



**HAL**  
open science

**Stimulation dichotique ou musicale pour l'aide aux  
apprenants de l'anglais L2, dyslexiques ou non :  
conception et validation d'un programme  
d'entraînement informatisé**

Margot Bouhon

► **To cite this version:**

Margot Bouhon. Stimulation dichotique ou musicale pour l'aide aux apprenants de l'anglais L2, dyslexiques ou non : conception et validation d'un programme d'entraînement informatisé. Sciences cognitives. Université de Lyon, 2022. Français. NNT : 2022LYSE1090 . tel-04125923

**HAL Id: tel-04125923**

**<https://theses.hal.science/tel-04125923v1>**

Submitted on 12 Jun 2023

**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.

N° d'ordre NNT : 2022LYSE1090



**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale 476**  
**(Neurosciences et Cognition)**

**Spécialité de doctorat : Sciences Cognitives**

Soutenue publiquement le 24/06/2022, par :

**Margot Bouhon**

---

**Stimulation dichotique ou musicale  
pour l'aide aux apprenants de l'anglais  
L2, dyslexiques ou non : conception et  
validation d'un programme  
d'entraînement informatisé**

---

Devant le jury composé de :

Rossetti, Yves	PUPH	CRNL - Université Lyon 1	Président
Ferragne, Emmanuel	MCF	CLILLAC- ARP – Université Paris Cité	Rapporteur
Habib, Michel	PH	LNC-CHU Marseille	Rapporteur
Colin, Cécile	PR	URNC - Université Libre de Bruxelles	Examinatrice
Tillmann, Barbara	DR	CRNL - CNRS	Co-directrice de thèse
Bedoin, Nathalie	MCF	CRNL – Université Lyon 2	Co-directrice de thèse
Ranvier, Marion		Contentsquare – Contentsquare Foundation	Invitée



A mon papa,

## Résumé

---

L'apprentissage de l'anglais est un défi pour les francophones, surtout en cas de dyslexie. Une série d'expériences teste de nouveaux dispositifs d'aide pour cet apprentissage tardif chez des adultes dyslexiques ou non. Leurs principes s'inspirent du modèle de l'asymétrie de l'échantillonnage temporel (AST) selon lequel l'hémisphère gauche (HG) et l'hémisphère droit (HD) participent au traitement temporel de façons complémentaires. L'HG offrirait des fenêtres d'analyse étroites et l'HD des fenêtres plus larges. Des difficultés liées au traitement de la durée pour de nouvelles consonnes et voyelles anglaises sont étudiées. En français, le VOT des occlusives sourdes est court, celui des occlusives sonores est long, mais c'est le contraire en anglais. L'efficacité d'un programme d'exercices perceptifs (2h30) ciblant cette particularité est testé (Expériences 2, 6, 8 et 9) avec des épreuves de perception catégorielle insensibles au test-retest (Expériences 1) et deux expériences de oddball passif en potentiels évoqués (Expériences 3 et 7). Les exercices présentaient des mots de façon binaurale ou dans une seule oreille (controlatérale à l'hémisphère supposé compétent pour la durée du VOT, bruit blanc dans l'autre oreille). Ce dernier mode de présentation (dichotique) devait optimiser l'effet de l'entraînement pour le voisement des consonnes, et la durée des voyelles /ɪ/ et /i:/ (Expériences 11 et 12) en perception et production (Expérience 10). Les résultats montrent l'amélioration de la perception catégorielle de consonnes opposées en voisement après l'entraînement perceptif chez les adultes dyslexiques ou non, encore plus avec le dispositif dichotique. La sensibilité pré-attentionnelle à l'occlusive sonore anglaise croît, et l'implication de l'HG et de l'HD change et devient plus adaptée aux VOTs après l'entraînement dichotique. Cela pourrait traduire l'établissement de nouvelles représentations phonémiques et le dépassement de l'assimilation au voisement en français. Ce changement phonologique concorde avec le modèle SLM-r pour qui l'apprentissage tardif de phonèmes est possible. L'entraînement perceptif améliore la perception des voyelles /ɪ/-/i:/ avec transfert en production même chez les dyslexiques, mais la stimulation dichotique ne renforce pas ces progrès.

Un programme d'entraînement très bref (45 min) par répétition de phrases ciblait par ailleurs la différence entre les intonations montante/descendantes des questions fermées/ouverte en anglais (Expériences 13 et 14). Les résultats montrent l'efficacité d'un amorçage rythmique global pendant l'entraînement (Expérience 13 et 14), surtout pour les adultes dyslexiques. L'effet positif des amorces est renforcé par un retour visuel sur les courbes d'intonation.

Les résultats nécessitent des répliques avant toute généralisation. Ils encouragent à explorer la piste de la stimulation dichotique pour l'aide à l'apprentissage en anglais, et pour la remédiation de déficits du traitement d'indices phonétiques temporels en cas de dyslexie.

**Mots clés : Attention temporelle ; Apprentissage d'une langue seconde ; VOT ; Prosodie ; Dyslexie**

**Title: Dichotic or musical stimulation to assist dyslexic and non-dyslexic L2 English learners: design and validation of a computer-based training program.**

## **Abstract**

---

Learning English is challenging for French speakers, especially in cases of dyslexia. There has been a series of experiments with new assistive devices for late learning in dyslexic and non-dyslexic adults. Their principles are based on the Asymmetric Sampling in Time (AST) model, according to which the left hemisphere (LH) and the right hemisphere (RH) participate in temporal processing in complementary ways. The LH would provide narrow windows of analysis and the RH would provide wider windows. Difficulties related to duration processing for new English consonants and vowels were studied. In French, the VOT of voiceless stop consonants is short, that of voiced consonants is long, but it is the opposite in English. The efficiency of a perceptual exercise program (2.5 hours) targeting this particularity is tested (Experiments 2, 6, 8 and 9) with test-retest insensitive categorical perception tests (Experiment 1) and two passive oddball evoked potential experiments (Experiments 3 and 7). The exercises presented words binaurally or in one ear only (contralateral to the hemisphere assumed to be competent for the duration of the VOT, white noise in the other ear). The latter presentation mode (dichotic) was intended to optimize the effect of training for consonant voicing, and vowel duration /ɪ/ and /i:/ (Experiments 11 and 12) in perception and production (Experiment 10). The results show improvement in categorical perception of opposite consonants in voicing after perceptual training in adults with and without dyslexia, even more so with the dichotic device. Pre-attentional sensitivity to English sound occlusion increases, and the involvement of HG and HD changes becomes more responsive to VOTs after dichotic training. This could reflect the establishment of new phonemic representations and the overcoming of assimilation to voicing in French. This phonological change is consistent with the SLM-r model for which late phoneme learning is possible. Perceptual training improves vowel perception of /ɪ-/i:/ with transfer to production even in dyslexics, but dichotic stimulation does not reinforce this progress.

A very brief (45 min) sentence repetition training program targeted the difference between rising/falling intonations of closed/open questions in English (Experiments 13 and 14). The results show the efficiency of global rhythmic priming during training (Experiments

13 and 14), especially for dyslexic adults. The positive effect of priming is reinforced by visual feedback on intonation curves.

The results require replications before any generalization. They encourage the exploration of dichotic stimulation as an aid to learning in English, and for the remediation of deficits in the processing of temporal phonetic cues in dyslexia.

**Keywords: Temporal attention ; Second language learning ; VOT ; Prosody ; Dyslexia**



## **Financement et laboratoire de rattachement**

---

Cette thèse s'est déroulée sous un contrat de Convention Industrielle de Formation par la Recherche (CIFRE) entre l'entreprise Contentsquare SAS et le Centre de Recherche en Neurosciences de Lyon (CRNL, Inserm U1028 - CNRS UMR 5292, Université Lyon 1).

## Remerciements

---

Je souhaite tout d'abord remercier mes trois formidables directrices : Nathalie Bedoin, Barbara Tillmann et Marion Ranvier.

Nathalie, je voudrais te remercier pour ces six années passées à tes côtés. Merci d'avoir toujours été présente pour moi, dans les bons et les mauvais moments. Tu es quelqu'un d'exceptionnelle et mon mentor. Je ne serais jamais arrivée jusqu'ici sans toi. Merci de m'avoir trouvé une thèse auprès de Marion, merci de m'avoir poussé ou écouté quand j'en avais besoin. Merci d'avoir toujours valorisé mon travail, même quand ce n'était pas fameux. Merci d'avoir toujours écouté mes propositions, interprétations et idées farfelues. Merci de m'avoir fait confiance et de me suivre dans mes millions de projets. J'ai pour toi, la femme, la chercheur et l'enseignante un respect incommensurable. Je suis pressée de pouvoir continuer à travailler à tes côtés encore de nombreuses années sur pleins de nouveaux projets !

Barbara, un immense merci pour ton aide et ton soutien tout au long de cette thèse. Merci pour ces conversations de travail toujours plus stimulantes et passionnantes, pour tes retours sur le manuscrit toujours plus rapide malgré le délai court. Merci d'avoir été là quand j'en avais besoin. Je te voue un profond respect, tu es une chercheur comme il en existe peu et je suis très fière d'avoir pu faire cette thèse sous ta direction.

Marion ... Je ne sais pas par où commencer ... Je ne me rendais pas compte, il y a de ça 4 ans, que la jeune femme que j'avais rencontré dans la salle polyvalente de Marcy l'Etoile, allait devenir une des personnes les plus importantes dans ma vie, et une auxquelles je ferais le plus confiance. Merci d'avoir toujours été à mes côtés, merci de m'avoir poussé, merci de m'avoir réconforté. Tu es mon modèle et quand je serais grande, je veux être comme toi. Une femme forte, intelligente, battante et par-dessus le marché une super maman. Je n'ai pas assez de mots pour te dire l'affection et le respect que j'ai pour toi... Merci de m'avoir convaincu de te suivre dans cette folle aventure de Contentsquare ! Merci de m'avoir donné la chance de te prouver ma valeur. J'aime travailler à tes côtés et j'espère que ça durera pendant très longtemps ! Enfin, merci de m'avoir aidé à construire mon avenir, nous allons faire de grandes choses ensemble ! Rendez-vous à la fondation !

Je me sens aujourd'hui très fière de dire que j'ai été encadrée pendant ces trois années par vous trois.

J'aimerais ensuite remercier les deux équipes qui m'ont accueilli, l'équipe Trajectoires et l'équipe CAP. J'aimerais en particulier remercier mes copains de bureaux, Julie, Anna, Aurélien et Raphaël. Merci de m'avoir accompagné et faire rire ! Je renouvelle mes excuses pour les odeurs dues à mes différents plats (Tacos, kebab, hamburger, etc.) qui embaumaient le bureau ! Julie, je suis heureuse de te compter parmi mes amies, ces années on les a faites ensemble et on s'est soutenues, je resterais présente pour toi jusqu'au bout ! Anna, I'll miss you so much when you go back to Australia. But I'm sooo happy for you, it gonna be amazing! I hope I can come and visit you soon, and, of course, I'll bring some wine! Aurél, je te souhaite tout le bonheur du monde et je l'espère, très vite une thèse ! Merci d'avoir été avec moi toutes ces années ! Raph', je te souhaite plein de bonheur et de réussite dans ta thèse, tu vas tout déchirer ! Merci aussi à toi, Marie, pour ces nombreuses conversations, déjeuner et café au soleil, tes conseils ont été précieux. Et surtout, d'avoir été patiente même si je repoussais nos rendez-vous 10 fois !

Je souhaite aussi remercier toutes les personnes du Centre de Recherche en Neurosciences de Lyon pour m'avoir permis de réaliser ce travail de thèse dans des conditions idéales.

Merci à Claire, Sandy, Anne-Lise, Elise et Margot, des étudiantes formidables que j'ai eu la chance de co-encadrer, et sans qui je n'aurais pas réussi tout ce travail. Je voudrais tout particulièrement te remercier, Clairette. Si tu n'avais pas été là, cette fin de thèse aurait pu être une catastrophe. Merci d'être la jeune femme que tu es, c'est un vrai bonheur de travailler avec toi. Tu es intelligente et déterminée. J'espère réellement pouvoir continuer à travailler avec toi, je ferais tout pour ! Mais avant ça, finissons ton mémoire !

Je voudrais aussi remercier mes gars sûrs, Kévin, Anthony et Antonin. Depuis Aidodys, en passant par AdapteMonWeb jusqu'à Contentsquare, vous n'avez eu de cesse que de me soutenir et m'encourager. Merci à Antonin d'avoir toujours pris du temps pour me faire répéter mes présentations, à Anthony d'avoir servi de cobaye pour tester mes programmes et à Kévin, pour ces marches de réflexions autour des bouts d'herbe place Antonin Poncet quand j'étais bloquée. Merci aussi à mes nouveaux copains du bureau lyonnais, vous n'êtes pas tous très agréable (coucou Valentin) mais je vous aime quand même beaucoup !

Je voudrais tout particulièrement remercier pour leur gentillesse, aide et patience Taze et Tash. Vous êtes deux personnes formidables, à vos côtés, je me sens forte et à ma place et ça, grâce à vous. Merci de me soutenir autant, d'être présent pour moi, de me protéger et

d'être aussi bienveillant envers moi. Vous êtes mes amis, et je suis très heureuse de travailler tous les jours avec vous ! (Ps : Tash, j'ai fait mon maximum pour rendre ce manuscrit le plus accessible possible).

Je voudrais aussi remercier ma nouvelle famille de la Contentsquare Foundation et la Csquad ! Je suis heureuse de venir travailler tous les jours et de rencontrer des personnes toujours plus formidables les unes que les autres. Je voudrais aussi remercier particulièrement Isa, je te souhaite tout le bonheur du monde, tu seras une maman formidable et je suis pressée de rencontrer le premier bébé de la CSF !

Je voudrais ensuite remercier mes amis, Hugo, Martintin, Sarah, Julien, Anna, Samy, Kim et Florian pour leur soutien infailible durant toutes ces années. Merci de me supporter !

Philipouille, tu es l'une de mes meilleures amies et une femme formidable. J'aime nos petits-déjeuners, nos déjeuners, dîners, soirées, j'aime tous les moments passés à tes côtés. Merci de m'avoir toujours écouté et fait rire quand j'en avais besoin. On s'est rencontré chez Boulanger pendant nos études, et nous sommes aujourd'hui deux jeunes femmes presque accomplies. On en a fait du chemin ensemble !

Jennifer, ma Jen. Depuis quelques mois, j'ai découvert une nouvelle facette de toi, celle d'une maman exceptionnelle. Tu es pour moi une personne formidable et une amie qui m'est très chère. Depuis ce jour, où l'on ne se connaissait pas, mais où tu m'as simplement réconfortée, jusqu'à aujourd'hui, tu as toujours répondu présente pour moi et je ne saurais comment te montrer ma gratitude. J'ai eu de la chance dans mon malheur, que ma route croise la tienne.

Jadou, mon soleil. Tu t'es enfui loin de moi (oui, je t'en veux toujours.), dans un pays où il pleut tout le temps et où on mange de frites, mais sans ça, nous n'aurions peut-être pas eu ces milliers d'heures au téléphone à tout et rien se dire. Je chéris notre amitié, et je veux que tu saches que malgré la distance, tu resteras toujours dans mon cœur. Merci d'être toujours douce et chaleureuse avec moi. Je me sens toujours mieux après t'avoir entendu. Je voulais juste que tu saches la place importante que tu as pour moi, dans ma vie.

Agathe, ma belle Agathe, ma forte Agathe. Tu es tellement généreuse et intelligente. J'aime être ton amie, j'aime passer ces moments formidables à tes côtés. Je te souhaite tellement... Tellement de bonheur, d'amour, de joie, de rire, d'épanouissement ! Merci

d'avoir été présente à chaque instant de ces dernières années, merci de m'avoir écouté me plaindre, rire et pleurer. Tout simplement merci d'être mon amie.

Coco, capitaine. Tu es mon meilleur ami, tu es un homme extraordinaire. Tu es drôle, gentil, sensible, prévenant et d'un soutien sans faille. Merci de toujours répondre présent quand j'en ai besoin. Merci.

Je voudrais aussi remercier mes amis des camions du cœur. Je suis arrivé avant ma thèse, vous avez donc suivi tout ce looong chemin. Merci tout particulièrement à Evelyne, Xavier, Philippe, Jojo et Iuri. Vous cinq avez été mes piliers, vous m'avez soutenu, poussé et vous me redonnez toujours confiance en moi pour aller toujours plus loin. Je vous souhaite à tous les cinq, le plus grand bonheur possible. Et promis, maintenant, je suis de retour !

Je souhaite aussi remercier ma famille, ma très grande famille. Celle que j'ai depuis ma naissance et celle que j'ai gagné depuis 10 ans. Merci à mes grands-oncles et tantes, tantes et oncles, cousins, cousines ! Je me rends compte de la chance que j'ai d'avoir une famille aussi formidable que la nôtre. Merci particulièrement de m'avoir nourri (pour les tatas) et fait rire toutes ces années. Merci à la famille Pillon de m'avoir accueillie comme si j'étais l'une des vôtres. Je suis très heureuse de vous compter parmi ma famille. Et merci tout particulièrement à toi, Agnès pour toutes nos conversations toujours plus intéressantes et de me soutenir dans mes projets. J'ai gagné une deuxième maman ! Je voudrais aussi avoir une pensée pour Mamé, qui nous regarde d'en haut, j'aurais aimé que tu sois présente, toi qui m'as toujours encouragée.

Je voudrais dire un merci particulier à vous deux, Clémie et Camo. Avec tout ce qu'on a vécu, vous m'aimez toujours. Vous êtes bien plus que mes cousines, vous êtes les sœurs que je n'ai pas eues. Je vous aime du plus profond de mon cœur et je suis heureuse, chaque jour, de vous avoir près de moi.

Merci aussi à ma marraine, Ninie. Merci de me nourrir, de me supporter depuis ma naissance et d'être aussi formidable avec moi. Merci d'être ma marraine et mon guide.

Maman, j'aimerais te dire à quel point je t'aime et je suis fière d'être ta fille. Malgré toutes les épreuves, tu as toujours été forte. Tu es là pour moi à chaque instant, pour me pousser, me motiver, me consoler et me faire rire. Tu es la meilleure maman dont on puisse rêver et je mesure chaque jour l'immense chance que j'ai de t'avoir. Je ne te remercierais jamais assez de tout ce que tu fais pour moi, à chaque minute de chaque jour. Je t'aime.

Baba, mon baba, mon grand frère, le meilleur des grands frères. Tu es mon roc et mon modèle dans la vie. La vie ne nous a pas épargné, mais tu es toujours resté là, près de moi, à me soutenir et m'encourager (et me disputer parfois). Tu as amené dans nos vies Anne Saucisse, la meilleure belle-sœur dont je pouvais rêver. Merci de me l'avoir fait connaître. Merci à toi Anne-So de prendre soin de lui, de nous. Je vous aime tellement tous les deux, et je suis tellement heureuse de l'arrivée prochaine de ma nièce ! Papa serait fier de toi Baba, tout comme je le suis.

Je voudrais aussi avoir une pensée pour mon papa, lui dédier ce manuscrit n'était pas suffisant. Il était un homme formidable, de qui je tiens mon caractère de battante. Il m'a toujours poussée à vouloir plus, et mieux. J'espère que tu es fière de moi depuis là-haut, ta sauterelle va devenir docteur.

Et enfin, le meilleur pour la fin, l'amour de ma vie, Rémy. Tu me fais rire, tu me fais pleurer et tu me rends dingue, mais merci de m'aimer comme tu le fais chaque jour. Tu es à mes côtés depuis presque 10 ans. Tu étais là pour chacune de mes victoires, chacune de mes défaites. Merci de m'avoir pris dans tes bras lorsque je désespérais, merci de me faire sentir comme la femme la plus formidable et forte du monde. Merci d'être toi avec moi. Ces années difficiles sont derrière nous et promis j'arrête les études. On va enfin pouvoir construire notre vie à deux. Je t'aime.

# Table des matières

<b>Résumé</b>	<b>iiiv</b>
<b>Abstract</b>	<b>vi</b>
<b>Financement et laboratoire de rattachement</b>	<b>viiiviii</b>
<b>Remerciements</b>	<b>lix</b>
<b>Table des matières</b>	<b>xiv</b>
<b>Préambule</b>	<b>1</b>
<b>Chapitre 1. Cadre théorique et expérimental de la recherche</b>	<b>3</b>
1.1. Séquenciation du signal de parole et attention temporelle	3
1.2. Oscillations cérébrales et traitement temporel du signal de parole	6
1.2.1. Théorie de l'échantillonnage asymétrique (AST)	6
1.2.2. Arguments pour l'asymétrie hémisphérique dans le modèle AST	10
1.2.3. Mécanismes de synchronisations des rythmes cérébraux au rythme de la parole	16
1.2.4. Fonctionnement hiérarchisé et couplages des oscillations cérébrales	19
1.2.5. Implication de traitements de haut niveau dans l'analyse temporelle	27
1.3. Pathologies neurodéveloppementales du langage : explications basées sur l'attention temporelle et les oscillations cérébrales	31
1.3.1. Théorie de l'échantillonnage temporel et dyslexie	33
1.4. Sensibilité à la durée : mesures en potentiels évoqués (ERP)	37
1.4.1. Les ondes pré-attentionnelles	37
1.4.2. Les ondes attentionnelles	40
1.5. Objectifs de la thèse	42
1.6. Méthodologie de la thèse	48
<b>Chapitre 2. Perception catégorielle</b>	<b>50</b>
2.1. Définitions : perception catégorielle, allophonique, et frontière catégorielle	50
2.2. Epreuves d'évaluation de la perception catégorielle : identification et discrimination	51
2.2.1. Méthodes d'évaluation de la perception catégorielle	53
2.3. Mise en place développementale lente de la perception catégorielle	54
2.4. Position de la frontière catégorielle pour le voisement des occlusives	54
2.5. Supports neuroanatomiques de la perception catégorielle	55
2.6. Perception catégorielle et dyslexie	56
2.7. Manque d'outils d'évaluation en français	59

2.8. Le voisement des consonnes occlusives .....	59
2.8.1. Définition du voisement.....	59
2.8.2. Indices de durée pour le voisement et asymétrie hémisphérique.....	60
2.9. Objectif.....	62
2.10. Expérience 1 : Evaluation comportementale de la perception catégorielle.....	63
2.10.1. Principe d'expérience et hypothèses .....	63
2.10.2. Evaluation de la perception catégorielle de consonnes en français.....	63
2.10.3. Evaluation de l'effet test-retest et de la fidélité .....	65
2.10.4. Analyse des données.....	65
2.10.5. Résultats.....	67
2.10.6. Evaluation de la spécificité et de la sensibilité : Courbes ROC.....	70
2.10.7. Analyse des données.....	70
2.10.8. Résultats.....	72
2.11. Discussion sur les tests de perception catégorielle en français (Expérience 1).....	74
2.11.1. Qualités diagnostiques des tests et paire de consonnes pour évaluer la PC.....	74
2.11.2. Défis de l'évaluation de la perception catégorielle chez des adultes dyslexiques et normo-lecteurs à l'université.....	76
<b>Chapitre 3. Remédiation de la perception du voisement en français pour adultes dyslexiques... 79</b>	
3.1. Outils contemporains pour la rééducation phonologique .....	79
3.2. Objectif.....	82
3.3. Expérience 2 : effet de l'entraînement sur la perception catégorielle .....	84
3.3.1. Principe de l'expérience et hypothèses.....	84
3.3.2. Méthode .....	84
3.3.3. Analyses de données.....	88
3.3.4. Résultats.....	89
3.4. Expérience 3 : effet de l'entraînement sur les potentiels évoqués .....	91
3.4.1. Principe de l'expérience et hypothèse .....	91
3.4.2. Méthode .....	92
3.4.3. Analyses de données.....	94
3.4.4. Résultats.....	95
3.5. Discussion sur l'effet d'un entraînement en français chez des adultes dyslexiques (Expériences 2 et 3).....	99
3.5.1. Différences phonologiques entre adultes dyslexiques et normo-lecteurs .....	99



3.5.2.	Effets de l'entraînement perceptif par stimulation dichotique sur la perception du voisement en français : données comportementales.....	103
3.5.3.	Effets de l'entraînement perceptif par stimulation dichotique sur la perception du voisement en français : données en potentiels évoqués .....	105
3.5.4.	Intérêts et limites d'un entraînement bref à la phonologie française chez des adultes dyslexiques .....	107
<b>Chapitre 4. Evaluation de la perception du voisement en anglais en L2 : conception d'outils ..</b>		<b>109</b>
4.1.	Difficultés d'apprentissage de l'anglais L2 : le constat.....	109
4.2.	Les modèles de l'apprentissage d'une seconde langue.....	110
4.3.	Le cas de l'anglais L2 pour les francophones .....	114
4.4.	Dyslexie et apprentissage de l'anglais L2.....	118
4.5.	Objectifs.....	122
4.6.	Expérience 4 : Evaluation de qualités psychométriques des épreuves de perception catégorielle en anglais .....	123
4.6.1.	Principe de l'expérience et hypothèses .....	123
4.6.2.	Méthode .....	123
4.6.3.	Analyses des données .....	124
4.6.4.	Résultats.....	124
4.7.	Expérience 5 : Comparaison des performances en anglais des adultes dyslexiques et des adultes normo-lecteurs.....	129
4.7.1.	Principe d'expérience et hypothèses .....	129
4.7.2.	Méthode .....	129
4.7.3.	Analyses de données.....	130
4.7.4.	Résultats.....	130
4.8.	Discussion sur les tests de perception catégorielle anglaise (Expériences 4 et 5) .....	133
4.8.1.	Qualités diagnostiques des tests.....	133
4.8.2.	Différence de PC entre dyslexiques et normo-lecteurs à l'âge adulte : une réalité ? .....	136
<b>Chapitre 5. Aide à l'apprentissage de la perception du voisement anglais .....</b>		<b>139</b>
5.1.	Dispositifs pour l'apprentissage de nouvelles consonnes en L2.....	139
5.2.	Objectif.....	145
5.3.	Expérience 6 : effets des entraînements par stimulation dichotique vs. binaurale sur la PC anglaise chez des adultes normo-lecteurs .....	147
5.3.1.	Principe de l'expérience et hypothèses .....	147
5.3.2.	Méthode .....	148
5.3.3.	Analyses des données .....	153

5.3.4.	Résultats .....	153
5.3.5.	Discussion intermédiaire.....	157
5.4.	Expérience 7 : effet de l'entraînement dichotique sur les potentiels évoqués chez des adultes non-dyslexiques .....	157
5.4.1.	Principe de l'expérience et hypothèses .....	157
5.4.2.	Méthode .....	158
5.4.3.	Analyse de données .....	160
5.4.4.	Résultats.....	161
5.4.5.	Discussion intermédiaire.....	166
5.5.	Expérience 8 : effet de l'entraînement dichotique en anglais sur la PC en anglais chez des adultes dyslexiques.....	167
5.5.1.	Principe de l'expérience et hypothèses .....	167
5.5.2.	Méthode .....	167
5.5.3.	Analyses de données.....	168
5.5.4.	Résultats.....	168
5.5.5.	Discussion intermédiaire.....	170
5.6.	Expérience 9 : effet de l'entraînement en anglais sur la PC en français chez des adultes normo-lecteurs et dyslexiques .....	171
5.6.1.	Principe de l'expérience et hypothèses .....	171
5.6.2.	Méthode .....	171
5.6.3.	Analyse de données .....	171
5.6.4.	Résultats.....	171
5.6.5.	Discussion intermédiaire.....	173
5.7.	Discussion sur l'aide à l'apprentissage de la perception du voisement en anglais (Expériences 6 à 9).....	173
5.7.1.	Difficulté d'apprentissage tardif de l'opposition de voisement en anglais : un argument supplémentaire en ERP .....	174
5.7.2.	Efficacité d'un entraînement perceptif bref (binaural) pour la phonologie chez des apprenants tardifs de l'anglais L2 .....	178
5.7.3.	Efficacité de la stimulation dichotique pour optimiser l'effet d'entraînement perceptif au voisement en anglais L2 .....	180
5.7.4.	Aspects cognitifs et neuronaux de l'effet du dispositif dichotique sur l'apprentissage du voisement anglais .....	185
5.7.5.	La stimulation dichotique comme aide à l'apprentissage du voisement en anglais L2 chez des adultes dyslexiques .....	187
5.7.6.	La stimulation dichotique : un entraînement favorisant le traitement des sons courts anglais.....	189
5.7.7.	Processus cognitifs sous-jacents : mécanismes <i>top-down</i> et <i>bottom-up</i> ?.....	191

5.7.8.	Processus cognitifs sous-jacents : une amélioration des traitements temporels ?.....	192
5.7.9.	Limites et résultats inattendus .....	194
<b>Chapitre 6. Aide à l'apprentissage de la perception des durées des voyelles anglaises.....</b>		<b>197</b>
6.1.	La différence entre voyelles courtes et longues en anglais .....	197
6.2.	Entraînements disponibles sur la durée des voyelles .....	200
6.3.	Objectifs.....	203
6.4.	Expérience 10 : Evaluation de qualités psychométriques des épreuves de perception et de production de mots en anglais .....	206
6.4.1.	Principe de l'expérience et hypothèses .....	206
6.4.2.	Méthode .....	206
6.4.3.	Analyse des données.....	209
6.4.4.	Résultats .....	209
6.4.5.	Discussion intermédiaire.....	215
6.5.	Expérience 11 : effets des entraînements par stimulation dichotique vs. binaurale sur la perception et la production de voyelles anglaises chez des adultes non-dyslexiques .....	216
6.5.1.	Principe de l'expérience et hypothèses .....	216
6.5.2.	Méthode .....	216
6.5.3.	Analyses des données .....	221
6.5.4.	Résultats : comparaison des groupes GD et GB à T1.....	221
6.5.5.	Résultats : effet de l'entraînement par stimulation dichotique .....	223
6.5.6.	Résultats : effet de l'entraînement par stimulation binaurale .....	225
6.5.7.	Discussion intermédiaire.....	227
6.6.	Expérience 12 : effets de l'entraînement dichotique sur la perception et la production de voyelles anglaises chez des adultes dyslexiques.....	229
6.6.1.	Principe de l'expérience et hypothèses .....	229
6.6.2.	Méthode .....	229
6.6.3.	Analyses des données .....	230
6.6.4.	Résultats : Comparaison des performances entre les groupes NL et DYS avant l'entraînement (à T1).....	231
6.6.5.	Résultats : Effet de l'entraînement dichotique pour le groupe DYS .....	234
6.6.6.	Discussion intermédiaire.....	237
6.7.	Discussion sur l'aide à l'apprentissage de perception des durées des voyelles anglaises (Expériences 10, 11 et 12).....	240
6.7.1.	De nouveaux tests d'évaluation de la perception et de la production des voyelles anglaises.....	241

6.7.2.	Apports des entraînements perceptifs pour l'apprentissage du contraste de durée des voyelles anglaises chez les adultes normo-lecteurs.....	244
6.7.3.	Effet de l'entraînement dichotique chez les adultes normo-lecteurs.....	247
6.7.4.	Effet de l'entraînement dichotique chez les adultes dyslexiques.....	248
<b>Chapitre 7. Aide à l'apprentissage de la prosodie anglaise .....</b>		<b>250</b>
7.1.	La prosodie : une nouvelle entrave à l'apprentissage de l'anglais L2.....	250
7.2.	Musique, oscillations cérébrales et prosodie.....	254
7.2.1.	Synchronisation des oscillations cérébrales aux rythmes de la musique et de la parole . .....	254
7.2.2.	Régularité du rythme musical et aides aux traitements langagiers .....	257
7.3.	Dyslexie, prosodie et grammaire .....	260
7.4.	Entraînements rythmiques et utilisation d'un retour visuel comme aides au langage.....	262
7.4.1.	Effets de ces entraînements sur les traitements dans la langue maternelle .....	262
7.4.2.	Effets de ces entraînements sur les traitements dans une L2.....	264
7.5.	Objectifs.....	266
7.6.	Expérience 13 : comparaison de trois entraînements chez des adultes normo-lecteurs .....	269
7.6.1.	Principe de l'expérience et hypothèses .....	269
7.6.2.	Méthode .....	270
7.6.3.	Analyse des données.....	274
7.6.4.	Résultats.....	275
7.6.5.	Discussion intermédiaire.....	278
7.7.	Expérience 14 : Effet de l'entraînement avec amorçage rythmique et retour visuel chez des adultes dyslexiques.....	279
7.7.1.	Principe de l'expérience et hypothèses .....	279
7.7.2.	Méthode .....	280
7.7.3.	Analyses de données.....	281
7.7.4.	Résultats : Effet de l'entraînement MR chez les adultes dyslexiques – comparaison T1 – T2 .....	281
7.7.5.	Discussion intermédiaire.....	284
7.8.	Discussion sur l'aide à l'apprentissage de la prosodie anglaise (Expériences 13 et 14) .....	285
7.8.1.	Une certaine compétence initiale pour la prosodie des questions anglaises .....	287
7.8.2.	Apprentissage implicite d'une particularité en prosodie anglaise.....	287
7.8.3.	Effet d'amorçage musical rythmique sur la prosodie : liens avec les effets sur l'analyse grammaticale .....	289
7.8.4.	Un amorçage rythmique basé sur la hiérarchie des oscillations .....	290

7.8.5.	Un dispositif adapté aux apprenants les plus en difficulté .....	293
7.8.6.	Association efficace de l’amorçage rythmique et du retour visuel dans le cadre théorique PRISM.....	294
7.8.7.	Un dispositif d’amorçage et de retour visuel adapté aux apprenants dyslexiques .....	295
7.8.8.	Limites et perspectives pour des dispositifs d’aide à l’apprentissage brefs .....	297
<b>Chapitre 8. Discussion générale .....</b>		<b>299</b>
8.1.	Rappel des objectifs, des choix théoriques et justification des choix méthodologiques .....	299
8.2.	Synthèse de la confrontation des hypothèses générales aux principaux résultats.....	302
8.3.	Apprentissage de phonèmes anglais par des adultes francophones normo-lecteurs : origine et dépassement des difficultés .....	304
8.3.1.	Origine des difficultés .....	304
8.3.2.	Dépassement des difficultés .....	305
8.3.3.	Contribution aux modèles d’apprentissage d’une L2 et aux modèles de traitement temporel de la parole .....	307
8.4.	Apprentissage de phonèmes par des adultes francophones dyslexiques : origine et dépassement des difficultés .....	311
8.4.1.	Origine des difficultés .....	311
8.4.2.	Dépassement des difficultés .....	312
8.4.3.	Contribution aux modèles d’apprentissage d’une L2 et de déficit des traitements temporels dans la dyslexie .....	314
8.5.	Pré-ajustement de l’attention temporelle pour le traitement d’information prosodiques .....	315
8.5.1.	Intérêt d’un amorçage rythmique global pour le traitement du langage .....	315
8.5.2.	Rôle d’une cascade oscillatoire dans l’amorçage rythmique sur la prosodie .....	316
8.5.3.	Apport du retour visuel pour un apprentissage prosodique en anglais .....	320
8.6.	Limites.....	321
8.7.	Perspectives pour de futures recherches et applications .....	325
8.8.	Conclusion .....	328
<b>Références</b>	.....	<b>330</b>
<b>Annexes</b>	.....	<b>376</b>

## Préambule

---

L'apprentissage de l'anglais en tant que seconde langue (L2) est un réel défi pour les apprenants francophones, d'autant plus en cas de dyslexie. Cet apprentissage est pourtant nécessaire, car plus de la moitié des Français utilisent l'anglais au quotidien. Cette thèse, réalisée dans le cadre d'un contrat CIFRE, présente une série d'expériences testant l'efficacité de nouveaux dispositifs d'aide pour cet apprentissage tardif chez des adultes dyslexiques ou normo-lecteurs. Leurs principes s'inspirent de modèles issus de la recherche fondamentale. Il s'agit notamment du modèle de l'asymétrie de l'échantillonnage temporel (AST) selon lequel les deux hémisphères cérébraux jouent des rôles complémentaires dans le traitement temporel du signal sonore. L'hémisphère gauche (HG) offrirait des fenêtres d'analyse étroites et l'hémisphère droit (HD) des fenêtres plus larges. Cette proposition, cohérente avec la répartition asymétrique d'oscillations cérébrales rapides et lentes dans le cerveau, est valide sur le plan du neuronal.

Ce cadre théorique sera le fil conducteur des études présentées, tant du point de vue de la méthodologie que de l'interprétation des résultats. C'est pourquoi, le Chapitre 1 présente une revue de question faisant une synthèse des idées et des résultats expérimentaux au sujet de l'attention temporelle, mais aussi de la mécanique des oscillations cérébrales, au-delà de la simple opposition entre des fenêtres d'analyse de durées distinctes dans les cortex auditifs gauche et droit (HG/HD). Ce cadre théorique permet d'établir un lien entre les méthodologies utilisées, l'apprentissage d'une L2 et les troubles du langage avec l'exemple de la dyslexie.

Le lecteur de ce travail pourra suivre différents parcours de lecture en fonction de ses intérêts. Un lecteur intéressé par les mécanismes sous-tendant l'attention temporelle et les oscillations cérébrales pourrait ainsi s'adresser directement au Chapitre 1. Un lecteur cherchant surtout à en savoir plus sur des applications de ces modèles à l'apprentissage de l'anglais L2 pourra plutôt s'orienter vers les Chapitres 4 à 7. Enfin, un lecteur intéressé par la dyslexie et les remédiations de leur système phonologique pourra se tourner vers les Chapitres 1 (partie 1.3) à 3 pour la L1, et les Chapitres 4 à 7 pour l'apprentissage de la L2. Après le Chapitre 1, au début de chaque chapitre, des points théoriques spécifiques aux expériences décrites sont présentés.

Les objectifs de chaque étude sont synthétisés dans le Chapitre 1 (partie 1.5) et la méthodologie générale à toute la thèse se trouve dans la partie 1.6. L'annexe 3 sera aussi utile

pour comprendre comment les participants étaient répartis dans les expériences. Pour faciliter la lecture, les expériences sont en effet décrites de façon indépendante, alors qu'elles portent parfois, en tout ou partie, sur de mêmes participants. Enfin, la discussion générale (Chapitre 8) reviendra sur les modèles, avant de présenter les limites et les perspectives envisagées.

# Chapitre 1. Cadre théorique et expérimental de la recherche

---

## 1.1. Séquenciation du signal de parole et attention temporelle

En perception auditive, la séquenciation du flux sonore est essentielle. Il s'agit d'un ensemble de mécanismes perceptifs qui permet d'extraire des unités dans une succession d'évènements acoustiques et de les stocker en mémoire à court terme. La séquenciation est un processus temporel complexe qui consiste à intégrer l'information à la fois dans des unités (*chunks*), mais aussi dans une structure d'ensemble, car les unités sont souvent imbriquées. Pour se dérouler efficacement, la séquenciation implique des mécanismes cognitifs variés, comme la prise en compte d'éléments contextuels et l'anticipation perceptive, la première guidant en partie la seconde. Le découpage temporel du signal est, en effet, adapté s'il prend en compte le rythme antérieur de la séquence, pour attendre des instants précis où surviendront des évènements sonores importants. Cela suppose que l'auditeur attend une certaine régularité dans la séquence qu'il traite. Cette séquenciation doit néanmoins être flexible pour s'adapter aux imperfections du rythme du signal et rester cohérente : sa temporalité doit s'ajuster, se synchroniser au rythme de l'information entrante. Des influences de haut niveau peuvent participer au processus de séquenciation, à travers l'activation de connaissances sur les unités dans le domaine considéré (musique ou langage), leur durée, et les structures habituelles au domaine. Ce processus perceptif/cognitif de séquenciation participe à son tour à un traitement de plus haut niveau. Il permet en effet l'élaboration d'une représentation en mémoire à court terme temporellement plus régulière que le flux acoustique. Il offre alors une base apte à faciliter les traitements de nature interprétative.

Lorsqu'il s'agit du signal de parole, le processus de séquenciation est indispensable à l'analyse linguistique. Pour que l'analyse soit adaptée au langage, les unités à détecter sont aussi bien de longs groupes de mots que des unités lexicales, des syllabes, des phonèmes, ou des traits phonétiques. La séquenciation est alors complexe, car elle suppose une sensibilité et un principe actif à plusieurs échelles, aussi cohérentes entre elles que le sont les unités linguistiques emboîtées. Elle participe ainsi au stockage mnésique d'une information structurée temporellement sous la forme d'une représentation perceptive hiérarchisée et cohérente.

Sur le plan cognitif, la séquenciation du signal de parole mobilise des compétences en attention temporelle. Ce terme renvoie aux règles de distribution des ressources de traitement



dans le temps, par analogie avec l'attention spatiale qui rassemble les mécanismes de distribution des ressources dans l'espace. L'attention temporelle implique des mécanismes de synchronisation entre le rythme de diffusion des ressources de traitement cognitif et le rythme du signal. À partir du contexte, ces ajustements permettent à leur tour des anticipations sur ce qui se produira ensuite et surtout sur le moment précis où surviendront les événements. Dans le cas du signal de parole, les anticipations peuvent être modulées par des influences de haut niveau, telles que des connaissances sur des phonèmes, des enchaînements possibles de sons de parole, des probabilités de survenue de certains accords morphosyntaxiques à la fin des mots en fonction de ceux qui précèdent. Autant de connaissances et compétences issues d'un apprentissage le plus souvent statistique à partir des expériences linguistiques.

Du point de vue psychologique, il existe un cadre théorique sur la manière dont la répartition de l'attention se fait dans le temps. Il invite à la réflexion sur la séquenciation du signal de parole. La Théorie de l'Attention Dynamique (DAT, Jones, 1976, 2008 ; Jones & Boltz, 1989 ; Jones *et al.*, 2002), initialement conçue pour expliquer le traitement temporel de la musique, présente la répartition de l'attention comme fondamentalement cyclique. Autrement dit, une fois déclenchée, la diffusion des ressources d'analyse se poursuivrait de façon rythmique. Cette distribution non-linéaire donnerait lieu à une succession discontinue d'instantanés où l'information serait traitée de manière optimale, entrecoupée d'instantanés où elle serait moins bien perçue. La distribution discontinue de l'attention permettrait donc au système perceptif d'agir sur la chaîne acoustique en extrayant des informations plus profondément à certains instants, au rythme d'une succession rapide d'engagements et de désengagements de l'attention. Cela permettrait la séquenciation, c'est-à-dire le découpage du signal avec extraction et intégration d'unités d'information. Idéalement, une bonne cohérence temporelle entre le rythme attentionnel et la durée des unités à extraire est souhaitable, car elle permet de capturer les unités d'information et/ou de favoriser la sensibilité à leur début pour une séquenciation adaptée.

Aujourd'hui, un lien est envisagé entre le rythme de distribution de l'attention et les rythmes internes au cerveau, qui constitueraient leur support électrophysiologique. Dans le cerveau, des groupes de neurones se synchronisent pour émettre spontanément des signaux électromagnétiques à un rythme commun. L'énergie ainsi produite en une succession de phases peut se concevoir comme une onde, définie par une fréquence. Une libération de potentiels d'action se produit lors du pic de chaque phase et la sensibilité du substrat neuronal s'élève, en alternance avec des instants où l'énergie est retenue, économisée. Pour faire le lien

avec une conception psychologique de la perception, le pic dans la phase d'un rythme cérébral correspondrait à un instant où les ressources attentionnelles sont diffusées pour un traitement optimal (Schroeder *et al.*, 2008). Lorsqu'un rythme cérébral produit de manière endogène, implique suffisamment de neurones coordonnés pour déterminer les variations d'excitabilité du support neuronal, il est décrit comme une oscillation cérébrale. Les oscillations cérébrales façonneraient les opérations perceptives et attentionnelles rythmant le cycle engagement/désengagement de l'attention (Vanrullen & Dubois, 2011).

Idéalement, une bonne cohérence temporelle entre le rythme cérébral et la durée des unités à extraire est souhaitable. Elle permet alors de découper le signal en capturant les unités d'information et/ou en favoriserait la sensibilité à leur début, pour une séquenciation adaptée. Lorsqu'il s'agit de langage, cette harmonie est possible si les unités d'information se présentent de façon régulière, conformément aux attentes (Giraud & Poeppel, 2012), grâce à un débit de parole peu fluctuant. Des anticipations sont alors possibles sur l'instant où devrait survenir les prochaines informations pertinentes. Cette cohérence est aussi favorisée par les capacités d'ajustement des oscillations cérébrales, dont le rythme poursuivrait celui de l'information entrante, dans une tentative permanente de synchronisation (mécanisme d'*entrainment*), jusqu'à l'ajustement de phase. La parole n'étant jamais très régulière, des modifications des rythmes cérébraux seraient sans cesse opérées, d'une manière fluide selon certains chercheurs, à travers des *resets* plus abrupts selon d'autres (Giraud & Poeppel, 2012) mais toujours de manière très dynamique.

Le cerveau produit simultanément des oscillations différentes dans des aires distinctes, offrant une variété de rythmes. Le cerveau des mammifères est animé par des réseaux d'oscillations répartis dans cinq bandes de fréquence (des plus lentes aux plus rapides : delta (0.1-4 Hz), thêta (4-7 Hz), alpha (7-13 Hz), bêta (13-30 Hz), gamma (> 30 Hz)). L'attention peut ainsi s'orienter dans le temps à partir de plusieurs générateurs, rendant possible le traitement de la parole à plusieurs niveaux. Au-delà d'une simple complémentarité des rythmes endogènes, les modèles de traitement de la parole s'intéressent à une potentielle organisation hiérarchique de ces rythmes, avec des relations d'emboîtement qui concorderaient avec l'organisation et la durée des unités linguistiques. La concordance entre des oscillations cérébrales lentes et rapides, d'une part, et les durées typiques des unités linguistiques (notamment les syllabes et les phonèmes), d'autre part, est l'idée centrale du modèle de l'échantillonnage asymétrique développé à partir de 2003 par Poeppel. Ces inter-synchronisations intra-cérébrales s'ajoutent à la tentative de synchronisation des oscillations

aux informations externes. Des imperfections de ce fonctionnement hiérarchisé, de même que des déficits de synchronisation des oscillations au signal acoustique, font l'objet d'hypothèses d'explication neuropsychologiques au sujet de troubles neurodéveloppementaux du langage.

## **1.2. Oscillations cérébrales et traitement temporel du signal de parole**

### **1.2.1. Théorie de l'échantillonnage asymétrique (AST)**

Dans le cadre de recherches sur la perception de la parole, Poeppel propose en 2003 la théorie de l'échantillonnage asymétrique (*Asymmetric Sampling in Time*, AST), qui suppose un traitement bilatéral du signal de parole dans ses premières étapes, avant les opérations linguistiques qui impliquent des combinaisons et des prédictions de nature véritablement linguistique (Poeppel, 2014). Ce traitement bilatéral serait néanmoins asymétrique dans les analyses réalisées, car une différence fonctionnelle fondamentale existerait entre le cortex auditif des deux hémisphères. Cette opposition les rend complémentaires pour traiter des événements selon leur durée : la détection de changements rapides dans la parole (e.g. transition de formants indiquant le lieu d'articulation des consonnes occlusives, ou très brefs changements de fréquence indiquant le début d'une unité prosodique) serait mieux réalisée dans le cortex auditif gauche, alors que des événements plus longs (e.g. groupe de mots, syllabe, indice phonétique long comme l'aspiration des consonnes occlusives sourdes anglaises) seraient mieux traités par le cortex auditif droit. Cette conception n'allait pas de soi lorsqu'elle a émergé en 2003, dans un contexte où l'hémisphère gauche était considéré comme extrêmement prioritaire pour le traitement du langage, en particulier pour la phonologie.

Avant cela, une nuance avait tout de même été proposée par une autre théorie – elle aussi qualifiée d'acoustique – de l'asymétrie hémisphérique pour le traitement de la parole : les aspects temporels et spectraux feraient l'objet d'une dominance respective des hémisphères gauche (HG) et droit (HD) qui expliquerait globalement la spécialisation de l'HG pour le langage (Zatorre, 1997). Pour la parole, l'HG présenterait un avantage dans le traitement d'indices temporels, même si cela ne suffit pas pour un traitement optimal du langage, car sa résolution fréquentielle est plus faible que celle de l'HD (Liégeois-Chauvel *et al.*, 1999, 2001). Un point fort de ce modèle est de rendre compte des dominances hémisphériques opposées, de façon générale, pour le traitement du langage et de la musique à partir de caractéristiques acoustiques de bas niveau, et non d'une différence de domaines

(Zatorre *et al.*, 2002). Les traitements du langage et de la musique sont respectivement plus dépendants des caractéristiques temporelles et spectrales du signal, et une dominance de l'HG pour les traitements temporels dans leur ensemble expliquerait sa dominance pour le langage. Les difficultés spécifiques à traiter les aspects temporels (par opposition à spectraux) du signal auditif chez des patients atteints de lésions gauches (Robin *et al.*, 1990) apportent un premier argument d'ordre neurologique à cette théorie. Celle-ci s'appuie aussi sur les résultats d'expériences présentant des stimuli non-linguistiques variant sur le plan temporel ou spectral. Des données en Tomographie par Émission de Positron (TEP, Zatorre & Belin, 2001) et en IRMf (Jamison *et al.*, 2006 ; Schönwiesner *et al.*, 2005 ; Warrier *et al.*, 2009) montrent ainsi une augmentation de l'activité dans le cortex auditif primaire gauche face à des variations temporelles, mais une modulation de l'activité dans des aires associatives droites en cas de variations spectrales. En opérant séparément des dégradations graduelles des caractéristiques temporelles ou spectrales de phrases et de mélodies, une expérience récente a confirmé les spécialisations complémentaires des deux hémisphères pour des variations spectro-temporelles distinctes (Albouy *et al.*, 2020). Les auteurs ont d'ailleurs conclu que cela permet au langage et à la musique d'exploiter avec efficacité les extrêmes sur un continuum de variations spectro-temporelles. Une expérience en TEP montre aussi qu'en présence de même stimuli CVC (i.e. Consonne-Voyelle-Consonnes), une forte latéralisation droite de l'activité cérébrale se produit s'il est demandé de juger leurs différences de hauteur (pitch, déterminé par des différences spectrales), alors que l'activité est latéralisée à gauche si les jugements portent sur des différences de lieu d'articulation (Zatorre *et al.*, 1992). Notons que dans ce dernier cas, les analyses perceptives nécessaires pour la tâche portent sur des changements temporels très rapides (30-40 ms).

