J. Exp. Psychol. -Learn. Mem. Cogn.* 20, 621-638.

Morais, J., Cary, L., Alegria, J., Bertelson, P., 1979. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7, 323-331.

Norris, D., 1994. Shortlist : a connectionist model of continuous speech recognition. *Cognition* 52, 189-234.

Perea, M., Lupker, S.J., 2004. Can CANISO activate CASINO ? Transposed-letter similarity effects with nonadjacent letter positions. *J. Mem. Lang.* 51, 231-246.

Perea, M., Lupker, S.J., 2003. Does jugde activate COURT ? Transposed-letter similarity effects in masked associative priming. *Mem. Cognit.* 31, 829-841.

Peressotti, F., Grainger, J., 1999. The role of letter identity and letter position in orthographic priming. *Percept. Psychophys.* 61, 691-706.

Radeau, M., Morais, J., Segui, J., 1995. Phonological priming between monosyllabic spoken words. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1297-1311.

Read, C., Yun-Fei, Z., Hong-Yin, N., Bao-Qing, D., 1986. The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition* 24, 31-44.

Shillcock, R., 1990. Lexical hypotheses in continuous speech, in: *Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives*. MIT Press, p. 24.

Spinelli, E., Alario, F.-X., 2002. Gender context effects on homophone words. *Lang. Cogn. Process.* 17, 457-469.

Tabossi, P., 1988. Accessing lexical ambiguity in different types of sentential contexts. *J. Mem. Lang.* 27, 324-340.

Toscano, J.C., Anderson, N.D., McMurray, B., 2013. Reconsidering the role of temporal order in spoken word recognition. *Psychon. Bull. Rev.* 20, 981-987.

Warren, R.M., 1970. Perceptual restoration of missing speech sounds. *Science* 167, 392-393.

Zwitserslood, P., 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32, 25-64.

Résumé

Le processus inconscient de sélection d'une forme mentale abstraite à partir du signal physique de parole est un domaine d'étude de la psycholinguistique. Les chercheurs recueillent des données sur le traitement cognitif des mots oraux à l'aide de techniques comportementales et essaient de les organiser au sein d'un modèle cohérent. Avant d'opter pour la bonne entrée lexicale, le cerveau envisagerait plusieurs hypothèses sur la base d'un code infra-lexical ajusté aux exigences du signal. Les mécanismes d'appariement et de sélection standards suggèrent que le jeu de candidats, issu de différents points de l'input, est trié en fonction de l'actualisation des données perceptives et de l'influence de connaissances lexicales intégrées. Tous les modèles classiques s'accordent sur le principe d'un appariement séquentiel en totale conformité avec l'ordre du signal mais aucun n'explique de façon précise et réaliste comment notre cerveau parvient à cette invariance temporelle. Des chercheurs (Toscano et al., 2013), inspirés d'expérimentations menées à l'écrit, ont récemment mesuré une activation de candidats différant du mot entendu par l'ordre de leurs phonèmes, suggérant un codage de la position des segments plus grossier que ce qui est actuellement envisagé par les modèles de référence.

Dans la lignée d'un précédent mémoire d'orthophonie, notre travail vise à collecter des preuves en français de l'effet découvert par l'équipe de Toscano et à montrer, comme elle, qu'il ne témoigne pas d'un simple effet pré-lexical, mais bien de la considération d'hypothèses non conformes à l'ordre du mot traité. Avec un paradigme d'amorçage phonologique, nous avons mesuré l'activation de mots aux phonèmes permutés tout en appréciant l'influence de la fréquence relative entre amorces et cibles. Les données recueillies avec ce protocole confirment notre hypothèse de recherche : l'effet d'amorçage est uniquement détectable lorsque l'amorce est moins fréquente que la cible, ce qui lui confère un locus lexical. Cette interprétation des résultats encourage la formulation de nouvelles hypothèses concernant le codage de l'ordre temporel des segments dans la reconnaissance des mots parlés.

Mots-clés : orthophonie, reconnaissance des mots, parole, ordre temporel, phonème, phonologie, amorçage phonologique