Le modèle AST reste compatible avec le modèle basé sur l'opposition temporel/spectral, en ce sens qu'il reconnaît à l'HG un rôle particulièrement exigeant sur le plan temporel : traiter les éléments les plus rapides. Il conçoit cependant les traitements temporels comme partagés entre les hémisphères, et associe l'HD à une compétence pour la séquenciation avec une fenêtre temporelle plus longue. Cette asymétrie hémisphérique, basée sur les durées, est non seulement compatible avec la participation de l'HD aux traitements syntaxiques et prosodiques, mais elle permet aussi de reconnaître l'intervention de l'HD dans l'analyse élémentaire acoustique/phonétique de la parole. De manière encore plus claire que ne le proposent Zatorre et ses collègues, le modèle AST est aussi compatible avec le fait que les lésions péri-sylviennes gauches entraînent des conséquences plus graves sur le langage

que les lésions droites. Le modèle AST insiste pour cela sur la haute résolution temporelle dont seul l'HG est capable, et qui est indispensable à la perception d'indices rapides dont le signal de parole foisonne, pour identifier correctement les phonèmes. Des données en pathologie appuient cette hypothèse d'un traitement temporel bilatéral, mais distinct à gauche et à droite. L'analyse des difficultés de patients cérébro-lésés atteints d'agnosie verbale sont ainsi plus facilement expliqués si l'on admet que les mécanismes réalisés par l'HG permettent avant tout d'extraire l'information à travers une fenêtre d'intégration de 25-50 ms, alors que ceux de l'HD le font surtout à travers une fenêtre plus longue, de 150-250ms (Poeppel, 2001). En-dehors d'une pathologie, c'est la rapidité des changements qui impose le plus grand défi à l'auditeur du signal de parole, même lorsqu'il s'agit de changements rapides de nature spectrale, comme la transition de formants qui matérialise acoustiquement une opposition comme /ba/-/da/ (Giraud & Poeppel, 2012). Ainsi s'expliquerait la dominance générale classique de l'HG dans le traitement du langage, selon le modèle AST.

Pour fonder son modèle, Poeppel est parti des observations suivantes : 1) le signal de parole contient des informations temporelles importantes dont la saisie relève de différentes échelles temporelles ; 2) des fenêtres d'intégration temporelles variées sont vraisemblablement nécessaires pour organiser un tableau temporel complet des informations phonologiques ; 3) le langage est certes majoritairement dépendant de l'HG, mais bon nombre de données montrent que l'analyse du signal de parole, au niveau phonétique/phonologique, est sous-tendue bilatéralement par le cortex temporal supérieur (Hickok & Poeppel, 2000, 2007). Cela donne du sens aux activations dans le lobe temporal droit et le sillon temporal supérieur droit pour des traitements phonétiques (Belin *et al.*, 2000 ; Buchman *et al.*, 1986 ; Burton *et al.*, 2000 ; Hickok & Poeppel, 2000 ; Poeppel *et al.*, 2004). Le modèle AST découle de ces remarques, en expliquant l'implication de l'HD dans le traitement de la parole par sa compétence élevée pour la séquenciation lente, utile à l'intégration temporelle du signal. Le cortex temporal supérieur permettrait, par son implication bilatérale, la construction de niveaux de représentation et d'analyse distincts, particulièrement aux niveaux acoustique, phonétique et phonologique (Poeppel, 2001). Dans le cortex auditif gauche, les informations seraient extraites à partir d'une fenêtre temporelle étroite (de l'ordre de 20-50 ms) ; dans le cortex auditif droit, elles le seraient à travers une fenêtre d'intégration plus grande (150-250 ms). Dès 1980, Schwartz et Tallal décrivaient l'HG comme particulièrement efficace pour les indices d'une durée de 40 ms et non ceux de 80 ms. Cette distinction est particulièrement pertinente pour la recherche développée dans cette thèse, puisqu'elle renvoie à des durées

d'indices acoustiques présents dans les phonèmes qui seront étudiés. La précision de cette description temporelle est confortée par des données électrophysiologiques et psychophysiques antérieures sur lesquelles le modèle AST s'appuie (Näätänen, 1992 ; Theunissen & Miller, 1995).

Les oscillations cérébrales serviraient d'infrastructure aux traitements du signal de parole, en permettant de découper la chaîne acoustique continue en unités (*chunks*) indispensables aux analyses ultérieures. C'est la similitude entre la durée des unités imbriquées dans la parole et la durée des phases des oscillations disponibles dans les aires auditives qui rend crédible l'idée du rôle médiateur des oscillations dans la séquenciation à échelles multiples (multiplex sampling) pour l'encodage de la parole. La parole implique en effet une hiérarchie de composantes : les unités du niveau prosodique coïncident avec le rythme endogène delta (0,5-2 Hz), les syllabes sont compatibles avec le rythme thêta (4-8 Hz), et la durée de beaucoup de phonèmes coïnciderait avec le rythme gamma (> 30 Hz). En bref, la présence intrinsèque d'oscillations neuronales lentes (delta et thêta) et rapides (gamma), conjointes, mais parallèles dans le cortex auditif primaire, confère une réalité physiologique à deux types de fenêtres temporelles – ou fenêtres d'intégration perceptive – dont l'existence était soupçonnée depuis longtemps dans les études en psychophysique (Viemeister & Wakefield, 1991). Ces rythmes cérébraux intrinsèques donnent aux décharges neuronales déclenchées par les stimulations auditives, une organisation dynamique adaptée à la séquenciation de la parole, car ils sont temporellement cohérents avec l'organisation hiérarchique des unités de celle-ci (Schroeder *et al.*, 2008).

Un point important dans les modèles de séquenciation à échelles multiples du signal de parole est le statut privilégié d'un rythme cérébral lent (thêta) : il serait déclencheur et coordinateur dans l'ajustement rythmique cérébral au signal. Le rôle déterminant des oscillations thêta, par ailleurs importantes pour l'extraction des syllabes, serait fondé sur des caractéristiques physiologiques fondamentales, liées aux composantes articulatoires (motrices) du langage. Par exemple, des stimulations audio-visuelles fournissant des informations articulatoires ont déclenché une activité dans le cortex auditif en correspondance avec l'activité du cortex moteur gauche à une fréquence syllabique (thêta) et non à une fréquence phonémique (Morillon *et al.*, 2010). L'analyse syllabique serait donc fortement ancrée dans un processus perceptivo-moteur basique, qui coïncide avec l'idée évolutionniste d'une émergence du langage à partir de la combinaison de vocalisations syllabiques, de gestes des mains (Morillon *et al.*, 2010), et de mouvements mandibulaires. Ces derniers sont liés à la

mastication sur un rythme similaire à celui des syllabes (Giraud *et al.*, 2007). L'analyse phonémique de la parole serait quant à elle un processus secondaire acquis au cours du développement. Au-delà du rôle fondamental des ondes lentes en présence de signal de parole, l'activité cérébrale recherche activement le rythme de la chaîne acoustique et s'y ajuste à la fois à travers les changements de phase des ondes cérébrales lentes et les changements d'amplitude des ondes rapides (gamma) (Gross *et al.*, 2013). La combinaison des deux permettrait l'encodage temporel le plus complet, avec un échantillonnage à plusieurs échelles, elles-mêmes adaptées à la durée des unités importantes dans la parole (syllabes, phonèmes, traits phonétiques).

Les fenêtres étroites et larges pour l'analyse à plusieurs niveaux de granularité seraient elles-mêmes sous-tendues par les oscillations cérébrales selon leur disponibilité dans l'un ou l'autre hémisphère. Dans le cortex temporal supérieur, les oscillations dans l'HG seraient majoritairement de haute fréquence (Spironelli & Angrilli, 2010) (bêta 12-30 Hz, gamma > 30 Hz) et celles enregistrées dans l'HD relèveraient essentiellement de basses fréquences (delta < 3 Hz, thêta 4-8 Hz). Les oscillations delta, thêta et gamma prédominent dans le cortex auditif (Lakatos *et al.*, 2005). Bien que majoritairement latéralisées à droite, les oscillations thêta seraient présentes de façon bilatérale, ce qui leur permettrait un rôle de coordination entre les aires cérébrales, particulièrement dans le traitement des phonèmes (Morillon *et al.*, 2012).

L'intelligibilité du discours dépend de la séquenciation, elle-même possible grâce à la synchronisation temporelle des réponses du cortex auditif à l'enveloppe des stimuli dans la parole continue. La netteté (*sharpness*) des fluctuations temporelles issues de la cochlée serait un indice important pour s'aligner au rythme des enveloppes syllabiques. Ces événements auditifs saillants renforceraient les oscillations lentes (delta et thêta) et augmenteraient leur tendance à s'adapter au rythme de l'information entrante (Gross *et al.*, 2013). Des fluctuations temporelles amples et/ou brusques inciteraient des oscillations lentes à s'engager dans une recherche de synchronisation (*entrainment*) au rythme des syllabes, ce qui est propice aux étapes de décodage plus élaborées. Ces événements abrupts augmenteraient aussi le couplage entre des rythmes différents de façon intra-cérébrale, favorisant la séquenciation du signal à des échelles multiples et cohérentes (Gross *et al.*, 2013).

### **1.2.2. Arguments pour l'asymétrie hémisphérique dans le modèle AST**

La théorie AST a ouvert la voie à un ensemble de propositions globalement rassemblées dans



les modèles à résolution temporelle multiple ou modèles d'analyse à échelles multiples (*Multi-Time-Resolution Model*, Hickok & Poeppel, 2007). Plus exactement, ces modèles supposent une sensibilité potentiellement bilatérale des deux cortex auditifs aux modulations acoustiques rapides. Les modulations lentes feraient par contre l'objet d'une spécialisation plus ferme du cortex auditif droit. Le tout résulterait en une prise en charge opposée et complémentaire des analyses temporelles selon les échelles.

Au sujet de cette asymétrie hémisphérique pour les traitements temporels, il existe différentes sources d'arguments, qui peuvent être évoqués comme soutien au modèle AST.

#### **1.2.2.1. Arguments neuroanatomiques et cellulaires**

L'asymétrie hémisphérique fonctionnelle pour les traitements temporels selon leur durée est cohérente avec des différences anatomiques entre les deux hémisphères dans les aires auditives. Une différence de taille moyenne a été observée entre les cortex temporaux supérieurs gauche et droit (Geschwind & Levitsky, 1968), ceci même pour le cortex auditif primaire, de plus grand volume dans l'HG (Penhune *et al.*, 1996). Cette aire sensorielle primaire se singularise aussi dans l'HG par son organisation cellulaire : elle contient plus de matière blanche que dans l'HD, ce qui favorise ses connexions avec d'autres aires, notamment le planum tempore gauche. Ce dernier est d'ailleurs lui-même de taille particulièrement grande à gauche et davantage pourvu en matière grise dans cet hémisphère (Anderson *et al.*, 1999 ; Dorsaint-Pierre *et al.*, 2006), ce qui lui permet de recevoir plus de fibres afférentes. La couche de cellules pyramidales est aussi plus épaisse dans le cortex auditif primaire gauche (Hutsler & Galuske, 2003), garantissant cette fois une transmission optimale de l'information par de larges axones vers d'autres régions du cortex auditif, et offrant de bonnes conditions pour une analyse temporelle plus fine du signal. Cet ensemble de qualités pourrait expliquer les capacités des aires auditives de l'HG à prendre en charge le traitement temporel de stimuli auditifs (Penhune *et al.*, 1996), ceci surtout s'ils réclament une haute précision parce qu'ils sont brefs. Pour revenir à la dimension dynamique, Giraud et Poeppel (2012) soulignent que les cellules pyramidales, plus abondantes dans le cortex auditif gauche, permettent la génération d'oscillations gamma dont la rapidité est adaptée à la séquenciation fine du signal et aux traitements temporels rapides.



### **1.2.2.2. Arguments basés sur des données électrophysiologiques observées chez le bébé et le jeune enfant**

Une étude combinant la spectroscopie proche infra-rouge fonctionnelle (fNIR) et l'électroencéphalographie (EEG) chez des nouveau-nés a récemment montré que des aires cérébrales différentes s'activaient déjà à l'écoute de modulations d'amplitude lentes ou rapides dans du français (activations bilatérales pour ce qui est lent, mais latéralisées à gauche pour les changements rapides) (Cabrera & Gervain, 2020). Cette spécialisation hémisphérique précoce établirait les fondations de l'apprentissage du langage dans sa complexité temporelle. Avec des stimuli auditifs non-verbaux mais fluctuants comme de la parole, une étude combinant elle aussi EEG et fNIR a montré que des nouveau-nés traitaient préférentiellement les modulations de 25 à 50 ms (Telkemeyer *et al.*, 2009), durée pertinente pour la perception des phonèmes et des traits phonétiques (Rosen, 1992). Conformément aux modèles à échelles temporelles multiples, les réponses corticales étaient pour cela bilatérales, alors que les réponses à des changements acoustiques lents (160 and 300 ms) étaient latéralisées à droite (Telkemeyer *et al.*, 2009). Deux hypothèses ressortent de cet ensemble : 1) de façon extrêmement précoce au cours du développement chez l'être humain, les réponses cérébrales aux fluctuations temporelles sont bilatérales, 2) mais l'émergence d'une asymétrie en fonction des durées se produit très tôt dans le développement, suggérant que la mise en place des processus de perception de la parole est liée aux capacités fondamentales des processus auditifs.

Chez des jeunes enfants, entre 3 et 5 ans, même l'enregistrement des oscillations cérébrales endogènes au repos montre une asymétrie conforme à celle décrite par le modèle AST (Poeppel, 2003). De plus, chez ces enfants, les oscillations les plus rapides sont d'autant plus latéralisées à gauche dans le cortex auditif qu'ils perçoivent efficacement les mots en contexte bruité, et sont bien équipés pour traiter la parole dans des environnements naturels et complexes (Thompson *et al.*, 2016).

### **1.2.2.3. Arguments basés sur des données électrophysiologiques observées chez l'adulte**

#### Données favorables à la latéralisation droite pour l'analyse temporelle lente

Chez l'adulte, des données en imagerie cérébrale confirment que les cortex auditifs gauche et droit sont sensibles à la structure temporelle des événements auditifs non-verbaux. Certaines expériences révèlent surtout le rôle de l'HD en cas d'évènements acoustiques lents,

conformément au modèle AST. Par exemple, les réactions du sillon temporal supérieur droit sont décrites comme plus fortes que celles du sillon gauche en réponse à des modulations acoustiques lentes (150-300 ms) (Boemio *et al.*, 2005). De même, lorsque le contenu spectral d'une information acoustique varie de manière répétitive et régulière, plus cette modulation est temporellement lente (entre 20 et 305 ms), plus il faut de temps et assez lente, aux fenêtres temporelles pour saisir cette régularité, et plus cela s'accompagne d'une latéralisation droite des activations dans le planum temporale, la partie antérieure du gyrus temporal supérieur et le sillon temporal supérieur (Overath *et al.*, 2008).

#### Données favorables à la latéralisation gauche pour l'analyse temporelle rapide

D'autres travaux soutiennent, de leur côté, l'autre aspect de l'asymétrie hémisphérique supposée par le modèle AST : la dominance gauche pour l'échantillonnage fin du signal. Par exemple, la présence de stimuli auditifs de complexité spectrale variable déclenche des réponses cérébrales dans les hautes fréquences (gamma, 25-60 Hz) plus nettes dans l'HG que dans l'HD (Poeppel *et al.*, 2000). Lorsque des tâches actives sont requises sur le signal, des données en TEP suggèrent aussi un biais vers l'HG pour l'analyse de changements spectraux rapides dans les sons (Belin *et al.*, 1998), pour la détection auditive dans des stimuli (mots ou syllabes) contenant des changements temporels rapides plutôt que dans des voyelles longues (Fiez *et al.*, 1995), et pour la discrimination de sons purs contenant des changements spectraux rapides plutôt que lents (Johnsrude *et al.* 1997). En IRMf, cette fois, des activations restreintes aux cortex auditifs primaire et secondaire gauches ont été observées en cas de perception d'informations temporelles rapides, pour des stimuli artificiels comme pour de la parole naturelle (Zaehle *et al.*, 2004).

#### Données favorables à la complémentarité hémisphérique selon la durée des événements

Certaines données apportent un soutien plus complet à la complémentarité des hémisphères supposée par le modèle AST pour les traitements temporels. Cette complémentarité a été enregistrée en EEG et en IRMf : des oscillations gamma se produisent lors de l'activation du cortex auditif gauche, alors que des oscillations thêta se produisent lors de l'activation du cortex auditif droit (Giraud *et al.*, 2007). Des études en MEG confirment aussi que les transitions acoustiques lentes sont traitées de façon plus robuste par le cortex auditif droit, alors qu'un signal contenant des transitions acoustiques rapides est davantage traité dans le cortex auditif gauche (Shtyrov *et al.*, 2000). Toujours en MEG, il est apparu que lorsque des stimuli non-verbaux ont une structure temporelle, il faut qu'elle corresponde à la durée supposée des fenêtres d'analyse selon le modèle AST (autour de 25 ms ou de 200 ms) pour

provoquer une recherche d'ajustement de la part des oscillations gamma (rapides) et thêta (lentes) (Luo & Poeppel, 2012). Les données montrent aussi que cette recherche active est conduite de manière bilatérale par le rythme gamma, mais plus nettement à droite par le rythme thêta. D'autres expériences en MEG ont également montré qu'une synchronisation tente de s'établir avec la parole dans le cortex auditif à partir de la phase des oscillations lentes (delta, thêta) et à partir de l'amplitude des oscillations rapides. Or, le phénomène d'ajustement de phase est décrit comme plus fort dans le cortex auditif droit et l'ajustement d'amplitude comme plus fort dans le cortex auditif gauche (Gross *et al.*, 2013).

#### 1.2.2.4. Arguments issus d'expériences en écoute dichotique

Il est possible d'étudier des asymétries hémisphériques en audition à partir de données comportementales recueillies à travers les performances de participants dans des tâches d'écoute dichotique. L'auditeur écoute deux stimuli verbaux différents présentés simultanément, l'un à l'oreille droite, l'autre à l'oreille gauche (Kimura, 1967). Il doit répéter à voix haute ce qu'il a entendu. Cette technique crée une concurrence entre les informations, et celle-ci renforce la relation prioritaire entre chaque oreille et le cortex auditif de l'hémisphère controlatéral. Cela permet d'obtenir des indices sur le type d'information particulièrement bien traité dans l'un et l'autre hémisphère. Dans une étude en français (Bedoin *et al.*, 2010), les membres de la paire de mots commençaient tous par une consonne occlusive et ils se distinguaient seulement par cette consonne. Les réponses des participants ont montré une dominance de l'oreille droite (*Right Ear Advantage*, REA, témoignant d'une meilleure performance de l'HG, à cause d'une prédominance des voies neuronales controlatérales) pour le traitement de stimuli qui se distinguaient seulement par le lieu d'articulation, et donc par des indices acoustiques brefs. En revanche, la dominance de l'oreille droite disparaissait pour les couples de stimuli qui différaient par le voisement, lorsque la consonne sonore était présentée à l'oreille gauche et la sourde à l'oreille droite. L'inverse se produisait dans une expérience équivalente en anglais : l'avantage de l'oreille droite disparaissait si la consonne sourde anglaise était présentée à l'oreille gauche et la sonore à l'oreille droite (Rimol *et al.*, 2006).

En français, les occlusives sonores contiennent un indice acoustique long et les sonores un indice court. Il s'agit de leur délai de voisement (*Voice Onset Time*, VOT), c'est-à-dire de l'intervalle séparant le bruit d'explosion et le son provenant de la vibration des plis vocaux. En français, ce VOT est long pour les occlusives sonores et bref pour les sourdes, et c'est l'inverse en anglais. Au final, les diminutions de dominance de l'oreille droite, qui se

produisaient de façon opposée dans les deux langues, sont cohérentes avec les prédictions du modèle AST. En effet, dans les deux langues le REA diminuait si le stimulus adressé prioritairement à l'HD contenait un indice long. Une interprétation possible est que l'HD s'implique fortement pour traiter un phonème contenant un trait phonétique de longue durée. Comme le REA était au contraire présent dans les deux langues pour les couples de consonnes se distinguant par le lieu d'articulation, c'est-à-dire par des indices courts, les données étaient cohérentes avec l'implication forte de l'HG pour les événements acoustiques brefs. Conformément au modèle AST, la prise en charge du traitement des indices acoustiques des phonèmes se ferait selon l'hémisphère offrant la fréquence d'échantillonnage la plus adaptée.

L'écoute dichotique a aussi permis de montrer que la spécialisation de l'HG pour les indices acoustiques courts (ici, indices de lieu d'articulation) se développe entre 6 et 11 ans (Bedoin *et al.*, 2011). Cependant, cette spécialisation peut être affaiblie en cas d'épilepsie bénigne avec décharges électriques dans l'HG (Bedoin *et al.*, 2011). Cet affaiblissement est aussi documenté en cas de dyslexie (Johnson *et al.*, 2013) aussi bien chez l'enfant (Helland *et al.*, 2008) que chez l'adulte (Iliadou *et al.*, 2010).

#### **1.2.2.5. Limites du lien exclusif entre oscillations lentes/rapides et HD/HG**

Chez l'adulte, il existe beaucoup d'arguments pour un traitement préférentiel des unités acoustiques longues par l'HD et un traitement préférentiel des unités courtes par l'HG. Les modèles d'analyse temporelle à échelles multiples présentent bien cette complémentarité, basée sur des préférences décrites de manière nuancée. L'asymétrie hémisphérique entre les compétences pour la séquenciation du signal à une échelle fine ou large est une question de dominance relative et non de spécialisation stricte.

Une expérience en IRMf chez l'adulte a montré que la perception de sons non-verbaux contenant des modulations d'amplitude lentes ou rapides active les mêmes régions corticales (les gyrus et sillon temporaux supérieurs), et c'est surtout la rapidité de ces activations qui varie selon le rythme des modulations du signal (Giraud *et al.*, 2000). Lorsqu'une différence hémisphérique est observée dans ce type d'expériences, il arrive que les réponses du gyrus de Heschl soient bilatérales pour les modulations d'amplitude lentes (8 Hz dans cette expérience) et préférentiellement à gauche pour les plus rapides (Belin *et al.*, 1998 ; Liégeois-Chauvel *et al.*, 2004). La même chose a été observée chez des nouveau-nés (Cabrera & Gervain, 2020). D'autres résultats, évoqués plus haut, montrent au contraire que la tentative de synchronisation des rythmes cérébraux au signal acoustique est réalisée de manière bilatérale

par le rythme gamma, mais plus nettement à droite par le rythme thêta (Luo & Poeppel, 2012). L'ensemble des données évoque bien une asymétrie hémisphérique au niveau des cortex auditifs, mais celle-ci est relative, comme dans les propositions des modèles d'analyse temporelle à échelles multiples au sujet de la voie ventrale du traitement de la parole (e.g. Hickok & Poeppel, 2008). Cette perspective nuancée permet d'être optimiste au sujet de patients atteints de lésions cérébrales, car des compensations sont envisageables. Une asymétrie non-strictement est cohérente avec la souplesse de prise en charge d'indices courts et longs par l'HD, en cas de nécessité. C'est ce que montrent les performances en reconnaissance de mots chez des enfants épileptiques dont le foyer de décharges paroxystiques est pourtant centro-temporal gauche. Leurs performances dans cette épreuve de langage sont dans la norme, malgré une perte de dominance de l'HG pour traiter les indices acoustiques rapides (i.e., les indices de lieu d'articulation) d'après une épreuve d'écoute dichotique (Bedoin *et al.*, 2011). Il existe une certaine souplesse dans la prise en charge des indices de différentes durées par l'un ou l'autre hémisphère, au-delà de l'asymétrie des traitements préférentiels.

Une autre limite est apparue à la suite de travaux récents. Les tentatives d'ajustement des rythmes cérébraux à l'enveloppe temporelle de la parole mettent en jeu des mécanismes basés sur des oscillations cérébrales qui répondent aux fluctuations du signal, de façon bottom-up. L'ajustement de ces rythmes cérébraux est globalement latéralisé en fonction des durées auxquelles ils se synchronisent. Les tentatives d'ajustement impliqueraient cependant aussi des mécanismes descendants, basés sur des signaux provenant de sources situées en dehors des aires auditives (dans le cortex antérieur et liées à la motricité). Lorsque ces derniers mécanismes sont dominants dans une tâche, par exemple quand le traitement sémantique est requis, la dominance hémisphérique classique droite pour l'ajustement au rythme syllabique (bande thêta) s'atténue fortement (Assaneo *et al.*, 2019). Cela incite à la réflexion sur le choix des tâches pour étudier l'asymétrie hémisphérique liée au traitement des durées dans le signal acoustique.

### **1.2.3. Mécanismes de synchronisations des rythmes cérébraux au rythme de la parole**

Le modèle AST décrit deux grands principes de synchronisation.

D'une part, les rythmes cérébraux tenteraient activement (*entrainment*) de s'ajuster au rythme d'informations externes discontinues jusqu'à l'obtention d'un alignement temporel

(i.e. *phase-locking*) (Lakatos *et al.*, 2005) qu'il faut constamment réajuster à cause de la régularité imparfaite du signal de parole. Des événements saillants dans le signal (*edges*) déclencheraient prioritairement une réinitialisation (*reset*) des rythmes thêta et gamma pour stimuler ces efforts d'alignement (Gross *et al.*, 2013). Qu'elle soit large ou étroite, la fenêtre d'analyse alors définie glisserait dans le temps, et se réinitialiserait à l'occasion d'événements saillants.

D'autre part, plusieurs rythmes cérébraux s'impliquent dans ce processus d'*entrainment*, et ils seraient organisés en hiérarchie : les oscillations lentes exerceraient une influence sur de plus rapides. Ces influences sont de deux types : 1) un rythme lent peut moduler un rythme plus rapide en faisant en sorte qu'un nombre fini de phases de ce dernier s'intègrent dans l'une de ses phases (ajustement phase-phase, représenté par l'action d'une roue sur une autre dans la Figure 3 proposée pour synthétiser les relations inter-fréquences) ; 2) un rythme lent peut aussi moduler l'amplitude des phases d'un rythme plus rapide pour que cette amplitude s'accroisse à une certaine étape de sa phase, typiquement avant son pic (ajustement phase-amplitude). L'ensemble crée de la cohérence à travers des relations de nichage (*nesting*) entre des rythmes cérébraux hiérarchisés (Ghitza, 2011).

La description proposée par Giraud et Poeppel (2012) pour la séquenciation du signal de parole prend en compte ces deux types de modulations de l'activité cérébrale rythmique, et propose cinq étapes (Fig. 1). La première se produirait dans la couche d'entrée du cortex auditif, qui répond de façon analogique aux variations du signal de parole en produisant des décharges quasi-régulières fidèles à la temporalité du signal. Un codage plus abstrait s'enclencherait lorsqu'un événement saillant dans ce signal (*edge*) déclenche, plus en profondeur, la réinitialisation des oscillations thêta et gamma (i.e. *phase reset*). Les ondes thêta s'engageraient alors dans la deuxième étape : la poursuite active du signal externe pour tenter de synchroniser leur phase à son rythme, essentiellement syllabique (i.e. *envelope tracking*). Selon l'étape 3 de la description de Giraud et Poeppel, ces ondes lentes exerceraient dès lors, une relation d'emboîtement sur les ondes gamma en modulant leur amplitude selon l'état de sa phase, créant une cohérence inter-rythmique intra-cérébrale : l'amplitude des phases gamma varierait régulièrement au sein de chaque phase du rythme thêta (i.e. *nesting*). En quatrième étape, à la sortie du cortex auditif, cela donnerait lieu à un train de décharges temporellement plus régulier qu'à l'entrée, mais fidèle au rythme moyen du signal de parole que thêta poursuit sans relâche pour des réajustements. La série de décharges dans les couches de sortie du cortex auditif primaire serait déterminée par ces variations d'amplitude du rythme





dynamique et de l'organisation des oscillations, même si un réseau se dessine, les caractéristiques de l'activité neuronale sont devenues prioritaires dans cette description conduite sous un angle neurolinguistique.

L'organisation hiérarchisée de ces mécanismes en cascade fait l'objet d'une réflexion conduisant Murphy et Benitez-Burraco (2019) à l'idée de « *language oscillome* ». Cette réflexion reste cohérente avec le modèle AST. Deux types d'oscillations seraient impliqués dans cette hiérarchie. 1) Certaines oscillations cérébrales sont dites *évoquées*, parce que leur phase est synchronisée à celle de la stimulation externe, en réponse à cette stimulation. 2) D'autres, dites *induites*, seraient générées par des processus de plus haut niveau guidés par des connaissances et l'attention. Ces oscillations induites visent le *binding*, c'est-à-dire une neuro-synchronisation susceptible de déboucher sur une perception intégrant différentes dimensions. Une telle synchronisation peut aussi donner lieu à un état de conscience particulier favorisant les influences top-down, ou l'intérêt pour des indices guidant les anticipations, ou encore la prise de conscience, la concentration... L'émergence des oscillations induites ne serait pas directement provoquée par la présence d'un stimulus mais, par des mécanismes plus autonomes, et leur influence porterait surtout sur l'amplitude des autres rythmes cérébraux. Leur intervention dans la dynamique oscillatoire qui sous-tend le traitement temporel du signal de parole est l'objet de débats actuels. Dans ce qui suit, nous proposons une synthèse et un schéma autour de la question des différents couplages entre rythmes cérébraux, en ne perdant pas de vue la question de l'asymétrie hémisphérique du traitement des durées courtes/longues, fil conducteur dans cette thèse.

#### **1.2.4. Fonctionnement hiérarchisé et couplages des oscillations cérébrales**

Le décodage du signal de parole composé d'éléments hiérarchisés serait possible par une réaction en cascades d'oscillations cérébrales déclenchée par un rythme d'entrée. Ghitza a proposé en 2011 un modèle – *Tempo* – (Fig. 2) expliquant comment différents rythmes fonctionnent ensemble pour un traitement en parallèle mais harmonieux des informations, comme le suppose le modèle AST. Dans le modèle *Tempo*, le postulat est que le rythme thêta (4-8 Hz) est le point de départ de la réaction en cascade oscillatoire. Comme le montre la partie inférieure du schéma de l'auteur (Fig. 2), les autres rythmes (bêta et gamma) se caleraient sur celui-ci. Les oscillations thêta seraient les premières à rechercher le rythme entrant (Luo & Poeppel, 2007), à caler leur phase sur lui et à découper le signal de parole selon ce rythme qui est adapté aux syllabes (pour une revue des arguments, voir Peelle &



Davis, 2012). La séquenciation syllabique ainsi réalisée est connue pour contribuer à l'intelligibilité du discours (Houtgast & Steeneken, 1985).

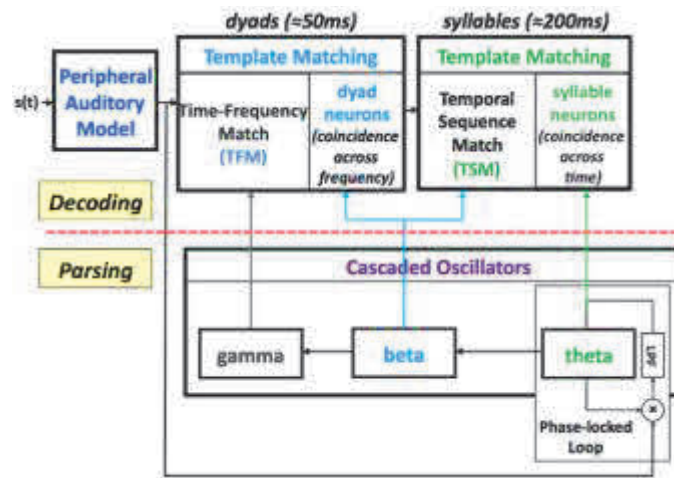





Figure 2. Modèle Tempo. Reproduit à partir de «Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm», par Ghitza, O., 2011, *Frontiers in Psychology*, 2, 1-13.

Le couplage le plus décrit pour la parole correspond à la modulation que les oscillations thêta exercent sur des oscillations dans la partie basse de la bande gamma (30-90 Hz). Il s'agit du couplage phase-amplitude exposé par Giraud et Poeppel (2012) spécifiquement pour la perception du langage. Dans ce couplage, les oscillations thêta feraient augmenter l'amplitude des ondes gamma à un instant précis de chaque phase thêta. Ce rythme gamma permettrait une séquenciation complémentaire, plus fine, adaptée à l'analyse des transitions rapides dans le signal de parole, à l'extraction de traits phonétiques brefs et de phonèmes courts (transitions de formants, VOT brefs, consonnes occlusives, voyelles courtes...). Une étude a montré que l'influence du rythme thêta sur le rythme gamma s'exerce plus fortement dans l'HG (Gross *et al.*, 2013). C'est dans l'HG que le rythme cérébral rapide gamma devient particulièrement puissant et permet une analyse fine du signal à des instants déterminés par le rythme thêta. Comme ce dernier se synchronise à l'organisation syllabique de la parole, une analyse raffinée de l'attaque des syllabes serait par exemple possible, grâce à l'augmentation d'amplitude des oscillations gamma à l'instant où elles surviennent.

Nous proposons un schéma (Fig. 3) basé sur l'idée d'engrenage pour illustrer les relations dynamiques et hiérarchiques entre les rythmes thêta, gamma, mais aussi bêta et delta. Il distingue les deux types de relations (voir légende de la Figure 3) : l'action d'une roue crantée sur une autre représente l'influence d'un rythme sur la phase d'un autre rythme

plus rapide, alors qu'un amplificateur (en violet) représente la modulation exercée par un rythme sur la puissance générale d'un autre rythme ou sur ses variations d'amplitude.

### Légende du schéma

	Amplificateur	Amplificateur symbolisant une relation hiérarchique phase-amplitude d'un rythme cérébral sur un autre
	Roue d'engrenage	Relation entre des roues crantées symbolisant une relation hiérarchique entre les phases de deux rythmes cérébraux.
	Reset	Ligne symbolisant le <i>reset</i> des rythmes cérébraux thêta et gamma sur le rythme de la stimulation externe

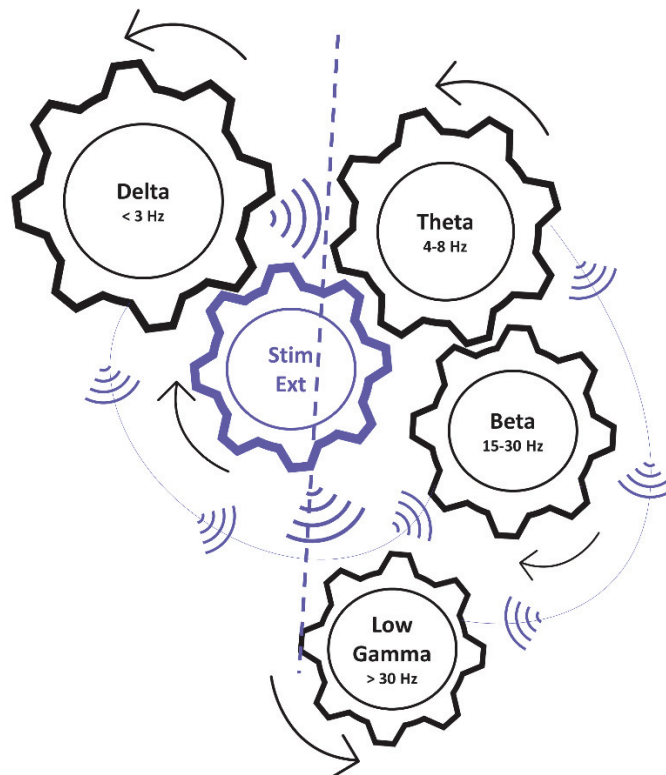


Figure 3. Réactions en cascade des différents rythmes cérébraux.

Dans ce schéma, le rythme du signal de parole est représenté par une roue crantée de couleur violette. Les rythmes cérébraux influencés par le rythme externe sont aussi représentés par des roues. Le schéma cherche à illustrer le fait que les rythmes cérébraux sont si rapidement en action en présence d'un signal qu'ils deviennent acteurs de la synchronisation. Une fois le rythme thêta synchronisé, la cascade d'oscillations recherchait les caractéristiques du signal (*tracking*). Des modulations présentes dans le signal de parole produiraient d'abord le *reset* décrit par Giraud et Poeppel (2012) sur les oscillations thêta et gamma, mais elles exerceraient aussi, en plus, des actions séparées sur les rythmes delta, thêta et gamma.

Pour ce qui est du rythme gamma, dans une expérience comparant l'écoute d'une histoire et l'écoute de parole inversée, Gross *et al.* (2013) ont montré que des changements de 3 à 7 Hz dans l'enveloppe de la parole modulent directement l'amplitude de l'activité gamma (35-45 Hz). L'implication des oscillations gamma d'une manière harmonieuse avec les phases du rythme thêta apporte ainsi des informations complémentaires sur l'enveloppe de la parole, qui précisent celles qu'extrait thêta. Le rythme gamma recevrait cette modulation directe de la part du signal dans les deux hémisphères, mais davantage dans le cortex auditif gauche. Cette asymétrie est cohérente avec le modèle AST : les oscillations gamma qui sous-tendent un découpage fin du signal sont particulièrement modulées par ce signal et adaptées à celui-ci dans l'HG.

Pour ce qui est du rythme delta (1-2 Hz), sa phase est elle aussi modulée par les variations temporelles lentes dans le signal de parole (Gross *et al.*, 2013). Les ajustements de phases des rythmes delta et thêta au signal externe ont pour point commun d'opérer significativement plus fort dans l'HD que dans l'HG. D'après Gross *et al.*, une dissociation existe cependant entre les deux. Ils notent par exemple que la latéralisation droite des oscillations delta est particulièrement vraie dans les aires frontales et pariétales, alors que la latéralisation droite des oscillations thêta concerne surtout les aires temporales supérieures, où elles réaliseraient un *phase-locking* pour le codage des syllabes. La dissociation entre les ajustements de phase des rythmes delta et thêta au signal externe participe sans doute à leurs retentissements distincts sur le plan linguistique. Cette différence est confirmée par l'absence d'ajustement de phase entre delta et thêta. Des données en MEG apportent aussi des arguments pour défendre l'idée d'oscillations cérébrales recherchant dans le signal acoustique des éléments distincts, de façon indépendante Cogan et Poeppel (2011). Les rythmes thêta et delta sont toutefois en relation, mais seulement par la modulation d'amplitude exercée par

delta sur thêta dans les aires temporales. Cette relative dissociation conforte le rôle central du rythme thêta, fortement lié au rythme gamma, pour le traitement acoustique-phonétique du signal de parole.

Concernant l'asymétrie hémisphérique, l'ajustement de phase entre les ondes thêta des deux cortex auditifs a été mesuré pour évaluer la coordination inter-hémisphérique. Lorsque l'alignement sur le rythme de la parole augmente, l'ajustement de phase entre les ondes thêta diminue entre les deux cortex auditifs (Gross *et al.*, 2013). Pour ces auteurs, cela témoignerait d'une recherche d'ajustement au rythme lent de la parole plus active de la part du rythme thêta du cortex auditif droit. Ce point est cohérent avec le modèle AST, parce que l'ajustement de phase est significativement plus fort à droite qu'à gauche dans le cortex auditif pour des modulations de l'ordre de 50 à 100 ms, compatibles avec la fréquence thêta et que, même au repos, ce rythme thêta est dominant à droite (Morillon *et al.*, 2012).

A cet ensemble s'ajoute le rythme bêta qui, selon Ghitza (2011), serait l'objet d'un ajustement de phase de la part du rythme thêta, d'où la relation entre les roues crantées représentant ces deux types d'oscillations dans le schéma (Fig. 3). L'oscillateur thêta réglerait le rythme oscillatoire des ondes bêta afin qu'elles soient un multiple de sa propre fréquence. Leurs débuts de phase se coordonneraient et bêta serait niché dans thêta. Les oscillations bêta permettraient une analyse plus fine des événements sonores dans la syllabe, au niveau d'une dyade (reflet acoustique d'un geste par lequel un articulateur passe d'un phonème au suivant). Ghitza décrit enfin un ajustement de phase entre ce rythme bêta et le rythme gamma, dont les phases pourraient se nicher dans une phase bêta. De telles relations nichées sont cohérentes avec les propriétés des oscillateurs décrites plus généralement en neuroscience au sujet du fonctionnement cérébral (Schroeder & Lakatos, 2009). Ces relations d'emboîtement ont l'avantage d'être en accord avec l'échelle temporelle de la hiérarchie des unités linguistiques dans la parole. La communication parlée a sans doute évolué en s'appuyant sur ces caractéristiques neuronales dans le cortex auditif et, nous y reviendrons, dans le cortex moteur (Lieberman & Whalen, 2000).

Il existe deux sources de renforcement des relations de couplage entre les rythmes cérébraux.

L'une est de très bas niveau en ce sens qu'elle vient du signal : la survenue de changements d'amplitude importants (*edges*) dans le signal opère la remise à jour des oscillations (*reset*) en exerçant plusieurs effets (Gross *et al.*, 2013). Les *edges* dans le signal

1) augmentent l'alignement entre les oscillations auditives et l'enveloppe de la parole, 2) ils sont suivis d'un accroissement des couplages inter-fréquences entre les oscillations impliquées dans les traitements auditifs, et 3) ils accentuent la précision des relations de nichage coordonnant les oscillations à différentes échelles. Peelle et Davis (2012) décrivent ce *reset* et cet alignement comme instantanés. Le début d'un mot ou d'une syllabe peut le produire et déclencher immédiatement, à partir des ondes thêta et gamma alignées, la mécanique des oscillations hiérarchisées servant de guide à la perception.

Si l'on se replace dans la situation de perception d'un mot, grâce au *reset*, tout se passe comme si des fenêtres d'analyse temporelle adaptées – assez larges pour des événements acoustiques longs dans l'HD (rythme thêta), suffisamment étroites pour les indices les plus courts dans l'HG (rythme gamma), et intermédiaires (rythme bêta) – s'offraient alors instantanément pour analyser, par exemple, une voyelle longue ou courte, ou un délai de voisement long ou court dans des consonnes. Cependant, l'ensemble de l'organisation dynamique des oscillations guidant l'analyse temporelle est généralement étudié dans la situation de perception de parole continue (*ongoing speech*). Le rythme du locuteur au début d'une phrase comme « je dis... » règle alors déjà l'ensemble de la cascade oscillatoire et permet des prédictions temporelles efficaces pour le mot qui suit. Dans les situations habituelles d'apprentissage d'une langue non-native (L2), par exemple lors d'exercices de vocabulaire, l'enseignant ne produit pas toujours ce début de phrase et ce phénomène d'amorçage n'est pas forcément présent. Il en est de même lorsque la prononciation d'un mot est écoutée à partir d'un fichier-son sélectionné sur Internet : le mot est isolé. La régularité de diffusion des mots nouveaux donnés à entendre dans un exercice pourrait en partie pallier le fait qu'il ne s'agit pas de parole connectée, en imprimant une certaine régularité à laquelle les oscillations peuvent tenter de se synchroniser. Cependant, un retard dans cette synchronisation, soupçonné en cas de dyslexie (Goswami, 2011), peut mettre certains apprenants en difficulté. Dans ce type de situation, l'apprenant doit surtout compter sur une capacité neuronale à réaliser le *reset* immédiat pour qu'un mot isolé soit perçu dans de bonnes conditions d'analyse temporelle. Là encore, des différences inter-individuelles existent dans les relations entre les durées des événements acoustiques et la phase des rythmes cérébraux engagés dans le *reset* (Peelle & Davis, 2012). En cas de dyslexie, les événements acoustiques saillants, comme l'attaque d'une syllabe par exemple dans l'enveloppe d'amplitude (*rise-times*) sont mal traités (Goswami *et al.* 2002 ; Leong & Goswami, 2014a ; Van Hirtum *et al.*, 2019a). Ils provoquent une synchronisation anormalement faible des oscillations thêta dans

les 150-650 ms qui suivent dans les aires auditives (Lizarazu *et al.*, 2021). C'est pourquoi, dans les conditions d'apprentissage phonétique basées sur la présentation de mots isolés, le traitement temporel du signal a besoin d'être facilité. Augmenter l'intensité des changements d'amplitude dans le signal pour favoriser le *reset* des oscillations améliore la perception de parole en cas de dyslexie (Van Hirtum *et al.*, 2019b). L'objectif de la thèse est de proposer de nouveaux dispositifs d'aide (présentation dichotique et amorçage rythmique) en conservant l'usage de stimuli de parole naturelle. Il s'agit pour cela de se baser sur les connaissances présentées ici à propos du support oscillatoire neuronal des traitements temporels du signal de parole.

L'autre source de renforcement des relations de couplages oscillatoires implique des aspects de traitement de plus haut niveau, c'est-à-dire des traitements non-sensoriels impliquant des aires cérébrales en-dehors des aires auditives primaires, ainsi que des connaissances sur la langue et bien souvent l'attention. La cascade des couplages, dont les ondes thêta seraient le capitaine selon Ghitza, serait renforcée par des modulations phase-amplitude (Gross *et al.*, 2013) représentées sur le schéma (Fig. 3) par un amplificateur. Selon Arnal *et al.* (2015), un couplage phase-amplitude entre delta et bêta concourrait par exemple au maintien des relations hiérarchiques. L'intervention de ce rythme très lent ouvre aussi une fenêtre d'intégration temporelle propre à opérer des groupements (*chunking*) en unités importantes pour la prosodie ou l'analyse syntaxique. Il opère ainsi sur les suites de syllabes formant des phrases ou non, sans être déterminé par le contenu sémantique (Boucher *et al.*, 2019). Son intervention est alors bien descendante, et module des traitements temporels élémentaires. Ce phénomène ferait en effet intervenir des oscillations induites (venant de sources propres, *genuine*) et non simplement évoquées par le signal. De telles oscillations sont par exemple le rythme delta issu d'aires antérieures du cortex, et correspondant avec les systèmes moteur et sensoriel. L'aspect dynamique de ces relations en cascade, que nous avons symbolisé par les engrenages, ne sous-tendrait pas simplement une séquenciation en réponse (évoquée) au signal, mais permettrait aussi, particulièrement grâce aux oscillations induites, des prédictions temporelles (Zoefel, 2018). Ces prédictions consistent à amplifier la sensibilité neuronale pour les informations survenant au moment attendu (Morillon & Schroeder, 2015) ce qui améliore la perception des événements auditifs (Jones *et al.*, 2017). Si de telles oscillations induites participent, cela permet d'inclure des influences de haut niveau (i.e. non strictement dirigées par les données du signal acoustique) dans l'analyse temporelle de la parole. Appuyé sur une mécanique oscillatoire hiérarchisée répondant de façon bottom-up au

signal d'entrée, et sur une activité cérébrale proactive et prédictive de type top-down, le décodage de la parole apparaît alors comme robuste et capable de se réaliser même en contexte bruité (Golumbic *et al.*, 2012).

Pour Arnal & Giraud (2010), des prédictions temporelles précises seraient possibles parce que le système cérébral moteur prendrait part au contrôle des oscillations dans le temps. L'implication du cortex moteur est cohérente avec l'amélioration de la perception de la parole observée si l'aire motrice des lèvres est stimulée par Stimulation Magnétique Transcrânienne (TMS) (D'Ausilio *et al.*, 2009). Le couplage phase-amplitude delta-bêta sous-tendrait la précision de telles prédictions : la puissance des ondes bêta est maximale à l'arrivée d'un indice acoustique important et prédit. Le rythme thêta prendrait aussi part à ces prédictions, car il est présent dans les aires auditives mais aussi motrices, et correspond à la fois au rythme des syllabes et au rythme de base de la mandibule (Giraud *et al.*, 2007). L'ensemble soutient un modèle dans lequel l'orchestration de la dynamique oscillatoire entre les systèmes sensoriels et la dynamique oscillatoire de plus haut niveau, incluant les aires motrices, est exploitée pour sélectionner à des échelles multiples (*multiplexing*), mais aussi prédire (Zoefel, 2018) avec précision, les informations sensorielles dans le temps.

Pour illustrer la participation de connaissances de haut niveau au traitement temporel élémentaire de la parole, une expérience montre que les changements de basse fréquence présents dans le signal de parole modulent les phases d'ondes lentes et l'amplitude d'ondes rapides plus fortement pour l'écoute d'une histoire que pour une séquence de parole inversée (Gross *et al.*, 2013). Les auteurs interprètent cette différence. Pour eux, ce ne sont pas forcément des connaissances sémantiques activées par l'histoire qui retentissent directement sur les phases et l'amplitude d'ondes impliquées dans l'analyse temporelle élémentaire. Des connaissances plus générales sur l'organisation possible des suites de sons dans la parole, altérée dans la parole inversée, ou le simple fait d'avoir repéré qu'il s'agit de langage, sont des aspects de haut niveau de traitement. Ces aspects pourraient expliquer l'influence différente de l'histoire et de la parole inversée sur les ondes cérébrales impliquées dans l'analyse temporelle. En résumé, la séquenciation du signal de parole n'est pas seulement déterminée par les caractéristiques du signal, ni par les réponses évoquées dans un système oscillatoire hiérarchisé : ce dernier n'est pas passif, mais animé par ses contacts avec des processus non-sensoriels.

Dans ce qui suit, il s'agit justement de contextualiser la cascade des oscillations, la séquenciation et les anticipations dans le contexte plus large du réseau cérébral du langage,



mais aussi du réseau des traitements temporels. Nous verrons que le système moteur, qui peut générer des simulations de mouvements, revient à plusieurs niveaux dans cette réflexion, et influe sur les aires sensorielles, notamment auditives (Nelson *et al.*, 2013), pour moduler les prédictions temporelles.

### **1.2.5. Implication de traitements de haut niveau dans l'analyse temporelle**

Sur le plan neuro-fonctionnel, il convient de replacer les traitements temporels opérés sur le signal de parole dans deux cadres plus vastes : le réseau cérébral du langage et celui du traitement du temps.

#### Réseaux des traitements du langage

Les supports neurologiques du langage ne sont plus décrits comme restreints à l'aire de Broca et à l'aire de Wernicke associée au gyrus angulaire, reliés par le faisceau arqué. Un réseau plus vaste est évoqué, le long de voies dites ventrale et dorsale. Cette distinction entre deux voies reprend une idée préalablement développée en neuropsychologie de la vision. Leurs rôles respectifs sont surtout étudiés pour des tâches complexes recourant à des traitements lexicaux, syntaxiques et sémantiques. L'articulation entre les deux voies est aussi évoquée – plus marginalement – pour expliquer comment des connaissances et traitements relatifs à la langue (lexique, syntaxe, sémantique) modulent les analyses de plus bas niveau, comme le traitement temporel du signal étudié ici. La voie ventrale traiterait ce signal en vue de la compréhension du message. Elle serait fonctionnellement bilatérale et réaliserait des analyses spectrales et temporelles (séquenciation) différentes dans l'HG et dans l'HD. La voie dorsale, impliquant des aires plus latéralisées à gauche, permettrait à la parole d'être analysée en partie à travers un réseau articulo-moteur en région frontale (Hickok & Poeppel, 2007). En l'absence de traitement du langage (i.e. au repos), les oscillations dans les bandes de fréquence utiles à la perception de la parole sont latéralisées seulement dans les aires auditives, sensorimotrices, motrices articulo-motrices et dans le cortex pariétal inférieur, mais pas dans les aires de Broca et de Wernicke (Morillon *et al.*, 2010). Cela suggère que cette latéralisation dans la première série de régions sous-tend des traitements essentiellement temporels, utiles au langage, mais non spécifique à celui-ci. La latéralisation de l'activité cérébrale gauche dans les aires classiques du langage (Broca, Wernicke) est associée à la réalisation du traitement du langage, mais elle est absente au repos, ce qui fait supposer qu'elle est indépendante des phénomènes d'*entraînement* au rythme du signal.



Dans une conception théorique où des influences top-down peuvent moduler des traitements élémentaires, cette organisation cérébrale plus large a son importance. L'objectif plus ou moins clairement linguistique d'une tâche, ou le fait de savoir dans quelle langue les prochains stimuli seront proposés sont des exemples de connaissances de haut niveau sur la situation pouvant exercer des influences top-down précises. En bref, les traitements acoustiques/phonétiques du signal de parole dans le cortex auditif primaire ne sont pas totalement encapsulés, des aires de plus haut niveau jouent un rôle.

Certaines influences de haut niveau viennent de la voie dorsale. Des lésions pariétales inférieures gauches produisent ainsi parfois des déficits d'identification de phonèmes lorsqu'il faut prendre en compte un indice temporel comme leur VOT (Voice Onset Time, trait phonétique temporel déterminant pour le voisement des consonnes) (Blumstein *et al.*, 1977b). Des lésions du gyrus supra-marginal et de l'opercule pariétal voisin perturbent aussi les traitements acoustiques/phonétiques (Caplan *et al.*, 1995). De même, la simple discrimination de phonème – qui ne demande pas explicitement d'utiliser les connaissances sur les phonèmes – est très mal réalisée en cas de lésions frontales gauches (Blumstein *et al.*, 1977a) plus connues pour détériorer la production de parole. Ainsi, même si l'encodage du signal de parole ne doit pas être nécessairement linguistique pour réaliser une discrimination (pareil-différent), cette tâche implique non seulement la partie la plus dorsale du gyrus temporal supérieur, mais aussi les aires auditives associatives (i.e. des connaissances sur des phonèmes) et une connexion entre perception et action (Gross *et al.*, 2013). Si la tâche requiert explicitement l'accès à la représentation de phonèmes, des données confirment que cet encodage n'est pas strictement acoustique-phonétique et active la voie dorsale du langage (Hickok & Poeppel, 2000). Des interactions auditivo-motrices interviennent alors, car les éléments phonologiques seraient interprétés en termes de gestes articulatoires (Poeppel *et al.*, 2007). De plus, Giraud *et al.* (2007) montrent que la fluctuation de puissance de l'activité thêta et gamma dans les cortex auditifs droit et gauche est corrélée à l'activité dans l'aire cérébrale prémotrice relative à la bouche. Cela traduit un couplage entre les propriétés temporelles de la perception et de la production de la parole. Comme le résumait Hickok et Poeppel (2007), en perception de parole, la voie dorsale traduirait automatiquement le signal acoustique en représentations articulatoires, grâce à une interface sensori-motrices en région temporo-pariétale, et avec la participation du lobe frontal et des régions pré-motrices.

L'encodage du signal de parole peut aussi activer la voie ventrale dans la jonction temporo-pariéto-occipitale, surtout si la tâche demande l'accès au lexique mental (Hickok &

Poeppel, 2000). Dans la voie ventrale, les parties médiane et postérieure du sillon temporal supérieur peuvent aussi intervenir dans des traitements phonologiques. Elles s'activent bilatéralement, mais présentent tout de même une certaine dominance gauche. Cette dernière peut masquer un traitement du signal plus élémentaire bilatéral, comme la séquenciation, au cœur de cette étude.

### Réseaux des traitements temporels

Les supports neurologiques des traitements temporels et rythmiques sont aujourd'hui décrits comme un large réseau plutôt que comme deux systèmes séparés. Un des systèmes, centré sur le cervelet, était censé permettre les traitements temporels simples comme les estimations d'intervalles. L'autre, centré sur le striatum, était supposé soutenir les traitements temporels complexes avec estimations de durées relatives.

Des expériences en neuropsychologie comparant les déficits de patients atteints d'atrophie cérébrale dans l'une ou l'autre de ces régions ont permis à Teki et ses collègues (2012) de montrer que ces deux aires interviennent dans les deux types de traitements temporels. Elles font partie d'un plus large réseau dont l'efficacité repose sur la cohérence. Des études en neuro-imagerie ont confirmé que plusieurs structures cérébrales jouent un rôle dans la perception du temps, telle qu'utile pour traiter le langage. On retrouve alors le cervelet (Ivry & Keel, 1989) et les ganglions de la base, en particulier le striatum (Artieda *et al.*, 1992). D'autres régions faisant par ailleurs parties de la voie dorsale du langage seraient impliquées, comme les aires motrices pré-supplémentaire et supplémentaire (Grahn, 2009 ; Halsaband *et al.*, 1993), ainsi que le cortex pré-moteur et moteur (Oshio, 2011). Il existe ainsi une superposition partielle entre le réseau de traitement du langage et celui du temps.

Teki *et al.* (2011) proposent un modèle unifié de la perception du temps (Fig. 4) : un réseau striato-thalamo-cortical et olivocérébelleux dont l'activation synchronisée améliorerait la précision d'analyse du signal temporel. Cette précision serait d'autant plus grande qu'une personne a de bonnes capacités de synchronisation auditivo-motrice (Kern *et al.*, 2021). Elle peut alors tirer parti du fait que les régularités rythmiques de la parole concordent avec les propriétés rythmiques endogènes des systèmes auditif et moteur. Si elle a pour cela besoin d'aide, il semble possible de déclencher un traitement temporel basé sur tout le réseau en suscitant l'implication de la motricité, par exemple à travers le déclenchement de mouvements à l'écoute d'une musique bien rythmée.

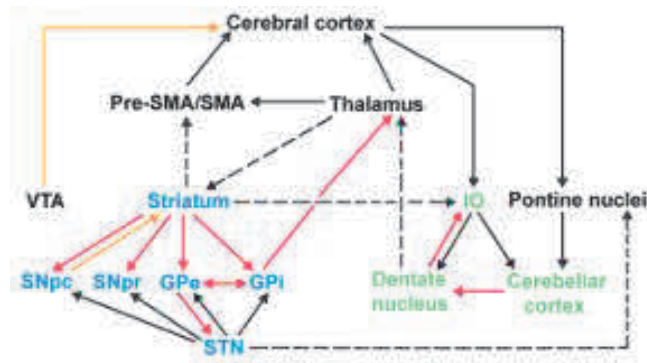


Figure. 4. Modèle unifié de la perception du temps. Reproduit à partir de « A unified model of time perception accounts for duration-based and beat-based timing », par Teki, S., Grube, M., Griffiths, T.D., 2012, *Frontiers in Integrative Neuroscience*, 5, 1-7.

En remplaçant la séquenciation de la parole dans l'ensemble plus vaste des réseaux de traitement du langage et de traitement du temps, la dimension temporelle dans le traitement de la parole apparaît particulièrement centrale. On peut mieux comprendre l'origine de certaines influences top-down, notamment motrices, sur un traitement aussi élémentaire que l'extraction d'unités de durées multiples dans le signal. Des données montrent par exemple que, au repos, un rythme endogène delta-thêta adapté aux syllabes est dominant à gauche dans les parties du cortex moteur correspondant à la main (dominante droite). Des données montrent que cette latéralisation des ondes lentes dans des aires motrices est corrélée à l'asymétrie hémisphérique gauche de l'activité des aires auditives lors de la perception de parole (Morillon *et al.*, 2010). Les auteurs imaginent qu'il est alors possible qu'un input venant des aires motrices de la main ou des lèvres puisse contribuer, directement ou par le biais du cortex sensori-moteur, à l'analyse syllabique de la parole dans le cortex auditif. En utilisant la stimulation magnétique transcrânienne, d'autres recherches montrent aussi que la stimulation de l'aire motrice correspondant à la langue et aux lèvres facilite la perception de la parole (ici, la discrimination de phonèmes) (d'Ausilio *et al.*, 2009).

Dans cette thèse, l'expérience 14 (Chapitre 7) concerne l'aide à l'apprentissage de caractéristiques prosodiques de l'anglais appris tardivement. Elle consiste à exploiter ce type de caractéristique en diffusant une amorce auditive dotée d'un rythme lent (2 Hz) propre à renforcer un rythme endogène lent dans les aires motrices. Celles-ci seraient alors sollicitées pour leur rôle dans le réseau des traitements temporels. L'accentuation de la puissance de ces oscillations lentes pourrait influencer des oscillations elles aussi lentes dans le cortex auditif, grâce à un phénomène d'intégration auditivo-motrice. L'hypothèse est que, immédiatement après, le cortex auditif serait en mesure de diriger plus facilement la séquenciation d'un signal de parole à cette échelle lente. Des unités prosodiques pourraient ainsi être mieux extraites et

traitées avec plus d'attention, pour un meilleur apprentissage de particularités prosodiques d'une langue non-native.

### **1.3. Pathologies neurodéveloppementales du langage : explications basées sur l'attention temporelle et les oscillations cérébrales**

La dyslexie est un trouble neurodéveloppemental décrit pour la première fois par Pringle-Morgan comme une « cécité verbale congénitale » en 1896. Il s'agit d'une difficulté sévère d'apprentissage de la lecture en dépit d'une intelligence et d'une motivation normales, d'un enseignement conventionnel et d'opportunités socioculturelles adéquates (Rutter, 1978). Selon l'Observatoire National de la Lecture (1998, p. 171), un dyslexique est un mauvais lecteur « chez qui le déficit résulte, en partie en tout cas, d'une anomalie de la capacité d'identification des mots écrits. L'origine de cette anomalie se trouve dans les structures cérébrales et cognitives qui sous-tendent cette capacité ». Des déficits phonologiques contribuent souvent à son origine (Schulte-Körne *et al.*, 1999 ; Serniclaes *et al.*, 2001 ; Snowling, 1981, 1987 ; Sprenger-Charolles L., Serniclaes W., 2004), avec des représentations phonémiques floues (Bedoin & Krifi, 2009 ; Serniclaes *et al.*, 2004) et/ou difficiles d'accès (Boets *et al.*, 2013).

Sur le plan neuro-anatomique, Galaburda (1979 et 1985, cité par Habib & Robichon, 1996) fut le premier à décrire, post-mortem, des anomalies morphologiques (ectopies, dysplasies, polymicrogyri) associées à la dyslexie développementale. Elles sont situées dans la région périsylvienne gauche, déterminante pour le langage. Leonard *et al.* (1993) et Heiervang *et al.* (2000) ont par ailleurs relaté une absence d'asymétrie hémisphérique du planum temporale et/ou du lobe pariétal inférieur, régions cruciales pour la phonologie et l'intégration de différentes sources d'information. Le corps calleux serait aussi plus épais en cas de dyslexie (Habib, 2000 ; Robichon & Habib, 1998), ce qui est cohérent avec une faible asymétrie hémisphérique. Des travaux révèlent aussi un lien entre les capacités de lecture du père et la qualité de la substance blanche dans les aires ventrales gauches de l'enfant, suggérant une part d'hérédité dans la dyslexie (Vandermosten *et al.*, 2017).

Des anomalies cérébrales non plus anatomiques mais fonctionnelles sont aussi observées. Les dyslexiques souffrent aussi d'une désorganisation des phases de sommeil, sans doute défavorable à la consolidation dans l'apprentissage (Bruni *et al.*, 2009). De manière plus spécifique au langage écrit, des études en TEP ont observé des anomalies de l'asymétrie

hémisphérique fonctionnelle, avec une réduction de l'activité cérébrale dans l'HG pendant la lecture chez les dyslexiques, surtout dans le gyrus temporal médian (Paulesu *et al.*, 2001). Au-delà du lobe temporal, en étudiant la connectivité de différents réseaux par IRM de diffusion (DTI) et IRMf, Vandermosten *et al.* (2012, 2016) ont montré des connexions atypiques entre les lobes temporal et frontal chez de jeunes dyslexiques ainsi qu'une faible intégrité du faisceau arqué gauche. Ces aires cérébrales sont impliquées, nous l'avons vu, de façon top-down dans les relations entre les rythmes cérébraux participant à l'analyse temporelle de la parole. Le caractère atypique de leurs connexions n'est sans doute pas étranger à un traitement acoustique-phonétique imparfait de la parole perturbant l'élaboration du système phonologique.

LLinas, dès 1993, défendait déjà deux idées compatibles avec cette perspective : 1) les modifications des propriétés dans les circuits neuronaux seraient responsables de perturbations des aspects temporels de la cognition chez les personnes dyslexiques, 2) et cela s'expliquerait plus fondamentalement par un dysfonctionnement au niveau cellulaire. Ce dysfonctionnement entraînerait une désynchronisation des relations entre les oscillations, mais surtout une inaptitude à réajuster des rythmes oscillatoires pour suivre des rythmes externes proches en termes de fréquence.

La perturbation des traitements temporels des personnes dyslexiques a souvent été présentée sous un seul angle : celui de leurs difficultés à traiter des informations acoustiques brèves se succédant rapidement. Il s'agit de la perspective défendue par Tallal *et al.* (1996) qui montrait chez les enfants dyslexiques l'amélioration de la perception de phonèmes s'ils étaient artificiellement étirés dans le temps. Des données en français ont confirmé ce point en perception de phonèmes (Habib *et al.*, 2002) comme en jugement d'ordre temporel dans des séquences de phonèmes (De Martino *et al.*, 2001 ; Rey *et al.*, 2002). Aujourd'hui, les anomalies des traitements temporels considérées pour étudier les déficits sous-jacents à la dyslexie sont plus variées. Cela permet de tenir compte d'une variété de difficultés liées au traitement du temps dans cette pathologie (Habib, 2021). Le point commun aux anomalies décrites serait de reposer sur des disconnexions dans un large réseau crucial pour les apprentissages et très dépendant des aspects temporels du traitement de l'information. Le modèle proposé par Habib prend en compte les dernières découvertes sur les difficultés des personnes dyslexiques à réaliser des couplages précis entre des oscillations cérébrales topographiquement variées, de même qu'entre celles-ci et les caractéristiques temporelles du signal de parole. Cela déboucherait sur une variété de difficultés d'apprentissage. Ces

anomalies oscillatoires, relevant des neurosciences, peuvent être mises en lien avec les anomalies de l'attention dans une approche conforme à la psychologie cognitive, grâce au cadre de réflexion offert par la théorie de l'échantillonnage temporel au sujet de la dyslexie.

### 1.3.1. Théorie de l'échantillonnage temporel et dyslexie

La théorie de l'échantillonnage temporel (*Temporal Sampling Framework*, TSF, Fig. 5) a été proposée par Goswami dans sa revue d'opinion de 2011. Cette théorie repose sur un lien entre attention temporelle et oscillations cérébrales. Elle part du principe que l'encodage temporel est un aspect très important pour coder les informations au niveau cérébral (Buzaki & Draguhn, 2004 ; Schroeder *et al.*, 2008). L'activité synchrone des réseaux de neurones oscillatoires à différentes fréquences découpant le signal auditif en unités de durées différentes serait un élément-clé de la perception de parole. Goswami exprime clairement, pour la première fois au sujet des troubles neuro-développementaux du langage, l'idée selon laquelle la modulation et l'alignement de phase des rythmes cérébraux, induits par le stimulus sur des fréquences oscillatoires lentes, sont déterminants pour la perception syllabique (Poeppel *et al.*, 2008) et prosodique (Ghitza & Greenberg, 2009). L'originalité de son cadre théorique est d'explorer ce point pour dessiner une nouvelle perspective d'explication de la dyslexie. Nous avons vu que, par la suite, des recherches ont permis de détailler les nombreuses relations entre oscillations pour le traitement temporel, et cela donne encore plus de pertinence à son approche. Ce modèle intègre les difficultés de perception des changements brusques d'amplitude au début d'une enveloppe (*rise-time*), et des changements plus lents dans l'enveloppe d'amplitude (Soltész *et al.*, 2013) : ces déficits d'échantillonnage temporel impliqueraient les mécanismes de basses fréquences (Thêta et Delta) (Goswami & Leong, 2013 ; Soltész *et al.*, 2013).

Le TSF s'appuie aussi sur le modèle du traitement temporel à échelles multiples de la parole (MTRM, Poeppel *et al.*, 2008) expliqué plus haut. Le réseau d'échantillonnage thêta localisé essentiellement dans l'HD permettrait de découper le signal de parole d'une manière utile à l'analyse supra-segmentale, alors que le réseau gamma, quant à lui bilatéral, permettrait un découpage en unités phonémiques et infra-phonémiques. Selon Goswami, les deux réseaux seraient déficitaires en cas de dyslexie, ce qui dépasse l'hypothèse, plus ancienne, de difficultés limitées aux éléments temporels très brefs (Tallal *et al.*, 1996). Soroli *et al.* (2010) ont même décrit en cas de dyslexie un plus grand déficit au niveau supra-segmental que segmental. Ce déficit temporel se confirme à l'aide de tâches d'identification de sons du langage et de sons non-langagiers (Vandermosten *et al.*, 2010).

Sur le plan cérébral, une latéralisation hémisphérique atypique associée au traitement de la durée des événements est relatée chez les personnes dyslexiques. Leur synchronisation aux indices temporels lents serait imparfaite. Cette anomalie a été décrite comme liée à une faible latéralisation droite des oscillations thêta (Cutini *et al.*, 2016 ; Di Liberto *et al.*, 2018 ; Goswami, 2011 ; Lizarazu *et al.*, 2015). L'ajustement de phase à un rythme de 2 Hz, impliquant cette fois la synchronisation d'oscillations delta, a été décrit comme anormalement réparti entre les hémisphères chez ces personnes. Il est, en effet, bilatéral lorsqu'il est mesuré dans un groupe d'adultes dyslexiques, alors qu'il s'opère avec une dominance droite chez les normo-lecteurs (Hämäläinen *et al.*, 2012). Les oscillations rapides seraient quant à elles trop faiblement latéralisées à gauche en cas de dyslexie (Dushanova *et al.*, 2020 ; Lehongre *et al.*, 2013).

Du fait de ces anomalies à plusieurs échelles (et pas seulement à l'échelle la plus fine), les phases d'excitabilité du substrat neuronal ne se feraient pas correctement. Cela créerait un ajustement oscillatoire inadapté et, par suite, un mauvais traitement du signal de parole chez l'enfant et l'adulte dyslexique (Soltész *et al.*, 2013 ; Thomson *et al.*, 2006 ; Thomson & Goswami, 2008). Dans leur cas, la synchronisation corticale au signal de parole serait défectueuse ou trop faible pour repérer correctement l'enveloppe du signal de parole (Lizarazu *et al.*, 2021). Même lorsqu'il ne s'agit pas de parole, mais de musique, pourtant temporellement plus régulière, les adultes dyslexiques se singularisent par une difficulté de synchronisation de leurs oscillations cérébrales aux stimulations externes, surtout si le rythme est lent. Ils réalisent un tel couplage avec le rythme musical si celui-ci est extrêmement régulier, mais seuls les normo-lecteurs tentent une recherche de couplage à l'écoute de rythmes musicaux moins réguliers (Fiveash *et al.*, 2020). L'efficacité de certaines aides apportées aux adultes dyslexiques peut aussi être interprétée comme un argument favorable au cadre de réflexion TSF. Van Hirtum *et al.* (2019a) ont par exemple montré que les étudiants dyslexiques percevaient mieux la parole grâce à des modifications artificielles du signal, avec mise en valeur de l'attaque (souvent brève) des enveloppes d'amplitude, mais aussi l'accentuation de la structure temporelle de l'ensemble de l'enveloppe. L'accent est en tout cas mis, dans le modèle TSF, sur le retentissement particulièrement délétère des anomalies des rythmes cérébraux lents.

Un exemple de rythme lent, important pour le traitement du langage et de musique, est le rythme à 2 Hz. Des analyses acoustiques ont montré que des événements sonores importants surviennent dans le signal de parole toutes les 500 ms (2 Hz) environ. Ils peuvent



normalement entrer en résonnance avec les oscillations delta, et ce, particulièrement en anglais (Soltész *et al.*, 2013). L'ajustement des oscillations cérébrales et l'analyse temporelle de la parole à un tel rythme sont pertinents notamment pour le traitement prosodique. Or, en cas de dyslexie, une activité atypique est observée dans le cortex auditif droit, où les rythmes cérébraux lents sont en principe dominants (Ghitza & Greenberg, 2009 ; Giraud *et al.*, 2007 ; Luo *et al.*, 2010 ; Luo & Poeppel, 2007 ; Poeppel *et al.*, 2008). C'est aussi le cas chez les personnes dysphasiques (Gauger *et al.*, 1997 ; Heim *et al.*, 2003). La réponse en EEG à des modulations à 2 Hz est aussi plus faible chez les adultes dyslexiques que chez des contrôles (McAnally, & Stein, 1997). Sur le plan comportemental, les personnes dyslexiques présentent d'ailleurs des anomalies lorsqu'il faut synchroniser des mouvements à la perception d'un rythme à 2 Hz. Qu'ils soient enfants ou adultes, les dyslexiques échouent dans des épreuves de tapping à 2 Hz (Goswami, 2011 ; Thompson et Goswami, 2008). Lorsqu'elles traitent le langage, les personnes dyslexiques présentent aussi des difficultés à sélectionner les propriétés rythmiques de la parole, aussi bien en français (Muneaux *et al.*, 2004) qu'en anglais (Goswami *et al.*, 2002).

Le modèle TSF insiste aussi beaucoup sur l'importance d'un autre rythme permettant une séquenciation lente : le rythme thêta, qui sous-tendrait le découpage syllabique (Luo & Poeppel, 2007 ; Poeppel *et al.*, 2008). Leong et Goswami (2014a) ont montré que la séquenciation basée sur les syllabes est le mécanisme le plus perturbé en cas de dyslexie. Que ce soit en perception ou en production, les dyslexiques se trouvent pour cela en difficulté. En perception, au niveau des oscillations, leur ajustement au rythme syllabique est trop lent et, sur le plan comportement, l'épreuve de tapping au rythme des syllabes accentuées en anglais est réalisée en tapant trop tôt, c'est-à-dire avec des anticipations attentionnelles imprécises. Les auteurs font remarquer que l'altération de la perception des syllabes accentuées en anglais engendre des effets en cascade sur la structure hiérarchique de la phonologie. Le traitement du contexte temporel pour représenter les syllabes faibles et les phonèmes est en effet perturbé.

Une autre hypothèse est proposée par le TSF, cette fois à propos des rythmes cérébraux rapides chez les personnes dyslexiques. Selon Goswami (2011), les mécanismes d'analyse rapides supportés par les oscillations gamma seraient davantage préservés que les mécanismes d'analyse temporelle lente. De ce fait, les dyslexiques pourraient développer une perception allophonique, accordant trop d'importance à des traits acoustiques non-pertinents phonologiquement auxquels ils devraient être devenus sourds. Cela les conduirait à percevoir des différences trop saillantes entre des exemplaires d'une même catégorie phonémique. Cette



perception allophonique se ferait au détriment de la perception catégorielle des phonèmes, qui garantit un bon traitement de la parole ainsi que l'apprentissage de la lecture. Des données expérimentales soutiennent cette hypothèse d'une hyper-perception des différences acoustiques au sein d'une catégorie phonémique (Serniclaes *et al.*, 2001, 2003). Cette anomalie n'empêcherait pas la communication parlée, mais entraverait l'apprentissage des correspondances graphème-phonème pour la lecture (Serniclaes *et al.*, 2004, 2005).

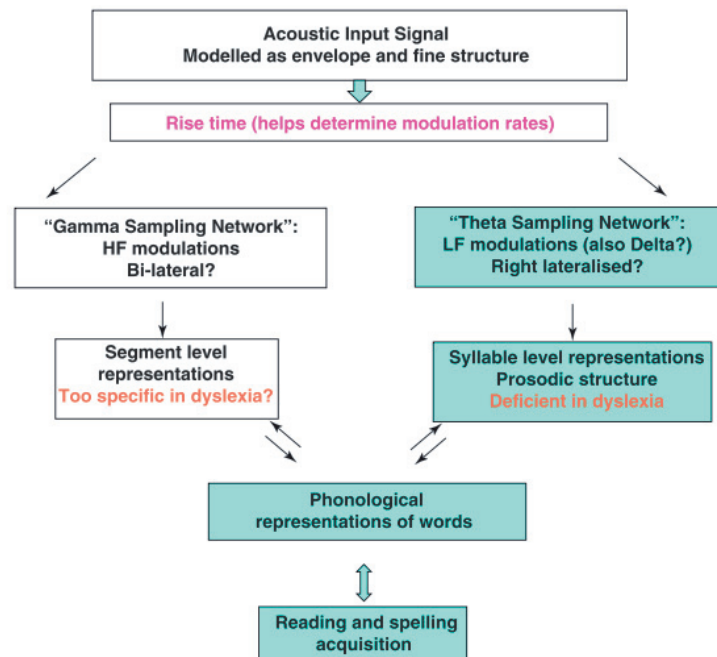


Figure 5. La théorie de l'échantillonnage temporel (*Temporal Sampling Framework*, TSF). Reproduit à partir de «A temporal sampling framework for developmental dyslexia.», par Goswami, U., 2011, *Trends in Cognitive Sciences*, 5(1), 3-10

Le modèle TSF a eu enfin un certain retentissement sur la recherche appliquée et les pistes de rééducation. Goswami (2011) incite à stimuler l'ajustement rythmique des enfants en vue d'améliorer le traitement du langage oral et, par suite, écrit. Des chercheurs ont suivi cette incitation et ont montré que la musique peut contribuer à de nouvelles stratégies de rééducation, surtout quand elles sont basées sur le rythme. Un impact est attendu sur les représentations phonologiques (Bhide *et al.*, 2013) et sur la sensibilité aux phonèmes (Thomson *et al.*, 2013). Des prises en charge de quelques semaines par la musique ont apporté des résultats encourageants, par exemple avec des progrès en conscience phonémique (Thomson *et al.*, 2013), en lecture (Flaugnacco *et al.*, 2015 ; Habib *et al.*, 2016), en perception catégorielle, en perception des caractéristiques temporelles de la parole, en attention auditive et en répétition de pseudo-mots (Habib *et al.*, 2016). Exploiter les similitudes musique/langage sur le plan du rythme pourrait être particulièrement bénéfique en cas de dyslexie. La musique pourrait améliorer la sensibilité supra-segmentale. Dans cette thèse,

nous nous appuyons aussi sur le cadre TSF et les modèles d'analyse de la parole à échelles multiples pour défendre l'hypothèse selon laquelle le traitement de petites unités linguistiques (phonèmes et traits acoustiques) pourrait aussi bénéficier d'une aide, mais ceci à travers une stimulation intense et ciblée offrant, grâce à l'écoute dichotique, une occasion de découpage temporel adapté à la durée des unités à traiter.

#### **1.4. Sensibilité à la durée : mesures en potentiels évoqués (ERP)**

La sensibilité passive et automatique aux catégories phonémiques peut être observée au moyen de l'électroencéphalographie (EEG), souvent avec le paradigme *oddball*. Avec cette technique de présentation, un stimulus déviant est présenté aléatoirement dans environ 20 % des items d'une série de stimuli standards (80%). Les réactions électrophysiologiques induites par ces deux types de stimuli sont multiples. Elles sont étudiées comme des composantes (ondes) dans les potentiels évoqués (ERP) reflétant différents mécanismes ou processus cognitifs. Ces composantes varient avec l'âge (Cheour *et al.*, 2001). Chez l'adulte, il est classique de distinguer des ondes précoces pré-attentionnelles, parmi lesquelles la P200 et la *Mismatch Negativity* (MMN), et les ondes attentionnelles, notamment la P300 et la *Late Discriminative Negativity* (LDN) qui reflètent des processus plus tardifs et des traitements de plus haut niveau.

##### **1.4.1. Les ondes pré-attentionnelles**

###### **1.4.1.1. La Mismatch Negativity (MMN)**

L'onde MMN est très étudiée en perception. Elle est négative et pointe en général entre 100 et 250 ms après la survenue d'une déviance perceptive (Näätänen & Picton, 1987). Elle révèle une sensibilité pré-attentionnelle à un événement perceptif traité (éventuellement non consciemment) comme différent du contexte. Elle a été décrite pour la première fois par Näätänen *et al.* (1978) dans une tâche d'écoute dichotique à la suite de l'article de Hillyard *et al.* Picton (1973). Ces chercheurs la dépeignent comme une onde négative visible environ 200 ms après un stimulus inattendu (Näätänen, *et al.*, 1989), souvent suivie d'une onde positive pointant environ 300 ms après la déviance (P300). Autrement dit, elle reflète la détection de tout changement discriminable dans une séquence de stimuli auditifs répétitifs (Pulvermüller & Shtyrov, 2006). Pour le calcul de son amplitude, la MMN est le résultat de la soustraction entre la négativité observée après les déviants et celle – plus petite ou absente – observée dans la même fenêtre temporelle après les standards. Elle est particulièrement ample quand elle est

recueillie sur les électrodes frontales et centrales chez l'adulte, et parfois pariétales chez l'enfant (Cheour *et al.*, 1996, 1998 ; Leppänen *et al.*, 1997). De plus, plus la détection de la déviance est facile, plus la latence du pic de la MMN est courte (Näätänen *et al.*, 1989 ; Sams *et al.*, 1985). Une des conditions d'apparition de la MMN est qu'une quantité suffisante de stimuli standards aient été présentés pour que l'auditeur construise une représentation de ceux-ci (i.e. une catégorie mentale) et que la survenue du déviant vienne enfreindre ce qui définit cette représentation. Elle témoigne d'une mémoire échoïque du stimulus standard et d'une anticipation perceptive à propos du stimulus suivant. Autrement dit, la MMN reflète l'existence d'une représentation mentale, et elle peut indiquer par exemple qu'une représentation phonologique d'un phonème existe bien chez l'auditeur. Elle est en effet d'autant plus ample que la différence standard-déviant fait partie du système de représentation de l'auditeur, et s'il s'agit de langage, de son système phonologique (Sams, *et al.*, 1985). Elle permet aussi bien d'étudier les aspects neuronaux de la perception auditive centrale que les différentes formes de mémoires auditives et les processus contrôlant l'accès aux entrées sensorielles auditives (Näätänen *et al.*, 2007).

L'apparence de la MMN se développe avec l'âge et sa latence diminue entre l'enfance et l'âge adulte (Cheour *et al.*, 2001). Par ailleurs, il est possible de modifier son amplitude, sa latence ou sa durée à l'aide d'entraînements perceptifs (Atienza *et al.*, 2002 ; Tamminen *et al.*, 2015 ; Tremblay *et al.*, 1997). Dans le cadre de l'apprentissage d'une L2, plusieurs travaux ont pu montrer qu'un changement de la MMN était un bon témoin de l'acquisition de nouveaux phonèmes d'une L2 (Heidlmayr *et al.*, 2021 ; Krzonowski, 2020 ; Menning *et al.*, 2002 ; Saloranta *et al.*, 2022).

#### **1.4.1.2. L'onde P200 (P2)**

La P2 est une onde positive survenant aux environs de 150 ms après l'apparition d'un stimulus, et elle précède directement la MMN. Cette onde très précoce reflète un mécanisme non-conscient qui est avant tout une réaction à une stimulation, par exemple acoustique. Des données montrent qu'elle est d'autant plus ample que le stimulus est traité de manière approfondie, focalisée, en inhibant les distracteurs (Smith *et al.*, 2004). Son amplitude est généralement liée à la saillance de la stimulation sensorielle.

Cette onde a aussi la particularité de diminuer en amplitude dans certaines circonstances, et l'interprétation de cette variation est sujette à débat. Une explication pour sa diminution en réponse à certains stimuli est qu'elle traduit un filtrage précoce de

l'information, appelé *sensory-gating*. Une telle diminution est, en effet, mise en évidence avec le paradigme d'écoute passive de doubles-clics, dans lequel deux stimuli auditifs identiques se suivent. L'onde P2 est moins ample après le deuxième stimulus, car sa similitude avec le précédent inciterait le cerveau à envoyer un message aux aires sous-tendant des traitements plus élaborées, message les enjoignant de ne pas traiter le stimulus de manière approfondie (Lijffijt *et al.*, 2009a). Des données en pathologie concordent avec cette idée, en montrant qu'il n'y a pas de diminution d'amplitude de l'onde P2 pour des stimuli n'apportant pas d'information nouvelle chez des patients atteints de troubles de l'inhibition, par exemple en cas de schizophrénie (Boutros *et al.*, 2004), de trouble bipolaire (Lijffijt *et al.*, 2009b) ou de TDAH (Broyd *et al.*, 2005 ; Lazzaro *et al.*, 2001 ; Oades *et al.*, 1996). L'étude du *sensory-gating*, à travers les modifications de l'onde P2 est très intéressante sur le plan de l'aide au diagnostic. L'absence de *sensory-gating* permet ainsi de déceler certaines perturbations des processus d'inhibition chez des patients pour qui les tests comportementaux classiques échouent à le faire. C'est par exemple le cas chez des consommateurs d'ecstasy (Roberts *et al.*, 2013).

Ce phénomène d'inhibition très précoce affaiblirait les ressources allouées à l'approfondissement du traitement, ce qui éviterait à des régions cérébrales impliquées dans des analyses de haut niveau – notamment frontales – d'être submergées par des informations sans intérêt (Venables, 1964). Par exemple, dans des épreuves de GoNogo auditives, où il s'agit d'inhiber la réponse motrice associée aux stimuli standards lorsqu'un déviant survient, la diminution d'amplitude de l'onde P2 est liée à la réussite de l'inhibition de réponse (Bedoin *et al.*, 2019 ; Benikos *et al.*, 2013). Il s'agit donc bien du corrélat d'un mécanisme inhibiteur. En bloquant l'approfondissement du traitement d'informations considérées comme banales, ce mécanisme de filtrage précoce est adaptatif (Lijffijt *et al.*, 2009a). En effet, organiser et hiérarchiser la pertinence des informations sensorielles en provenance d'un environnement en constante évolution sont primordiales pour ajuster le comportement. Cela protège de manière globale les traitements cognitifs de haut niveau. L'information doit être filtrée et le contrôle inhibiteur passe pour cela par des mécanismes bottom-up et top-down. Sur le mode bottom-up, le cerveau est en mesure de moduler sa sensibilité à l'information entrante et inhibe automatiquement certaines informations sensorielles non-pertinentes, par exemple parce qu'elles sont répétitives. Souvent étudiée dans le complexe d'ondes P50-N1-P2, la P2 reflèterait ainsi plus particulièrement un filtrage de la distribution des ressources attentionnelles (Lijffijt *et al.*, 2009a).

Dans le cadre de l'apprentissage d'une deuxième langue, la survenue d'un phonème déviant présentant une différence saillante avec les standards dans un paradigme de *oddball* pourrait donner lieu à l'apparition d'une P2. Des données montrent que lorsque la différence acoustique est infime, la P2 ne se produit qu'après un entraînement perceptif (Tremblay *et al.*, 2001). Lorsque la différence est au contraire saillante et inhabituelle dans la langue maternelle de l'apprenant, on peut faire l'hypothèse qu'un entraînement perceptif puisse réduire l'amplitude initiale de la P2. Une telle diminution traduirait le fait que la nouvelle différence phonémique est devenue familière, et n'est plus assez étrange pour nécessiter une analyse approfondie.

## **1.4.2. Les ondes attentionnelles**

### **1.4.2.1. L'onde P300**

La P300 est une onde positive survenant autour de 300 ms après l'apparition d'un stimulus déviant, tout de suite après une MMN. Cette onde traduirait, quant à elle, des processus plus tardifs d'évaluation ou de catégorisation de nouveautés. Elle est plus marquée sur les électrodes pariétales. Elle a été décrite dès 1964 par Chapman et Bragdon.

Il existe deux types de P300, une plus précoce (P3a) et l'autre plus tardive (P3b), traduisant différents traitements. La P3a survient entre 200-650 ms après le début de la déviance (Polich, 2007). La composante P3a suit souvent la MMN, et indique un changement d'attention non-volontaire et potentiellement non-conscient (Polich, 2007) vers les sons déviants. Contrairement à la MMN, et en complément de celle-ci, la P3a est supposée refléter un processus de haut niveau consacré à la détection d'événements qui nécessitent un traitement ultérieur (Horváth, *et al.*, 2008). Elle serait associée à une activité d'alerte émanant du cortex frontal (Wronka *et al.*, 2009). Deux sources peuvent moduler indépendamment la P3a : (1) une violation de la régularité qui a été détectée par des processus sensoriels de bas niveau, (2) une activation cérébrale spécifique provoquée par des souvenirs à long terme préexistants (c'est-à-dire sémantiques ou émotionnels, catégories, concepts, phonèmes...).

Anciennement appelée P3b (Squires *et al.*, 1975), la P300 serait provoquée par la survenue d'un événement improbable, inattendu. Ainsi, sur le plan perceptif, moins l'événement est prédictible, plus l'amplitude de la P300 est grande (Donchin, 1981). La latence de cette onde est aussi sensible à la prédictibilité du stimulus. Par exemple, lorsque les stimuli d'une série sont présentés d'une manière temporellement régulière, la discrimination de traits dans ces stimuli est mieux réussie que si la présentation est irrégulière. La

prédictibilité du stimulus est ainsi élevée, ce qui améliore son traitement et fait décroître la latence de l'onde P3b associée (Schmidt-Kassow *et al.*, 2009).

La P300 est aussi interprétée comme le reflet d'une attraction de l'attention par un élément acoustique du stimulus, et par l'intérêt que lui porte l'auditeur. Par exemple, Fosker et Thierry (2004) ont réalisé une expérience de décision lexicale (i.e. la tâche attire l'attention sur la lexicalité) et seuls 20% des mots ne commençaient pas par le même phonème que les autres. Ces déviants ont suscité l'augmentation de la P300 chez des adultes normo-lecteurs, mais pas chez des dyslexiques, ce qui traduirait l'incapacité de ces derniers à détecter un changement phonologique lorsqu'ils ne le recherchent pas pour réaliser la tâche. Les mêmes auteurs ont publié en 2005 une autre expérience dans laquelle la P300 apparaît dans les deux groupes, parce que la tâche demande explicitement de détecter les mots initialisés par un phonème différent des standards (Fosker & Thierry, 2005). Selon Wronka *et al.* (2009), la P3b serait issue d'un réseau temporo-pariétal lorsqu'une comparaison s'opère entre le stimulus et une connaissance de façon attentionnelle. Ce point est intéressant pour l'étude de la participation de connaissances stables en mémoire à long terme dans l'identification de phonèmes nouvellement appris. Cela souligne les liens entre cette onde et l'attention, mais suggère aussi que l'absence de modulation de cette onde par un changement acoustique pourrait rendre compte d'un traitement peu approfondi de ce changement chez des personnes dyslexiques.

Bien que la P300 suive la MMN, ces deux ondes ne sont pas fortement couplées. Elles sont considérées comme les reflets de mécanismes dissociés (Rinne *et al.*, 2006). La P300 peut ainsi émerger ou croître en amplitude sans augmentation simultanée de la MMN, ce qui suggère des mécanismes cognitifs fonctionnant au moins en partie de façon parallèle (Horváth *et al.*, 2008). Alors que la MMN est considérée comme un indice privilégié de la sensibilité pré-attentionnelle à une différence, la P3b se distinguerait par sa grande sensibilité à la focalisation attentionnelle, comme en témoigne son amplitude plus grande dans les expériences de détection active que dans les expériences de oddball passif (Wronka *et al.*, 2008).

Comme la P300 accompagne une étape de traitement souvent consciente, donnant par exemple des indications sur la discrimination subjective de stimuli auditifs chez des patients sourds porteurs d'implants cochléaires (Jacquemin *et al.*, 2019), cette onde apparaît aussi dans des tâches de catégorisation où il faut, par exemple, détecter volontairement certains déviants comme n'appartenant pas à la catégorie de référence. En ce cas, l'amplitude de la P300 est

d'autant plus grande que le stimulus dévie beaucoup par rapport aux standards (Wronka *et al.*, 2009). C'est le cas lorsque le nombre de traits par lesquels le déviant diffère des standards est grand. Par exemple, l'amplitude de la P3 est plus élevée pour les phonèmes déviants qui diffèrent du standard par un plus grand nombre de traits phonétiques, ou par des traits particulièrement saillants (Bedoin *et al.*, 2013). C'est pourquoi l'étude de cette réponse évoquée par un phonème d'une langue non-native peut aussi fournir des informations sur des modifications de perception de celui-ci après un entraînement perceptif.

#### **1.4.2.2. Late Discriminative Negativity (LDN)**

La LDN est une autre onde générée au niveau fronto-central en réponse à un stimulus déviant parmi des standards. Elle se produit pour des sons de parole ou des sons non-verbaux, mais son amplitude est plus grande en cas de sons de parole. Elle présente un pic d'amplitude entre 300 et 600 ms après le début du stimulus. Différents points de vue coexistent quant à son interprétation dans la littérature. Elle pourrait refléter l'évaluation du stimulus à un niveau cognitif attentionnel. Du fait qu'elle présente une latence plus élevée que la MMN, elle ne devrait pas refléter des traitements purement sensoriels. Bien que certaines études portent sur cette onde chez l'adulte (Jakoby *et al.*, 2011 ; Shulte-Körne *et al.*, 1999b), elle est le plus souvent analysée chez l'enfant (Alonso-Bua *et al.*, 2005 ; Korpilahti *et al.*, 1995 ; Moreno *et al.*, 2015 ; Shulte-Körne *et al.*, 1998), car il semblerait que son amplitude décroisse avec l'âge (Bishop *et al.*, 2011 ; Cheour *et al.*, 2001). Cette onde présente un intérêt particulier, car il semblerait qu'elle soit présente en cas de dyslexie et pourrait donc être une autre mesure possible en cas d'absence de MMN (Shulte-Körne *et al.*, 1998, 1999, 2001). Elle n'est cependant pas étudiée chez les normo-lecteurs.

### **1.5. Objectifs de la thèse**

L'objectif principal de cette thèse est la création et l'évaluation d'outils numériques pédagogiques pour aider à l'apprentissage de certains aspects de la phonologie et – plus marginalement – de la prosodie en anglais L2 chez de jeunes étudiants. Ces outils d'aide à l'apprentissage ont été imaginés à partir du cadre théorique présenté, centré sur les notions d'attention dynamique et d'asymétrie des rythmes cérébraux sous-tendant la séquenciation du signal de parole. Nous avons mené nos études avec du matériel français et anglais auprès d'étudiants normo-lecteurs et dyslexiques francophones natifs. Notre réflexion a conduit à six études répondant à cet objectif à travers l'analyse de performances (perception et production) et l'analyse de données EEG (enregistrement de potentiels évoqués).



Les hypothèses générales suivantes ont été formulées :

**Hypothèse générale 1.** Il est possible, à l'âge adulte, d'améliorer le traitement d'un nouveau contraste phonémique entre des consonnes ou entre des voyelles nouvelles grâce à un entraînement qui le cible en apprentissage tardif d'une L2.

**Hypothèse générale 2.** Un dispositif de stimulation dichotique, conçu pour favoriser un traitement précis de nouveaux phonèmes en les adressant plus directement à l'hémisphère cérébral le plus compétent (selon le modèle AST) pour traiter les indices de durée courts ou longs qu'ils comportent, pourrait optimiser l'effet positif d'un entraînement perceptif ciblant ces phonèmes et les contrastes phonologiques qui les opposent en L2 chez des apprenants adultes.

**Hypothèse générale 3.** Le modèle AST (Poeppel, 2003), qui suppose des fenêtres d'attention temporelle offrant des capacités d'analyse de la parole à échelles multiples, offertes par des oscillations cérébrales rapides dans le cortex auditif gauche et plus lentes dans le cortex auditif droit, est une proposition théorique valide, car efficacement exploitable par la stimulation dissociée de l'un ou l'autre hémisphère selon leurs capacités d'analyse temporelle supposées complémentaires. Exploiter cette technique sur la base de ce modèle pourrait déboucher sur une aide à l'apprentissage de la phonologie dans une L2.

**Hypothèse générale 4.** Un entraînement perceptif ciblé sur un contraste phonémique peut aussi permettre un meilleur traitement de phonèmes en L1 ou en L2 en cas de pathologie du langage (dyslexie développementale) chez des adultes, bien que l'entraînement soit bref.

**Hypothèse générale 5.** La stimulation par un amorçage rythmique musical couplée à un retour visuel dans des exercices de répétition de phrases, peut permettre l'apprentissage de particularités prosodiques dans une L2 apprise tardivement.

**Hypothèse générale 6.** Le traitement du signal de parole sur des unités syllabiques portant des modulations prosodiques peut être réalisé plus efficacement et donner lieu à l'apprentissage de particularité prosodiques dans une L2, si l'extraction de ces unités syllabiques est favorisée par un amorçage rythmique à 2 Hz. Cette prédiction s'appuie sur une conception de l'attention selon laquelle celle-ci serait distribuée de manière cyclique (modèle DAT, Jones, 1976 ; Barnes & Jones, 2000), et sous-tendue par les rythmes cérébraux capables de synchroniser la distribution des ressources au rythme du signal de parole (Giraud & Poeppel, 2012). Selon cette perspective théorique, la synchronisation de l'attention sur un



rythme lent (2 Hz) devrait engager une cascade de synchronisations pour d'autres rythmes plus rapides (Ghitza, 2011) permettant une analyse suffisamment fine pour des syllabes et leur contenu prosodique. Les prédictions des recherches en neurosciences cognitives sur les oscillations cérébrales et leur potentiel explicatif pour le fonctionnement cognitif (Murphy & Benitez-Burraco, 2019) et une conception de l'attention comme sous-tendue par ces oscillations le font espérer. En effet, plus précisément, la stimulation d'oscillations endogènes delta par une amorce rythmique à 2 Hz pourrait renforcer la synchronisation de ces oscillations et les modulations d'amplitude qu'elles exercent sur les oscillations thêta, adaptées aux syllabes, agissant elles-mêmes sur l'amplitude des oscillations gamma induisant un traitement très précis des détails intra-syllabiques utiles à l'apprentissage de particularités prosodiques d'une L2.

### **Étude 1 : Perception catégorielle de consonnes françaises**

Les tests de perception catégorielle sont des outils permettant d'évaluer l'existence et la solidité d'un système phonologique. Nous avons vu qu'ils sont aussi connus pour donner lieu à des performances déficitaires en cas de dyslexie. L'évaluation de la dyslexie à l'âge adulte est aujourd'hui difficile à cause du manque d'outils de diagnostic normés pour mesurer le déficit de mécanismes cognitifs sous-jacents à la pathologie chez cette population.

Deux épreuves de perception catégorielle (i.e. identification et discrimination) sur des consonnes occlusives opposées par le voisement en français ont été créées afin de tenter de répondre à ce besoin. **L'Expérience 1** visait à évaluer l'effet test-retest de ces nouveaux tests comportementaux de jugement explicite chez de jeunes adultes, en vue de leur utilisation pour évaluer l'efficacité de programmes de remédiation phonologique. Secondairement, d'autres qualités psychométriques de ces épreuves ont été évaluées (fidélité, sensibilité, spécificité) pour une réflexion sur leur usage dans le cadre de diagnostics.

### **Étude 2 : Remédiation de la perception du voisement en français pour adultes dyslexiques**

À la suite de l'évaluation des épreuves de perception catégorielle en français de l'Étude 1, l'Étude 2 avait pour objectif d'évaluer l'effet d'un entraînement intensif bref (5 jours, 30 min/jour) portant sur les oppositions de voisement des consonnes occlusives françaises, au niveau comportemental (**Expérience 2**) et en potentiels évoqués (**Expérience 3**).

L'innovation dans cet entraînement perceptif était l'utilisation de l'écoute dichotique comme moyen de stimulation permettant de favoriser l'association hémisphère – son la plus compétente en français pour la durée des stimuli, c'est-à-dire HG (oreille droite) – occlusives sourdes et HD (oreille gauche) – occlusives sonores. Dans les occlusives françaises, le délai de voisement (VOT) est en effet un indice acoustique temporel important pour distinguer les consonnes sourdes (court VOT positif) des sonores (long VOT négatif). Cet entraînement avait pour but de tenter de renforcer le système phonologique des adultes dyslexiques, toujours instable à leur âge.

### **Étude 3 : Perception du voisement en anglais L2**

L'Étude 3 nous a permis d'évaluer la perception catégorielle de consonnes anglaises opposées par le voisement chez des adultes francophones (dyslexiques ou non). Pour cela, des épreuves de perception catégorielle ont été créées en anglais et l'évaluation a été réalisée au niveau comportemental (**Expérience 4**) pour mesurer un possible effet test-retest ainsi que la fidélité des épreuves. Cela nous a permis de proposer un début de norme en termes de moyenne, écart-type et intervalle de confiance pouvant être utilisé à posteriori dans le cadre d'évaluation de niveau dans le monde académique. L'**Expérience 5** nous a ensuite permis de comparer les performances (mesures comportementales) des adultes normo-lecteurs et dyslexiques, avant un entraînement utilisant le dispositif dichotique en anglais, sur la perception catégorielle anglaise, afin de mesurer une potentielle différence entre ces deux populations

### **Étude 4 : Aide à l'apprentissage de la perception du voisement en anglais L2**

L'Étude 4 regroupe quatre expériences dont l'objet était d'évaluer l'efficacité d'un entraînement bref intensif (5 jours, 30 min/jour) en anglais portant sur les consonnes occlusives chez des étudiants avec et sans dyslexie. L'**Expérience 6** visait à montrer l'intérêt de l'utilisation de l'écoute dichotique dans un entraînement perceptif pour améliorer les performances en perception catégorielle de consonnes opposées en voisement, par rapport à l'écoute binaurale chez des adultes normo-lecteurs. Les deux conditions d'écoute étaient comparées. L'écoute dichotique était supposée placer l'apprenant dans de meilleures conditions pour traiter efficacement les traits temporels dans les consonnes entendues. Pour cela, des mots anglais initialisés par une occlusive sonore comportant un court VOT étaient adressés à l'oreille droite pour être traités prioritairement par l'HG, supposé plus compétent pour analyser des indices brefs. Les mots initialisés par une occlusive sourde comportant un long VOT étaient adressés à l'oreille gauche-HD. Une concurrence était systématiquement

assurée par un bruit blanc dans l'autre oreille. L'intérêt supposé de l'écoute dichotique était de faire en sorte que, sur le plan de l'asymétrie hémisphérique associée au traitement des durées longues et courtes d'après le modèle AST. En effet, avec ce dispositif, les francophones ne traiteraient pas les occlusives sourdes et sonores comme ils le font en français avec un investissement plus fort de l'HG pour les sourdes et de l'HD pour les sonores. Ce dispositif dichotique visait à permettre un traitement optimal des consonnes anglaises au moment de l'écoute, et potentiellement à apprendre, en contexte anglophone, à systématiser l'association des consonnes sonores et sourdes anglaises à l'hémisphère le plus compétent pour les prendre en charge. Par exemple, le graphème « b » est identique dans les deux langues, mais le phonème correspondant contient un VOT long en français, mais un VOT court en anglais ; le graphème « p » est aussi utilisé dans les deux langues, mais le phonème correspondant se réalise avec un VOT court en français et un VOT long en anglais. L'objectif de l'Etude 4 était de tenter d'aider implicitement l'auditeur à différencier l'implication asymétrique des hémisphères cérébraux en fonction des durées propres aux langues française vs. anglaise, et non des noms (graphèmes) des phonèmes. **L'Expérience 7**, réalisée en EEG, était destinée à évaluer l'effet de l'entraînement par stimulation dichotique en potentiels évoqués, cette fois en condition d'écoute passive, avec la possibilité de mesurer des effets sur des mécanismes pré-attentionnels. Cet entraînement en anglais a aussi été évalué auprès d'une population d'adultes dyslexiques francophones, au niveau comportemental, en perception catégorielle de consonnes anglaises (**Expérience 8**). **L'Expérience 9** nous a permis de contrôler que l'entraînement en anglais n'avait pas de répercussion sur la phonologie en français aussi bien chez les normo-lecteurs, que chez les dyslexiques, dont le système phonologique est instable (même si renforcé par la remédiation en français des Expériences 2 et 3).

### **Étude 5 : Aide à l'apprentissage de la perception des durées des voyelles en anglais L2**

L'Étude 5 observait l'effet d'un entraînement bref intensif (3 jours, 20 min/jour), lui aussi basé sur une stimulation dichotique conforme à l'asymétrie hémisphérique pour le traitement des durées, sur un autre indice phonologique de la langue anglaise : la durée de voyelle. Son intérêt réside dans le fait que la durée des voyelles n'est pas pertinente sur le plan phonologique en français, et qu'il est nécessaire pour un francophone apprenant l'anglais en L2 de créer un nouveau système de représentation phonologique propre à la L2 pour cette notion. **L'Expérience 10**, nous a permis d'évaluer l'effet test-restes et la fidélité de 4 tests (2 en perception de mots et 2 en production de mots) ensuite utilisés pour évaluer et comparer l'effet des entraînements sur les voyelles. **L'Expérience 11** visait à montrer l'intérêt de

l'utilisation de l'écoute dichotique dans un entraînement perceptif pour améliorer les performances en perception et en production de voyelles anglaises, par rapport à l'écoute binaurale chez des adultes normo-lecteurs. Les deux conditions d'écoute étaient comparées. Comme dans l'étude précédente, l'écoute dichotique était supposée placer l'apprenant dans de meilleures conditions pour traiter efficacement les traits temporels dans les voyelles entendues. Cette notion de durée de voyelles n'existant pas en français, nous avons souhaité, par cet entraînement voir s'il était possible d'étendre les effets obtenus sur les consonnes vers d'autres phonèmes opposés par la durée que sont les voyelles anglaises. Les voyelles courtes étaient jouées dans l'oreille droite-HG et les voyelles longues dans l'oreille gauche-HD. La concurrence était ici aussi assurée par un bruit blanc joué parallèlement dans l'oreille opposée. Dans cette expérience, les mesures ont été faites au niveau comportemental. Des adultes dyslexiques ont aussi pu bénéficier de l'entraînement par stimulation dichotique et ses effets ont été mesurés au niveau comportemental (**Expérience 12**).

### **Étude 6 : Aide à l'apprentissage de la prosodie en anglais L2**

La dernière étude de cette thèse portait sur un autre aspect de l'anglais, difficile à apprendre en tant que francophone : la prosodie. À travers deux expériences comportementales, nous avons retenu le cas des questions en anglais, car ces dernières peuvent présenter deux types d'intonations à la fin, porteuses d'informations sur la nature de la réponse attendue (questions ouvertes vs. fermées). Une différence existe en français entre des questions avec intonation finale montante ou descendante, mais de façon très atténuée par rapport à l'anglais. Aider les apprenants à devenir plus sensibles à cette différence est important pour leur compréhension en anglais. La perception de tels indices supra-segmentaux, souvent de durée assez longue, étant déficitaire en cas de dyslexie, cette étude a porté sur deux groupes de jeunes adultes, les uns dyslexiques, les autres non. L'entraînement était particulièrement court et proposait cette fois une activité de production, en une séance de 45 min. Les exercices avaient ceci d'innovant qu'ils combinaient 1) un retour visuel direct sur la qualité acoustique de la répétition de phrases et 2) l'utilisation d'amorces rythmiques supposées pré-ajuster le rythme attentionnel de manière favorable à l'extraction de changements relativement lents de hauteur de voix en fin de phrase. **L'Expérience 13** évaluait séparément l'effet des deux aspects de l'aide apportée (retour visuel, amorçage rythmique, combinaison des deux) chez des adultes sans pathologies. En se fondant sur les résultats de l'Expérience 13, le meilleur entraînement a été proposé à des adultes dyslexiques (**Expérience 14**).

## 1.6. Méthodologie de la thèse

Cette thèse s'est articulée autour de protocoles d'entraînements réalisés pour la plupart sur plusieurs jours. Nous avons suivi une méthodologie d'étude longitudinale courte. Nous avons vu les participants deux à trois fois, à plusieurs jours d'intervalle. Pour les entraînements musicaux, des lignes de conduite à adopter ont été formalisées afin d'améliorer la transparence et la spécificité des articles sur le sujet (Robb *et al.*, 2011), et nous avons tenté de les adopter pour l'entraînement avec amorçage dans l'Étude 6. Dans la littérature scientifique, de telles règles ne sont malheureusement pas souvent suivies pour la description des entraînements perceptifs, souvent peu détaillée, difficile à reproduire ou relater. Cela rend plus difficile la comparaison entre les entraînements perceptifs dichotiques et binauraux que nous avons utilisés dans les Études 2, 4 et 5 et les travaux analogues dans le domaine.

Cette thèse s'étant déroulée dans un contexte sanitaire particulier lié à la COVID-19, presque tous les protocoles d'entraînement ont pu être adaptés. Les entraînements avaient été initialement conçus pour être réalisés en autonomie, au domicile du participant, mais les épreuves de pré- et post-entraînements devaient, quant à elles, être réalisées en présence d'un expérimentateur. La COVID-19 a été un réel défi pour les travaux présentés dans cette thèse. Ici, elle a demandé de s'adapter aussi bien au niveau matériel que personnel, pour que toutes les étapes de l'étude longitudinale se fassent à distance, hormis les enregistrements EEG. Par ailleurs, le projet de thèse initial a dû être modifié, car l'expérimentation envisagée dans les collèges et lycées n'était plus possible. Un nouveau protocole d'entraînement auprès d'adultes a été créé pour compenser (Étude 6 : Expériences 14 et 15). Afin que les protocoles puissent être réalisés à distance, nous avons créé une fiche de passation (Annexe 1 et 2) qui était envoyée aux participants lors de la prise de rendez-vous qui expliquait toutes les manipulations à réaliser depuis l'installation des logiciels jusqu'à l'envoi des fichiers réponses. En plus de cette fiche, nous les suivions par téléphone à chaque étape des tests et nous nous assurons que les entraînements avaient été faits en vérifiant les fichiers réponses. Ce suivi rapproché a été primordial pour que les données soient de la meilleure qualité possible. Lors du dernier appel téléphonique, nous validons la réception des données par mail et nous proposons aux participants un débriefing de l'expérience et l'explication de celle-ci. Ils ont aussi reçu, pour leur participation, un cadeau d'une valeur de 10€ (un livre ou un jeu de société, parmi une sélection) pour les expériences avec mesures comportementales et 40€ par virement bancaire pour les expériences en EEG. Cependant, certains protocoles (Expérience 12 en EEG et une partie des Expériences 14 et 15 où un retour visuel était nécessaire)

nécessitaient du matériel spécifique et devaient être réalisés en présentiel. Ils ont été repoussés jusqu'à ce que les expérimentations en laboratoire redeviennent possibles selon les consignes sanitaires. Cette période a demandé une redéfinition des priorités dans le calendrier de thèse.

Grâce aux ajustements imposés par ces défis, nous avons récolté un grand nombre de données à travers 6 études. C'est pourquoi lors de la rédaction de ce manuscrit et dans un souci de clarté, nous avons découpé les différentes études en plusieurs expériences, même si les participants étaient parfois les mêmes. Cela est rappelé dans la partie méthode de chaque expérience (Annexe 3).

Tout le travail expérimental a été approuvé par le Comité de Protection des Personnes. Tous les participants ont reçu une notice d'information (Annexe 4) et ont signé un formulaire de consentement (Annexe 5) éclairé.

## Chapitre 2. Perception catégorielle

---

### 2.1. Définitions : perception catégorielle, allophonique, et frontière catégorielle

D'une manière générale, la perception catégorielle (PC) (i.e. non-linéaire) s'oppose à la perception de changements graduels prenant la forme de réponses linéaires à des changements eux aussi linéaires dans une série de stimuli physiques.

Dans le domaine du langage, selon la version stricte de la perception phonologique catégorielle (Liberman *et al.*,1957), deux stimuli linguistiques seraient perçus comme différents acoustiquement seulement s'ils appartiennent à deux catégories distinctes dans le système phonologique de l'auditeur. Il s'agit alors de sons qui relèvent de phonèmes distincts avec au moins un trait phonologique qui diffère. Il peut s'agir par exemple du lieu d'articulation (e.g. /p/-/t/) ou du voisement (e.g. /p/-/b/). Pour une différence acoustique équivalente, mais entre deux exemplaires d'une même catégorie, la différence ne serait pas perçue : les stimuli ne sont pas discriminés.

Ce mode de perception n'est pas le seul possible. Ainsi, lorsque les exemplaires d'une catégorie ne sont pas catégorisés de façon homogène comme appartenant sans conteste à une seule des deux catégories proposées, la perception est dite « allophonique », et non catégorielle.

Beaucoup d'expériences de perception de sons de parole constituant des intermédiaires régulièrement séparés par des différences équivalentes entre deux phonèmes montrent que les réponses données pour les identifier se distribuent de manière discontinue. Un changement brusque du nom de catégorie choisi se produit à un endroit précis du continuum : cela reflète la PC phonémique. S'il faut par exemple identifier les événements d'un continuum entre /p/ et /b/, la courbe des réponses des auditeurs montre qu'un nombre restreint d'évènements sont perçus comme ambigus (les réponses ne diffèrent alors pas du hasard). Les autres événements sont au contraire perçus sans hésitation comme relevant de l'une des catégories. L'ensemble de la courbe de réponse est alors non-linéaire. Une grande partie des stimuli sont par exemple catégorisés dans près de 100% des réponses comme des /p/ et les autres comme des /b/. Une frontière catégorielle abrupte démarque alors les réponses entre deux blocs de stimuli perçus sans hésitation comme des exemplaires de l'une des catégories phonémiques.

Les phonèmes seraient ainsi délimités par des frontières dont la netteté reflète la clarté de l'organisation du système phonologique. Cette frontière est floue chez l'enfant, plus précise chez l'adulte (Aslin *et al.*, 1981) et se détériorerait chez la personne âgée (Strouse *et al.*, 1998).

La précision de la frontière dépend de l'appartenance des phonèmes au système phonologique de l'auditeur, car la perception catégorielle implique l'intervention de connaissances phonologiques spécifiques à la langue. Pour être efficace, la représentation d'un phonème doit agir comme un attracteur puissant qui rassemble de façon très homogène dans une même catégorie des exemplaires variés (phones) censés être perçus comme dénués de différence acoustique, ou avec des différences très faibles selon la version moins stricte, et plus réaliste, de la perception catégorielle (Mcmurray & Aslin, 2005).

## **2.2. Épreuves d'évaluation de la perception catégorielle : identification et discrimination**

Il existe deux épreuves d'évaluation de la PC en tâches comportementales : l'identification et la discrimination. Celles-ci utilisent un continuum de stimuli variant de manière linéaire selon un seul paramètre acoustique entre deux prototypes de phonèmes. Les intermédiaires peuvent varier de l'un à l'autre par une augmentation graduelle du paramètre acoustique (e.g. pour le voisement, la durée du VOT peut varier par pas de 10 ms). Une différence de 10 ms n'est pas perçue consciemment, mais des données en IRMf montrent que le système neuronal est sensible à une telle différence (Blumstein *et al.*, 2005).

L'épreuve d'identification demande de choisir entre les noms de deux phonèmes cibles pour une série d'items du continuum présentés aléatoirement. Ce type de tâche est utilisé dès l'âge de 5 ans (Boets *et al.*, 2011). Les pourcentages de réponses désignant un des phonèmes permettent de dessiner la courbe d'identification sur laquelle chaque item (avec, par exemple, une valeur croissante de VOT) est replacé en abscisses sur le continuum. Une perception catégorielle se traduit par une sigmoïde dont la pente est quantifiée pour son aspect plus (perception très catégorielle) ou moins abrupt (perception moins catégorielle). La sigmoïde est délimitée par deux plateaux reflétant la faible sensibilité aux différences intraphonémiques (Fig. 6, exemple du couple /b-p/). Le seuil d'identification (point de la courbe pour 50% d'identification) doit être conforme à celui de la langue étudiée (e.g. pour l'opposition de voisement : VOT = 0 à 10 ms en français, et 25 à 30 ms en anglais).



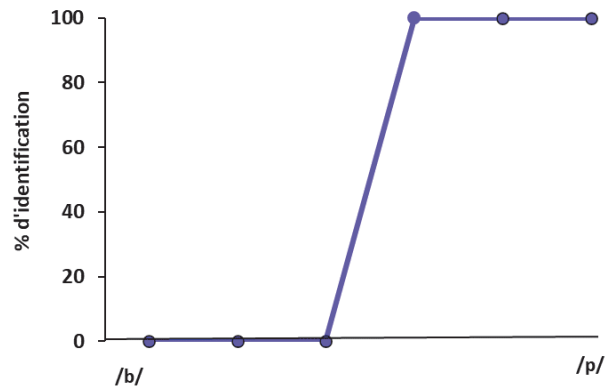


Figure 6. Courbe d'identification entre les phonèmes /b/ et /p/.

Dans l'épreuve de discrimination, elle aussi utilisée à partir de 5 ans (par exemple en français, Bogliotti *et al.*, 2002) les stimuli du continuum séparés de deux pas (e.g. VOT 10 et 30) sont présentés par couples. L'auditeur doit décider s'il y a une différence acoustique entre les deux membres du couple. Cela ne demande pas explicitement de mobiliser les représentations phonémiques. Dans 50% des cas, le même stimulus est présenté deux fois. La perception catégorielle est une illusion perceptive par laquelle une même différence acoustique est mieux détectée (pic de discrimination) si les deux stimuli sont de part et d'autre de la frontière phonémique définie par le test d'identification (Fig. 7). Un pic ample traduit une frontière nette. Cette tâche permet aussi de repérer la persistance de frontières catégorielles non-pertinentes pour la langue lorsque plusieurs pics se dessinent, ce qui est un indice d'allophonie. La perception est aussi dite catégorielle si le pic de discrimination correspond au milieu de la sigmoïde en identification (perception catégorielle relative, PCR).

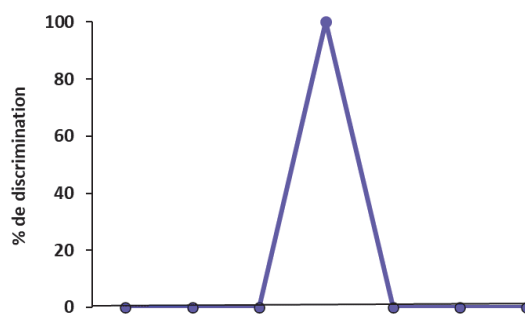


Figure 7. Courbe de discrimination.

Une autre approche de la PC consiste à étudier la discrimination pré-attentionnelle de contrastes entre des sons de parole qui appartiennent ou non à la même catégorie phonémique et qui sont présentés en série en utilisant la méthodologie d'EEG. Une sensibilité passive plus élevée est attendue si le contraste est phonémique. Cela se traduit en potentiels évoqués (ERPs) par une onde MMN plus ample que pour une différence entre deux exemplaires d'une

même catégorie (e.g. pour une différence phonémique basée sur le voisement, Sharma & Dorman, 1999). Cette augmentation de la MMN est plus nette si la différence phonémique est pertinente dans la langue parlée par l'auditeur (Näätänen *et al.*, 1997). Une augmentation de la MMN lors d'un changement correspondant à une différence phonémique est aussi considérée comme un indice d'apprentissage d'un nouveau phonème, par exemple à la suite d'un entraînement ciblé (Tamminen *et al.*, 2015).

### **2.2.1. Méthodes d'évaluation de la perception catégorielle**

Afin d'évaluer PC d'un individu dans une langue donnée, il est possible d'utiliser trois méthodes (Bogliotti, 2005) utilisant les tâches d'identification et de discrimination : la précision catégorielle (PrC), l'effet de la frontière phonémique (EFP), et la perception catégorielle relative (PCR). Selon Simon et Fourcin (1978), la PrC est représentée par une frontière précise et une faible discrimination entre deux exemplaires appartenant à une même catégorie. L'EFP prend en compte uniquement les résultats en discrimination. Celle-ci compare les scores obtenus pour les couples de sons appartenant à la même catégorie (i.e. qui sont du même côté de la frontière) aux scores obtenus pour des sons appartenant à deux catégories distinctes (Wood, 1976), dans une épreuve qui contient aussi des couples de stimuli acoustiquement identiques. Enfin, la perception catégorielle relative (PCR), quant à elle, ne prend pas uniquement en compte la précision de la frontière. Selon Liberman *et al.* (1975), la PCR est la cohérence entre les réponses en identification et en discrimination de phonèmes. La frontière catégorielle définie par les réponses en identification est pour cela considérée comme un prédicteur de la position du pic dans les scores en discrimination pour les stimuli du même continuum. Si les deux concordent, les réponses en identification et en discrimination sont considérées comme homogènes, ce qui serait le signe d'une perception vraiment catégorielle.

Sur le plan clinique, le choix des épreuves de PC et des traits phonétiques pour les continua n'est pas anodin pour faire ressortir des différences entre des personnes dyslexiques et leurs contrôles. Maasen *et al.* (2001) ont en effet montré que des enfants dyslexiques avaient de moins bonnes performances que leurs contrôles en discrimination, que la différence porte sur le lieu d'articulation ou le voisement, alors qu'en identification seul le contraste de voisement donnait lieu à une différence entre les groupes. Ce dernier trait, en grande partie basé sur un indice de durée (le VOT), est pour cela particulièrement intéressant aussi dans une perspective clinique.

### **2.3. Mise en place développementale lente de la perception catégorielle**

La PC rendrait l'auditeur cognitivement sourd à des indices acoustiques non-pertinents dans une langue donnée. En effet, il s'habituerait rapidement à ne plus tenir compte de certaines nuances acoustiques, ce qui le conduirait à donner des réponses homogènes et unanimes pour la catégorisation d'un ensemble d'exemplaires très large, parfois distants, peu typiques de la catégorie phonémique. Alors que la même quantité de différence acoustique autour de la frontière amène l'auditeur à catégoriser les deux sons comme appartenant à deux catégories différentes. Ce mode de perception est très utile pour traiter la parole dans une langue donnée. Elle permet de ne pas être limité par des particularités de prononciation des locuteurs, ni par des variations intra-individuelle dues au contexte des phonèmes, au débit, et aux émotions.

Les performances en PC varient beaucoup avec l'âge. Elles augmentent entre 5 et 10 ans selon Bogliotti *et al.* (2002), et même jusqu'à 12 ans pour l'identification de phonèmes (Hazan & Barrett, 2000). Elles croissent encore ensuite selon Elliot *et al.* (1981), avec par exemple l'augmentation de la précision de la frontière catégorielle entre 8-11 ans et l'âge adulte. Chez les enfants, les scores de discrimination sont aussi trop faibles par rapport à ce que peuvent faire attendre les scores en identification (Medina & Serniclaes, 2005).

### **2.4. Position de la frontière catégorielle pour le voisement des occlusives**

Des expériences évaluant la PC ont permis de préciser la position de la frontière catégorielle entre des consonnes occlusives opposées en voisement (/b-d-g/ - /p-t-k/) en français : elle se situerait autour de 0-10 ms. Elle est communément admise à 0 ms, mais certaines études la placent aux alentours de +10 ms (+7.3 ms, selon Médina *et al.*, 2010). La position de la frontière peut être influencée par la voyelle qui suit, ainsi que par le lieu d'articulation des consonnes opposées en voisement (Lisker & Amramson, 1967 ; Serniclaes, 1975). En anglais, elle se situe autour de +25 ms (Aslin, *et al.*, 1981) ou +30 ms (Jusczyk, *et al.*, 1989 ; Lisker 1978). Cette sensibilité à des intervalles temporels précis, importante pour structurer le système phonologique, serait basée sur des durées pour lesquelles le système auditif des mammifères, et particulièrement celui du nourrisson humain, est physiologiquement sensible (Simos, *et al.*, 1997).

## 2.5. Supports neuroanatomiques de la perception catégorielle

La PC des phonèmes implique l'activation de connaissances phonologiques sur des consonnes et des voyelles : ce traitement rapide, qui s'automatise avec le développement, implique des aspects de traitement cognitif. Sur le plan neuroanatomique, la PC des phonèmes implique des aires cérébrales sous-tendant des traitements qui ne sont pas seulement perceptifs. Une méta-analyse à partir de 23 expériences en IRMf (Turkeltaub & Coslett, 2010) met ainsi en évidence, d'une part, la participation bilatérale du cortex auditif primaire et des aires à proximité dans le gyrus temporal supérieur pour l'analyse acoustique de la parole ou de stimuli auditifs non linguistiques. Elle confirme d'autre part l'intervention de la voie dorsale pour l'analyse phonétique et phonologique de la parole, avec le sillon temporal supérieur médian-postérieur gauche (ou médian-antérieur selon Liebenthal *et al.* 2005) qui ne s'active que pour des stimuli verbaux associés à des catégories phonémiques apprises (Liebenthal *et al.*, 2005) dont l'auditeur prend conscience (Desai *et al.*, 2008). Le gyrus temporal médian serait lui aussi activé seulement pour des stimuli répartis de part et d'autre d'une frontière phonémique (Joanisse, *et al.*, 2007). Le lobe pariétal inférieur gauche serait quant à lui indispensable à la détection de différences entre les phonèmes (Joanisse *et al.*, 2007). Ces aires de la voie dorsale sont moins activées pour la discrimination de phonèmes en cas de dyslexie (Dufor *et al.*, 2007). La latéralisation gauche dans cette voie dorsale correspondrait à l'implication de connaissances phonologiques.

Le gyrus frontal médian postérieur gauche interviendrait également dans la PC (Turkeltaub & Coslett, 2010) et permettrait des traitements phonologiques de haut niveau. Ce rôle est cohérent avec les travaux sur l'organisation dynamique des oscillations cérébrales, renforcée de façon top-down par des oscillations lentes issues de cette aire frontale lors de l'analyse du signal de parole. En cas de dyslexie chez l'adulte, la connectivité entre ce gyrus frontal et les cortex auditifs bilatéraux serait entravée (Boets *et al.*, 2013). Toujours dans le cortex frontal, les aires pré-motrice et motrice supplémentaire inférieures gauches s'activent d'autant plus que la PC est réussie chez des adultes normo-lecteurs. En revanche, chez les dyslexiques, plus ces aires sont activées et plus la perception allophonique est intense (Dufor *et al.*, 2009). Cela suggère que le codage moteur de la parole affine généralement le codage de manière appropriée au système phonologique, mais de manière inadaptée et déconnectée des connaissances phonologiques en cas de dyslexie. Enfin, le réseau participant à la PC s'élargirait aussi à l'activation des gyri supramarginal et angulaire gauches.

## 2.6. Perception catégorielle et dyslexie

En cas de dyslexie, la perception des phonèmes est souvent déficitaire car moins catégorielle (Godfrey *et al.*, 1981 ; Sprenger-Charolles & Serniclaes, 2004 ; Werker & Tees, 1987). Ces déficits ont été montrés dans des expériences longitudinales proposant des entraînements. La lecture de mots nouveaux (pseudo-mots) s'est en effet améliorée après un entraînement ciblant la PC (avec RapDys©), comparé à un entraînement exerçant une compétence visuo-attentionnelle (MAEVA ©) (Zoubrinetzky *et al.*, 2019). Lorsqu'il est présent, le déficit de PC chez les personnes dyslexiques se manifeste à travers trois phénomènes.

(1) En identification de phonème, la pente de la sigmoïde est moins abrupte que chez les contrôles (e.g., en espagnol, Serniclaes *et al.*, 2021). Cela signifie qu'une plus grande quantité d'exemplaires sont perçus comme phonémiquement ambigus sur le continuum en cas de dyslexie, même à l'âge adulte (Noordenbos *et al.*, 2013).

(2) La frontière catégorielle relevée chez les personnes dyslexiques n'est pas décrite comme mal positionnée dans les épreuves d'identification (Serniclaes *et al.*, 2021). Le seuil d'identification ne serait donc pas incorrect. Par contre, il est décrit comme imprécis, faible, ce qui se traduit surtout par un pic de discrimination peu élevé (Serniclaes *et al.*, 2001) ; une méta-analyse le confirme (Noordenbos & Serniclaes, 2015). Chez les adultes dyslexiques aussi, l'imprécision des frontières catégorielles est aussi documentée. L'amplitude de la MMN est, en effet, anormalement faible à l'écoute d'un changement de consonnes (Schulte-Körne *et al.*, 2001), surtout si les exemplaires diffusés présentent de petites différences de hauteur de voix (variations du F0) comme dans les situations de perception naturelle (Virtala *et al.*, 2020).

(3) La troisième anomalie découverte chez des personnes dyslexiques est une trop grande sensibilité aux différences intra-catégorielles, qui retirerait son caractère exceptionnel à la sensibilité aux différences inter-catégorielles au niveau de la frontière phonémique (Serniclaes *et al.*, 2001). Cette perception, dite « allophonique », serait répandue en cas de dyslexie (pour des revues, voir Serniclaes, 2018 ; Serniclaes & Seck, 2018). Dans la tâche d'identification, ce type de perception se traduit par des asymptotes anormalement étroites (Serniclaes *et al.*, 2021). Dans la tâche de discrimination, elle apparaît chez les enfants faibles lecteurs de 10 ans à travers une discrimination accrue entre les exemplaires d'une même catégorie, alors que le pic de discrimination au niveau de la frontière phonémique est trop peu marqué (Bogliotti *et al.*, 2002). Ce phénomène est aussi observé chez des enfants à risque de dyslexie (Noordenbos *et al.*, 2012) et des enfants dyslexiques (e.g. en français, Serniclaes *et*

*al.*, 2004). Cette perception allophonique disparaît des données comportementales des enfants à risque de dyslexie lorsqu'ils grandissent, mais des traces demeurent dans les enregistrements neurophysiologiques (Noordenbos *et al.*, 2012), y compris lorsqu'ils sont adultes (Noordenbos *et al.*, 2013). Chez les adultes dyslexiques, une perception allophonique est difficile à mettre en évidence avec des données comportementales. Elle a en revanche été décrite à travers une MMN d'amplitude aussi forte pour des stimuli auditifs d'une même catégorie phonémique que pour des stimuli représentant des catégories différences (Noordenbos *et al.*, 2013).

Cette catégorisation inadaptée des sons de parole implique qu'au sein d'une même catégorie plusieurs exemplaires sont distingués les uns des autres sur la base de petites différences acoustiques. Une hypothèse serait que les prédispositions perceptives présentes à la naissance et non pertinentes pour la discrimination des mots dans la langue d'une personne dyslexique ne soient pas assez désactivées (Bogliotti *et al.*, 2002). La prédisposition permettant aux très jeunes enfants de percevoir l'ensemble des contrastes phonétiques (Werker & Tees, 1984a) aurait le défaut, en cas de dyslexie, de ne pas s'atténuer avec l'âge selon la pertinence des catégories phonémiques de la (ou des) langue(s) parlée(s). Par exemple, pour un francophone dyslexique, [k] et [k<sup>h</sup>] ne sont pas identifiés comme deux exemplaires d'une seule catégorie, le phonème /k/ correspondant à la lettre « k » et [k<sup>h</sup>] comme une autre catégorie. La différence acoustique entre les deux serait perçue avec une anormale saillance par la personne dyslexique, qui n'apprendrait pas à rester insensible à cette différence inutile dans la langue française (Fig. 8). Le cas le plus connu est celui des VOT -30 ms et +30 ms, frontières phonémiques possibles pour le voisement dans différentes langues du monde, localisées par les enfants de moins de 6 mois (Hoonhorst *et al.*, 2009). Ces frontières ne devraient plus être perçues dans des langues où la frontière appropriée est autour de 0 ms, comme le français ou l'espagnol. Chez des enfants dyslexiques français, la persistance d'un pic de discrimination pour le VOT -30 ms a été décrite (Serniclaes *et al.*, 2004), et le maintien des deux frontières -30 ms et +30 ms a été mis en évidence chez des enfants dyslexiques espagnols (Serniclaes *et al.*, 2021).

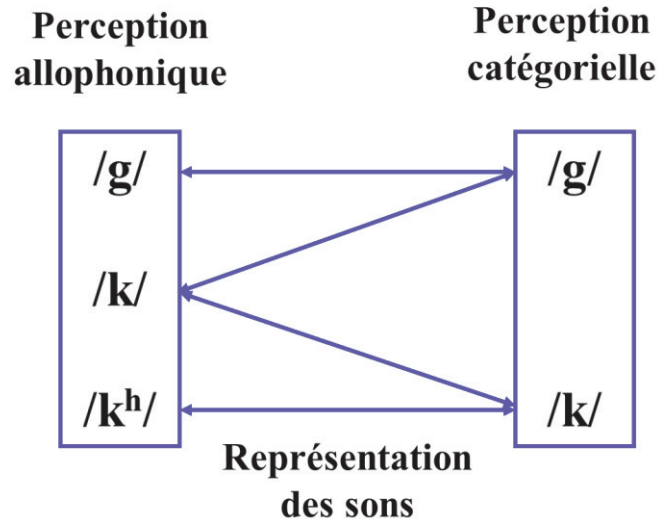


Figure 8. Relation entre perception allophonique et perception catégorielle.

La perception allophonique pourrait en partie trouver son origine dans des réponses instables du cerveau à des stimuli de parole, y compris au niveau du tronc cérébral (Hornickel & Kraus, 2013). Cette instabilité affaiblirait le développement de représentations phonologiques robustes. Cela entraînerait un trouble de la conscience phonologique, déficit très répandu en cas de dyslexie et perturbant la manipulation mentale volontaire de phonèmes dont les frontières ne sont pas assez claires. Cette conscience phonologique est elle-même une condition pour apprendre et appliquer de façon simple les règles de correspondance graphème-phonème utiles au processus de décodage en lecture. Bogliotti (2005) suppose aussi que la perception allophonique contribue au déficit de mémoire à court terme phonologique accompagnant souvent la dyslexie, à cause d'un nombre trop élevé d'unités à décoder entraînant une surcharge mnésique.

Une autre hypothèse sur l'origine de la difficulté à acquérir des catégories phonémiques en cas de dyslexie est un déficit d'apprentissage de catégories auditives. Il s'agirait alors d'une difficulté générale, non-spécifique au langage. Ce déficit se manifeste par exemple dans un jeu vidéo mobilisant l'apprentissage procédural, et dans lequel les adultes dyslexiques acquièrent moins bien de nouvelles catégories auditives non-verbales et catégorisent moins bien, après cet apprentissage, des exemplaires familiers et nouveaux (Gabay & Holt, 2015). Cette hypothèse est cohérente avec les difficultés générales d'apprentissage procédural des personnes dyslexiques (Lum *et al.*, 2013 ; Nicolson & Fawcett, 2011), en lien avec des perturbations dans des aires cérébrales associées à ce type d'apprentissage, comme le cervelet (Nicolson *et al.*, 1999 ; Pernet *et al.*, 2009).

Enfin, plus récemment, une hypothèse d'explication neuronale de la perception allophonique s'intègre particulièrement bien au *Temporal Sampling Framework* de Goswami (2011). Cette hypothèse suggère que la perception trop fine de détails acoustiques trouve son origine dans un couplage insuffisant entre des oscillations corticales rapides. Des oscillations autour de 33 Hz, adaptées à la détection des frontières catégorielles situées aux VOT -30 ms et +30 ms, devraient normalement subir une inhibition produite par des oscillations deux fois plus lentes (autour de 17 Hz). En-dehors des pathologies, un couplage entre les deux types d'oscillations atténuerait la sensibilité aux frontières -30 ms et +30 ms au profit d'une frontière intermédiaire, autour de 0 ms. Celle-ci deviendrait alors privilégiée comme il se doit pour le français ou l'espagnol (Serniclaes *et al.*, 2021). L'absence d'inhibition d'oscillateurs très rapides serait ainsi la conséquence de couplages déficitaires dans la hiérarchie des rythmes cérébraux, au niveau des plus rapides d'entre eux (bêta-gamma).

## **2.7. Manque d'outils d'évaluation en français**

Il existe encore aujourd'hui peu d'outils cliniques facilement disponibles pour évaluer la PC en français. Il est pourtant important de le faire pour cette langue étant donné que ses frontières phonémiques ont des propriétés singulières. Il est nécessaire d'utiliser un outil développé spécifiquement pour la langue ciblée. En orthophonie, les professionnels utilisent des outils standardisés pour évaluer les capacités d'identification et de discrimination des phonèmes, mais il ne s'agit pas de véritables tests de PC au sens où l'entend la recherche scientifique. Beaucoup de tests orthophoniques centrés sur les phonèmes sont d'ailleurs avant tout des épreuves de conscience phonologique, ce qui est différent.

## **2.8. Le voisement des consonnes occlusives**

### **2.8.1. Définition du voisement**

Le voisement est une catégorie de traits phonologiques déclinée le plus souvent en deux possibilités pour les consonnes (voisée/non-voisée, autrement dit sonore /b-d-g/ - sourde /p-t-k/). Pour les consonnes occlusives dans un contexte consonne-voyelle (CV), la valeur du trait de voisement est associée, sur le plan articulatoire, au moment où se produit la vibration des plis vocaux par rapport à l'ouverture brusque de l'appareil vocal (relâchement de l'occlusion supra-glottique). Sur le plan acoustique, le voisement d'une consonne occlusive est associé à plusieurs indices. Le principal est le délai de voisement (*Voice Onset time*, VOT), durée qui



sépare le début du bruit émis par la voix et le bruit d'explosion (*burst*) (Lisker & Abramson, 1964).

En français, cet indice acoustique temporel oppose les occlusives sonores /b/, /d/, /g/ aux occlusives sourdes /p/, /t/, /k/. Bien que d'autres indices acoustiques subtils, comme la durée du premier formant (Stevens & Klatt, 1974) en français, et, en anglais, la hauteur de l'attaque de ce premier formant dans la voyelle qui suit (Lisker *et al.*, 1977) et l'intensité du bruit d'explosion en français, contribuent à distinguer les consonnes occlusives sourdes et sonores, le VOT est considéré comme l'indice le plus fondamental. Il est privilégié dans l'étude du voisement. En français, on parle de VOT nul ou légèrement positif pour les occlusives sourdes et de « pré-voisement » ou VOT négatif pour les occlusives sonores, car la vibration des cordes vocales commence avant le bruit d'explosion. Il est possible de représenter un continuum des VOTs des exemplaires possibles entre deux prototypes de consonnes opposées seulement en voisement (Fig. 9).

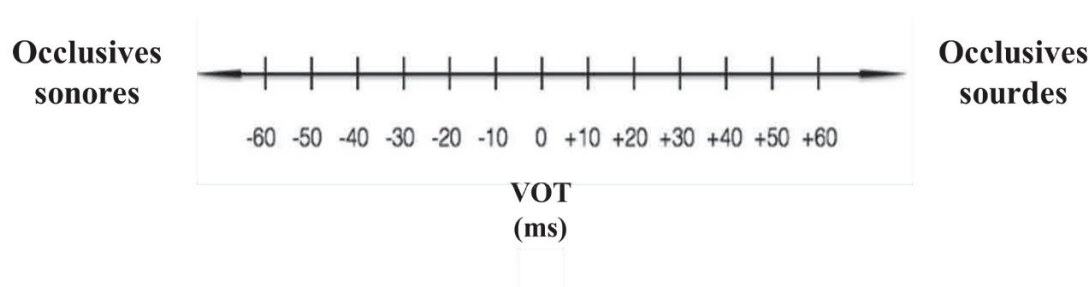


Figure 9. Continuum de VOT entre des occlusives sonores et des occlusives sourdes.

### 2.8.2. Indices de durée pour le voisement et asymétrie hémisphérique

La durée des occlusives sonores et sourdes, opposées par le voisement, varie en français. Les sonores ayant un long VOT négatif, leur durée est plus longue que celle des sourdes caractérisées par un court silence suivant le bruit d'explosion (court VOT positif). Les francophones peuvent identifier facilement une occlusive sourde en cas d'indice acoustique bref et une sonore en cas d'indice long.

Au niveau cérébral et selon les modèles à résolution temporelle à échelles multiples, les oscillations rapides dans le cortex auditif gauche (les ondes dites *gamma*) permettraient de détecter plus facilement les indices acoustiques brefs comme le VOT des sourdes. Les oscillations plus lentes dans le cortex auditif droit (les ondes dites *thêta*) permettraient de détecter et d'intégrer de longs indices comme le VOT des sonores, grâce à des fenêtres d'intégration temporelle respectivement plus longues.

Cette complémentarité hémisphérique pour le traitement des consonnes occlusives trouve un argument dans les résultats d'une étude comportementale en français (Bedoin *et al.*, 2010). Dans cette étude en écoute dichotique, les auteurs présentaient simultanément dans chaque oreille des syllabes différentes dont les consonnes initiales étaient opposées par le lieu d'articulation ou par le voisement, et l'auditeur rappelait ce qu'il entendait. Pour les couples opposés par le lieu d'articulation, un avantage classique de l'oreille droite (*Right Ear Advantage*, REA) a été observé, cohérent avec le fait que le trait important était alors un indice acoustique bref (rapide transition formantique). En revanche, pour les couples opposés en voisement, lorsque le mot commençant par l'occlusive sonore (VOT long) était présenté à l'Oreille Gauche – HD et la sourde (VOT court) à l'Oreille Droite - HG (/bas/-/pas/), le REA disparaissait. Ce résultat suggère que l'HD traite facilement le VOT long d'une occlusive sonore française. Un effet opposé obtenu en anglais, langue où les liens entre VOTs court/long et voisement sont opposés à ceux du français, est cohérent avec cette interprétation. En effet, Rimol *et al.* (2006) ont montré que le REA le plus faible a été observé en anglais pour les couples de mots adressant l'occlusive sourde (VOT long) à l'Oreille Gauche – HD et l'occlusive sonore (avec VOT court) à l'Oreille Droite – HG.

Dans l'expérience en français de Bedoin *et al.* (2010), lorsque les consonnes initiales des deux mots différaient par le lieu et le voisement, beaucoup de réponses ne correspondant à aucun des mots étaient produites en combinant le lieu d'articulation de la consonne initiale de l'un et le voisement de la consonne initiale de l'autre (*blended responses*, e.g., /baʁ/-/caʁ/ → réponse /gaʁ/). Cependant, cela ne se produisait, là encore, que lorsque la consonne contenant un VOT long était adressée à l'Oreille gauche – HD : l'HD prenait apparemment en charge prioritairement un VOT long. Ces phénomènes sont en accord avec le modèle AST (Hickok & Poeppel, 2000 ; Poeppel, 2003). Elles le sont aussi avec les données en IRMf qui soutiennent la dominance de l'HD pour les modulations lentes du signal de parole (e.g. Boemio *et al.*, 2005) en particulier au début des stimuli (McKibbin *et al.*, 2003). L'implication forte de l'HD dans le traitement du voisement de manière générale, sur laquelle Simos *et al.* (1997) insistent à partir de divers arguments comportementaux, neuropsychologiques et électrophysiologiques, pourrait s'expliquer non par la spécialisation de l'HD pour la catégorie phonologique de voisement, mais par sa spécialisation pour les indices phonétiques longs présents dans les tests impliquant le voisement.

## 2.9. Objectif

La perception catégorielle (PC) est une des compétences permettant d'assurer la maîtrise des phonèmes d'une langue. En cas de dyslexie, elle serait souvent déficitaire (Noordenbos & Serniclaes, 2015 ; Serniclaes *et al.*, 2001). Les catégories phonémiques seraient mal délimitées, créant des troubles phonologiques entraînant des répercussions moins massives sur la compréhension orale que sur l'apprentissage de la lecture. En effet, les capacités en lecture et la perception catégorielle étant liées (Delahaie *et al.*, 2004). Chaque langue ayant ses propres frontières phonémiques, il est nécessaire que des tests d'évaluation soient adaptés aux caractéristiques de cette langue pour évaluer la PC.

L'objectif principal de cette partie de la recherche était de créer un nouvel ensemble de tests de PC en français, à partir d'épreuves d'identification et de discrimination, et de mesurer certaines de ses qualités psychométriques en lien avec la dyslexie à l'âge adulte. Certaines caractéristiques des tests étaient déterminées par l'objectif principal de cette thèse, qui était de les utiliser pour évaluer les effets de deux types d'entraînement perceptif réalisé dans cette recherche (Chapitres 3 et 5). C'est pourquoi, la PC a été évaluée pour deux paires de consonnes. Toutes étaient occlusives et opposées par le voisement. L'une des paires contenait des consonnes vélares (/g-k/), l'autre contenait des consonnes alvéolaires (/d-t/). La paire de consonnes vélares devait permettre d'évaluer si la PC de certaines consonnes présentées pendant l'entraînement évoluait après celui-ci. La paire de consonnes alvéolaires devait permettre d'évaluer un effet de transfert sur des consonnes non présentées pendant l'entraînement. Étant donné l'objectif prioritaire, la principale qualité recherchée pour cet ensemble de tests était d'être aussi peu sensible que possible à l'effet test-retest. Il était en effet important que d'éventuels progrès observés après l'entraînement ne soient pas entièrement explicables par un effet de la pratique. La fidélité des épreuves a aussi été évaluée par l'étude des corrélations entre les performances des deux sessions.

L'objectif secondaire était de contribuer à la réflexion sur l'élaboration de tests de PC pouvant être utiles pour d'autres usages. Étant donné qu'un déficit en PC participe souvent à l'origine de la dyslexie, il semblait intéressant sur le plan clinique de détecter cette particularité chez les patients, et d'en évaluer l'intensité. En particulier, dans le cas d'adultes qui souffrent de difficultés persistantes en lecture qui n'ont pas été diagnostiquées, cela pourrait permettre de découvrir une des origines de leurs difficultés, et d'orienter vers un bilan complet pour le diagnostic d'une potentielle dyslexie. Beaucoup d'adultes atteints de cette pathologie sont encore non diagnostiqués. Une fois qu'ils le sont, des tests de PC pourraient

aussi permettre une étude plus approfondie de la nature de leurs déficits en phonologie, en vue de proposer des aides adaptées. Des tests de PC pourraient également servir au suivi des progrès liés à une prise en charge orthophonique.

Ce dernier objectif se heurte cependant à un obstacle. S'il s'agit d'adultes dyslexiques, des compensations ont souvent été développées et le diagnostic dans son ensemble est difficile. Pour ce qui est de la PC chez les adultes dyslexiques, des indices de perturbations sont relevés dans des expériences électrophysiologiques, mais nous avons vu que les données comportementales révèlent rarement des déficits. Nous évaluerons dans quelle mesure les nouvelles épreuves confirment cette difficulté. Pour cela, leur sensibilité et leur spécificité seront mises à l'épreuve au moyen du calcul de courbes ROC incluant les données d'un groupe d'adultes normo-lecteurs et d'un groupe d'adultes dyslexiques. Il s'agira d'évaluer à quel point les performances permettent de catégoriser des adultes comme dyslexiques ou non-dyslexiques. Bien que la sensibilité à un trouble de la PC soit un défi pour des tests comportementaux chez l'adulte, les données recueillies contribueront à l'élaboration de premières normes à partir des groupes d'adultes normo-lecteurs ayant réalisé les tâches.

## **2.10. Expérience 1 : Evaluation comportementale de la perception catégorielle**

### **2.10.1. Principe d'expérience et hypothèses**

L'Expérience 1 visait à évaluer les tests de perception catégorielle (PC) en français. En effet, ils ont pour vocation à être ensuite utilisés comme outils pour l'évaluation de l'effet de différents entraînements (Chapitre 3 et 5). L'hypothèse était qu'il y aurait un faible effet test-retest, ainsi qu'une bonne fidélité pour les deux épreuves d'identification et de discrimination proposées. Nous avons également voulu évaluer la sensibilité et la spécificité de ces tests dans le but de l'élaboration de premières normes à des fins d'utilisation de ces tests dans d'autres domaines.

### **2.10.2. Evaluation de la perception catégorielle de consonnes en français**

Deux épreuves de PC (i.e. identification et discrimination) ont été construites. Elles portaient sur les paires /gi-ki/ et /di-ti/. Deux continua (un pour chaque paire) ont été construits à l'aide du logiciel Praat (Boersma & Weenink, 2016) à partir de syllabes naturelles produites par un homme bilingue natif français/anglais, par modification progressive du VOT par pas de 10 ms. La voyelle et le bruit d'explosion, qui contiennent d'autres indices du voisement restaient stables pour tous les stimuli du continuum. Pour le continuum /gi-ki/, le prototype de /gi/ avait

une barre d'énergie en basse fréquence de 60 ms avant l'explosion (VOT -60 ms), qui a été progressivement raccourcie pour créer les intermédiaires jusqu'au VOT 0. Pour créer les /ki/, le court VOT positif du /ki/ a été ajouté au burst du /gi/ et allongé positivement par pas de 10 ms grâce à l'ajout d'une aspiration extraite d'une production anglaise du locuteur. Les continua /di-ti/ ont été construits sur le même principe (13 stimuli/continuum) (Fig. 10).

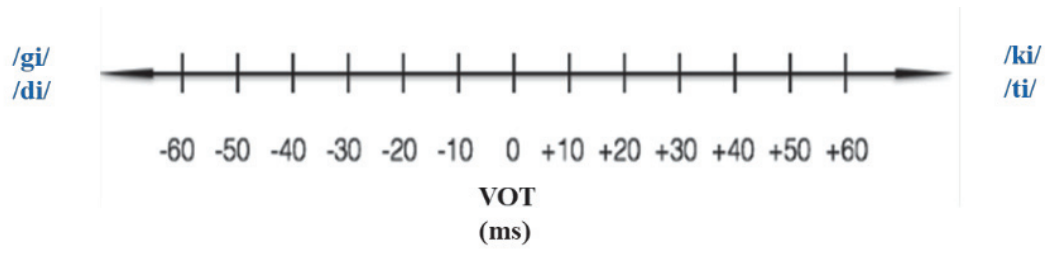


Figure 10. Continuum de VOT en français utilisé lors des épreuves de perception catégorielle dans cette langue.

Les épreuves se déroulaient toujours dans le même ordre avec la discrimination réalisée avant l'identification afin de ne pas influencer les résultats par des informations sur les noms des catégories (Jamieson & Morosan, 1986). L'évaluation sur la paire /di-ti/ était toujours réalisée avant la paire /gi-ki/.

#### 2.10.2.1. Epreuve de discrimination

Dans cette première épreuve, le participant entendait deux stimuli et devait décider si, selon son jugement, les deux stimuli étaient identiques ou différents (Fig. 11, gauche). La consigne était : deux sons devaient être considérés comme identiques si l'auditeur pensait que le même fichier-son avait été produit deux fois sinon ils devaient les considérer comme différents. Il répondait en appuyant le plus vite possible sur l'une des deux touches de réponse. Les membres du couple étaient présentés avec un intervalle de 20 ms entre les deux. Chaque couple était diffusé 10 fois : 5 fois dans un sens et 5 fois dans l'autre (e.g. VOT 0 ms puis VOT 20 ms ; VOT 20 ms puis VOT 0 ms). Les couples de stimuli identiques étaient aussi présentés 10 fois chacun (e.g. VOT 0 ms et VOT 0 ms). Toutes ces paires de stimuli étaient présentées de manière pseudo-aléatoire pour éviter les successions de couples identiques, et limiter les successions de plus de 3 réponses identiques. Cette épreuve durait 40 minutes.

#### 2.10.2.2. Epreuve d'identification

Durant l'épreuve d'identification, le participant n'entendait qu'un seul stimulus par essai et devait déterminer la catégorie à laquelle il appartenait parmi un choix forcé (i.e. /gi/ ou /ki/ ; /di/ ou /ti/). Il avait pour cela deux touches à sa disposition et devait répondre aussi vite que possible (Fig. 11, droite). Chaque exemplaire des continua était présenté 10 fois dans un ordre

aléatoirement contrôlé selon des règles conformes à celles suivies pour l'épreuve de discrimination. Cette épreuve durait 20 minutes.



Figure. 11 Procédure de l'épreuve de discrimination (gauche) et d'identification (droite).

### 2.10.3. Evaluation de l'effet test-retest et de la fidélité

#### 2.10.3.1. Participants

Vingt jeunes adultes monolingues francophones (16 femmes,  $M = 22.3$  ans ;  $ET = 1.9$ ), sans troubles des apprentissages ont participé à cette étude. Ils déclaraient avoir une bonne audition et une vision normale ou corrigée et étaient droitiers ( $\geq 80\%$  de réponses « main droite » au test d'Edinburgh de Oldfield, 1971). Ils ont réalisé les deux épreuves de PC (i.e. identification et discrimination) sur deux paires de consonnes deux fois (T1-T2) à 5 jours d'intervalle, sans aucune autre intervention proposée entre les deux sessions.

#### 2.10.3.1. Equipement

La passation des deux épreuves de PC s'est faite à distance en raison de la crise sanitaire due à la COVID-19. Le participant utilisait son propre ordinateur (Chromebook exclu car non compatible avec le logiciel OpenSesame) et son casque audio stéréo. Les épreuves ont été programmées avec OpenSesame (Mathôt, *et al.*, 2012).

### 2.10.4. Analyse des données

**Identification.** La pente, qui témoigne de la précision de la frontière entre les phonèmes, a été évaluée en utilisant la toolbox Palamedes de Matlab (Kingdom & Prins, 2010) et la fonction PAL\_PFML\_Fit. Cette dernière permet l'ajustement entre les paramètres de seuil ( $\alpha$ ) et de pente ( $\beta$ ) d'une fonction logistique aux données d'identification. Le seuil était aussi évalué et indiquait la valeur de VOT correspondant au taux de réponses de 50% (niveau du hasard) dans l'épreuve.

**Discrimination.** Les réponses ont été converties en  $d'$  (indice de sensibilité issu de la Théorie de la Détection du Signal ( $z(\text{Hits}) - z(\text{Fausses Alarmes})$ )) en calculant, pour chaque couple

contenant deux sons différents (e.g. VOT +20 et +40 ms), la différence entre le  $z$  score du taux de différences correctement détectées par le participant (Hits) et la moyenne des  $z$  scores du taux de fausses alarmes pour les paires identiques correspondantes (ici, VOT+20 et +20 ms et VOT+40 et +40 ms), conformément à la recommandation de Hary et Massaro (1982). Selon eux, l'évaluation de la discrimination est alors plus indépendante d'un biais de réponse vers *même* ou *différent*. Avant le calcul du  $z$ , les taux de discrimination (de 0 à 1) ont été ajustés en enlevant 0.042 aux valeurs supérieures à 0.5, et en ajoutant 0.042 aux valeurs inférieures à 0.5 selon la préconisation de Macmillan et Creelman (2005) (cité par Noordenbos *et al.*, 2013). Cela permet notamment d'éviter les taux égaux à 0 et 1 conduisant à des valeurs infinies lors de la conversion en  $z$  score. La localisation du pic de discrimination sur le continuum de VOT était définie comme le VOT de la paire avec le score moyen le plus haut. Les traitements statistiques ont porté sur ce pic.

Toutes les analyses statistiques ont été réalisées avec logiciel R-Studio, en respectant le seuil de significativité de  $p \leq .05$ .

Afin d'évaluer l'effet test-retest, des tests  $t$  de Student intra-individuels bilatéraux avec correction de Bonferroni ainsi que des tests  $t$  avec permutations ont été appliqués pour comparer deux séries de données. Avec cette dernière technique, le test de l'hypothèse n'est plus soumis à l'ajustement des données à une distribution théorique de la statistique pour évaluer la significativité, mais à l'ajustement à une distribution empirique. Celle-ci est construite exclusivement à partir des données elles-mêmes. Cette opération est effectuée un très grand nombre de fois – idéalement pour toutes les combinaisons possibles permettant de constituer deux groupes – et ces combinaisons sont ensuite ordonnées pour constituer une distribution de référence. Celle-ci a l'avantage d'être exclusivement basée sur des valeurs existant dans les données, y compris des valeurs aberrantes. Rien n'est donc transformé ou supprimé, toutes les données sont prises en compte. Avec R, seul un échantillon du nombre de permutations possibles est effectué, en employant des approches de type Monte Carlo sur la base de la génération d'une série de nombres aléatoires, mais la quantité de permutations est tout de même très grande (environ 10 000). Le test avec permutations n'est alors pas qualifié d'exact, mais d'approximatif. Ensuite, cette distribution empirique sert de base au calcul d'un  $t$  sur les données observées. La valeur  $p$  obtenue est celle d'un test bilatéral. Selon le plan de l'analyse, un test avec permutation pour données appariées était effectué.

Des analyses supplémentaires à l'aide du test  $t$  de Student ont été réalisées avec calcul de la taille d'effet (i.e.  $d$  de Cohen) afin de pouvoir comparer l'intensité de différents effets.

La fidélité a quant à elle été évaluée par l'étude des corrélations (test de Pearson) avec une estimation de leur taille d'effet (i.e.  $r$ ) afin d'évaluer le lien entre les performances à T1 et T2. Les tailles des effets ( $d$ ,  $r$ ) étaient interprétées selon Cohen (1988) comme faible  $> .20$ , moyenne  $> .50$  ou grande  $> .80$ .

## 2.10.5. Résultats

### 2.10.5.1. Evaluation des effets test-retest

**En discrimination**, il n'avait pas de différence de hauteur du pic entre T1 et T2, que ce soit pour la paire /gi-ki/ ( $t(19) = 0.65$ ,  $p = .523$ ,  $d = .18$  ;  $Z = 0.66$ ,  $p = .509$ ) ou la paire /di-ti/ ( $t(19) = .01$ ,  $p = .993$ ,  $d = .00$  ;  $Z = 0.01$ ,  $p = .992$ ) (Fig. 12).

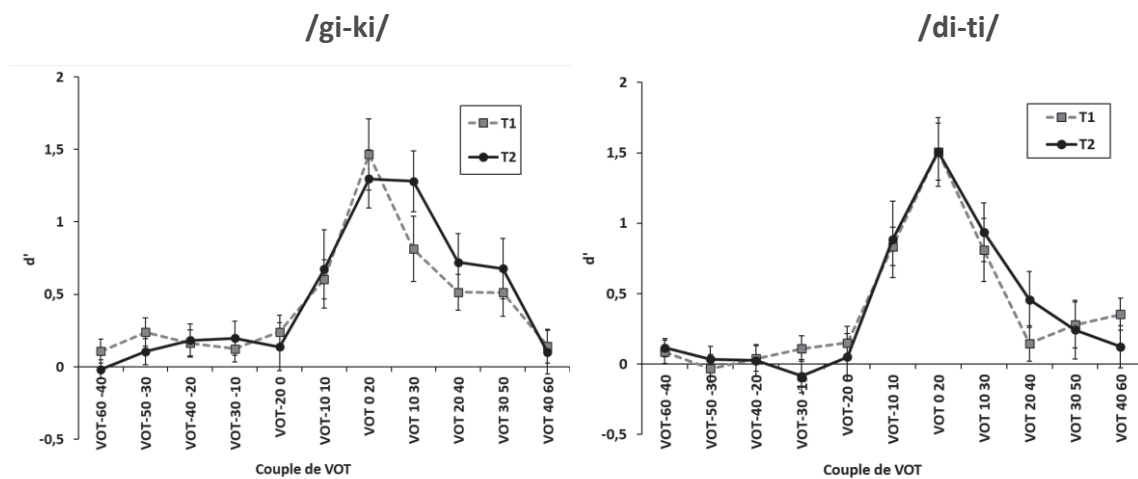


Figure 12.  $d'$  en français à T1 (pointillés) et T2 (plein) pour les paires /gi-/ki/ (à gauche) et /di-/ti/ (à droite) pour le groupe sans entraînement (GSE). Les barres d'erreurs représentent l'erreur-type.

**En identification**, les courbes sont représentées en figure 13, et il n'y avait pas de différence de seuil ou de pente entre T1 et T2 (Tableau 1).



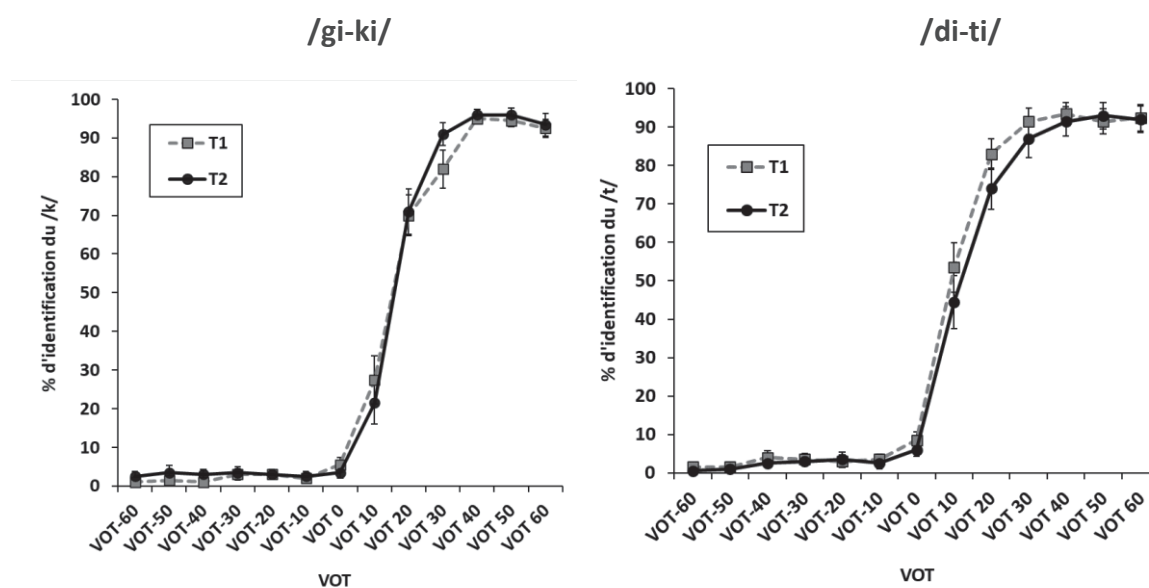


Figure 13. Pourcentage d'identification en français à T1 (pointillés) et T2 (plein) pour les paires /gi-ki/ (à gauche) et /di-ti/ (à droite). Les barres d'erreurs représentent l'erreur-type.

		Test t de Student		Test t avec permutation		
		<i>t</i>	<i>p</i>	<i>d</i>	<i>p</i>	
/gi-ki/	<i>Seuil</i>	0.03	.980	.00	0.03	.979
	<i>Pente</i>	0.08	.938	.02	0.08	.935
/di-ti/	<i>Seuil</i>	1.45	.163	.22	1.41	.157
	<i>Pente</i>	0.56	.585	.14	0.57	.572

### 2.10.5.1. Evaluation de la fidélité

**En discrimination**, l'étude des corrélations a permis de montrer que les performances étaient fortement corrélées entre T1 et T2 au niveau du pic de discrimination pour la paire /di-ti/ ( $r(18) = .69, p = .001$ ), mais pas pour la paire /gi-ki/ ( $r(18) = .20, p = .393$ ) (Fig.14).

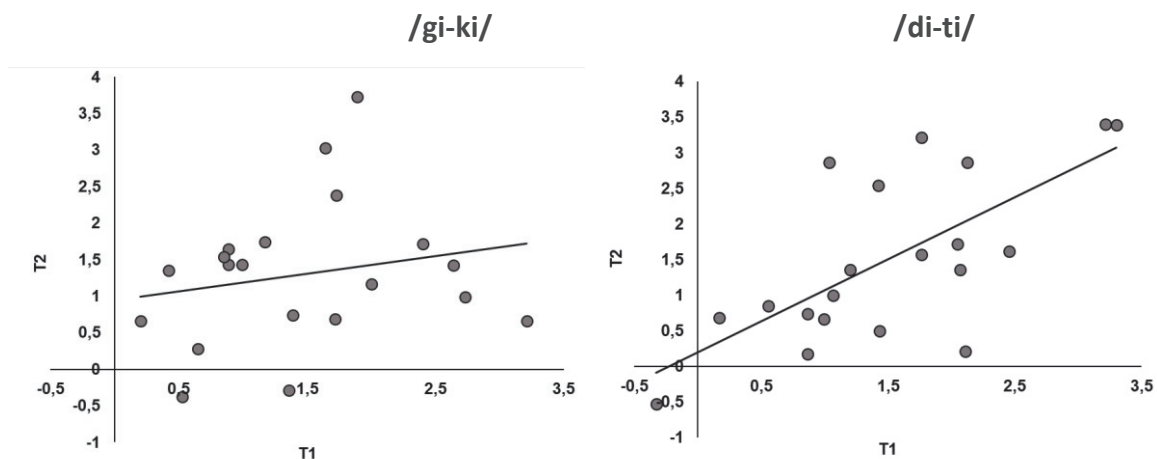


Figure 14. Nuage de points représentant la corrélation entre T1 et T2 pour la valeur de  $d'$  au niveau du pic de discrimination de la paire /gi-ki/ (à gauche) et de la paire /di-ti/ (à droite). La barre noire représente la droite de régression.

L'étude des corrélations **en identification** a montré une différence entre les deux paires. En effet, pour la paire /gi-ki/, il n'y avait pas de lien significatif entre les performances à T1 et à T2, que ce soit pour la valeur de seuil (avec toutefois une taille estimée comme moyenne) ( $r(18) = .35, p = .129$ ) ou de pente ( $r(18) = .24, p = .319$ ). Concernant la paire /di-ti/, les performances étaient moyennement corrélées pour la valeur de seuil d'identification ( $r(18) = .77, p < .001$ ), mais non significativement (avec une corrélation estimée comme de taille moyenne) pour la valeur de pente ( $r(18) = .38, p = .097$ ) (Fig. 15). Sans la donnée aberrante (outlier) observée sur la figure 15 (gauche), la corrélation sur la valeur de seuil de la paire /di-ti/ restait significative mais de taille moyenne ( $r(18) = .56, p = .013$ ).

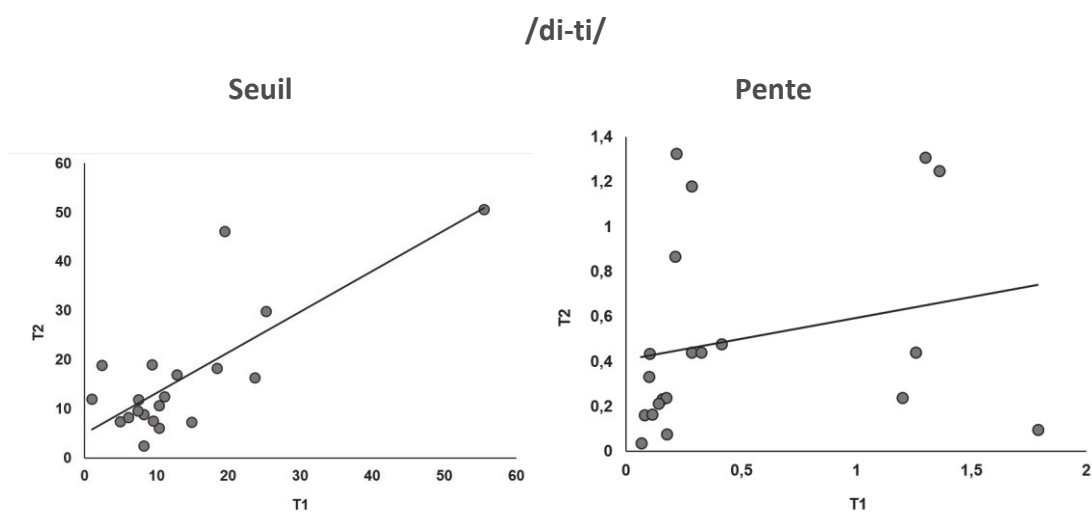


Figure 15. Nuage de points représentant la corrélation entre T1 et T2 pour la valeur de seuil (à gauche) et de pente (à droite) de la paire /di-ti/. La barre noire représente la droite de régression.

## 2.10.6. Evaluation de la spécificité et de la sensibilité : Courbes ROC

### 2.10.6.1. Participants

Les données sont les performances de 74 (dont 20 de l'Expérience 1) jeunes adultes monolingues francophones (50 femmes,  $M = 21.4$  ans ;  $ET = 2.3$ ) sans troubles des apprentissages et 27 jeunes adultes monolingues français (19 femmes,  $M = 21.9$  ans ;  $ET = 2.9$ ) ayant été diagnostiqués dyslexiques durant l'enfance (les adultes dyslexiques représentent 35% de l'échantillon global) (Annexe 3). Ils déclaraient avoir une bonne audition et une vision normale ou corrigée. Tous les participants ont réalisé les deux épreuves de perception catégorielle (i.e. identification et discrimination) décrites ci-dessus, et les données étaient recueillies à T1. Les valeurs déviantes de  $\pm 1,5$  écart-type de la moyenne du groupe (normo-lecteur ou dyslexiques) étaient retirées des calculs (Tableau 2).

Tableau 2. Nombre de participants retirés des analyses pour chaque mesure et chacune de paires pour le groupe normo-lecteurs et le groupe dyslexiques.

		Normo-lecteurs		Dyslexiques	
		/gi-ki/	/di-ti/	/gi-ki/	/di-ti/
Discrimination	$d'$	4	4	4	3
Identification	Seuil	3	2	3	1
	Pente	0	0	2	3

### 2.10.7. Analyse des données

Les données en identification et en discrimination des participants de cette partie de l'expérience ont été traités comme décrit dans la partie 2.10.4 de ce chapitre. Puis ont été soumis à différentes analyses décrites ci-dessous pour l'évaluation de la sensibilité/spécificité des tests.

Les courbes ROC, les indices  $J$  de Youden ainsi que les différents seuils d'intérêts ont été calculés à l'aide du logiciel R et des packages *PlotROC* (M. Sachs, Version 2.2.1). Ces calculs ont porté sur la valeur du pic de discrimination exprimée en  $d'$ , ainsi que les valeurs de seuil et de pente en identification.

L'évaluation de la sensibilité/spécificité de ces épreuves a été réalisée après avoir vérifié l'absence d'un effet test-retest, et après avoir estimé la fidélité des tests. Il s'agissait de déterminer les *cut-off*, c'est-à-dire les valeurs-seuils à partir desquelles une performance pouvait être considérée comme pathologique, en tenant compte des priorités (sensibilité et/ou spécificité) pour le diagnostic. La taille des échantillons étant encore assez faible, ces seuils seront à considérer avec prudence.

La courbe ROC (*Receiver Operating Characteristic Curve*) a permis de représenter graphiquement la relation entre la sensibilité et la spécificité du test pour toutes les valeurs seuils possibles, ceci pour chacune des variables dépendantes. Le seuil était sélectionné à partir de la prise en compte de la meilleure combinaison entre la sensibilité et la spécificité du test. Pour une valeur-seuil X, la sensibilité du test était évaluée par la proportion de vrais positifs chez les patients (i.e. proportion de patients qui seraient effectivement catégorisés comme atteints de la pathologie testée si leurs performances franchissaient le seuil X). La spécificité du test était évaluée par la proportion de vrais négatifs chez les contrôles (i.e. proportion de contrôles qui seraient catégorisés comme exempts de la pathologie testée si les performances franchissaient le seuil X). La meilleure combinaison entre sensibilité et spécificité a été calculée afin de proposer un seuil optimal pour le diagnostic. Il est parfois plus pertinent d'être plus tolérant au sujet de la sensibilité ou au contraire de la spécificité du test, et d'autres seuils sont alors à prendre en considération. Il y a plusieurs façons d'estimer l'intérêt diagnostique d'un test et du seuil choisi (Bewick *et al.*, 2004). Tout d'abord, l'examen de la courbe ROC donne des indications. Plus elle est concentrée en haut à gauche, loin de la diagonale entre le point inférieur gauche et le point supérieur droit de la figure et plus l'intérêt diagnostique du test est grand. En d'autres termes, plus l'aire sous la courbe (*Area Under the Curve*, AUC) est grande et plus le test est informatif. Par convention, un AUC de 0,5 à 0,7 est peu informatif, il l'est moyennement de 0,7 à 0,9 et il est très informatif à partir de 0,9 (un AUC de 1 étant parfait).

Il existe plusieurs indices combinant en une même expression la sensibilité et la spécificité du test pour le seuil choisi. L'indice J de Youden (Sensibilité + Spécificité – 1) en est un exemple : s'il est proche de +1, la méthode de diagnostic est précise. Il n'est pas parfait, mais s'avère utile lorsqu'il s'agit de comparer la précision de plusieurs tests (Böhning *et al.*, 2008). Cet indice guide pour choisir le seuil le plus approprié si l'on recherche un seuil (*cut-off*) correspondant au meilleur équilibre entre la sensibilité et la spécificité. La valeur J peut être interprétée comme la probabilité du hasard, pour une personne vraiment atteinte de la pathologie, d'être détectée comme pathologique lorsque son score dépasse le seuil correspondant.

Les données des adultes aux épreuves de PC ont fait l'objet d'un début de calcul de normes, qui devraient être ultérieurement complétées en agrandissant l'échantillon d'adultes, et en étendant à des groupes d'enfants.

### 2.10.8. Résultats

**En discrimination.** L'examen des courbes ROC présentées en Figure 16 suggérait que la paire /di-ti/ dans l'épreuve de discrimination permettait des mesures plus pertinentes que la paire /gi-ki/ pour le diagnostic d'une dyslexie chez l'adulte. En effet, pour /di-ti/, la courbe s'arquait suffisamment pour s'éloigner de la diagonale entre le point où Sensibilité et Spécificité = 0 (en bas à gauche) et celui où Sensibilité et 1 - Spécificité = 1 (en haut à droite). Cela suggérait une assez bonne capacité de classification des adultes dyslexiques et des adultes normo-lecteurs.

Pour la paire /di-ti/, l'aire sous la courbe (AUC) était de .72 (moyennement informatif). Les *cut-off* déterminés selon le meilleur équilibre entre sensibilité – spécificité permettaient d'établir les valeurs seuils au-delà desquelles les mesures reflèteraient une pathologie :  $d' < 1.04$  (sensibilité = 0.72, spécificité = 0.70, précision : J de Youden = .42). Dans un souci d'aide au diagnostic, nous avons proposé des résultats maximisant la sensibilité et la spécificité. Lorsque la **sensibilité était maximisée**, le *cut-off* pour le  $d'$  était  $< 0.70$  (sensibilité = 0.85, spécificité = .52, précision : J de Youden = .37). Lorsque la **spécificité était maximisée**, le *cut-off* pour le  $d'$  était  $< 1.33$  (sensibilité = 0.50, spécificité = .83, précision : J de Youden = .33).

Pour la paire /gi-ki/, l'AUC était de .64 (peu informatif), avec un *cut-off* de  $d' < 1.73$  respectant seulement une très bonne spécificité (sensibilité = 0.29, spécificité = 1, précision : J de Youden = .29).

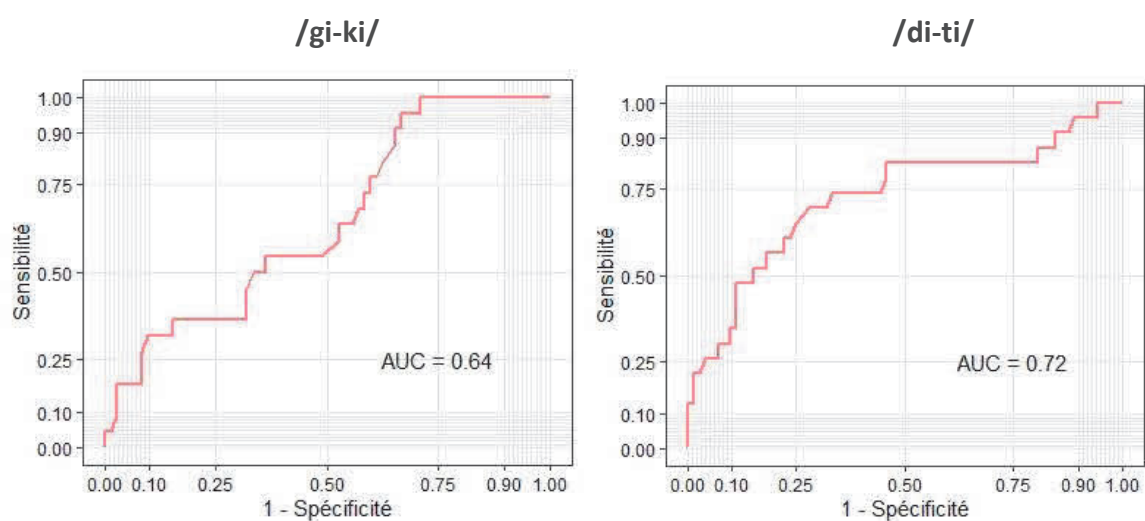


Figure 16. Courbes ROC (Receiver Operator Characteristic) pour les  $d'$  (participants normo-lecteurs, dyslexiques) pour les paires /gi-ki/ (gauche) et /di-ti/ (droite).

**En identification**, pour la paire /di-ti/ (Fig. 17), l'AUC pour les valeurs de seuil d'identification était de .54 (peu informatif), avec *cut-off* < 12.31 (sensibilité = 0.67, spécificité = 0.48, précision : J de Youden = .15). Pour les valeurs de pente, l'AUC était de .57 (peu informatif), avec un *cut-off* < 0.28 où seule la spécificité était correcte (sensibilité = 0.39, spécificité = 0.91, précision : J de Youden = .31).

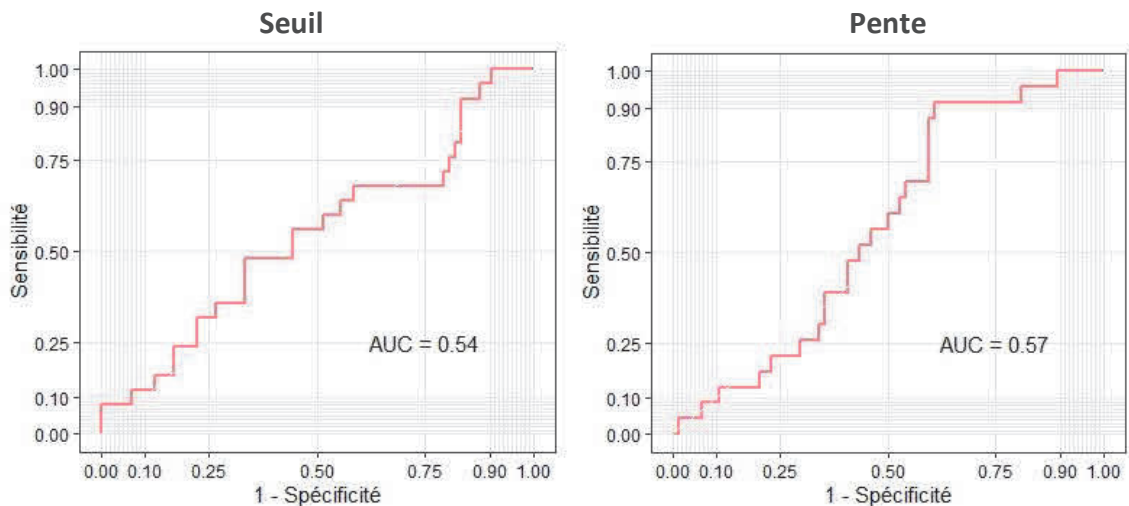


Figure 17. Courbes ROC (Receiver Operator Characteristic) pour le seuil (gauche) et la pente (droite) d'identification (participants normo-lecteurs, dyslexiques) pour la paire /di-ti/.

Pour la paire /gi-ki/ (Fig. 18), l'AUC pour les valeurs de seuil d'identification était de .53 (peu informatif), avec un *cut-off* < 13.25 où seule la sensibilité était correcte (sensibilité = 0.76, spécificité = 0.39, précision : J de Youden = .15). Pour les valeurs de pente, l'AUC était de .61 (peu informatif), avec un *cut-off* < 0.27 où seule la spécificité était correcte (sensibilité = 0.43, spécificité = 0.83, précision : J de Youden = .27).

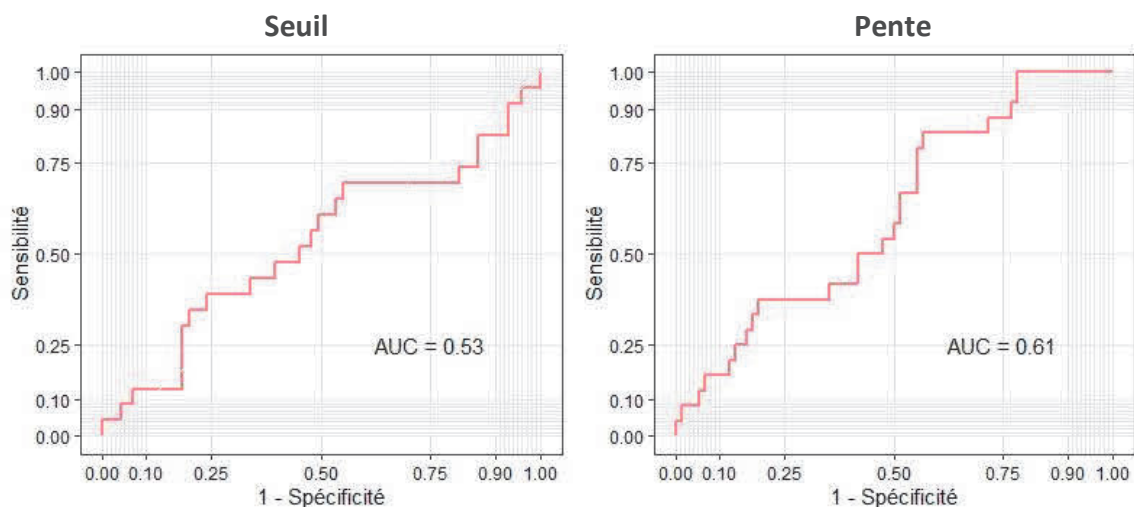


Figure 18. Courbes ROC (Receiver Operator Characteristic) pour le seuil (gauche) et la pente (droite) d'identification (participants normo-lecteurs, dyslexiques) pour la paire /gi-ki/.

## **2.11. Discussion sur les tests de perception catégorielle en français (Expérience 1)**

Ce chapitre avait pour objectif principal de créer un nouvel ensemble de tests de PC en français en vue de disposer d'outils pour évaluer les effets de deux programmes d'entraînement perceptifs en sessions pré- et post-entraînement (Chapitres 3 et 5). Pour cela, la principale qualité recherchée était un faible effet test-retest. Les épreuves d'identification et de discrimination utilisaient deux paires de consonnes occlusives opposées par le voisement (/d-t/ et /g-k/) et présentant des lieux d'articulation distincts. Secondairement, d'autres qualités psychométriques des tests ont été évaluées, en lien avec le diagnostic d'une dyslexie à l'âge adulte : leur fidélité et leur sensibilité/spécificité pour catégoriser des adultes avec ou sans dyslexie. L'effectif était faible pour ce type d'évaluation, et il s'agissait simplement ici de repérer les épreuves et les paires les plus adaptées à la mise en place d'éventuelles normes, afin de contribuer par un nouvel outil au difficile diagnostic des déficits sous-jacents à la dyslexie chez l'adulte.

La discussion de ce chapitre commentera les qualités et les limites de ces épreuves de PC pour la première utilisation envisagée (l'évaluation de l'effet de programmes de remédiation), puis plus modestement pour la deuxième piste concernant l'aide au diagnostic de déficits sous-jacents à la dyslexie chez l'adulte.

### **2.11.1. Qualités diagnostiques des tests et paire de consonnes pour évaluer la PC**

Les quatre épreuves de PC en français ont toutes présenté la qualité recherchée pour leur utilisation de cette recherche : elles ne sont pas sensibles à l'effet test-retest. En effet, il n'y a pas eu de modification significative des performances, que ce soit pour le *d'* en discrimination ou les valeurs de seuil et de pente d'identification, lorsque les participants normo-lecteurs réalisaient deux fois chaque test à 5 jours d'intervalle. Cette qualité s'est observée pour les deux paires évaluées. Il y a peu d'effet d'apprentissage de la tâche dû à la pratique. La vérification de cette qualité encourage à utiliser les quatre épreuves pour évaluer les effets de programmes d'entraînement en testant les mêmes individus avant et après l'entraînement. Les potentiels progrès seront alors attribuables à l'entraînement et non simplement à un effet de pratique de la tâche. Cette vérification sur deux paires de phonèmes permettra d'utiliser l'une d'elles pour évaluer l'effet sur des phonèmes perçus pendant l'entraînement, et l'autre pour tester si un transfert s'est produit à des phonèmes non perçus pendant celui-ci.

Pour l'évaluation de la fidélité, les deux paires se sont avérées de qualité inégale. La paire /d-t/ présentaient une bonne fidélité générale. En effet, les valeurs de *d'* en



discrimination et le seuil en identification étaient largement corrélées entre la première et la deuxième passation. La corrélation pour valeur de pente n'était certes pas significative, mais sa taille pouvait être estimée comme modérée. Cette dernière valeur est cependant la moins fidèle pour l'identification de cette paire de consonnes. Cette moindre qualité pourrait être liée à la méthode de calcul de la pente. Le calcul était automatique, quelle que soit la qualité générale de la sigmoïde attendue, et dans le cas d'une courbe très éloignée d'une sigmoïde, un ajustement était tout de même fait, surévaluant alors la qualité de la pente. La valeur d'ajustement de la pente n'est donc pas la plus fiable pour les résultats obtenus dans l'étude. Dans l'ensemble, les tests étaient moins fidèles pour la paire /g-k/ : aucune des trois mesures de corrélation entre les sessions 1 et 2 n'atteignait le seuil de significativité, et seule la fidélité de la valeur du seuil d'identification pouvait être estimée comme de taille modérée.

Dans le cadre d'un suivi de patient ou pour toute approche longitudinale, l'utilisation de la paire /d-t/ serait donc recommandée pour l'absence d'effet test-retest et pour sa fidélité. La paire /g-k/ présente néanmoins une qualité intéressante avec son absence d'effet test-retest, même si sa moindre fidélité en fait une moins bonne paire que /d-t/.

Si l'objectif est différent est qu'il s'agit de contribuer à un diagnostic, et surtout de comprendre certains déficits sous-jacents à une dyslexie chez l'adulte, les critères liés à l'effet test-retest et à la fidélité sont moins déterminants, puisqu'une seule séance de test sera proposée. En ce cas, la sensibilité et la spécificité sont plus pertinentes. Il se trouve que, encore sur ce plan, la paire /d-t/ présentait de meilleures qualités dans notre étude. Elle était plus efficace que la paire /g-k/ pour classer les adultes comme normo-lecteurs ou dyslexiques, particulièrement avec l'épreuve de discrimination. Probablement car la paire /d-t/ présentait moins de variabilité inter-sujet que la paire /g-k/.

Il est difficile d'expliquer cette meilleure qualité de la paire d'alvéolaires pour évaluer la PC en français. Il est possible qu'une part de l'effet peut être due à l'ordre dans lequel les paires ont été présentées dans les épreuves mais ça ne peut être la seule explication au vu des plus nombreux résultats en faveur de cette paire. Un rapprochement peut aussi être fait avec l'utilisation de la même paire de consonnes, avec un contexte vocalique parfois, dans plusieurs expériences réalisées par Serniclaes et ses collègues. Ils ont par exemple montré que les paires /de/-/te/ ou /di/-/ti/ étaient mieux à même de révéler une différence entre des enfants dyslexiques et normo-lecteurs que la paire /b/-/p/ (Serniclaes *et al.*, 2021). Par exemple, le score de discrimination n'était significativement différent entre les groupes que pour les paires d'alvéolaires. Une différence entre la PC d'enfants faibles lecteurs et bons lecteurs était



déjà apparue dans une étude antérieure avec la paire /do/-/to/ (Bogliotti *et al.*, 2003). Serniclaes avait par ailleurs obtenu un effet du lieu d'articulation (alvéolaires vs. vélaires, comme dans la présente étude) des paires de consonnes opposées en voisements dans des épreuves de PC (Serniclaes *et al.*, 2004). Enfin, pour se rapprocher d'une étude visant elle aussi à évaluer l'impact de deux programmes d'entraînement, c'est aussi avec une paire d'occlusives alvéolaires que les progrès en discrimination et en identification de phonèmes ont été évalués, avec succès, par Zoubrinetzky *et al.* (2019). Cela se manifestait par un effet principal de la paire sur les performances et même une interaction entre le groupe (dyslexiques vs. normo-lecteurs) et la paire. Au final, les résultats des travaux de Serniclaes et ses collègues et ceux de la présente étude concourent à conseiller d'utiliser en priorité une paire d'occlusives alvéolaires /d/-/t/ plutôt que des paires d'occlusives labiales ou vélaires pour la PC et ses troubles dans la dyslexie. Par contre, dans le cadre d'une évaluation pré- et post-entraînement, il peut être intéressant d'avoir une seconde paire de phonèmes comme (/g-k/) pour tester un éventuel effet de transfert de compétence.

### **2.11.2. Défis de l'évaluation de la perception catégorielle chez des adultes dyslexiques et normo-lecteurs à l'université**

Les résultats des qualités psychométriques des quatre épreuves de PC en français en font de bons indicateurs pour l'évaluation avant/après d'une prise en charge. Ils conduisent cependant à insister sur leurs limites dans le cadre de l'aide au diagnostic chez l'adulte. En effet, elles ne se sont pas avérées très efficaces pour différencier les adultes dyslexiques des normo-lecteurs (l'une des paires étant peu informative et l'autre l'étant seulement moyennement). Ces tests ne sont pas recommandés comme outil de diagnostic. Ils peuvent contribuer à une analyse qualitative des difficultés d'un patient particulier sur le plan de l'organisation de son système phonologique, mais cela surtout dans le cadre d'une recherche de pistes pour la prise en charge.

Plusieurs explications peuvent être évoquées concernant l'origine des performances médiocres des épreuves pour la classification des participants. Un débriefing avec les participants nous a permis de constater que certains qualifiaient les tâches « peu habituelle » et « difficiles ». Ils n'ont pas l'habitude de répondre aux demandes de ce type de tâche dans le cadre académique et un effet de surprise a pu brouiller les effets. De plus, les épreuves avaient une durée totale d'1 heure (40 min de discrimination et 20 min d'identification). Cela demandait au participant une forte concentration et le maintien de l'attention est difficile dans une tâche répétitive et un peu ennuyeuse. En effet, les autres études utilisent des continua plus

court, réduisant la durée des passations chez l'enfant. Comme il s'agissait ici de tester des adultes, les tests proposait un plus grand nombre d'items. La vigilance était alors mobilisée et des variations inter-individuelles pour cette dernière compétence ont pu noyer certains effets, malgré les pauses régulières.

Une seconde explication renvoie à des difficultés déjà documentées : la dyslexie est globalement difficile à repérer et diagnostiquer chez les adultes, et les déficits en PC sont chez eux peu visibles à travers des données comportementales (Noordenbos & Serniclaes, 2015). Il y a à cela deux raisons principales : 1) ils ont développé des stratégies de compensation, 2) les troubles phonologiques sont moins saillants que les déficits métaphonologiques (conscience phonologique) chez les adultes dyslexiques. Les compensations, tout comme l'atténuation de la visibilité des anomalies phonologiques chez les participants dyslexiques de notre étude peuvent s'expliquer par l'efficacité des prises en charge reçues souvent pendant des années et ciblant prioritairement la phonologie. Par ailleurs, les participants dans cette étude avaient accédé aux études supérieures, phénomène assez rare puisqu'on estime à 1,4 % la part d'étudiants dyslexiques dans les études supérieures en France (Colé & Sprenger-Charolles, 2021). Les dyslexiques universitaires peuvent être considérés comme une population à part des dyslexiques tout-venant, car la capacité à développer de fortes compensations est probablement une de leurs particularités. Il est aussi possible que certains participants de l'étude soient atteints d'une forme de dyslexie dans laquelle les troubles sous-jacents sont essentiellement visuo-attentionnels, et la petite taille de l'échantillon peut avoir, par malchance, joué sur ce plan en notre défaveur. Nous n'avions en effet, par eu accès à leur bilan orthophonique. Il serait donc intéressant de compléter les normes avec des adultes dyslexiques et normo-lecteurs pus nombreux et « tout-venant », qui n'ont pas tous obtenu le baccalauréat et entamé des études. Cela permettrait de plus refléter la réalité de cette pathologie chez l'adulte.

L'étude des déficits du système phonologique lui-même, et non de la conscience phonologique ou de l'accès au lexique phonologique est un réel défi chez l'adulte dyslexique. De nombreuses études montrent des déficits de PC chez l'enfant dyslexique, avec de nombreux indices d'une perception allophonique des sons du langage (Godfrey *et al.*, 1981 ; Sprenger-Charolles & Serniclaes, 2004). En revanche, ce déficit est peu relaté chez l'adulte. Par exemple, Ramus *et al.* (2003) ont réalisé une batterie de tests incluant des épreuves de perception auditives et de catégorisation de sons chez 16 adultes dyslexiques anglophones, comparés à des normo-lecteurs. Ils n'ont pas trouvé de différence de performances entre les

groupes pour les tâches proposées. Les évaluations comportementales ont à cet égard des limites. Des recherches montrent que la PC est atypique chez les adultes dyslexiques d'après des enregistrements neurophysiologiques (Noordenbos *et al.*, 2013). Par exemple, l'imprécision de la frontière catégorielle persiste, selon l'amplitude anormalement faible de la MMN à l'écoute d'un changement de consonnes (Schulte-Körne *et al.*, 2001). Cette anomalie est encore plus nette si les exemplaires diffusés présentent de petites différences de hauteur de voix (variations du F0) comme dans les situations de perception naturelle (Virtala *et al.*, 2020). Cependant, Noordenbos *et al.* (2013) ont montré qu'il était plus difficile d'observer, chez les mêmes participants adultes, des signes de ces déficits dans des expériences comportementales. En proposant ici encore des épreuves de PC à des adultes dyslexiques, nous nous sommes heurtés à la même difficulté. C'est pourquoi, la suite de la recherche de cette thèse ne se limite pas au recueil de données comportementales, mais s'ouvre à des mesures électrophysiologiques pour évaluer les compétences perceptives ainsi que l'effet de remédiations portant sur la perception du contraste de voisement en français chez des adultes francophones dyslexiques.

## **Chapitre 3. Remédiation de la perception du voisement en français pour adultes dyslexiques**

---

Même si les performances aux tâches de perception catégorielle (PC) ne suffisent pas à elles seules pour différencier des adultes dyslexiques d'adultes normo-lecteurs, il est reconnu que la PC présente des anomalies chez des adultes dyslexiques. Ils ont souvent mis en place des compensations et progressé en phonologie, notamment grâce aux prises en charges orthophoniques, mais des données suggèrent que leur perception phonologique n'est pas aussi catégorielle qu'elle devrait l'être. C'est pourquoi, à travers le Chapitre 3, nous avons souhaité développer un outil de remédiation à la perception de phonèmes opposés en voisement en français pour des adultes dyslexiques. Il existe aujourd'hui très peu d'outils permettant la rééducation ciblée de la PC, et surtout à destination des adultes. L'effet de cette remédiation sera évalué avec les épreuves de PC construites et évaluées précédemment (Chapitre 2), ainsi qu'avec des mesures électrophysiologiques (ERP). L'objectif de cet outil d'entraînement à visée de remédiation est de pouvoir ensuite être proposé aux orthophonistes lors de prises en charge.

### **3.1. Outils contemporains pour la rééducation phonologique**

De nos jours, de plus en plus de matériel de rééducation orthophonique est informatisé et la plupart s'adressent à un public jeune. Il n'existe que très peu d'outils de remédiation de la phonologie en français pour l'adulte. Les outils disponibles se focalisent sur la phonologie en général, la conscience phonologique, les compétences sémantiques et lexicales, ainsi que sur la mémoire. Pourtant, un entraînement perceptif ciblé sur des compétences déficitaires précises dans la langue maternelle, en cas de dyslexie, peut s'avérer efficace (pour une revue voir Serniclaes *et al.*, 2015). Ici encore, la plupart des recherches menées à ce sujet sont réalisées chez des enfants, très peu s'intéressant à la dyslexie adulte.

Quelques études se sont intéressées au renforcement des capacités attentionnelles et d'analyses auditives pour améliorer le traitement phonologique. Temple *et al.* (2003) ont utilisé pour cela pendant 28 jours à raison de 5 jours/semaine (1h40 par jour) le programme d'entraînement *Fast ForWord Language* (FFW, Scientific Learning Corporation, Oakland, CA) auprès de 20 enfants dyslexiques anglophones (8-12 ans). Ce programme utilise sept exercices d'entraînement sur ordinateur qui mettent l'accent sur différents aspects du langage oral en insistant sur l'attention auditive pour améliorer la discrimination et le traitement

phonologique. Les exercices mobilisent des processus auditifs et attentionnels nécessaires au langage oral ainsi que la mémoire auditive. En comparant avant/après entraînement, les données en IRMf ont montré une augmentation de l'activité dans le cortex temporo-pariétal gauche et le gyrus frontal inférieur gauche, égalant alors une activation normale telle qu'observée chez des enfants sans pathologie. Les données montrent aussi une augmentation de l'activation dans les régions frontales et temporales droites et dans le gyrus cingulaire antérieur. Ce dernier est impliqué dans le contrôle attentionnel et dans l'apprentissage : les exercices semblent ainsi engager un apprentissage dans le domaine du langage. Les auteurs ont montré que l'activité cérébrale sous-tendant le traitement de la langue maternelle (L1) peut être modifiée, tout au moins chez l'enfant dyslexique.

Stevens *et al.* (2008) ont réalisé le même programme d'entraînement (FFW) pendant 6 semaines (1h40 par jour), cette fois chez 8 enfants atteints d'un Trouble Développemental du Langage (TDL) et 12 enfants normo-lecteurs (6 - 8 ans). Ils ont fait des mesures en potentiels évoqués et évalué les mécanismes attentionnels. Ils ont montré une amélioration des mécanismes neuronaux sous-tendant l'attention sélective (par une augmentation de la taille de la N1 après l'entraînement), soulignant ainsi les liens entre les mécanismes attentionnels et le traitement du langage aussi bien chez les enfants TDL que chez les contrôles. En recueillant cette fois des données en IRMf, Gaab *et al.* (2007) ont réalisé quant à eux le programme FFW encore plus longtemps (8 semaines) chez 22 enfants dyslexiques (moyenne = 10.5 ans). A la suite de cette prise en charge, les performances en langage oral et en lecture se sont améliorées. De plus, l'activation produite dans le cortex préfrontal gauche par les transitions formantiques rapides, qui étaient anormalement lente avant l'entraînement, s'est accélérée, ce qui a été interprété là encore comme un renforcement du lien entre langage et processus attentionnels. L'implication d'une activation dans le cortex préfrontal gauche peut être interprétée, dans le cadre de modèles associés au traitement de tels indices acoustiques spectraux dans l'HG (Belin *et al.*, 1998), mais mobilisant aussi la capacité à réaliser des traitements temporellement rapides.

Chez des enfants dyslexiques (9 enfants de 8 à 12 ans), Russo *et al.* (2005) ont testé l'efficacité d'Earobics (Diehl, 1999 ; Morrison, 1998). Par le biais de jeux informatiques interactifs, ce programme d'entraînement auditif (sélection d'images ou de mots représentés à l'écran en fonction de ce qui a été entendu) ambitionne de développer la conscience phonologique, le traitement auditif et le traitement du langage. Les chercheurs l'ont proposé pendant 35 à 40 séances d'1h réparties sur 8 semaines. Il leur a permis de produire une

amélioration significative de la synchronisation neuronale au niveau du tronc cérébral lors de la perception d'une syllabe. Comme les réponses du tronc cérébral reflètent l'encodage neuronal préconscient des caractéristiques acoustiques des syllabes, une meilleure synchronisation observée à ce niveau reflète une plus grande précision de l'encodage syllabique, déficitaire en cas de dyslexie (Leong & Goswami, 2014a). Il s'agit d'un progrès accompagné d'une amélioration de performances perceptives et même académiques.

Un autre programme d'exercices audio-visuels (Danon-Boileau & Barbier, 2000) ciblant l'amélioration des représentations phonologiques, et plus particulièrement cette fois le contraste de voisement, a été testé en français (Bedoin, 2003). Des animations visuelles (jeu de ballon) accompagnaient un entraînement à l'identification de consonnes opposées en voisement dans des syllabes CV diffusées dans un casque audio, et les enfants dyslexiques répondaient en sélectionnant l'une des deux consonnes écrites. Ils recevaient un feed-back sur la réussite et étaient entraînés 4 jours par semaine (30 min/jour) pendant 5 semaines. L'effet était mesuré sur l'implication de connaissances sur le voisement en lecture silencieuse. Avant l'entraînement, les enfants dyslexiques n'étaient pas sensibles au partage du trait de voisement par les consonnes d'un mot disyllabique écrit, alors qu'ils le sont devenus après l'entraînement, comme les enfants et adultes normo-lecteurs. L'implication plus systématique de connaissances sur le trait phonologique de voisement traduisait un transfert de compétences phonologiques vers la lecture silencieuse. Au bout de 10h d'un tel entraînement auditivo-phonologique, des effets positifs se sont aussi produits dans une tâche d'identification de mots écrits (Magnan *et al.*, 2004).

Il existe aujourd'hui un autre programme d'entraînement ciblant sur la perception du voisement des consonnes alvéolaires en français. Ce programme d'entraînement – RapDys (Collet *et al.*, 2017) – a été conçu dans le cadre d'une recherche sur la perception de la frontière catégorielle liée au voisement, en vue d'une rééducation des troubles spécifiques du langage écrit (dyslexie et TDL). Il s'adresse aux enfants présentant des « confusions sourdes-sonores » en lecture dès le début de l'apprentissage. Il pourrait aussi être proposé plus tôt dans le cas d'un contexte familial de dyslexie permettant de s'intéresser à des enfants à risque de dyslexie. Cet entraînement de 9h a permis à Collet *et al.* (2012) d'améliorer significativement la perception catégorielle (identification et discrimination) chez 18 enfants TDL francophones (de 6 à 11 ans). Proposé pendant 6h, ce logiciel a aussi permis à un groupe de 45 enfants dyslexiques d'accéder à une perception plus catégorielle de sons du langage et d'améliorer leurs compétences en conscience phonologique et en lecture (Collet *et al.*, 2017). Les effets de

RapDys ont aussi été comparés à ceux de MAEVA (Lobier *et al.*, 2012), un entraînement à la catégorisation de caractère, visant un élargissement de l’empan visuo-attentionnel. Les résultats suggèrent que les deux entraînements favorisent l’amélioration de la conscience phonologique et la lecture de mots réguliers. RapDys est cependant le seul à améliorer en plus la discrimination de phonème et la lecture de pseudo-mots, tandis que MAEVA améliore l’empan visuo-attentionnel et la lecture de mots irréguliers (Zoubrinetzky *et al.*, 2019). Dans le cadre de notre étude, le logiciel RapDys, basé sur la perception catégorielle, est un bon modèle pour la remédiation à la perception du voisement en français. Cependant, ce dernier n’a été utilisé que chez l’enfant, et il pourrait paraître insuffisant chez l’adulte.

Des modifications des traitements phonologiques semblent possibles grâce à des entraînements perceptifs, mais à notre connaissance très peu sont testés chez l’adulte. Tremblay et Kraus (2002) ont cependant montré chez de jeunes adultes sans pathologies (21 - 31 ans) des changements significatifs en potentiels évoqués après un entraînement perceptif de 2h40 (4 sessions de 40 min) sur deux exemplaires synthétiques de la même syllabe /ba/. Des modifications ont été enregistrées sur des ondes précoces – la P1, la N1 (dans l’hémisphère droit) et la P2 (bilatéral) – traduisant un changement de réaction corticale pré-attentionnelle à des sons de parole, qui ne se produisait pas dans le groupe contrôle sans entraînement. Il ne s’agissait pas de perception catégorielle, puisque les deux stimuli présentés commençaient par des exemplaires d’un même phonème. Ces résultats montrent toutefois que l’activité neuronale peut changer sous l’influence d’un entraînement perceptif, y compris chez l’adulte.

### **3.2. Objectif**

Dans l’étude précédente (Expérience 1), des outils d’évaluation de la PC ont été créés. Même s’ils n’ont pas montré un fort intérêt diagnostique pour des adultes dyslexiques, ils sont tout de même utilisables dans le cadre de l’évaluation pré- post- entraînement d’un outil de remédiation, car ils se sont avérés peu sensibles à l’effet test-retest. Beaucoup d’arguments issus d’expériences, la plupart du temps conduites auprès d’enfants, montrent que la PC est toujours déficitaire en cas de dyslexie chez l’adulte (Noordenbos et Serniclaes, 2015). Elle est basée sur un mode allophonique plutôt que catégoriel, c’est-à-dire que des différences entre des exemplaires d’un même phonème sont perçues avec trop d’acuité et que la frontière catégorielle n’est pas perçue de façon précise. Même si les déficits phonologiques sont difficiles à mesurer à l’âge adulte à cause des compensations mises en place, il n’en reste pas

moins que des anomalies demeurent, d'après quelques données électrophysiologiques (Noordenbos *et al.*, 2012, 2013).

La difficulté à recueillir des indices de ces anomalies est le principal verrou à lever dans l'étude de ce chapitre. Nous avons vu que des recherches chez l'enfant dyslexique suggèrent qu'il est possible d'améliorer la PC de phonèmes dans leur langue maternelle à l'aide d'entraînements perceptifs. Les études traitant cette question de remédiation chez l'adulte sont rares, mais laissent place à quelques espoirs, malgré une plasticité fonctionnelle du cerveau connue pour être moindre par rapport à celle de l'enfant. L'innovation de l'entraînement proposé dans l'étude longitudinale courte de ce chapitre, est d'utiliser l'écoute dichotique pour stimuler préférentiellement l'association hémisphère-son la plus compétente pour le voisement en français, en prenant en compte les caractéristiques temporelles de ce trait phonétique. Dans des exercices perceptifs, le dispositif consiste à présenter systématiquement les consonnes occlusives sourdes caractérisées par un VOT long à l'oreille gauche (pour favoriser un traitement préférentiel par l'HD), et occlusives les sourdes caractérisées par un VOT court à l'oreille droite (pour favoriser un traitement préférentiel par l'HG). Le dispositif est qualifié de « dichotique », car la priorité de la voie controlatérale est renforcée par la diffusion, en parallèle, d'un bruit blanc dans l'autre oreille. Comme dans les expériences d'écoute dichotique (Bedoin *et al.*, 2010 ; Rimol *et al.*, 2006), il s'agissait d'introduire la concurrence qui renforce cette priorité.

L'objectif de cette partie de la recherche (Expériences 2 et 3) était d'évaluer l'effet d'un entraînement bref ciblé sur le voisement des consonnes françaises auprès d'adultes dyslexiques. Cette remédiation était répartie sur 5 jours consécutifs à raison de 30 minutes par jour. Les effets de l'entraînement ont été évalués à travers des données comportementales (**Expérience 2**) mais aussi électrophysiologiques en potentiels évoqués (**Expérience 3**). Ce travail était guidé par le souhait de pouvoir proposer à des professionnels, particulièrement des orthophonistes, un nouvel outil de remédiation du voisement en cas de troubles du langage chez l'adulte à la suite de son évaluation.

Les résultats des deux expériences seront discutés et mis en relation dans la discussion générale de ce chapitre.



### **3.3. Expérience 2 : effet de l'entraînement sur la perception catégorielle**

#### **3.3.1. Principe de l'expérience et hypothèses**

L'Expérience 2 suivait deux objectifs. Le premier était de comparer les performances entre normo-lecteurs et dyslexiques en PC du français, leur L1. Le second était de mesurer l'effet d'un entraînement perceptif en français chez des adultes dyslexiques. Cet entraînement a été élaboré comme un outil de remédiation à la perception du voisement des consonnes occlusives françaises. Il utilisait la stimulation dichotique pour favoriser l'association Hémisphère – Son la plus compétente pour traiter la durée du VOT, indice acoustique important pour identifier le voisement de ces consonnes.

Les hypothèses suivantes ont été formulées :

- 1) Les performances en PC devraient être moins élevées pour le groupe d'adultes dyslexiques que pour le groupe d'adultes normo-lecteurs (*d'* plus faible en discrimination, seuil d'identification moins approprié à la langue française et pente d'identification moins abrupte).
- 2) L'entraînement perceptif avec dispositif de stimulation dichotique devrait permettre aux adultes dyslexiques d'améliorer leur PC (identification et discrimination) des consonnes occlusives françaises opposées en voisement ayant fait directement l'objet des exercices d'entraînement.
- 3) L'entraînement devrait aussi améliorer la PC de consonnes occlusives françaises opposées en voisement n'ayant pas été entendues pendant l'entraînement, témoignant ainsi d'un transfert de compétences.

#### **3.3.2. Méthode**

##### **3.3.2.1. Participants**

Les 27 participants (21 femmes et 6 hommes) de 18.5 ans à 26.3 ans ( $M = 21$  ans 9 mois;  $ET = 2$  ans 9 mois) étaient des étudiants monolingues francophones ayant reçu un diagnostic de dyslexie développementale par un orthophoniste durant l'enfance. Ils ne présentaient pas et n'avaient pas présenté d'autre trouble neurodéveloppemental (notamment un TDAH), ni de trouble articulaire ou bégaiement. Ils déclaraient avoir une bonne audition et une vision normale ou corrigée. Tous les participants étaient droitiers (80% de réponses « droite » dans la version abrégée de l'Edinburgh Handedness Inventory (Oldfield, 1971)).

Les participants normo-lecteurs (n = 20, 16 femmes, M = 21 ans et 5 mois, ET = 2 ans 6 mois) inclus dans cette expérience de PC étaient ceux du groupe contrôle de l'Expérience 6 (Chapitre 5) (Annexe 3). Ces participants normo-lecteurs n'effectuaient pas cet entraînement en français. La comparaison entre ces deux groupes avant les entraînements (données à T1) était motivée par l'investigation de potentielles lacunes en PC toujours visibles à l'âge adulte en cas de dyslexie.

### **3.3.2.2. Matériel et équipement**

La PC a été testée à l'aide du matériel décrit dans le Chapitre 2, partie 2.10.1.

Les exercices de l'entraînement en français ont été créés en 2016 (Moulay, 2016). Initialement programmés avec E-Prime (Psychology Software Tools, Inc., Sharpsburg, PA), ils ont été reprogrammés avec OpenSesame (Mathôt, S., *et al.*, 2012), logiciel open-source, pour faciliter leur diffusion auprès des participants. Un total de 255 mots et pseudo-mots ont été enregistrés par une francophone monolingue natif à l'aide du logiciel ROCme! (2012). Les mots et pseudo-mots étaient tous monosyllabiques et commençaient par une consonne occlusive (/p/, /b/, /k/ ou /g/) alors que les consonnes utilisées pour les épreuves-tests de PC étaient /k/, /g/, /t/ et /d/. Les voyelles qui suivaient les consonnes dans les mots des exercices d'entraînement étaient très variées (a, e, i, o, u). Pour introduire de la diversité, chaque mot pour lequel un exercice impliquait une illustration était associé à 3 ou 4 photos différentes (172) à travers les sessions. Elles étaient gratuites et libres de droit avec modifications possibles.

### **3.3.2.3. Procédure de comparaison des performances d'adultes dyslexiques et normo-lecteurs en perception catégorielle**

Pour comparer la PC entre les deux groupes (dyslexiques et normo-lecteurs), leurs performances dans la session T1, avant tout entraînement, ont été prises en compte.

### **3.3.2.4. Procédure d'évaluation pré- post-entraînement**

Afin d'évaluer l'efficacité de l'entraînement, les participants dyslexiques réalisaient deux fois les épreuves de PC décrites au Chapitre 2 (deux tâches d'identification et deux tâches de discrimination), à 5 jours d'intervalles (Fig. 19), c'est-à-dire en sessions pré-entraînement (T1) et post-entraînement (T2). Comme décrit précédemment, les paires /g-k/ et /d-t/ étaient utilisées pour ces épreuves de PC, la première pour évaluer l'effet de l'entraînement sur les mêmes consonnes que celles de l'exercice, et la seconde pour un possible effet de transfert,

avec généralisation à des consonnes non présentées dans l'entraînement. Il n'y avait pas de groupe contrôle (un groupe dyslexique testé deux fois en français sans entraînement (groupe contrôle passif) ou avec un autre type d'entraînement (groupe contrôle actif) entre les deux sessions), car dans le cadre plus général de cette étude, l'entraînement en français était considéré comme souhaitable pour renforcer et protéger les connaissances phonologiques en français de tous les dyslexiques, qui participait ensuite à un deuxième entraînement sur du matériel anglais.

Le participant réalisait aussi un test d'écoute dichotique en français (Bedoin et al., 2010) à T1 et T2 permettant de mesurer la dominance hémisphérique pour le traitement d'indices acoustiques courts à l'écoute d'un matériel verbal (144 paires de mots monosyllabiques diffusés de manière parfaitement synchronisée, un tiers des couples différaient seulement par le lieu d'articulation de leur consonne initiale). Les participants n'étaient pas informés que des mots différents leur étaient présentés dans l'une et l'autre oreilles (il leur était expliqué que le son de la voix du locuteur était volontairement altéré pour les besoins du test). Ils devaient focaliser leur attention sur un point placé en face d'eux, pour minimiser les biais attentionnels vers une partie de l'espace. Ils devaient répéter à haute voix ce qu'ils entendaient et l'expérimentatrice notait les réponses. Aucun feed-back n'était donné mais des encouragements étaient systématiquement fournis lors des trois pauses. Dans les tests d'écoute dichotique, il arrive que certains participants rappellent systématiquement un mot particulier quelle que soit l'oreille de présentation, ce qui constitue un biais pour une mesure précise de l'asymétrie hémisphérique (Voyer & Techentin, 2009). Les couples, pour lesquels cela se produisait, étaient systématiquement retirés avant le calcul de l'asymétrie, selon les recommandations de Wexler et Halwes (1983) et Zatorre (1989). L'indice lambda ( $\lambda = \ln\left(\frac{R+1}{L+1}\right)$ ) où  $\ln$  = logarithme naturel, R = nombre de réponses indiquant le mot de l'oreille droite, L = nombre de réponses indiquant le mot de l'oreille gauche) (Bryden & Sprott, 1981) a été calculé pour évaluer l'avantage de l'oreille droite (*Right Ear Advantage*, REA) traduisant un avantage de l'HG. Cet indice est considéré comme particulièrement fiable, car il ne dépend pas du niveau général d'efficacité (de Bode *et al.*, 2007 ; DiStefano *et al.*, 2004 ; Fernandez & Smith, 2000). Lorsqu'il est positif, le REA indique une dominance de l'oreille droite - HG, tandis qu'il indique une dominance de l'oreille gauche - HD (*Left Ear Advantage*, LEA) s'il est négatif. Ici, il était pris en compte pour les couples de mot différaient par le lieu d'articulation, trait phonologique associé à des indices acoustiques très brefs (transitions formantiques), propres à permettre le relevé d'un REA de grande taille.

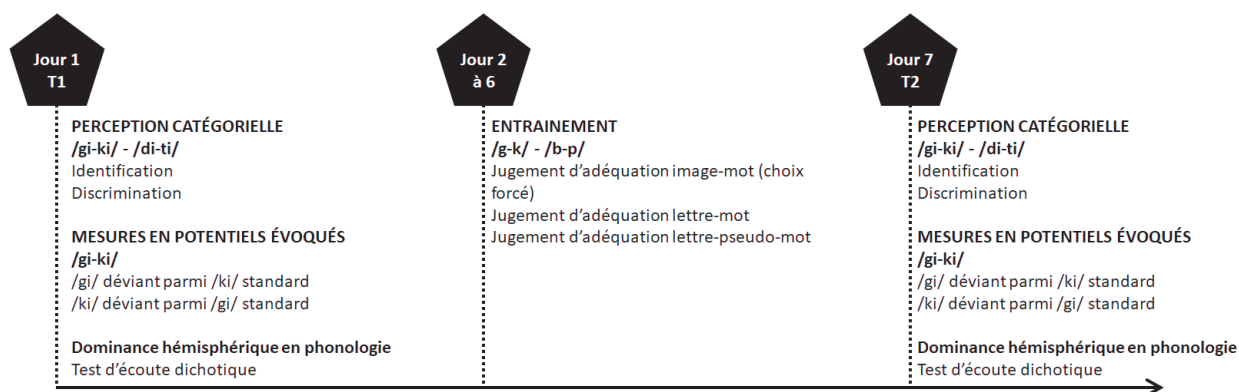


Figure 19. Procédure générale des Expériences 2 et 3.

### 3.3.2.5. Procédure d'entraînement en français avec dispositif dichotique

L'entraînement se déroulait durant 5 jours consécutifs à raison de 30 minutes par jour. Chaque jour, les participants réalisaient 3 types d'exercices (Fig. 20). Les sons étaient toujours présentés dans l'oreille controlatérale à l'hémisphère privilégié pour la durée du VOT de la consonne initiale. L'autre oreille recevait un bruit blanc simultanément. L'**Exercice de jugement 1** (20 min) présentait deux images à l'écran avec un mot français CVC écrit entre les deux et entendu à travers le casque. Il était demandé au participant d'appuyer sur la touche associée à l'image correspondant au mot écrit. L'exercice était facile et visait aussi à attirer l'attention sur le lien entre la forme de la lettre initiale (seule différence entre les mots représentés par les 2 images) et le stimulus sonore. Les deux images représentaient les deux exemplaires de la paire minimale à laquelle appartenait le mot cible. L'**Exercice de jugement 2** (5 min) était un jugement d'adéquation entre la consonne initiale du mot français présenté dans une oreille et la lettre présentée à l'écran ; les stimuli n'étaient plus des paires minimales et l'exercice était assez difficile car dépourvu de support imagé. L'**Exercice de jugement 3** (5 min) était comme l'Exercice 2, mais à partir de pseudo-mots : la difficulté croissait donc encore, sans possibilité d'évoquer des connaissances lexicales. Chaque jour, les trois exercices étaient réalisés, mais dans un ordre différent, et sur un matériel en partie différent pour maintenir la motivation. Les paires /g-k/ et /b-p/ faisaient l'objet de chacun des exercices.

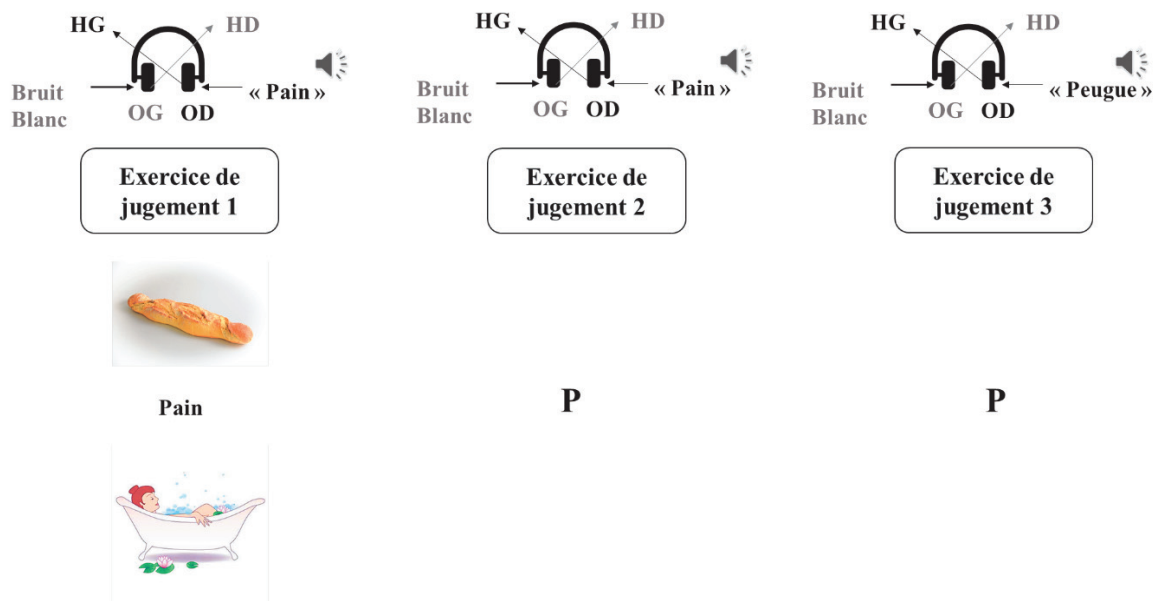


Figure 20. Déroulement de l'entraînement concernant les consonnes françaises en écoute dichotique, avec les trois types d'exercices proposés. Dans les trois types d'exercices, les stimuli auditifs étaient présentés selon le dispositif illustré en haut de la figure.

Lors de la réception des fichiers de résultats des entraînements, les fichiers de réponses étaient examinés pour vérifier que le participant n'avait pas appuyé sur d'autres touches que celles désignées pour les réponses, et qu'il n'avait pas produit plus de quatre temps supérieurs à 3 écart-types calculés à partir de ses données. Il n'y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Dans de cas où les réponses étaient globalement très lentes ou particulièrement rapides, les données du participant n'étaient pas conservées si cette lenteur ou rapidité extrêmement ne concernait que T1 ou que T2, car cela pouvait témoigner d'un manque de sérieux lors de l'une des étapes. Aucun participant n'a été exclu en respectant ces critères.

### 3.3.3. Analyses de données

Les données de l'épreuve de PC ont été traitées comme décrit dans le Chapitre 2, Expérience 1 partie 2.10.4. Les comparaisons entre le groupe normo-lecteurs (NL) et dyslexiques (DYS) ont été effectuées à l'aide de tests *t* de Student bilatéraux non-appariés avec corrections de Bonferroni sur chaque variable (*d'*, seuil, pente). L'effet de l'entraînement (entre T1 et T2) a été évalué à l'aide de tests *t* de Student bilatéraux appariés avec corrections de Bonferroni sur le *d'* en discrimination, les valeurs de seuil et de pente en identification et les valeurs lambda pour le REA en écoute dichotique. Les taille d'effet ont été interpréter selon Cohen (1988).

### 3.3.4. Résultats

**Comparaison des deux groupes à T1.** Les analyses n'ont montré aucune différence entre les deux groupes sur la hauteur du pic de discrimination, pour aucune des deux paires. En identification, seul le seuil de la paire /gi-ki/ présentait une différence de taille moyenne entre les deux groupes ( $t(30.88) = 2.03, p = .051, d = .52$ ). En effet, le groupe NL avait un seuil moyen à 16.12 ms et le groupe DYS à 23.75 ms. Le seuil du groupe NL était plus proche du seuil classiquement décrit pour le français. Aucune autre différence n'est ressortie dans l'épreuve d'identification.

**Effet d'entraînement du groupe DYS sur la discrimination.** Les tests  $t$  de Student ont montré une augmentation significative du pic de discrimination entre T1 et T2 pour la paire entraînée (/gi-ki/ :  $t(26) = 2.14, p = .042, d = .37$ ) et pas pour la paire non-entraînée (/di-ti/ :  $t(26) = 0.49, p = .629, d = .08$ ) (Fig. 21). Pour /gi-ki/, l'augmentation du pic s'est faite sans que ce soit au prix d'une dégradation de l'allure typique de la courbe pour les couples avec un VOT positif assez long (la courbe devenait même plus lisse du VOT 0-20 ms au VOT 40-60 ms, Fig. 21, gauche). L'observation des courbes montre qu'à T2 il n'y avait pas d'élargissement de la zone de VOT pour laquelle le  $d'$  était élevé, et donc pas de frontière devenant plus floue, ce qui aurait été signe d'une perception plus allophonique.

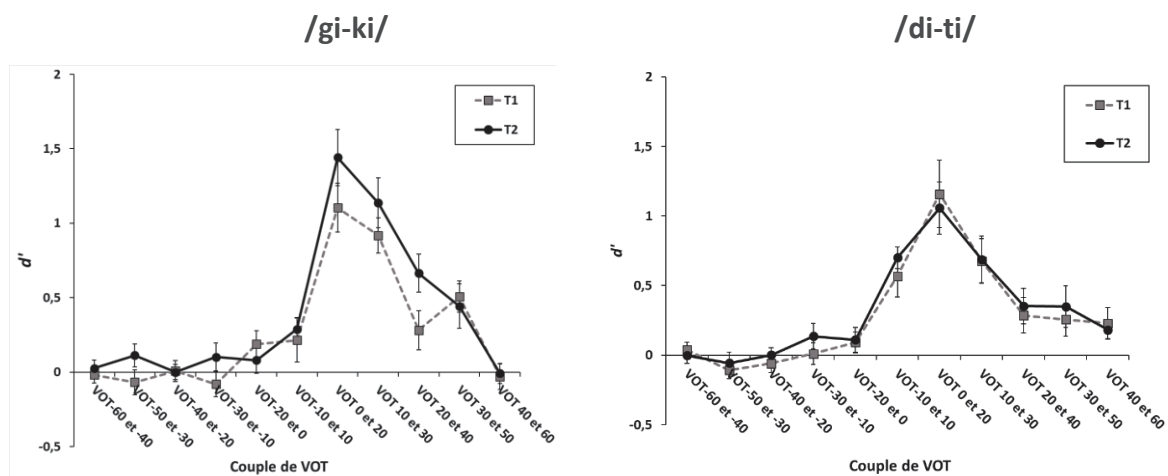


Figure 21. Courbes de discrimination pour les paires /gi-ki/ (gauche) et /di-ti/ (droite) à T1 (pointillés) et T2 (plein). Les barres d'erreurs représentent l'erreur-type.

**Effet d'entraînement du groupe DYS sur l'identification.** Au niveau des valeurs de seuil, les tests  $t$  de Student ont montré une tendance au déplacement du seuil entre T1 et T2 ( $t(26) = 1.75, p = .092, d = .27$ ) passant de 23.75 ms à 19.18 ms pour la paire /gi-ki/ dans le sens d'un rapprochement avec la frontière classique pour le français, mais pas de changement pour la paire /di-ti/. Aucun changement n'était significatif entre T1 et T2 pour les valeurs de pente.

Des analyses exploratoires complémentaires ont été réalisées sur les pourcentages d'identification selon des points d'intérêt observés sur la figure 22. En effet, il semblait se produire une amélioration de l'identification du /k/ entre T1 et T2 au niveau des VOTs 40 et 50 ms ainsi que pour l'identification du /t/ au niveau du VOT 40 ms. Les tests *t* de Student ont montré une augmentation significative du pourcentage d'identification du /k/ aux VOTs 40 ms ( $t(26) = 3.17, p = .004, d = .50$ ) et 50 ms ( $t(26) = 2.66, p = .013, d = .28$ ). L'amélioration du pourcentage d'identification du /t/ était quant à elle seulement tendancielle ( $t(26) = 1.88, p = .071, d = .42$ ). Les changements qui semblaient se dessiner au niveau des VOTs -10 et 0 ms pour /di-ti/ n'étaient pas significatifs (VOT -10 ms :  $t(26) = 1.21, p = .236, d = .30$  ; VOT 0 ms :  $t(26) = 1.72, p = .098, d = .37$ ).

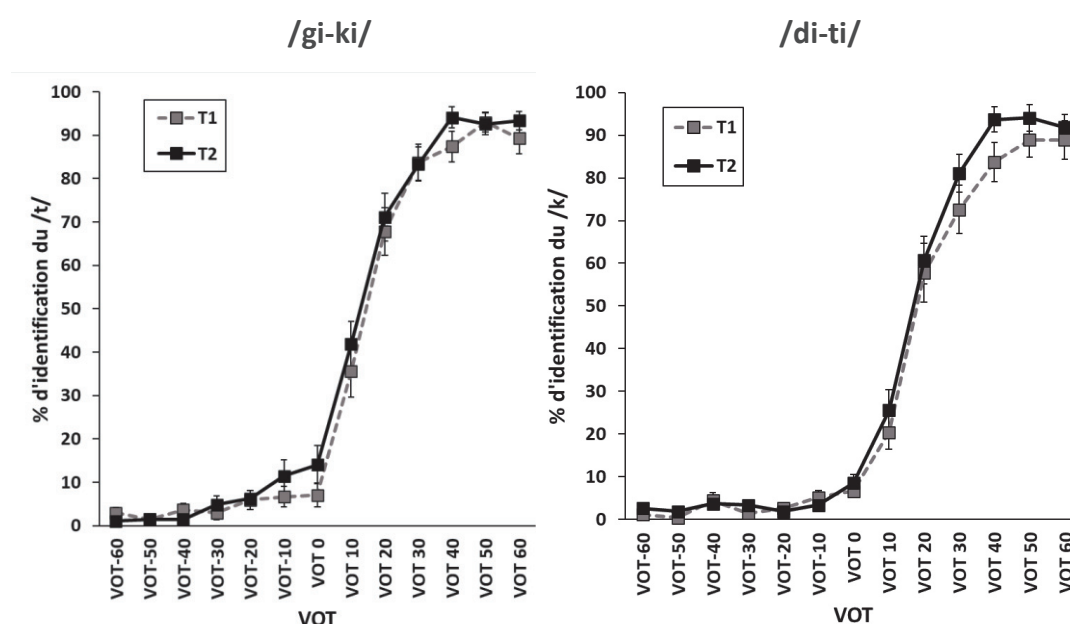


Figure 22. Courbes d'identification pour les paires /gi-ki/ (gauche) et /di-ti/ (droite) à T1 (pointillés) et T2 (plein). Les barres d'erreurs représentent l'erreur-type.

**Test d'écoute dichotique.** Le test *t* de Student a montré une augmentation du REA pour les couples de mots opposés en lieu d'articulation (paires les plus révélatrices d'un REA dans ce test) après l'entraînement ( $t(16) = -2.24, p = .034, d = .25$ ) passant de 0.72 à 0.94 (Fig. 23).

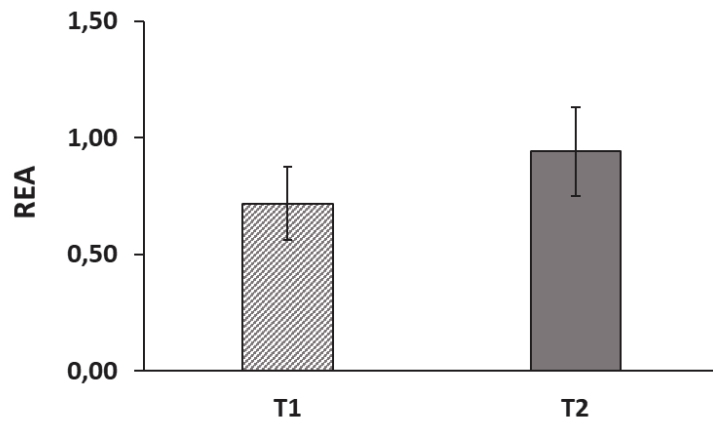


Figure 23. Histogramme des valeurs moyennes des REA à T1 (rayé) et T2 (plein). Les barres d'erreurs représentent l'erreur-type.

### 3.4. Expérience 3 : effet de l'entraînement sur les potentiels évoqués

#### 3.4.1. Principe de l'expérience et hypothèse

En complément des résultats comportementaux encourageants pour des dyslexiques adultes, obtenus avec l'entraînement perceptif dichotique pour améliorer le traitement du voisement des occlusives françaises, nous avons testé auprès de 12 des 27 participants de l'Expérience 2, si des modifications au niveau cérébral étaient observables lors de l'écoute passive de différence de voisement pour l'une des paires de consonnes entraînées. L'Expérience 3 avait pour but de mesurer un possible effet de l'entraînement dichotique en potentiels évoqués chez des adultes francophones atteints de dyslexie.

Les hypothèses suivantes ont été formulées :

A T1 :

- 1) Les participants devraient présenter une MMN avant l'entraînement pour l'écoute passive d'une consonne occlusive vélaire déviante sur le plan du voisement dans le contexte d'une série d'occlusives vélaire standards présentant l'autre valeur de voisement. La topographie de cette MMN devrait être typique de cette onde, enregistrée davantage avec les électrodes posées sur la partie antérieure du scalp plutôt que sur la partie postérieure.

A T2 :

- 1) L'amplitude de la MMN devrait croître lors de l'écoute passive de la consonne déviant en voisement.



- 2) La MMN devrait être plus ample sur les électrodes gauches pour les sourdes et sur les électrodes droites pour les sonores, conformément aux caractéristiques de durée des VOTs en français et à l'asymétrie hémisphérique supposée par le cadre théorique AST pour la perception d'indices temporels courts/longs.
- 3) Une LDN devrait apparaître ou son amplitude croître.
- 4) L'amplitude de l'onde P3 pour les stimuli déviants, devrait augmenter, témoignant elle aussi de l'augmentation de la sensibilité à la différence phonologique perçue.
- 5) L'onde P2 devrait diminuer pour les stimuli déviants, témoignant d'un filtrage précoce des stimuli qui, bien que surprenants par leur faible fréquence d'écoute, ont trouvé leur place dans le système phonologique plus stable et ne doivent donc plus être analysés en profondeur comme des stimuli très surprenants.

### **3.4.2. Méthode**

#### **3.4.2.1. Participants**

Douze participants dyslexiques (DYS) (8 femmes, M = 21 ans 6 mois ; ET = 2 ans 2 mois) de l'Expérience 2 (Annexe 3) ont accepté d'effectuer, en plus des épreuves de PC, une épreuve d'écoute passive dans des conditions d'enregistrement en EEG avant et après l'entraînement.

Seize participants normo-lecteurs (NL) (13 femmes, M = 21 ans et 6 mois ; ET = 2 ans 9 mois) ont été inclus dans cette expérience afin de comparer l'amplitude moyenne des ondes étudiées des deux groupes à T1. Ces participants étaient ceux qui avaient ensuite réalisé un entraînement dichotique (Chapitre 5, Expérience 6) (Annexe 3).

#### **3.4.2.2. Matériel et équipement**

Les deux stimuli /gi/ and /ki/ français étaient des syllabes produites par le même locuteur (un homme) bilingue natif anglais/français que pour les syllabes naturelles ayant servi de base aux continua dans les épreuves de PC des Expériences 1 et 2. Le /gi/ avait un VOT négatif de -60 ms et le /ki/, un VOT positif de +20 ms.

Un total de 506 items a été présenté dans un paradigme oddball passif (85% d'items standards, 15% d'items déviants) par le logiciel Presentation® (v. 14.9, [www.neurobs.com](http://www.neurobs.com)) à travers un casque audio (Beyerdynamics DT 770 Pro 250#). Un bonnet comportant 32 électrodes était posé (Electro-Cap International, INC., selon le système international 10–20), et le système Biosemi ActiveTwo EEG system—Version 5.36 (02-06-2006) fonctionnant à une fréquence d'échantillonnage de 512 Hz et une résolution de 24-bit a été utilisé. Sept

électrodes actives externes étaient placées sur le visage : quatre autour des yeux afin d'enregistrer les clignements et mouvements oculaires sur les deux axes (2 Vertical Electro-OculoGraphy (VEOG) pour les électro-oculogrammes verticaux ; 2 HEOG pour les électro-oculogrammes horizontaux), une électrode sur le nez, et deux électrodes sur les mastoïdes droite et gauche.

#### **3.4.2.3. Procédure d'évaluation de l'effet d'entraînement, en EEG**

Les mesures EEG et ERP ont été réalisées à T1 et T2 (i.e. avant et après l'entraînement dichotique, Figure 19). Chaque session était divisée en deux blocs, de sorte que /gi/ était le stimulus standard dans l'un d'eux, et /ki/ le stimulus standard dans l'autre, dans des ordres contrebalancés entre les participants. Les stimuli étaient présentés de manière binaurale. Ils ont été pseudo-randomisés pour s'assurer qu'au moins quatre standards étaient présentés avant chaque déviant. Le signal correspondant aux standards qui suivait un déviant était systématiquement exclu de l'analyse. Le participant était assis dans un fauteuil confortable, dans une salle insonorisée et anéchoïque. Il était seul dans le box, mais un dispositif permettait à l'expérimentateur de veiller visuellement à son bien-être pendant l'expérience. Il regardait un film silencieux non sous-titré et avait pour instruction de prêter attention au film, et non aux sons diffusés dans le casque. L'expérience était divisée en deux blocs de 10 minutes – un bloc pour chaque déviant possible –, et un temps de repos était proposé entre les blocs. L'enregistrement durait au total 25 minutes.

#### **3.4.2.4. Enregistrement électrophysiologique**

L'EEG a été enregistré en continu. Les électrodes ont été ajustées une à une pour obtenir une impédance stable inférieure à 20 mV. Les données ont été analysées hors ligne avec le module ERPLAB (v. 1.0.0.42) de la boîte à outils EEGLAB (v. 9.0.2.3.b) développée pour Matlab (v. 7.0.9.R2009b). Les données ont ensuite été filtrées avec un filtre passe-bande de 0.1 à 30 Hz (12 dB/oct). Elles ont été découpées de - 200 à 800 ms par rapport au début du stimulus. Pour la correction par la ligne de base, la tension moyenne dans l'intervalle de 200 ms avant le stimulus a été soustraite de chaque période. Un rejet automatique des artefacts a été pratiqué, excluant les époques dont le tracé s'écartait de  $\pm 100 \mu\text{V}$ . Pour chaque condition, les signaux ont été moyennés pour chaque participant, et référencés aux deux mastoïdes. Enfin, le signal des réponses aux stimuli standards a été soustrait des réponses aux stimuli déviants dans chacun des 2 blocs expérimentaux pour extraire les réponses MMN.

### 3.4.2.5. Procédure d'entraînement

L'entraînement réalisé par les participants était celui décrit dans l'Expérience 2 partie 3.3.2.4 de ce chapitre.

### 3.4.3. Analyses de données

La MMN a été calculée comme la réponse au stimulus déviant moins la réponse au stimulus standard. Les fenêtres d'analyse ont été choisies séparément pour les 2 blocs et selon l'étape (pré- et post-entraînement) à partir des courbes de moyennes calculées sur l'ensemble des participants. Le pic a pu être enregistré pour les deux déviants dans une même fenêtre de 120-220 ms. Les électrodes F3, F4, Fz, FC1, FC2, C3, C4, Cz, CP1, CP2, P3, P4, Pz ont été sélectionnées pour les analyses statistiques guidées par les hypothèses. Les analyses ont porté sur l'amplitude moyenne de l'onde.

Les analyses ont été réalisées en trois temps. Le premier set d'analyses a permis de comparer l'amplitude moyenne de la MMN des dyslexiques (à T1) à celle des normo-lecteurs (les deux groupes ayant déjà été comparés au niveau comportemental, Chapitre 3 partie 3.3.4 par un test *t* de Student bilatéral non-apparié avec correction de Bonferroni. Pour le deuxième set d'analyses, un test *t* de Student bilatéral apparié a d'abord été effectué pour déterminer si la MMN différait significativement de zéro, à T1 pour les deux groupes, à T2 pour le groupe DYS (seul groupe entraîné). Puis, l'amplitude moyenne de la MMN a été soumise à deux ANOVA à mesures répétées distinctes sur les données du groupe DYS :

- L'une portait sur l'ensemble des données, selon le plan 2 (session) × 5 (position) × 2 (voisement), avec session, position, voisement comme facteurs intra-sujet,
- L'autre portait sur le signal recueilli sur les seules électrodes posées à gauche ou à droite, pour évaluer l'effet de latéralité, selon le plan × 2 (session) × 5 (position) × 2 (voisement) × 2 (latéralité), avec session, position, voisement et latéralité comme facteurs intra-sujet.

La troisième analyse permettait de répondre à une question complémentaire, qui transparaît dans la littérature scientifique sur la présence ou non d'une LDN en cas de dyslexie comme l'avancent Jakoby *et al.* (2011) et Shulte-Körne *et al.* (1999b). Nous avons pour cela réalisé trois types d'analyses : 1) des tests *t* de Student à T1 pour comparer l'amplitude de la LDN à zéro et tester ainsi sa présence ou son absence, 2) des comparaisons entre T1 et T2 pour mesurer le possible effet de l'entraînement sur cette onde ; 3) des corrélations entre l'amplitude moyenne de la MMN et de la LDN, toutes données confondues.

Des analyses complémentaires ont été réalisées sur l'amplitude moyenne de l'onde P3 dans la fenêtre 200-400 ms et de l'onde LDN dans la fenêtre 300-800 ms.

#### 3.4.4. Résultats

**Comparaison de la MMN entre les deux groupes à T1.** Dans le groupe NL, il y avait sur l'ensemble des électrodes une MMN différente de zéro pour le déviant /gi/ ( $t(15) = 2.85, p = .012, d = 1.01$ ), mais pas pour le déviant /ki/ ( $t(15) = 0.49, p = .632, d = .17$ ). La comparaison de ces MMN avec celles du groupe DYS ne montrait pas de différence générale d'amplitude entre les deux groupes à T1 que ce soit pour le déviant /gi/ ( $t(26.00) = 0.41, p = .687, d = .15$ ) ou le déviant /ki/ ( $t(24.93) = 1.59, p = .124, d = .60$ ).

**Effet de l'entraînement dans le groupe DYS sur la MMN.** Les tracés correspondant à l'amplitude moyenne du signal (condition déviant moins condition standard) sont présentés dans la figure 24, par deux courbes distinctes lorsque le déviant était /gi/ et /ki/, séparément à T1 et à T2, sur les électrodes C3, Cz et C4. Les tests  $t$  de Student ont révélé que l'amplitude moyenne sur l'ensemble des électrodes était significativement différente de zéro et négative avant l'entraînement pour /ki/ ( $t(11) = 2.74, p = .019, d = 1.12$ ), mais pas pour /gi/ ( $t(11) = 1.73, p = .111, d = .71$ ) dans la fenêtre où se produisait la MMN. Même si l'amplitude moyenne de ces deux MMN n'était pas différente de celles des NL à T1 d'après l'analyse présentée dans le paragraphe précédent, le groupe DYS ne présentait pas de véritable MMN pour /gi/ avant l'entraînement, et ce phénomène le distinguait du groupe NL. De même, la MMN pour /ki/ à T1 chez les DYS était atypique par rapport au groupe NL.

A T2, une inversion s'est produite dans le groupe DYS. La MMN n'était plus significativement différente de zéro pour /ki/ ( $t(11) = 0.69, p = .504, d = .28$ ), mais le devenait pour /gi/ avec un effet de grande taille ( $t(11) = 3.59, p = .004, d = 1.46$ ).

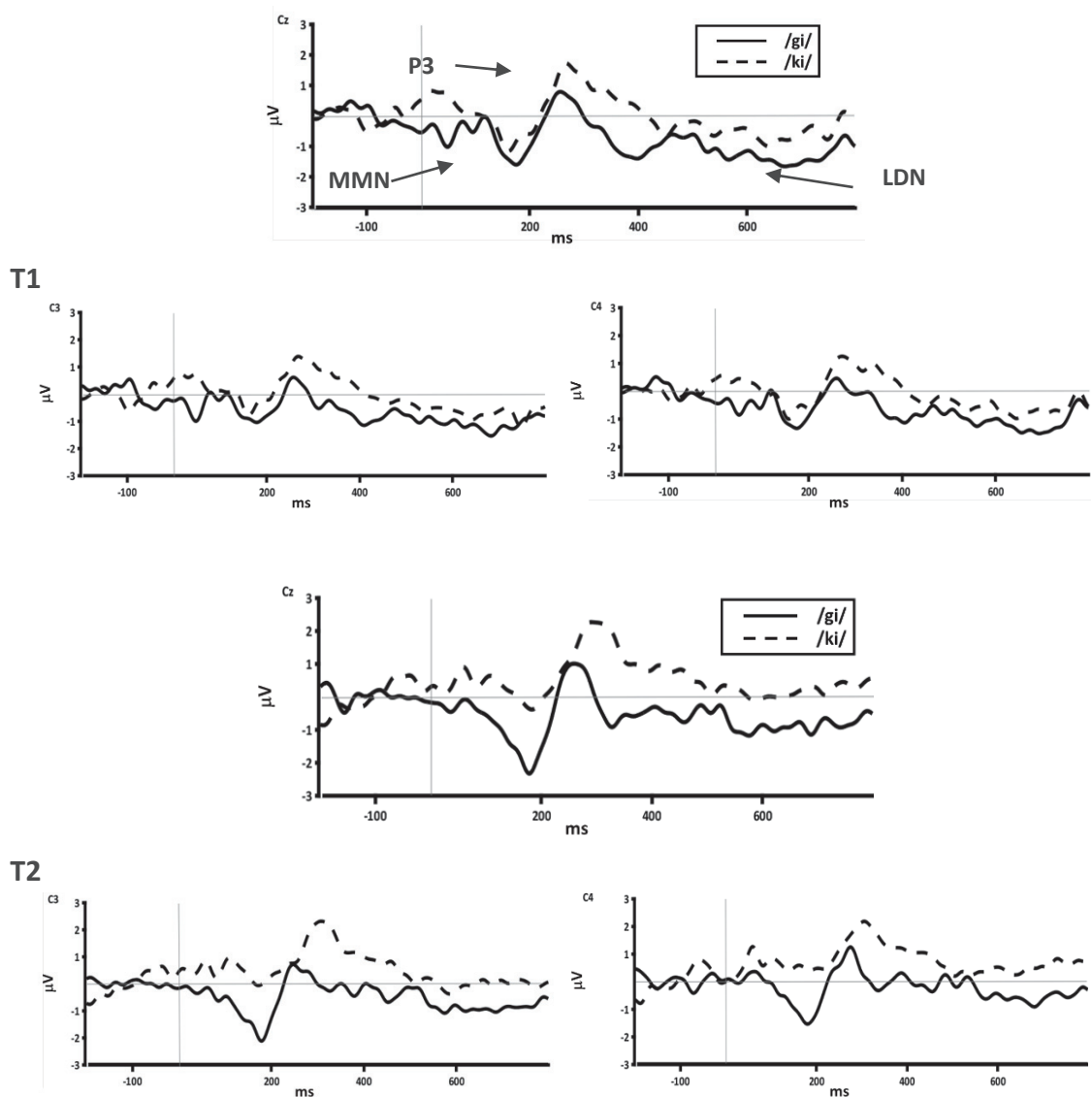


Figure 24. Grande moyenne ( $\mu\text{V}$ ) de la différence Déviant moins Standard, avant l'entraînement (T1, haut) et après l'entraînement en français (T2, bas), pour les déviants /gi/ (plein) ou /ki/ (pointillés) en français pour les électrodes Cz (haut), C3 (gauche) et C4 (droite).

Les résultats de la première ANOVA dans le groupe DYS indiquaient un effet principal du Voisement approchant le seuil de significativité ( $F(1,11) = 4.44, p = .059, \eta_p^2 = .29$ ). La MMN était en moyenne plus grande pour le déviant /gi/ ( $-1,12 \mu\text{V}$ ) que pour /ki/ ( $-0,19 \mu\text{V}$ ) lorsque les deux étapes (T1 et T2) étaient prises en compte (Fig. 25).

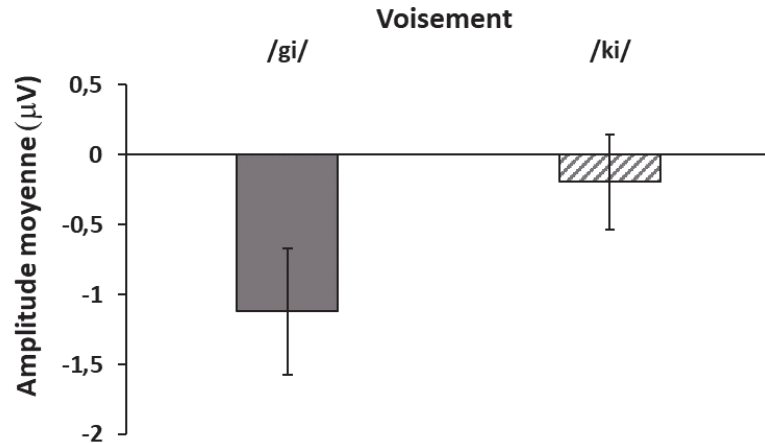


Figure. 25. Amplitude moyenne (µV) pour les déviants /gi/ (gauche) et /ki/ anglais (droite). Les barres d'erreurs représentent l'erreur type.

Comme illustré par la Figure 26, l'interaction Voisement × Etape était significative ( $F(1,11) = 5.12, p = .045, \eta_p^2 = .32$ ). L'observation de cette figure révèle que l'interaction s'expliquait par des effets de l'Etape, certes de petite taille, mais en directions opposées pour /gi/ et /ki/. Pour le déviant /ki/, la taille de la MMN diminuait à la suite de l'entraînement ( $t(11) = 2.24, p = .047, d = .85$ ). Pour le déviant /gi/, l'inverse se dessinait. L'amplitude de la MMN augmentait pour /gi/ à la suite de l'entraînement, non significativement, mais avec une taille d'effet non négligeable ( $t(11) = 1.46, p = .172, d = .49$ ).

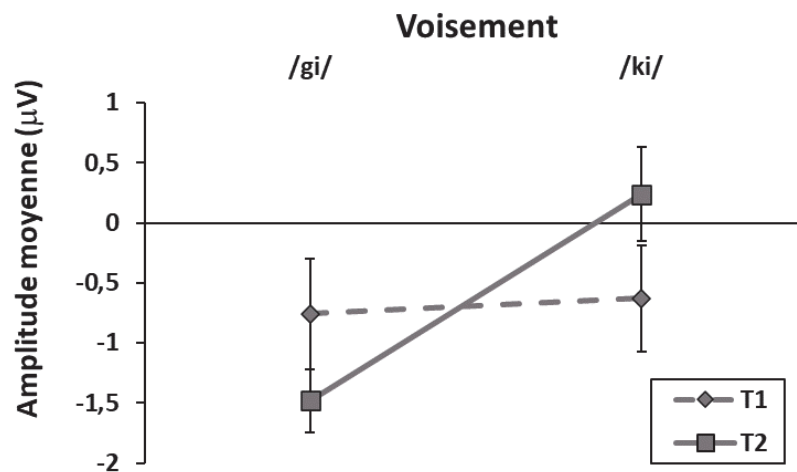


Figure. 26. Amplitude moyenne (µV) pour les déviants /gi/ (gauche) et /ki/ (droite) avant l'entraînement (T1, pointillés) et après l'entraînement (T2, trait plein). Les barres d'erreurs représentent l'erreur-type.

Comme illustré par la figure 27, l'interaction Voisement × Position ( $F(4,44) = 3.37, p = .017, \eta_p^2 = .23$ ) montre une MMN mieux captée sur les électrodes antérieures lorsque le déviant était /gi/, toutes étapes confondues. En effet, l'amplitude de la MMN ne différait pas significativement entre les positions frontale et frontocentrale où elle était la plus élevée ( $t(11) = 0.48, p = .640, d = .37$ ). La MMN la plus ample, enregistrée en position

frontocentrale, dépassait celles recueillies sur les électrodes plus postérieures : centrales ( $t(11) = -2.33, p = .040, d = .14$ ), centropariétales ( $t(11) = -2.07, p = .063, d = .22$ ) et pariétales ( $t(11) = -2.50, p = .030, d = .41$ ).

Pour le déviant /ki/, même si la Figure 27 semble montrer une MMN mieux captée sur les électrodes postérieures, il n’y avait pas de différence entre les positions. Même la différence de moyennes la plus grande, entre les électrodes pariétales et centrales, n’était pas significative ( $t(11) = 1.94, p = .078, d = .38$ ) et la variabilité inter-individuelle observée figure 27 l’explique.

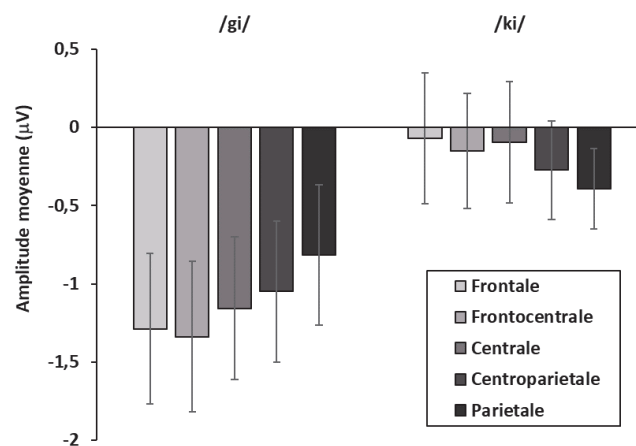


Figure. 27. Amplitude Moyenne ( $\mu\text{V}$ ) selon les positions des électrodes pour le déviant /gi/ (à gauche) et le déviant /ki/ (à droite). Les barres d'erreurs représentent les erreurs-type.

Il n’y avait pas d’effet principal de la Latéralité des électrodes ( $F(1,11) = 0.21, p = .657, \eta_p^2 = .02$ ), ni d’interaction Latéralité  $\times$  Voisement ( $F(1,11) = 2.37, p = .152, \eta_p^2 = .17$ ).

**Effet de l’entraînement dans le groupe DYS sur l’onde P3.** Afin d’investiguer plus amplement le phénomène de diminution de la MMN sur /ki/ entre T1 et T2, des analyses complémentaires ont été réalisées sur l’onde qui la suit, la P3. Le test  $t$  de Student a montré une augmentation d’amplitude de la P3 entre T1 et T2 approchant le seuil de significativité pour le déviant /ki/, modification de taille moyenne ( $t(11) = 2.11, p = .058, d = .72$ ). Il n’y avait pas de changement significatif entre T1 et T2 sur cette onde pour le déviant /gi/ ( $t(11) = 0.02, p = .985, d = .01$ ).

**Effet de l’entraînement dans le groupe DYS sur la LDN.** Comme l’illustre la figure 22, les tests  $t$  de Student ont révélé que l’amplitude moyenne de la LDN était significativement différente de zéro et négative avant l’entraînement pour /gi/ ( $t(11) = 2.22, p = .049, d = .91$ ), mais pas pour /ki/ ( $t(11) = 1.19, p = .261, d = .48$ ) dans la fenêtre où elle se produisait.

Il n'y a pas eu de changement significatif de l'amplitude moyenne de la LDN à T2 que ce soit pour le déviant /gi/ ( $t(11) = 0.59, p = .569, d = .22$ ) ou le déviant /ki/ déviant ( $t(11) = 1.78, p = .102, d = .76$ ).

L'étude des corrélations entre l'amplitude moyenne de la MMN et celle de la LDN a révélé que les deux mesures étaient moyennement corrélées pour /gi/ à T1 ( $r(10) = .73, p = .005$ ) et à T2 ( $r(10) = .63, p = .020$ ). Pour le déviant /ki/, elles étaient faiblement corrélées à T1 ( $r(10) = .46, p = .111$ ) et moyennement à T2 ( $r(10) = .64, p = .018$ ).

### **3.5. Discussion sur l'effet d'un entraînement en français chez des adultes dyslexiques (Expériences 2 et 3)**

L'objectif de ce chapitre était d'évaluer l'effet d'un entraînement à la perception du voisement en français chez des adultes dyslexiques. Ce futur nouvel outil a été créé avec le projet de le proposer à des orthophonistes lors de la prise en charge d'adultes dyslexiques. En effet, il existe aujourd'hui peu d'outils pour la rééducation des adultes et aucun à notre connaissance visant la phonologie française pour ce public. L'entraînement était réalisé pendant 5 jours consécutifs à raison de 30 minutes par jour, ce qui en fait un dispositif bref, ciblant le traitement phonologique des consonnes occlusives françaises.

La discussion de ce chapitre reviendra sur les différences entre les adultes normo-lecteurs et dyslexiques au niveau de la PC, d'après les mesures comportementales et électrophysiologiques recueillies. Elle permettra de commenter ensuite l'effet de l'entraînement sur la PC et sur des mesures en ERP. Nous terminerons en évoquant les qualités et les limites de cette intervention.

#### **3.5.1. Différences phonologiques entre adultes dyslexiques et normo-lecteurs**

Les déficits de PC sont répertoriés en cas de dyslexie, et particulièrement chez l'enfant (Godfrey *et al.*, 1981 ; Sprenger-Charolles & Serniclaes, 2004 ; Werker & Tees, 1987). En effet, il semblerait qu'à l'âge adulte, les troubles métaphonologiques persistent plus durement que les déficits phonologiques (Habib, 2002). Cependant, des recherches ont montré que les mesures comportementales ne sont pas toujours sensibles, dans ce domaine, à des anomalies subtiles, pourtant persistantes, car détectées en électrophysiologie (Noordenbos *et al.*, 2013). Les données recueillies dans l'Expérience 2 confirment la difficulté à recueillir des indices de troubles phonologiques dans un groupe d'adultes dyslexiques, à partir de données



comportementales émanant pourtant de deux types d'épreuves (identification et discrimination) et de deux paires de consonnes (occlusives alvéolaires et vélares).

### **3.5.1.1. Particularité des adultes dyslexiques en perception du voisement selon les données comportementales**

La seule différence significative observée à travers les données comportementales entre les deux groupes est un seuil d'identification positionné différemment pour la paire /g-k/, alors que d'autres mesures ont été explorées.

Ce seuil, qui donne une indication de la frontière catégorielle dans le système phonologique des participants, est situé au VOT 16.12 ms dans le groupe d'adultes normo-lecteurs. Bien qu'un peu tardive, cette frontière reste conforme à ce qui est classiquement décrit en français (entre les VOTs 0 et +10 ms). Cependant, le groupe d'adultes dyslexiques place celle-ci au VOT 23.75 ms. Cette position est anormalement tardive pour le français. Cela montre que l'organisation des représentations phonologiques de certains phonèmes, et la représentation mentale des différences de voisement qui les opposent, présentent une anomalie dans ce groupe. Sur le plan phonétique, l'anomalie est au moins en partie liée à une perception ou à une interprétation incorrecte d'indices de durée. Le sens de la déviation du seuil d'identification observée dans le groupe d'adultes dyslexiques par rapport à la norme suggère en effet que des VOTs positifs très courts, entre 0 et 20 ms, sont insuffisamment considérés comme des indices pertinents pour identifier une consonne occlusive française comme sourde. Comme il s'agit d'un indice très bref, la difficulté à le percevoir pourrait expliquer ce phénomène. Ce résultat est alors cohérent avec l'hypothèse de Tallal (1980) qui, dans une théorie proposait qu'un déficit d'audition de haut niveau (i.e. audition centrale), évoquait les difficultés majeures des personnes dyslexiques (et TDL) à percevoir des traits phonétiques brefs. Le succès des techniques de remédiation basées sur cette hypothèse, et consistant à entraîner les enfants à percevoir des différences de voisement (e.g. /apsa/ - /absa/) avec des stimuli de durée augmentée, apportait des arguments pour cette hypothèse. D'autres données soutiennent l'hypothèse de la difficulté à traiter les informations temporellement très courtes en cas de dyslexie (Tallal *et al.*, 1996). Pourtant, des théories récentes (Goswami, 2011) et des données mettent en avant les difficultés des personnes dyslexiques à traiter des unités linguistiques temporellement longues (Soroli *et al.*, 2010), à ajuster leurs rythmes cérébraux à des rythmes lents (Lizarazu *et al.*, 2021). Des anomalies de l'asymétrie hémisphérique de la distribution et de la synchronisation des oscillations cérébrales lentes dans le cortex auditif (Hämäläinen *et al.*, 2012) ont aussi été trouvées et ne doivent pas faire

oublier que les évènements brefs sont souvent mal traités dans cette pathologie particulièrement dues à une faible latéralisation gauche pour ces indices (Dushanova *et al.*, 2020 ; Lehongre *et al.*, 2013). Ces dernières difficultés sont suffisamment résistantes à la maturité et à la rééducation orthophonique pour transparaître dans des résultats d'expériences chez des adultes dyslexiques, comme ici.

Les travaux réalisés par Serniclaes et ses collègues ont aussi contribué à un certain désintérêt pour un déficit possible de traitement d'indices acoustiques courts chez les enfants dyslexiques. Leur hypothèse sur la perception allophonique et les expériences qui soutiennent celle-ci montrent en effet leur sensibilité exagérée à des aspects phonétiques présentés comme des détails auxquels ils ne devraient plus réagir compte tenu du système phonologique de leur langue (Serniclaes *et al.*, 2004). Cela ne signifie cependant pas que les indices acoustiques en question soient les plus brefs dans la parole. Ainsi, d'après ces travaux, les VOTs auxquels les enfants dyslexiques restent anormalement sensibles sont de durée -30 ms et + 30 ms (Bogliotti *et al.*, 2002 ; Serniclaes *et al.*, 2004, 2021). La difficulté à traiter des indices acoustiques encore plus rapides, comme le court VOT de moins de 30 ms après le bruit d'explosion dans les occlusives sources françaises a été mise en évidence par le seuil situé au VOT 23.75 ms chez les dyslexiques de notre groupe. Cela pourrait aussi contribuer à expliquer la difficulté des personnes dyslexiques à élaborer une représentation phonologique vraiment adaptée au français pour la différence de voisement. La conclusion de Serniclaes *et al.* (2021) va d'ailleurs dans ce sens, et permet de faire le lien avec la dynamique oscillatoire cérébrale. Tout en respectant le cadre proposé par le TSF (Goswami, 2011), les auteurs s'éloignent d'une explication des difficultés de traitement phonémique en cas de dyslexie essentiellement basée sur un ajustement imparfait des oscillations lentes (notamment thêta) avec les oscillations plus rapides (gamma) responsables de la délimitation des segments phonémiques. Serniclaes *et al.* font l'hypothèse que des anomalies du couplage entre des oscillations rapides (influence de bêta sur gamma) empêcheraient les oscillations bêta de regrouper les phases des oscillations gamma de manière à inhiber une segmentation hyper-rapide. Seul un groupement des phases des oscillations gamma ( $\approx 33$  Hz) par des oscillations deux fois plus lentes (bêta,  $\approx 17$  Hz) pourrait permettre de saisir des indices tels que le VOT court des occlusives sourdes en français ou en espagnol, et de rester insensible à des différences temporelles situées en dehors des bornes de la frontière catégorielle.

Sur le plan électrophysiologique, les données de l'Expérience 3 vont dans le sens des travaux de Noordenbos *et al.* (2012, 2013). Malgré l'absence de différence entre les

performances des adultes dyslexiques de leur étude et leur paire normo-lecteurs, elles relevaient des particularités liées à la pathologie lors de l'écoute de consonnes voisées labiales et alvéolaires (/b/ et /d/). Chez des participants plus jeunes (Noordenbos *et al.*, 2012), les dyslexiques présentaient une MMN plus petite que celle des contrôles si les sons relevaient de phonèmes différents, et une sensibilité accrue à des variantes d'une seule catégorie. La MMN des adultes dyslexiques ne traduisait quant à elle que cette dernière anomalie – une sensibilité intra-catégorielle trop élevée – et cette perception trop allophonique n'était pas détectable à partir de données comportementales. De même, dans la présente étude, chez des adultes dyslexiques, il y a peu d'indices d'une PC catégorielle imparfaite dans les données comportementales, mais les données en potentiel évoqués montrent une configuration de résultats atypique par rapport aux contrôles.

### **3.5.1.2. Particularité des adultes dyslexiques en perception du voisement selon les données en ERP.**

Avant l'entraînement, les adultes de notre groupe dyslexique présentaient un indice électrophysiologique d'une certaine sensibilité pré-attentionnelle à une différence de voisement. Il n'y avait pas de différence générale d'amplitude de la MMN toutes conditions confondues de l'épreuve d'écoute passive entre les deux groupes. Chez les adultes dyslexiques, la MMN était significative pour le déviant /ki/ parmi des /gi/, mais pas pour le déviant /gi/ parmi des /ki/.

Bien que partielle, cette sensibilité est un point positif pour le système phonologique des adultes dyslexiques testés ici. Pour ce type de population, une absence de MMN dans des conditions similaires a été décrite dans certains travaux (Shulte-Körne *et al.*, 1998, 1999, 2001), y compris pour des stimuli acoustiques sonores non-verbaux (Kujala *et al.*, 2000). Cela suggère que les participants de notre étude n'ont pas de déficit phonologique, ni de déficit plus générale d'audition centrale, aussi sévères que les participants dyslexiques inclus dans d'autres études, et pour une langue différente.

La présence d'une MMN en réponse à une différence phonémique perçue passivement est interprétée par Kraus *et al.* (1996) comme liée à une PC correcte lorsqu'elle est réalisée de manière volontaire, attentionnelle, comme dans les épreuves d'identification et de discrimination permettant de relever des données comportementales en PC. Ces chercheurs ont montré ce lien en cas de dyslexie et nos résultats sont globalement en accord avec ce lien. Nos participants dyslexiques ont montré de bonnes capacités de PC, car ils présentaient peu

de différences avec le groupe contrôle, hormis un seuil d'identification atypique pour les occlusives opposées en voisement. La cohérence entre la présence d'une MMN (Expérience 3) et une PC relativement correcte (Expérience 2) chez des adultes dyslexiques est donc compatible avec l'hypothèse défendue par Kraus *et al.* (1996) D'une réponse MMN liée à la PC.

L'examen plus précis des données de l'Expérience 3 montre cependant que l'absence de différence générale d'amplitude de la MMN entre les adultes dyslexiques et normo-lecteurs cache une distinction plus subtile. La sensibilité pré-attentionnelle au voisement s'exprimait de façon différente dans les deux groupes. En effet, chez les adultes dyslexiques, la MMN était significative pour le déviant /ki/ parmi des standards /gi/, mais pas pour le déviant /gi/ parmi des standards /ki/. Une MMN asymétrique entre les deux conditions était aussi relevée chez les adultes normo-lecteurs, mais dans le sens inverse. Chez les normo-lecteurs, la MMN était en effet significative pour le déviant /gi/ parmi des /ki/, mais pas pour le déviant /ki/ parmi des /gi/. Des analyses complémentaires pourraient aussi être réalisées sur la latence du pic afin d'investiguer plus profondément ces différences. La sensibilité pré-attentionnelle au contraste de voisement, incomplète dans chaque groupe, présentait ainsi une différence qualitative compatible avec l'hypothèse générale du maintien de caractéristiques atypiques dans l'organisation du système phonologique des personnes dyslexiques à l'âge adulte, donc malgré la maturité, l'expérience et la rééducation.

### **3.5.2. Effets de l'entraînement perceptif par stimulation dichotique sur la perception du voisement en français : données comportementales**

L'entraînement perceptif par stimulation dichotique a montré des résultats encourageants sur la PC d'adultes dyslexiques. En effet, l'un des principaux résultats est une augmentation du pic de discrimination pour la paire /g-k/. Après l'entraînement, la frontière catégorielle était perçue avec plus de sensibilité et de façon plus homogène. Cette augmentation du pic de discrimination pour des paires de consonnes dont les membres étaient des exemplaires de phonèmes différents n'affectait pas la qualité du reste de la courbe. Celle-ci devenait même plus lisse pour les VOTs plus courts et plus longs, suggérant une perception plus catégorielle.

Le deuxième argument en faveur d'une amélioration de la PC après l'entraînement est le déplacement du seuil d'identification dans la direction de la norme française, toujours pour la paire /g-k/. Le seuil passait d'un VOT 23.75 ms à un VOT 19.18 ms après celui-ci. D'après l'apparition d'un plateau sur la sigmoïde des réponses, correspondant à un élargissement de

l'asymptote du côté des VOTs positifs, cette amélioration était surtout explicable par une identification moins allophonique des exemplaires du phonème /k/ après l'entraînement. Ce type de modification dans le tracé de l'asymptote est interprété dans d'autres études comme témoignant d'une perception qui devient moins allophonique et plus catégorielle (Serniclaes *et al.*, 2021).

Le troisième résultat est une modification du REA en tâche d'écoute dichotique montrant un renforcement de la latéralisation hémisphérique vers une meilleure prise en charge des indices courts par l'HG après l'entraînement. Les indices acoustiques de lieux, qui sont toujours de courte durée dans les consonnes françaises, sont davantage pris en compte par l'hémisphère le plus adapté à cette durée (i.e. l'HG) après l'entraînement. Ce résultat est le seul à montrer un effet de transfert. Il semble que l'entraînement avec le dispositif dichotique ait fondamentalement entraîné l'HG à traiter les indices courts d'une manière optimale sur le plan de l'asymétrie hémisphérique pour la répartition des fenêtres d'analyse temporelles étroites et larges, et il semblerait plus fondamentalement sur le plan de la dominance des oscillations rapides dans le cortex auditif gauche et des oscillations plus lentes dans ce cortex à droite. Grâce à cette modification, les participants dyslexiques traitent le lieu d'articulation d'une manière plus adaptée sur le plan hémisphérique, ce qui montre un transfert de compétence depuis le traitement du voisement vers celui du lieu d'articulation.

Nous faisons l'hypothèse d'autres effets de transfert, cette fois vers la PC de consonnes opposées par le voisement, mais non-apprises durant l'entraînement. Cette hypothèse n'a pas été validée. La brièveté de l'entraînement a peut-être défavorisé l'établissement d'un apprentissage suffisamment profond permettant une telle généralisation des connaissances acquises.

L'entraînement perceptif par stimulation dichotique semble un bon moyen de renforcer la prise en charge des indices court par l'HG. En effet, le traitement des indices courts est déficitaire en cas de dyslexie à cause notamment d'une faible latéralisation gauche pour le traitement de ces derniers (Dushanova *et al.*, 2020 ; Lehongre *et al.*, 2013), qui nécessitent les fenêtres d'analyse temporelle étroites rendues possibles par les oscillations gamma dominant dans le cortex auditif gauche (Poeppel, 2003 ; Spironelli & Angrilli, 2010). Nous pouvons imaginer que la stimulation répétée de l'hémisphère le plus compétent pour le traitement des durées courtes ait favorisé un renforcement et la synchronisation des rythmes rapides permettant la détection et l'analyse des indices acoustiques brefs. Dans cette étude, l'absence d'un entraînement perceptif binaural chez des adultes dyslexiques empêche d'affirmer que ces

effets sont dus à la stimulation dichotique et non à l'entraînement perceptif seul. Le choix d'utiliser directement le dispositif de stimulation dichotique était guidé par notre souhait de proposer aux participants dyslexiques le programme potentiellement le plus efficace.

### **3.5.3. Effets de l'entraînement perceptif par stimulation dichotique sur la perception du voisement en français : données en potentiels évoqués**

L'entraînement a amélioré certains aspects de la PC des adultes dyslexiques au niveau des performances, et quelques indices en potentiels évoqués montrent aussi des changements sur le plan cérébral. L'interaction significative entre le voisement du déviant (/g/ ou /k/) et l'étape (pré- ou post-entraînement) sur l'amplitude de la MMN chez les dyslexiques s'explique par l'apparition d'une configuration de sensibilité pré-attentionnelle au voisement conforme à celle des normo-lecteurs après l'entraînement, alors qu'elle était inverse avant celui-ci. Avant l'entraînement, chez les adultes dyslexiques, la MMN était significative pour le déviant /ki/ et non pour le déviant /gi/. Après l'entraînement, la MMN était significative seulement pour le déviant /gi/, comme dans le groupe de normo-lecteurs. Prise dans son ensemble, la plus grande conformité de la configuration des résultats à ceux du groupe contrôle pour la MMN suggère que l'entraînement a atténué chez les adultes dyslexiques une particularité phonologique atypique. Le changement d'amplitude de la MMN entre les étapes pré- et post-entraînement traduit chez les adultes dyslexiques une perception plus phonémique de l'occlusive sonore /g/ après l'entraînement. Il s'agit d'une consonne qui, en tant que déviant, suscitait une MMN particulièrement classique, davantage captée par les électrodes antérieurs (Näätänen *et al.*, 1989). Cette augmentation de la MMN témoigne d'une représentation mentale phonémique du /g/.

Par ailleurs, pour le déviant /g/, une corrélation était observée entre l'amplitude de la MMN et celle d'une onde plus tardive (LDN) dès l'étape pré-entraînement. Cette corrélation s'est maintenue après l'entraînement. Elle n'existait pas avant l'entraînement pour le déviant /k/, qui provoquait pourtant une MMN significative. A notre connaissance, une telle corrélation n'est pas documentée dans la littérature. La LDN plus petite chez l'adulte que chez l'enfant (Bishop *et al.*, 2011 ; Cheour *et al.*, 2001) est décrite comme un autre indicateur tardif de la discrimination de phonèmes (Schülte-Korn *et al.*, 1999). Elle reste peu étudiée chez l'adultes, sans doute à cause de l'affaiblissement de son amplitude avec l'âge. Si l'on interprète la corrélation entre la MMN et la LDN comme le signe d'une cohérence des mécanismes cognitifs et cérébraux engagés dans le traitement du phonème, la perception du /g/ après l'entraînement semble mobiliser un ensemble de mécanismes à la fois précoces (pré-

attentionnels) et plus tardifs présentant une certaine homogénéité. Cette interprétation est spéculative, mais la cohérence qu'elle met en avant pour le traitement après l'entraînement pourrait être un signe de qualité. S'intéresser à ce nouveau couple pourrait être une mesure intéressante dans le cadre d'études neurophysiologiques de la dyslexie. Des processus de différents niveaux semblent ainsi s'impliquer dans le traitement de déviations phonémiques, au moins lorsqu'elles sont basées sur des indices temporels. Une hypothèse possible est que cela soit l'une des compensations mises en place par nos participants pour mieux traiter les différences de voisement. Cela serait cohérent avec l'amélioration des performances des adultes dyslexiques en PC après l'entraînement.

Concernant le déviant /k/, une MMN significative lui était associée chez les adultes dyslexiques avant l'entraînement, mais cela tranchait avec l'absence de MMN pour ce déviant dans le groupe contrôle. De plus, l'amplitude de cette MMN n'était pas corrélée à celle de la LDN. Deux aspects des résultats pour la perception du déviant /k/ sont encourageant après l'entraînement. Tout d'abord, la corrélation entre l'amplitude de la MMN et de la LDN est devenue significative, comme pour le déviant /g/, potentiellement signe d'une cohérence croissante entre les mécanismes de perception et d'analyse phonémique impliqués. Ensuite, l'amplitude de l'onde tardive P3 a augmenté pour le déviant /k/ après l'entraînement. Ce phénomène ne se produisait pas pour le déviant /g/. Cela ne traduit pas une augmentation de la sensibilité pré-attentionnel au phonème /k/, mais pourrait refléter l'implication nouvelle de mécanismes de vérification plus tardifs souvent décrit comme associés à l'onde P3 (Horváth *et al.*, 2008). Des traitements de haut niveau, basés sur des connaissances phonologiques, s'engageraient pour approfondir l'analyse. Une augmentation de l'amplitude de la P3 est associée à la détection d'une violation de régularité. Elle est décrite comme le corrélat d'une activation cérébrale provoquée par des connaissances préexistantes, notamment catégorielle, comme les phonèmes (Polich, 2007). Ainsi, la modification de la configuration d'ensemble MMN-P3 après l'entraînement pourrait correspondre à une nouvelle manière d'appréhender ce phonème avec un traitement encore imparfait et maladroit, mais déjà plus phonémique (i.e. catégoriel). Comme pour le /g/, certains aspects des données suggèrent ainsi un traitement plus phonémique de cette occlusive sourde dans la langue maternelle des adultes dyslexiques après l'entraînement.



### 3.5.4. Intérêts et limites d'un entraînement bref à la phonologie française chez des adultes dyslexiques

De nombreuses recherches se sont intéressées à la remédiation du langage en cas de dyslexie. La plupart ont été menées chez l'enfant et peu se sont adressées à l'adulte. Elles se différencient aussi par la durée des prises en charge proposées. Certains protocoles se déroulent sur plusieurs semaines (Flaunacco *et al.*, 2015 ; Stevens *et al.*, 2008), d'autres sur seulement quelques jours (Habib *et al.*, 2016), voire quelques heures (Collet *et al.*, 2012 ; Franceschini *et al.*, 2013). Dans ce type de perspective, même s'il s'agit globalement de phonologie, les chercheurs ne ciblent pas toujours les mêmes compétences : il s'agit parfois de renforcer des capacités attentionnelles utiles dans ce domaine (Gaab *et al.*, 2007 ; Temple *et al.*, 2003), des capacités de prise de conscience d'aspects phonologiques (i.e. métaphonologie) (Bonacina *et al.*, 2015 ; Habib *et al.*, 2016), ou plus directement d'améliorer la perception phonologique ou l'organisation du système de représentation sous-jacent (Collet *et al.*, 2012 ; Tremblay & Kraus, 2002). Dans ce chapitre, nous avons proposé un entraînement à la phonologie, chez l'adulte, pendant 5 jours (30 min /jour, 2h30 au total). Le programme combinait un procédé d'entraînement perceptif classique ayant déjà fait ses preuves et un dispositif plus innovant (la stimulation dichotique). Cette technique visait à renforcer l'association hémisphère-son la plus compétente selon des critères de durée. Chez les adultes dyslexiques, après les exercices et la stimulation pendant 2h30, des modifications de la sensibilité pré-attentionnelle au voisement se sont produites. Leur PC de la paire /g-k/ qui faisait partie des consonnes présentées pendant les exercices s'est aussi améliorée, sans toutefois que la nouvelle compétence se transfère à un couple de consonnes non étudiées pendant le programme. Cela constitue une limite de l'efficacité de l'entraînement.

Cette limite interroge sur la nécessité de prévoir des exercices portant sur les paires d'occlusives opposées en voisement à chacun des lieux d'articulation possibles. Le système phonologique des dyslexiques est peut-être tellement instable qu'il soit nécessaire de remédier à chacune des frontières, séparément pour chaque couple d'occlusive partageant le même lieu d'articulation, afin de reconstruire chaque catégorie une à une. Les limites du succès de l'entraînement conduisent aussi à réfléchir aussi au retentissement de la grande brièveté de la remédiation. Seulement 2h30 d'exercices, même renforcées par un dispositif dichotique, sont sans doute insuffisantes pour pallier les déficits en phonologie subsistant depuis des années chez l'adulte dyslexique. Allonger la durée de la prise en charge permettrait peut-être à des mécanismes perceptifs automatiques de se mettre en place, pour un transfert vers les autres



paires de consonnes. Une autre piste pour améliorer l'entraînement serait de prévoir plus de diversité dans les voix présentées. Une certaine diversité était assurée par la variété des contextes vocaliques des consonnes, mais la haute variabilité des voix utilisées dans d'autres entraînements a contribué à leur succès (Bradlow *et al.*, 1997 ; Cebrian & Carlet, 2014 ; Kangatharan *et al.*, 2021 ; Lively *et al.*, 1993).

## **Chapitre 4. Evaluation de la perception du voisement en anglais en L2 : conception d'outils**

---

D'après les résultats présentés dans les chapitres précédents (chapitres 2 et 3), les épreuves de perception catégorielle et la stimulation dichotique semblent deux méthodes dotées de qualités satisfaisantes pour respectivement évaluer et améliorer la qualité et la stabilité du système phonologique de la L1 de jeunes adultes dyslexiques. À travers le Chapitre 4, nous présenterons un test de perception catégorielle en anglais basé sur les mêmes principes que celui construit en français, pour disposer d'un outil permettant d'évaluer l'effet d'un entraînement au voisement anglais par stimulation dichotique. Une fois que certaines caractéristiques psychométriques du test de perception catégorielle auront été évaluées, le chapitre comparera les performances d'apprenants normo-lecteurs et d'apprenants dyslexiques pour explorer si l'enseignement conventionnel de l'anglais, dont ils ont bénéficié avec le système éducatif français, a eu des retentissements différents chez les uns et les autres.

### **4.1. Difficultés d'apprentissage de l'anglais L2 : le constat**

La pratique de l'anglais et sa compréhension font partie intégrante du quotidien des Français. Cette langue est utilisée au travail par 58% d'entre eux. Chez les jeunes actifs (20 - 44 ans), ce pourcentage monte à 81%<sup>1</sup>. Cependant, le classement de la France au niveau international sur l'échelle des compétences en anglais ne la place pas en bonne position. Selon une étude de 2021 de l'*EF English Proficiency Index*, ce pays est classé 31<sup>ème</sup>/112 au niveau mondial et 24<sup>ème</sup>/35 au niveau européen. Dans le milieu professionnel, avoir une mauvaise maîtrise de l'anglais est aujourd'hui un frein pour 50% des salariés français<sup>2</sup> et 28% reconnaissent que des lacunes dans cette langue leur a fait perdre des opportunités d'emploi<sup>3</sup>.

Plusieurs raisons peuvent expliquer ces difficultés d'apprentissage. En France, l'anglais fait l'objet d'un enseignement à part entière classiquement dès 11 ans. Cet apprentissage pourra amener dans certains cas à un bilinguisme qualifié de tardif (le bilinguisme précoce faisant référence à l'apprentissage simultané et spontané des deux langues lors de la petite enfance) (Kail & Fayol, 2000). Le bilinguisme tardif permet de maîtriser cette L2, mais cela impose une pédagogie active favorisant l'oral, des retours individualisés, l'immersion dans un pays anglophone, ainsi que des échanges avec des

---

<sup>1</sup> ABA English

<sup>2</sup> Sondage OpinionWay pour le jobbard Monster

<sup>3</sup> ABA English

anglophones permettant de diminuer la distance sociolinguistique (Jordan, 2004). Les effectifs élevés en classe et des freins économiques restreignent ces conditions idéales. Selon Rojas (2016), la peur de l'échec réduirait aussi les prises de parole durant les leçons d'anglais, empêchant l'élève de tirer profit des corrections des enseignants. Cependant, de nouvelles mesures ont été prises par le gouvernement français pour introduire l'apprentissage des langues plus tôt dans la vie scolaire<sup>4</sup>. Dès la moyenne section de l'école maternelle, l'éveil à la diversité linguistique offre un premier contact pour les jeunes élèves avec les langues vivantes étrangères. Une langue vivante est enseignée une heure et demie par semaine aux élèves de l'école élémentaire et est obligatoire dès le CP depuis le 8 juillet 2013. Cet enseignement est ensuite renforcé dès le collège. Cependant, même si des efforts sont réalisés, nous sommes encore loin d'un apprentissage optimal des langues étrangères.

#### **4.2. Les modèles de l'apprentissage d'une seconde langue**

Si l'on se place au niveau cognitif et plus seulement pédagogique, il existe aujourd'hui deux modèles prédominant dans les recherches scientifiques sur l'apprentissage d'une L2, qui se complètent et peuvent être appliqués à ce travail de recherche : le Speech Learning Model (SLM), aujourd'hui révisé (SLM-r), (Flege, 1995a ; Flege & Bohn, 2021) et le Perceptual Assimilation Model (PAM) (Best, 1995 ; Best *et al.*, 2001).

Le modèle SLM fait l'hypothèse selon laquelle les systèmes phonétiques utilisés dans la production et la perception des voyelles et des consonnes restent adaptatifs tout au long de la vie, et qu'ils se réorganisent en réponse aux sons rencontrés dans une L2. Cette adaptation se ferait, soit par l'ajout de nouveaux systèmes phonétiques ou de nouvelles catégories, soit par la modification des anciennes. Plus précisément, lors de l'apprentissage d'une L2, les apprenants peuvent avoir des difficultés à distinguer une différence phonétique entre deux sons de la L2 ou entre un son de la L1 et un son de la L2 parce qu'ils se recouvrent fortement au niveau phonétique. Un son de la L2 serait alors assimilé à la catégorie phonologique du son le plus proche dans la L1. Le son de la L2 n'est alors pas considéré comme un phonème à part entière mais comme un allophone distant d'un phonème de la L1, c'est-à-dire un exemplaire assez éloigné du prototype, mais appartenant toujours à la catégorie phonémique. Selon Flege (1995a), cette erreur de classification serait due au fait que le système phonologique de la L1 filtre certaines caractéristiques ou propriétés du son de la L2. Ce filtrage empêcherait la

---

<sup>4</sup> <https://www.education.gouv.fr/les-langues-vivantes-etrangeres-et-regionales-11249>

perception complète et le traitement approfondi du son, limitant ainsi l'extraction de propriétés importantes pour sa classification. Le traitement serait alors réalisé uniquement selon des critères de surface et amènerait l'auditeur à le classer de manière erronée dans une catégorie délimitée de sa L1. Un son d'une L2 totalement différent de tout son de la L1 aurait plus de chances d'être appris, bien perçu et correctement produit.

Le modèle SLM retient ce mécanisme d'assimilation, mais soutient aussi une autre notion importante : la possibilité de mouvement des catégories phonémiques. L'apprentissage tardif d'une L2 pourrait générer des modifications de l'espace des représentations phonologiques. Les systèmes phonétiques de la L1 et de la L2 coexisteraient dans un seul et vaste espace phonologique. Par exemple, si un segment phonétique de la L2 est très différent de ce qui existe dans la L1, il n'est pas assimilé à une catégorie existante et peut devenir une nouvelle catégorie. C'est le cas du [y] français souvent correctement perçu et produit par des apprenants américains qui ne l'ont pas dans leur L1, et qui lui donnerait le statut de nouvelle catégorie.

Flege et Bohn révisent le modèle SLM qui devient SLM-r en 2021. Ce dernier se donne toujours pour objectif de rendre compte de la façon dont les systèmes phonétiques se réorganisent au cours de la vie en réponse à l'exposition à une L2. Cet objectif étant assez large, seuls les points pertinents dans le cadre de cette recherche seront évoqués. L'un des changements majeurs entre les deux versions du modèle est que la première version partait du postulat que plus l'apprenant était âgé, plus les mécanismes nécessaires à l'apprentissage d'une L2 étaient limités, à cause du système phonologique fermement établi de la L1. C'est la structure du système phonémique qui rendrait la perception plus rigide (Ioup, 2008), et les représentations phonémiques encore instables de l'enfant et même de l'adolescent dans la L1 seraient moins figées que celles des adultes pour permettre d'acquérir des phonèmes en L2 (Baker *et al.*, 2008). Il proposait cela comme une alternative à *l'hypothèse de la période critique* pour l'apprentissage tardif des langues qui explique qu'il est plus difficile d'apprendre une langue en dehors des périodes critiques (période de la vie où l'apprentissage du langage est optimal) à cause d'une réduction de la neuroplasticité liée à l'âge. Dans son modèle révisé, la création de nouvelles catégories en L2 dépend moins de l'âge que de la ressemblance du son de la L2 avec des sons de la L1, de la précision du système de représentation phonétique de la L1, et de la qualité de l'input de la L2. Une autre notion intéressante apportée par ce modèle est que les mécanismes de formation de catégories perceptives en L1 restent intacts et accessibles pour la L2. Même si SLM-r reconnaît que

l'apprentissage de la L2 ne peut jamais être parfait à cause des interférences issues la L1, il propose tout de même que les catégories phonétiques de la L1 et de la L2 restent malléables toute la durée de la vie. Les mécanismes d'acquisition de la L1 ne disparaîtraient pas avec l'âge, ils seraient encore en place chez l'adulte, et c'est le mécanisme d'assimilation qui ne ferait que les inhiber, sans les remplacer totalement. De nombreuses données expérimentales montrent aujourd'hui que des années d'expériences avec un système phonologique de la L1 ne modifie pas les capacités sensorielles de bas niveau impliquées dans la discrimination phonétique (pour une revue, voir Strange & Shafer, 2008). C'est plutôt la surutilisation de processus de sélection de haut niveau, intégrant certains indices acoustiques plutôt que d'autres et façonnés par la L1, qui s'interpose de façon top-down. De ce fait, l'apprentissage tardif et les progrès des adultes en L2 ne sont pas exclus. Par exemple, avec l'augmentation de l'expérience en français L2, des apprenants anglophones produisent le /t/ français comme en anglais (avec une longue aspiration, ou long VOT positif) pour les moins expérimentés, mais avec un VOT positif moins long pour les plus expérimentés, et un VOT aussi court que celui des Français pour des anglophones résidents depuis longtemps à Paris (Flege, 1987). Des expériences d'entraînement à la différence de voisement en anglais par des locuteurs adultes chinois ayant débouché sur de réels progrès à partir de la répétition de tâches d'identification et de discrimination ont aussi convaincu l'auteur dans cette voie (Flege, 1995b). L'étude détaillée de la qualité des phonèmes acquis par des francophones ayant un bon niveau en anglais L2 montre que l'acquisition de nouveaux prototypes non-natifs est possible, et que c'est plutôt la qualité et la position précise des frontières catégorielles entre les nouveaux phonèmes qui restent fragile (Heidlmayr *et al.* 2021). Flege est un bon ambassadeur pour les entraînements mis en place à destination des apprenants tardifs, et il a montré à plusieurs reprises leur efficacité (pour une revue, voir Bradlow, 2008). Nos hypothèses sont en accord avec le SLM-r car il défend l'idée d'une possible réorganisation des catégories phonétiques à l'âge adulte lors de l'apprentissage d'une L2.

Alors que SLM-r accepte aussi bien l'assimilation d'un phonème de la L2 à une catégorie de la L1 que la possibilité de création de nouvelles catégories phonémiques spécifiques à la L2, le modèle PAM (Best, 1995) n'envisage que la première solution, pour laquelle il détaille davantage un éventail de mécanismes possibles. Il propose un point de vue légèrement différent, plus cristallisé autour du principe d'assimilation et moins ouvert à la possibilité pour un apprenant tardif de développer des catégories phonémiques propres à la L2. Il décrit cependant des degrés plus ou moins forts d'assimilation d'un son de la L2 à une

catégorie de la L1. C'est seulement dans une version plus récente – PAM-L2 – (Best & Tyler, 2007, cité par Strange & Shafer, 2008) que le modèle reconnaît que le poids pris par la L2 dans la vie de l'apprenant tardif joue un rôle, en plus du principe d'assimilation, dans la différence perçue pour un contraste phonétique propre à la L2.

Ce modèle se singularise aussi par la nature de ce qui est pris en compte pour la similarité, qui guide l'assimilation : la perception et l'assimilation d'un son de la L2 se feraient essentiellement en fonction de similarités articulatoires et non acoustiques avec des sons de la L1. L'auteur propose l'exemple d'une L1 comportant des consonnes sonores (avec un pré-voisement) et sourdes (avec aspiration). Le locuteur de cette L1 percevra une consonne occlusive éjective comme le [k'] (qui existe en coréen mais pas en français) comme plus proche de l'occlusive sourde [k] que de l'occlusive sonore [g], parce qu'elle est plus proche sur le plan articulatoire (Best, 1995). Ce sont les gestes articulatoires qui fonderaient l'assimilation plutôt que les similitudes acoustiques.

Une autre particularité de PAM est d'imaginer de façon séparée la façon dont agissent les phénomènes d'assimilation lors de la perception d'un son isolé de la L2 (ce qui correspond à une tâche d'identification) et lorsque deux sons contrastés sont entendus (ce qui correspond à une tâche de discrimination). Pour l'identification d'un son de la L2 isolé, le modèle PAM évoque trois possibilités : 1) le son de la L2 est assimilé à une catégorie de la L1 ; 2) le son de la L2 est considéré comme non catégorisable, c'est-à-dire qu'il ne correspond à aucune catégorie de la L1 (mais ne devenant pas une catégorie nouvelle comme dans SLM) ; 3) le son n'est pas catégorisé comme relevant du langage (par exemple un son d'environnement). Pour la discrimination, trois types d'assimilation perceptive seraient possibles : 1) si les deux sons de la L2 sont perçus comme des exemplaires aussi représentatifs d'une catégorie de la L1, la discrimination sera très difficile ; 2) si ces deux sons sont assimilés à des exemplaires d'une catégorie de la L1 mais diffèrent dans la capacité à bien représenter cette catégorie, la discrimination sera plus facile ; 3) si les deux sons de la L2 sont assimilés à des catégories distinctes dans la L1, leur discrimination sera excellente. Best *et al.* (2001) ont présenté des exemples de ces différences, mais on note que la logique repose entièrement sur la notion d'assimilation comparé au précédent modèle qui suppose aussi la création de nouvelles catégories.

En cherchant à adopter une perspective encore plus cognitive, davantage centrée sur les mécanismes d'apprentissage, Strange (2011) a proposé le modèle de la perception sélective automatique (*Automatic Selective Perception*, ASP). Strange (2011) insiste sur le fait

que les apprenants adultes ont appris et renforcé des principes de catégorisation phonémiques qui se sont automatisés : ce sont des routines de sélection perceptives (SPRs). Selon elle, l'adulte conserve l'aptitude à créer de nouvelles routines de ce type pour une L2. La différence avec les routines acquises plus tôt en L1 est que le poids accordé aux différents indices acoustiques disponibles pour la catégorisation phonémique n'est pas toujours les meilleurs, même chez les apprenants tardifs les plus expérimentés (Strange & Shafer, 2008).

### 4.3. Le cas de l'anglais L2 pour les francophones

Selon Flege (1995a), lors de l'apprentissage d'une L2, plus les phonèmes sont similaires entre la langue maternelle et la L2, plus il y a d'assimilation entre les langues en perception et en production de sons de parole. Le français et l'anglais partagent des phonèmes, codés par les mêmes lettres, mais dont les réalisations phonétiques varient entre les langues. Par exemple, l'opposition des consonnes occlusives existe aussi bien en français qu'en anglais, mais ceci à partir de règles phonétiques différentes, opposées sur le plan temporel, entre les deux langues.

En français, les consonnes occlusives sourdes (/p, /t/, /k/) se caractérisent par un bruit d'explosion (*burst*) suivi d'un bref silence avant la voyelle (court VOT positif) et les consonnes occlusives sonores (/b/, /d/, /g/) par un long bruit de voix (vibration des cordes vocales) avant le *burst* (long VOT négatif) immédiatement suivi de la voyelle. Les francophones identifient une occlusive sourde en cas d'indice acoustique bref et une sonore en cas d'indice long. En anglais, le VOT négatif n'existe pas. Les occlusives sourdes sont associées à une aspiration entre le *burst* et le début de la voyelle (long VOT positif), et les sonores à un court silence entre le *burst* et la voyelle (court VOT positif). Sur le plan temporel, les associations entre la durée du principal indice pour le voisement et l'opposition sourde/sonore sont inverses entre ces deux langues. Certes, le VOT n'est pas l'unique indice permettant d'opposer les sonores et les sourdes, car l'intensité du *burst* (Halle *et al.*, 1957) et la durée de transition du premier formant de la voyelle qui suit (Stevens & Klatt, 1974) sont aussi des indications, mais l'indice temporel constitué par le VOT est considéré comme le principal.

La situation d'un francophone apprenant l'anglais est donc très particulière. Les oppositions phonémiques /p/-/b/, /t/-/d/, /k/-/g/ existent dans les deux langues, mais l'application d'une règle phonétique basée sur les indices temporels habituels pour le français peut conduire à des confusions sur des sons anglais.

Si l'on retient le principe d'assimilation central dans les modèles d'apprentissage de la L2, PAM et SLM, un francophone qui ne prendrait en compte que la position du VOT par rapport au bruit d'explosion pour interpréter le voisement des occlusives anglaises et qui appliquerait la règle phonétique du français, pourrait assimiler les consonnes anglaises sonores et sourdes comme faisant partie de la catégorie des sourdes, car elles ont toutes deux un VOT positif. En d'autres termes, les sonores anglaises avec leur court VOT positif seraient par erreur rapprochées dans l'espace phonémique des sourdes françaises typiques, et les sourdes anglaises, avec leur long VOT positif, pourraient être assimilées comme des exemplaires distants mais appartenant malgré tout aux catégories des sourdes françaises. Par exemple, le /g/ anglais étant phonétiquement proche du /k/ français (VOT positif court), l'auditeur francophone pourrait être amené, concrètement, à mal interpréter certains mots. Or, en anglais, beaucoup de mots fréquents s'opposent seulement par le voisement d'une consonne. De plus, si le mot est appris seulement à l'oral, il risque d'être associé mentalement à une lettre inappropriée (e.g. avec la règle française, l'anglais [gɒd] est perçu [kɒd] et peut être transcrit COD), d'où de futures erreurs d'orthographe. Si l'apprentissage de l'anglais se faisait sans indications orthographiques, l'assimilation risquerait de conduire les apprenants francophones à ne pas percevoir et à ne pas produire de différence entre les occlusives sourdes et sonores anglaises.

Les occlusives sonores anglaises sont les plus problématiques : entendu sans support orthographique, un mot anglais commençant par une occlusive sonore peut être perçu comme s'il commençait par une occlusive sourde selon les critères du système phonétique/phonologique français. Les occlusives sonores anglaises contiennent en effet un court VOT positif, comme les occlusives sourdes françaises auxquelles elles pourraient donc être assimilées. Ces assimilations sont cependant freinées par l'orthographe, dont les apprenants tardifs ont connaissance dès les premiers cours d'anglais. Le poids de l'orthographe est tel que les productions qu'ils risquent de réaliser pour les occlusives sonores peuvent être de simples assimilations aux sonores françaises, avec un long VOT négatif qui donne un accent français très marqué. Cela n'est pas forcément très grave pour l'auditeur anglophone, qui ne peut pas confondre une telle production avec un autre phonème de l'anglais et qui peut interpréter cette erreur comme un défaut de prononciation sans grande incidence pour la compréhension. En revanche, en situation d'écoute de mots anglais en L2, un francophone risque de commettre une erreur de catégorisation de la consonne occlusive sonore qu'il entend et l'interpréter comme une occlusive sourde. Il est donc important que la



forme écrite du mot lui soit présentée pour qu'il rectifie cette catégorisation. Il doit en effet, en contexte anglais, catégoriser comme /b/ un son qu'il interpréterait comme un exemplaire de /p/ dans le système phonétique/phonologique français. Cela demande d'inhiber, en contexte anglophone, de routines de sélection perceptive (SPRs) telles que décrites par le modèle de la perception sélective automatique (ASP) (Strange, 2011). La création de nouvelles routines est cependant considérée comme possible à tout âge, selon ce modèle.

Sur le plan perceptif, ce changement est difficile à opérer car, dans les termes des modèles présentés SLM et PAM, l'assimilation d'une occlusive sonore anglaise à une catégorie sourde de la L1 peut être ici particulièrement automatique, à cause de la forte similitude des courts VOT positifs. L'élaboration d'une nouvelle catégorie phonémique, que le modèle SLM envisage comme possible, est compromise par cela. En effet, selon SLM, seul un son de L2 très différent des sons de la L1 peut donner lieu à la création d'un nouveau phonème. Ici, la situation est tellement chargée d'ambiguïté que l'objectif n'est pas vraiment de créer un nouveau phonème. Il est plutôt être d'aider l'apprenant à sortir du système de référence phonologique français. Il s'agirait de l'inciter à suivre, en anglais, de nouvelles règles phonétiques pour des couples de phonèmes dont la réalisation phonétique est radicalement différente du français, tout en sachant que les couples de phonèmes anglais opposés en voisement sont toutefois associés à des couples de lettres communs avec le français, sur le plan des correspondances grapho-phonémiques mais pas phonétique. Aider à développer des associations phonétiques / phonologiques nouvelles semble ici plus approprié que de se baser sur des assimilations trompeuses et déroutantes dès que les formes écrites des mots sont connues.

Pour les occlusives sourdes anglaises, il n'y a pas le même problème d'ambiguïté. Une expérience auprès d'anglophones apprenant le français montre que les débutants produisent les occlusives sonores françaises (qui ont un VOT négatif long) comme des sonores anglaises (sans VOT négatif) et rallongent ce VOT négatif progressivement avec leur expertise. Un phénomène analogue pourrait se produire pour les francophones face aux occlusives anglaises (qui ont un VOT positif long), et ils pourraient les produire avec un VOT court (par assimilation avec les sourdes françaises) pour progressivement insérer l'aspiration qui leur confèrera un VOT positif plus long, nouveau par rapport à ce qui existe en français. Ici, ce sont les auditeurs anglophones qui risquent de ne pas interpréter les productions du francophone peu expert selon ce qu'il voulait dire : ils risquent d'interpréter une occlusive sourde anglaises [p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>] comme une occlusive sonore si le locuteur francophone n'introduit

pas l'aspiration pour créer un VOT positif long. En perception, l'apprenant francophone pourrait, en revanche, être moins induit en erreur par une occlusive anglaise sourde (e.g. /p/) que par une sonore (e.g. /b/), car l'étrangeté du VOT positif long pourrait faire catégoriser l'occlusive sourde anglaise comme un exemplaire peu typique mais tout de même acceptable de la catégorie de l'occlusive sourde /p/ déjà établie dans sa L1.

Au final, il apparaît que le cas particulier de la réalisation phonétique de l'opposition de voisement pour les occlusives en français et en anglais place les apprenants dans une situation complexe. En particulier, le principe d'assimilation se heurte ici au fait que l'assimilation d'une occlusive sonore anglaise à la catégorie phonologique la plus proche en français est facile, mais source d'erreur. Une nouvelle opposition phonétique, à associer à une opposition phonologique sourde/sonore déjà connue en L1 est souhaitable. C'est pourquoi nous avons envisagé une technique d'aide à l'apprentissage perceptif destinée à la fois à favoriser la perception de la durée des VOTs propres à l'opposition sourde/sonore en anglais, grâce à une stimulation dissociée adressant les indices de durée plus directement à l'hémisphère le plus compétent pour la durée de VOT considérée. Il semblait aussi important d'intégrer dans cet apprentissage une information sur l'orthographe des mots, en vue d'éviter l'association des occlusives sonores anglaises entendues, à des lettres représentant des occlusives sourdes. L'intégration de ces informations orthographiques semblait essentielle dans les exercices d'entraînement à la perception du voisement en anglais.

L'apprentissage de la nouvelle opposition de voisement en anglais peut être aussi envisagé en termes de position des frontières catégorielles. La frontière catégorielle pour le voisement en français se situerait autour de 0-10 ms et celle de l'anglais serait autour de +25 ms (Aslin *et al.*, 1981) ou +30 ms (Jusczyk *et al.*, 1989 ; Lisker 1978). Autour des consonnes sourdes françaises et sonores anglaises, il pourrait exister une zone d'ambiguïté pour un apprenant francophone de l'anglais L2 (Fig. 28). La représentation mentale de la frontière catégorielle peut être évaluée à partir des performances dans des tâches de perception catégorielle (identification et discrimination). Si l'on se base sur les prédictions du modèle PAM, qui anticipe les performances en discrimination, les capacités des apprenants français pourraient être assez bonnes en discrimination dans des situations écologiques où seuls des exemplaires typiques des occlusives sourdes et sonores anglaises seraient produites, puisque le long VOT des occlusives sourdes anglaises est atypique pour un francophone et saillant. Toutefois, dans les tests de discrimination, les différences de durée entre les deux membres de la paire à discriminer sont faibles, et l'apprenant tardif de l'anglais L2 devrait avoir des

difficultés à situer une frontière précise. Il devrait en être de même en identification, car la langue française n'habituait pas à identifier deux types de VOT positifs différents.

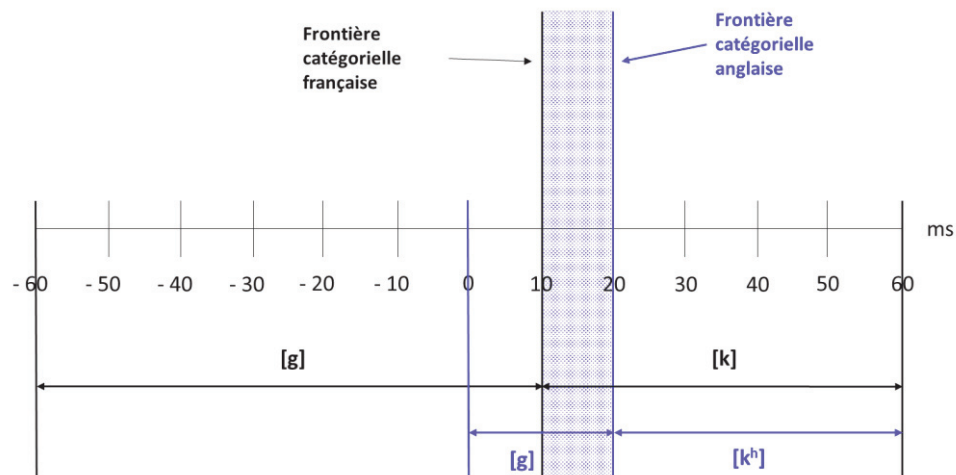


Figure 28. Continuum de VOT représentant le lieu des frontières catégorielles française et anglaise.

Au niveau cérébral et selon les modèles de résolution temporelle à échelles multiples (Poeppel *et al.*, 2008), les oscillations rapides dans le cortex auditif gauche (gamma et bêta) permettraient de détecter plus facilement les indices acoustiques brefs comme les VOTs des occlusives sourdes françaises et sonores anglaises, alors que les oscillations plus lentes dans le cortex auditif droit (e.g., oscillations thêta) permettraient de détecter et d'intégrer de longs indices comme les VOT des occlusives sonores françaises ou sourdes anglaises. Il serait souhaitable pour un francophone d'apprendre à utiliser ses hémisphères cérébraux de manière différente en français et en anglais, afin de les investir correctement pour un traitement optimal basé de la durée des indices phonétiques et non en se laissant guidé par les graphèmes. Ce changement pourrait être favorisé implicitement par des exercices enseignant du vocabulaire anglais et utilisant l'écoute dichotique, en vue de stimuler l'association hémisphère - durée du VOT la plus adaptée à l'acquisition du système phonétique/phonologique de la L2. La présence d'informations orthographiques dans ces exercices devrait aussi contribuer à échapper aux ambiguïtés.

#### 4.4. Dyslexie et apprentissage de l'anglais L2

Selon une étude de 2016 basée sur un questionnaire et réalisée auprès de 120 étudiants dyslexiques à l'Université de Lyon, les jeunes adultes se plaignent davantage de difficultés en apprentissage de l'anglais s'ils sont dyslexiques (Abadie & Bedoin, 2016 ; Mazur-Palandre *et al.*, 2016). Avec la difficulté à prendre des notes en cours, il s'agit des deux plaintes les plus récurrentes. Les dyslexiques ayant accédé aux études supérieures ont mis en place des

compensations, et les prises en charge orthophoniques les ont aidés. L'apprentissage tardif d'une L2 est cependant pour eux un nouveau déficit dans le domaine du langage.

Un petit groupe de chercheurs, autour de Ganschow et Sparks, s'est intéressé à leurs difficultés en L2, avec la volonté de conclure autrement que par la nécessité de renoncer à cet apprentissage, comme le faisait Dinklage (1971, cité par DiFino & Lombardino, 2004). Leurs travaux débouchent sur des recommandations aux enseignants et des pistes pour un enseignement alternatif, répondant notamment à leur besoin d'un accès multisensoriel à la nouvelle langue, avec des méthodes audio-visuelles, assorties d'un enseignement très structuré et explicite de la phonologie de la L2 (Sparks *et al.*, 1991), en opposition avec les approches par communication naturelle en vogue à l'époque. Des recommandations concrètes issues d'expériences réussies d'enseignement d'une deuxième langue à de petits groupes d'enfants dyslexiques, dans des conditions tenant compte de leurs particularités, avec au moins 2h de L2 par jour, l'introduction lente de nouveaux matériels linguistique, beaucoup de répétitions, et des enseignants formés aux troubles du langage, encouragent à ne pas abandonner l'idée de faire accéder cette partie de la population à un niveau minimum dans une L2 (Downey *et al.*, 2000).

Etant donné les difficultés en lecture qui définissent leur pathologie, les apprenants dyslexiques produisent des écrits de faible qualité en L2 et ils se démarquent par une lecture à voix haute lente et saccadée (DiFino & Lombardino, 2004). Ils ont souvent mis en place des compensations en lecture de leur L1, qui consistent notamment à ralentir pour être plus exacts, mais à l'âge où ils apprennent une L2, les demandes scolaires sont plus exigeantes et elles s'avèrent inadaptées (Sparks & Ganschow, 1991). Ils se montrent globalement plus à l'aise dans les activités d'écoute ou de production. C'est pourquoi, au moins pour les examens, il est recommandé de proposer des épreuves orales plutôt qu'écrites pour ces étudiants.

Cependant, l'oral en L2 devrait se faire en prenant certaines précautions, en tenant compte de la persistance de difficultés d'ordre phonologique même à l'âge adulte. Par exemple, la répétition rapide d'informations phonologiques complexes est échouée chez les collégiens dyslexiques (Catts, 1989). Une étude chez des élèves dyslexiques anglophones apprenant le français montre aussi que leurs difficultés ne se situent pas seulement à l'écrit, mais aussi à l'oral, en perception comme en production (Crombie, 1997). En répondant à des questions détaillées sur la nature de leurs difficultés en anglais, 60% des étudiants dyslexiques – mais seulement 30% des normo-lecteurs – déclarent qu'à l'occasion des cours d'anglais ils

ont fréquemment l'impression de confondre les sons (Mazur-Palandre *et al.*, 2016). Au-delà de leur ressenti, diverses recherches ont évalué chez les apprenants dyslexiques des troubles d'ordre phonologique persistants, pointés comme la principale entrave à l'apprentissage d'une L2. Des données montrent par exemple que le meilleur prédicteur du niveau d'acquisition d'une L2 au début de la deuxième année de cet apprentissage est la performance en discrimination de syllabes (Ganschow & Myer, 1988). Or, il subsiste chez une bonne proportion d'étudiants dyslexiques un trouble de discrimination auditive (Dinklage, 1971). Ce qui les perturberait le plus serait de faire face à un nouveau code phonologique sans pouvoir s'appuyer sur le sens ni sur l'écrit (Ganschow & Myer, 1988), et nous en tiendrons compte dans l'aide à l'apprentissage proposé dans cette thèse.

La nature exacte des troubles phonologiques à l'origine des difficultés en apprentissage de la L2 en cas de dyslexie fait débat. Des difficultés de discrimination d'un son nouveau dans une L2 sont rapportées (Crombie, 1997), mais pas toujours répliquées chez des adultes dyslexiques (Soroli *et al.*, 2010). Ils pourraient donc, dans certaines tâches de perception catégorielle, ne pas présenter de signe de pathologie évident. Il faut cependant se méfier des limites de ce que des mesures expérimentales révèlent sur les aptitudes des apprenants, car des représentations phonologiques correctement produites à partir d'images peuvent par exemple cacher des lacunes phonétiques importantes au niveau phonétique (Swan & Goswami, 1997). Évaluer la perception pré-attentionnelle des différences entre de nouveaux phonèmes pourrait peut-être offrir des mesures plus sensibles, et c'est une des raisons qui nous ont conduite à enregistrer des données en EEG en condition de *oddball* passif. Cependant, il est difficile en cas de dyslexie d'enregistrer ce genre de données car les ondes traduisant les capacités de catégorisation de sons (MMN), n'est pas toujours présente. Pour les chercheurs ne détectant pas d'anomalie majeure en perception de nouveaux phonèmes chez des dyslexiques, dont ces premiers ne nient pas les difficultés en L2, les déficits phonologiques responsables de ces difficultés concerneraient la conscience phonologique et la mémoire à court terme verbale (Soroli *et al.*, 2010). La persistance de déficits dans des tests de conscience phonologique a d'ailleurs été confirmée chez adultes dyslexiques anglophones (Pennington *et al.*, 1990) et des étudiants dyslexiques français (Abadie *et al.*, 2016).

La question du débit de parole est aussi importante. Chez les apprenants dyslexiques d'une L2, parler dans cette langue serait l'activité qu'ils réussiraient le mieux (Crombie, 1997). L'auteur l'explique par leur possibilité de choisir le rythme auquel ils parlent, alors qu'ils sont contraints par le débit du locuteur en perception. La vitesse du signal crée alors un

handicap, sans doute à cause de leurs difficultés à se synchroniser au rythme rapide du locuteur. Cette difficulté à traiter la parole rapide est majorée en cas de Trouble Développementale du Langage Oral (TDL, dysphasie) (Guiraud *et al.*, 2018) mais elle concerne aussi les personnes dyslexiques, qui signalent ce problème (Downey *et al.*, 2000), sans doute dû à des anomalies analogues de leur système d'ajustement des rythmes neuronaux au rythme du signal (Goswami, 2011). Des difficultés de séquenciation auditive sont effectivement documentées en L2 (Crombie, 1997 ; Thomson & Goswami, 2010) et ressenties : les apprenants dyslexiques disent ne pas savoir quand les mots commencent et finissent dans les phrases (Downey *et al.*, 2000).

La difficulté de traitement du rythme renvoie à des mécanismes d'analyse d'unités suprasegmentales, dont le traitement phonologique serait particulièrement perturbé chez les adultes dyslexiques (Soroli *et al.*, 2010). Un déficit de traitement de la pente des attaques des enveloppes d'amplitude (*rise time*) prendrait part à cette difficulté et empêcherait de construire de nouvelles représentations phonologiques parce que les caractéristiques phonétiques des phonèmes initiaux des syllabes seraient alors mal identifiées et donneraient lieu à des représentations sous-spécifiées (Thomson & Goswami, 2010). Ces derniers auteurs ont montré que les difficultés à traiter les attaques des enveloppes d'amplitude contribuent aux difficultés d'apprentissage de nouvelles représentations phonologiques de façon indépendante par rapport au déficit de conscience phonologique. Les difficultés pour ces unités de plus grande taille sont particulièrement préoccupantes pour les dyslexiques francophones apprenant l'anglais, car la prosodie et l'intonation suivent des règles différentes dans cette L2 et ne pas être armé pour l'analyse d'informations à ce niveau peut retentir sur toute la chaîne des analyses réalisées à des niveaux plus fins, notamment phonémiques.

Enfin, le système phonologique des apprenants dyslexiques reste malgré les prises en charge instable. Or, comme décrit dans le modèle SLM-r, plus le système phonologique est stable, plus la catégorisation de sons nouveaux d'une L2 sera facile. Avec la tendance à l'assimilation du système de la L2 à celui de la L1, les adultes dyslexiques vont aussi de ce fait théoriquement se retrouver en plus grandes difficultés face à l'apprentissage d'une L2 que leurs pairs normo-lecteurs (Crombie, 1997).

#### 4.5. Objectifs

Lors de l'apprentissage de l'anglais, le francophone se heurte à plusieurs difficultés, qu'elles soient d'ordre pédagogique ou linguistique. Les difficultés linguistiques résident particulièrement au niveau du traitement et de la réalisation phonétique de certains sons comme dans le cas des consonnes occlusives. En effet, ces dernières demandent au locuteur francophone apprenant l'anglais de se baser sur de mêmes représentations orthographiques mais sur des représentations temporelles des indices phonétiques opposées. Les consonnes sonores françaises étant caractérisées par un long VOT négatif et les consonnes sourdes par un court VOT positif, il semble normal que le francophone rencontre des difficultés lorsqu'il doit traiter ou produire en anglais une sonore avec un court VOT positif et une sourde avec un long VOT positif. En effet, dans le cadre des modèles SLM-r et PAM, si le locuteur francophone se trouve dans une situation d'écoute de phonèmes anglais sans appui orthographique, il pourra assimiler les deux consonnes anglaises comme appartenant à une seule et même catégorie de sa L1 : les occlusives sourdes, en l'absence de VOT négatif (le pré-voisement étant l'indice important pour la reconnaissance des sonores françaises). Ces difficultés pourraient être accrues lorsque la personne est aussi atteinte de dyslexie. Comme décrit dans le modèle SLM-r, plus le système phonologique est stable, plus la catégorisation de sons nouveaux d'une L2 sera facile. Le traitement de la temporalité des indices acoustiques étant déficitaire en cas de dyslexie, leur système phonologique de la L1 est instable et mal délimité, il pourrait donc être particulièrement difficile pour eux d'apprendre cette deuxième langue.

Il est pourtant indispensable aujourd'hui d'avoir une maîtrise suffisante de la langue anglaise pour sa vie personnelle comme professionnelle. Il est aussi de plus en plus utile de la maîtriser dès le plus jeune âge. Même si des efforts sont réalisés en ce sens, le niveau, en France, reste insuffisant.

Il paraît donc nécessaire de proposer de nouvelles méthodes d'aide à l'apprentissage de l'anglais pour les adultes francophones. Pour cela, il faut aussi mettre au point des outils d'évaluation, spécialement conçus pour les francophones, utilisables dans les écoles. Ils permettront d'estimer le niveau des étudiants et d'évaluer l'impact de différentes interventions pour l'apprentissage de l'anglais L2. Cela doit aussi prendre en compte les apprenants pouvant présenter un trouble des apprentissages du langage comme une dyslexie.

Ce chapitre a deux objectifs principaux. Le premier est d'évaluer les qualités d'un test de perception catégorielle en anglais qui permettra par la suite d'évaluer les possibles effets



d'un entraînement pour l'apprentissage du voisement en anglais L2. Le second est d'évaluer si la discrimination et en identification de consonnes anglaises diffère entre des adultes normo-lecteurs et des adultes présentant une dyslexie. Afin de répondre à ces deux objectifs, deux expériences ont été menées chez des étudiants francophones présentant ou non une dyslexie. **L'Expérience 4** visait à mesurer l'effet test-retest ainsi que la fidélité du test de perception catégorielle en anglais conçu sur le même modèle que celui pour le français (voir Expérience 1). **L'Expérience 5** compare les performances en perception catégorielle de l'anglais entre des adultes dyslexiques et des adultes normo-lecteurs, tous francophones, afin de vérifier si les plaintes exprimées concordent avec des mesures comportementales.

Les résultats des deux expériences seront discutés et mis en relation dans la discussion générale de ce chapitre

#### **4.6. Expérience 4 : Evaluation de qualités psychométriques des épreuves de perception catégorielle en anglais**

##### **4.6.1. Principe de l'expérience et hypothèses**

L'Expérience 4 visait à évaluer si les tests de perception catégorielle (PC) en anglais, construits sur le même modèle que ceux élaborés pour le français, sont de bons outils pour évaluer l'effet de différents entraînements en anglais. L'hypothèse était qu'il y aurait un faible effet test-retest, ainsi qu'une bonne fidélité pour les deux épreuves d'identification et de discrimination proposées. Nous avons également fourni des normes pour ce test en termes de moyenne, d'écart-type et d'intervalle de confiance, à partir d'un groupe de 60 jeunes adultes, taille d'échantillon encore faible, mais qui pourra être étoffé par la suite.

##### **4.6.2. Méthode**

###### **4.6.2.1. Participants**

Un total de 60 participants normo-lecteurs a réalisé ces épreuves de PC en anglais : les 20 participants de l'Expérience 1 à T1 (après les épreuves de PC en français) (16 femmes,  $M = 22.3$  ans ;  $ET = 1.9$ ) et 40 autres participants (34 femmes, 6 hommes) âgés de 18 à 29 ans ( $M = 24,1$  ans,  $ET = 1$  ans et 3 mois) qui ont réalisé ces mêmes épreuves de PC à T1 avant un autre entraînement, en anglais (Chapitre 5). Une partie d'entre eux ( $n = 20$  de l'Expérience 1) ont réalisé la PC deux fois, pour estimer les effets test-retest et la fidélité (Annexe 3).



#### 4.6.2.2. Matériel et équipement

Les deux continua (l'un portant sur la paire /gi-ki/, l'autre sur la paire /di-ti/) utilisés dans les épreuves de PC (i.e. identification et discrimination) en anglais ont été construits de la même façon que ceux en français (Chapitre 2, partie 2.10.1) à partir de syllabes naturelles anglaises du même locuteur bilingue natif français/anglais. Cependant, en anglais il n'y avait que 7 stimuli par continuum (VOT positif uniquement) (Fig. 29).

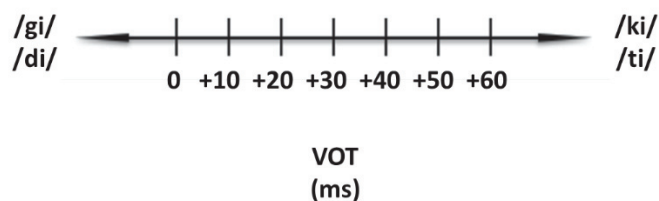


Figure 29. Continua de VOT en anglais utilisés lors des épreuves de perception catégorielle.

#### 4.6.2.3. Procédure d'évaluation

La même procédure d'évaluation que celle décrite pour la PC en français (Chapitre 2 partie 2.10.1 – Expérience 1) a été appliquée pour les épreuves en anglais, avec la discrimination toujours réalisée avant l'identification, et la paire /di-ti/ évaluée avant la paire /gi-ki/. Les deux passations pour évaluer l'effet test-retest des épreuves en anglais ont été réalisées à 5 jours d'intervalle. L'évaluation totale durait 30 minutes.

#### 4.6.3. Analyses des données

Le traitement de données en identification et en discrimination a été réalisé comme décrit dans le Chapitre 2, partie 2.10.4 (Expérience 1). Pour l'évaluation de l'effet test-retest et la fidélité du test, les analyses statistiques décrites au Chapitre 2, partie 2.10.4 (Expérience 1) ont été réalisées sur les données de 20 participants. Pour un début de création de normes, des valeurs de moyenne, écart-type et intervalle de confiance ont été calculées sur les 60 participants.

#### 4.6.4. Résultats

##### 4.6.4.1. Evaluation de l'effet test-retest

**En discrimination**, il n'avait pas de différence de hauteur du pic entre T1 et T2, que ce soit pour la paire /gi-ki/ ( $t(19) = 0.14, p = .894$  ;  $Z = 0.14, p = .890$ ) ou pour la paire /di-ti/ ( $t(19) = 1.53, p = .143$  ;  $Z = 1.48, p = .139$ ) (Fig. 30).

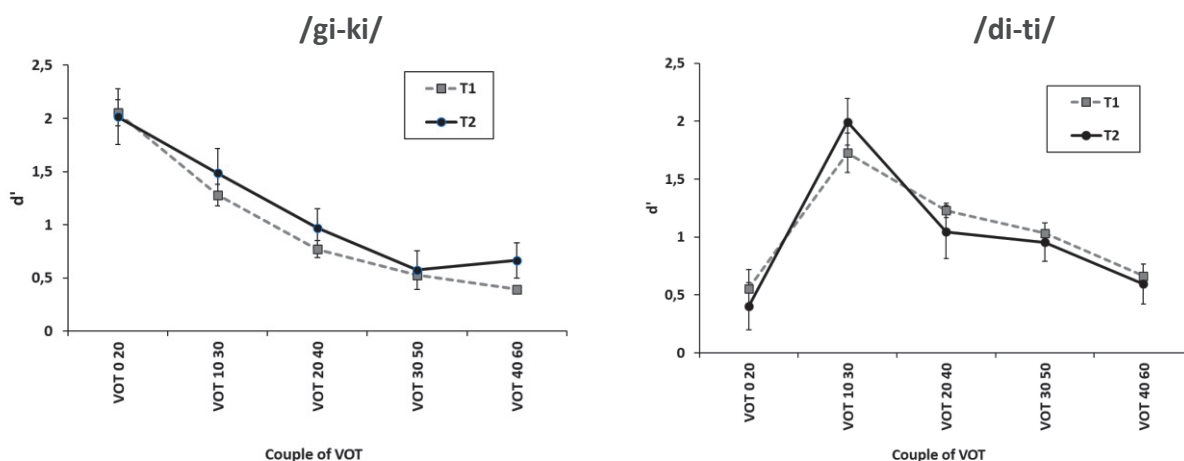


Figure 30. Courbes des  $d'$  en anglais à T1 (pointillés) et T2 (plein) pour les paires /gi-ki/ (à gauche) et /di-ti/ (à droite). Les barres d'erreurs représentent l'erreur-type.

**En identification** (Fig. 31), il n'y avait pas de différence entre les deux paires de sons étudiées à T1, que ce soit pour les valeurs de seuil ( $t(19) = 1.83, p = .083$  ;  $Z = 1.73, p = .083$ ) ou de pente ( $t(19) = 0.61, p = .552$  ;  $Z = 0.62, p = .538$ ). Il n'y avait pas non plus de différence de seuil ou de pente entre T1 et T2 (Tableau 3)

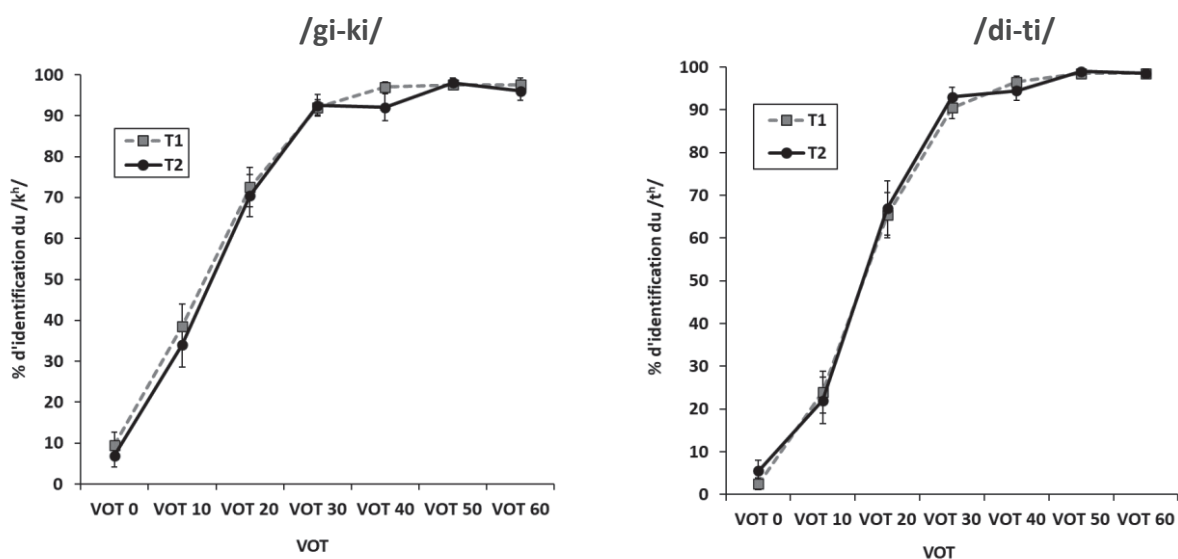


Figure 31. Pourcentage d'identification en anglais à T1 (pointillés) et T2 (plein) pour les paires /gi-ki/ (à gauche) et /di-ti/ (à droite). Les barres d'erreurs représentent l'erreur-type.

Tableau 3. Résultats aux tests  $t$  de Student et  $t$  avec permutations pour les valeurs de seuil et de pente en identification entre T1 et T2 pour les deux paires testées.  $ddl = 19$

		Test t de Student		Test t avec permutation	
		$t$	$p$	$Z$	$p$
/gi-ki/	Seuil	0.90	.380	0.90	.366
	Pente	0.07	.948	0.07	.946
/di-ti/	Seuil	0.07	.946	0.07	.944
	Pente	1.53	.142	1.49	.138

#### 4.6.4.2. Evaluation de la fidélité

**En discrimination**, l'étude des corrélations a permis de montrer que les performances entre T1 et T2 étaient moyennement corrélées au niveau du pic de discrimination pour la paire /di-ti/ ( $r(18) = .61, p = .005$ ), mais pas pour la paire /gi-ki/ ( $r(18) = .38, p = .101$ ) (Fig. 32).

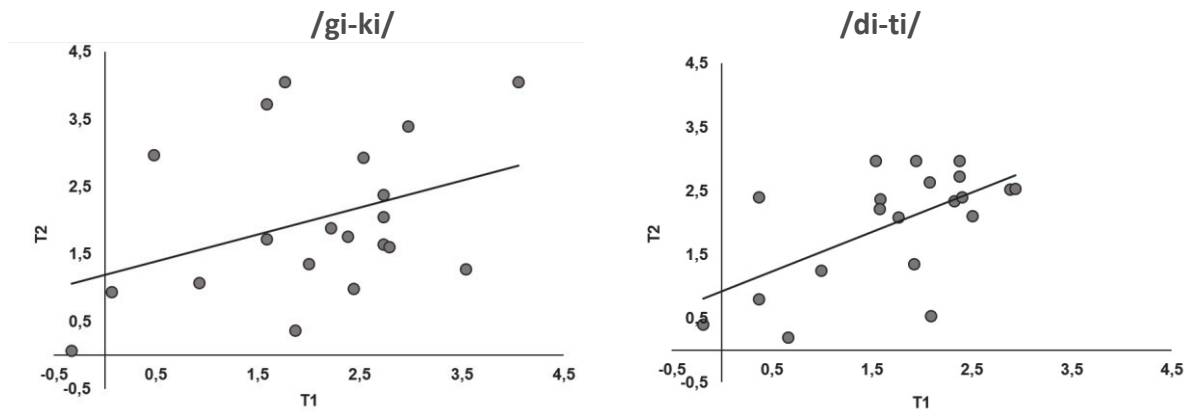


Figure 32. Nuage de points représentant la corrélation entre T1 et T2 pour la valeur de  $d'$  au niveau du pic de discrimination de la paire /gi-ki/ (à gauche) et de la paire /di-ti/ (à droite). La barre noire représente la droite de régression.

L'étude des corrélations **en identification** a permis de montrer que, pour les valeurs de seuil, les performances entre T1 et T2 étaient moyennement corrélées pour la paire /di-ti/ ( $r(18) = .65, p = .002$ ), mais seulement faiblement la paire /gi-ki/ ( $r(18) = .45, p = .046$ ) (Fig. 33).

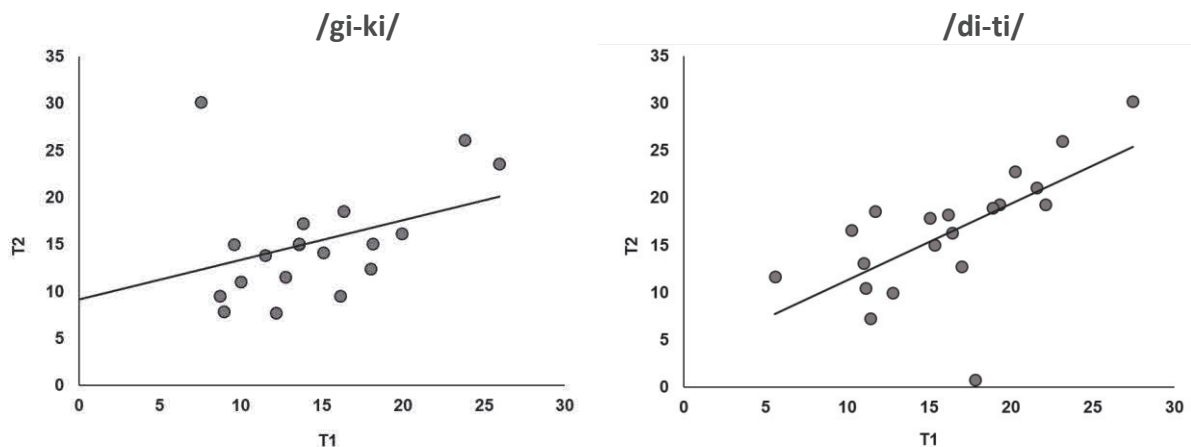


Figure 33. Nuage de points représentant la corrélation entre T1 et T2 pour la valeur de seuil de la paire /gi-ki/ (à gauche) et de la paire /di-ti/ (à droite). La barre noire représente la droite de régression.

Pour la pente, la corrélation entre T1 et T2 était plus élevée pour la paire /gi-ki/ ( $r(18) = .68, p < .001$ ) que pour la paire /di-ti/ ( $r(18) = .25, p = .282$ ). Cependant, les données n'étaient plus corrélées entre T1 et T2 en l'absence des 3 points externes droite de la figure 34 gauche ( $r(18) = .06, p = .836$ ). L'analyse qualitative du nuage de points pour /di-ti/ sur la figure 34 (à

droite) montre que les trois quarts des participants conservaient tout de même le même niveau à T1 qu'à T2 : les participants ayant de faibles performances à T1 gardaient des performances aussi faibles à T2, et les participants avec de très bonnes performances à T1 les conservaient à T2. Un quart de l'échantillon faisait exception : ils s'amélioraient entre T1 et T2 et un participant régressait.

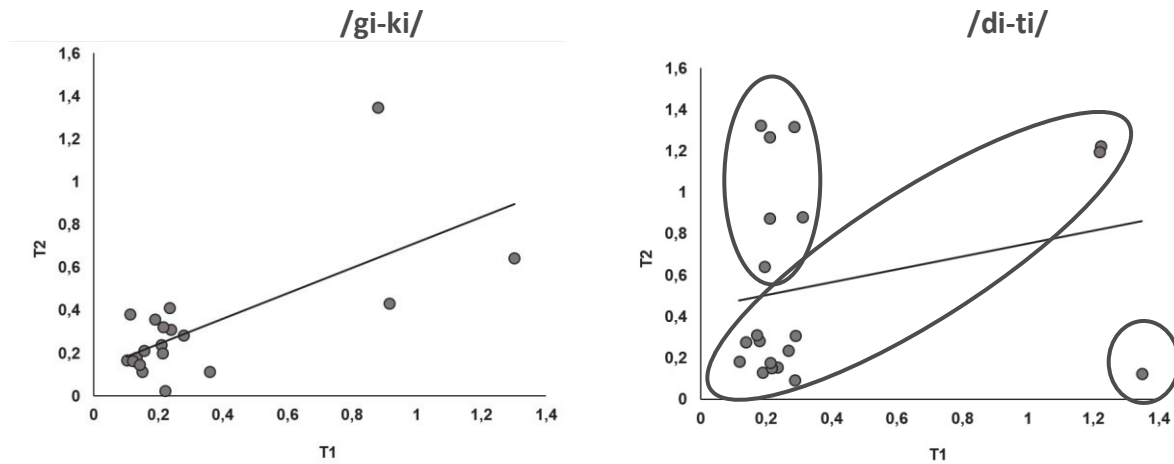


Figure 34. Nuage de point représentant la corrélation entre T1 et T2 pour la valeur de pente de la paire /gi-ki/ (à gauche) et de la paire /di-ti/ (à droite). La barre noire représente la droite de régression.

#### 4.6.4.3. Vers la création des normes

**Discrimination.** Les résultats des calculs pour de premières normes (ici, sur 60 participants) sur les valeurs de  $d'$  du pic de discrimination en termes de moyenne, d'écart-type, d'erreur-type et d'indice de confiance à 95% et 99% sont indiqués dans le Tableau 4 (Fig. 35).

Tableau 4. Résultats de normes de l'épreuve de discrimination en anglais sur les valeurs de  $d'$  du pic à T1. (M = Moyenne, Méd. = Médiane, Err-T = Erreur-type, IC = Intervalle de Confiance)

	M	Méd.	Err-T	IC (95%)	M + IC (95%)	M - IC (95%)	IC (99%)	M + IC (99%)	M - IC (99%)
/gi-ki/	1.66	1.75	0.14	0.26	1.92	1.39	0.35	2.01	1.31
/di-ti/	1.56	1.60	0.12	0.23	1.80	1.33	0.31	1.87	1.26

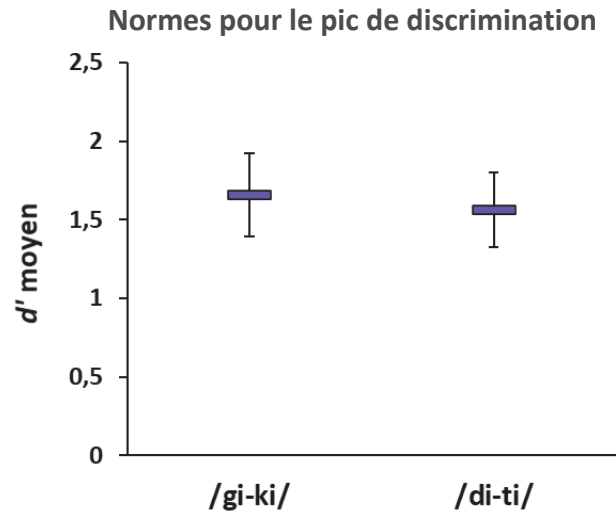


Figure 35. Moyennes des  $d'$  du pic de discrimination pour les paires /gi-ki/ et /di-ti/. Les barres d'erreurs représentent l'intervalle de confiance à 95% autour de la moyenne.

**Identification.** Les résultats des calculs pour de premières normes (sur 60 participants) sur les valeurs de seuil et de pente en termes de moyenne, d'écart-type, d'erreur-type et d'indice de confiance à 95% et 99% sont indiqués dans le Tableau 5 (Fig. 36). On remarque que, pour les valeurs de pente, les intervalles de confiance pour les deux paires ne se recouvrent pas.

Tableau 5. Résultats de normes de l'épreuve d'identification en anglais sur les valeurs de seuil et de pente à T1. (M = Moyenne, Méd. = Médiane, Err-T = Erreur-type, IC = Intervalle de Confiance)

		M	Méd.	Err-T	IC (95%)	M + IC (95%)	M - IC (95%)	IC (99%)	M + IC (99%)	M - IC (99%)
/gi-ki/	Seuil	12.50	12.42	0.87	1.71	14.21	10.80	2.24	14.74	10.26
	Pente	0.35	0.21	0.04	0.08	0.43	0.26	0.11	0.45	0.24
/di-ti/	Seuil	16.99	17.40	0.81	1.58	18.57	15.41	2.08	19.07	12.91
	Pente	0.41	0.24	0.05	0.10	0.51	0.31	0.13	0.54	0.28

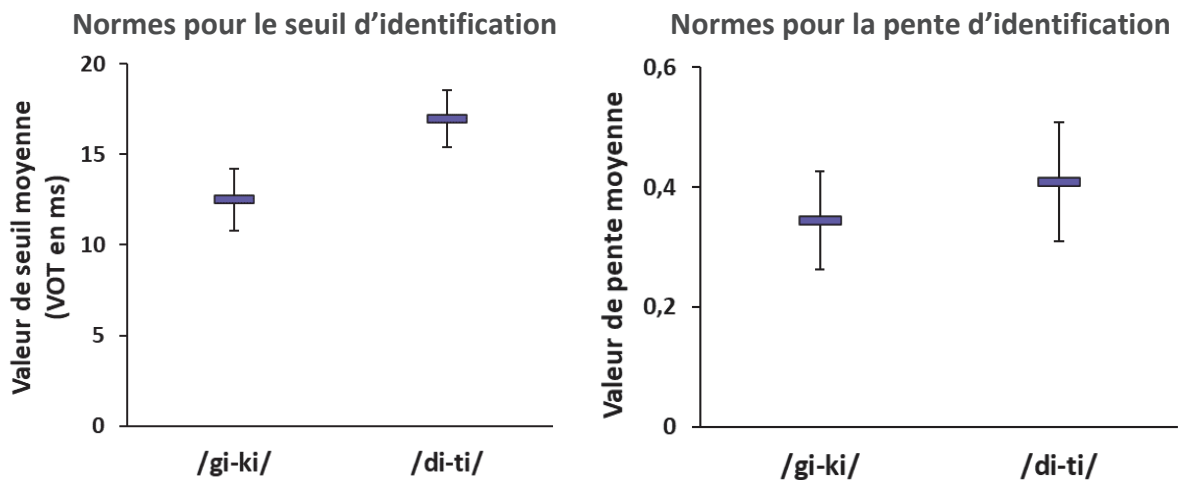


Figure 36. Valeurs de seuil (gauche) et de pente (droite) moyennes pour les paires /gi-ki/ et /di-ti/ en identification. Les barres d'erreurs représentent l'intervalle de confiance à 95%.

## **4.7. Expérience 5 : Comparaison des performances en anglais des adultes dyslexiques et des adultes normo-lecteurs**

### **4.7.1. Principe d'expérience et hypothèses**

Le principe de cette expérience était de comparer les performances de groupes d'adultes normo-lecteurs et dyslexiques dans des épreuves de perception catégorielle (PC) afin d'observer si la plainte exprimée par les adultes dyslexiques selon différentes études au sujet de la perception des sons en anglais était cohérente avec leur niveau de performance. L'une des hypothèses était que les performances en anglais en discrimination et en identification des participants dyslexiques seraient moins élevées que celles des participants normo-lecteurs. L'autre hypothèse était que les différences d'expression phonétique des contrastes de voisement en anglais et en français ne seraient pas traitées de la même façon par les deux groupes, cela avant tout entraînement.

### **4.7.2. Méthode**

#### **4.7.2.1. Participants**

Dans cette expérience, les 21 participants dyslexiques faisaient partie du groupe de 27 dyslexiques ayant réalisé l'entraînement dichotique en français (Chapitre 3, Expérience 2 partie 3.3.2.1.). En effet, 21 d'entre eux avaient accepté d'effectuer l'entraînement en anglais (et les tests de PC dans cette langue) après l'entraînement en français (pour les résultats de l'entraînement en anglais, voir Chapitre 5) (Annexe 3). Les performances en PC en anglais à T1 ont été recueillies avant le premier entraînement (en français).

Les participants normo-lecteurs (20) inclus dans cette expérience de PC étaient ceux du groupe contrôle des 21 dyslexiques ayant ensuite réalisé un entraînement en anglais, voir Expérience 8 (pour les normo-lecteurs, il n'y avait pas d'entraînement préalable en français) (Annexe 3). La comparaison entre ces deux groupes avant les entraînements était aussi motivée par le souhait de connaître leurs éventuelles différences de base avant l'entraînement en anglais après lequel ils allaient encore être évalués (pour l'effet de cet entraînement, voir Chapitre 5).

#### **4.7.2.2. Matériel et équipement**

Le matériel et l'équipement étaient les mêmes que ceux décrits dans l'Expérience 4 dans la partie 4.6.2.2 de ce chapitre.

#### 4.7.2.3. Procédure d'évaluation

Les épreuves d'évaluation étaient la même que celles décrites dans l'Expérience 4 dans la partie 4.6.2.3 de ce chapitre. Seules les performances à T1 ont été analysées ici entre les groupes NL et DYS.

#### 4.7.3. Analyses de données

Afin de comparer, à T1, les performances en anglais du groupe dyslexique (DYS) à celles du groupe normo-lecteur (NL), une série de tests  $t$  de Student bilatéraux avec correction de Bonferroni ont été réalisés. Puis, d'autres tests  $t$  comparaient, pour chaque groupe, les performances en anglais par rapport à leurs performances en français. Par ailleurs, les moyennes du groupe DYS en anglais ont été calculées pour les valeurs de  $d'$  pour la discrimination, de seuil et de pente en identification, afin de les comparer à la norme établie dans l'Expérience 4 dans la partie 4.6.4.3 de ce chapitre. Une observation qualitative a aussi été réalisée afin d'observer les différences comportementales des participants du groupe DYS par rapport au groupe sans pathologie plus large ( $n = 60$ ) constituant le début d'une norme (Expérience 4 - partie 4.6.2.1).

#### 4.7.4. Résultats

**Discrimination.** Les analyses ont montré qu'il n'y avait pas de différence significative de l'indice de sensibilité  $d'$  au niveau du pic de discrimination entre les deux groupes pour la paire anglaise /gi-ki/ ( $t(37.15) = 0.35, p = .732, d = .11$ ). Cependant, pour cette paire, la figure 37 (gauche) montre que le  $d'$  du groupe DYS se singularisait par sa faiblesse par rapport à celui du groupe NL au niveau du VOT 20 ms (condition VOT 10-30 ms). Bien que non-significative, cette différence était de taille modérée ( $t(37.61) = 1.63, p = .112, d = 0.51$ ). Le pic de discrimination pour la paire /gi-ki/ se situait toujours au VOT 10 ms.

Pour la paire anglaise /di-ti/, le pic de discrimination se situait au VOT 20 ms, et il était significativement moins élevé chez les DYS que chez les NL, l'effet étant de grande taille ( $t(37.60) = 5.00, p < .001, d = 1.55$ ) (Fig. 37, droite).

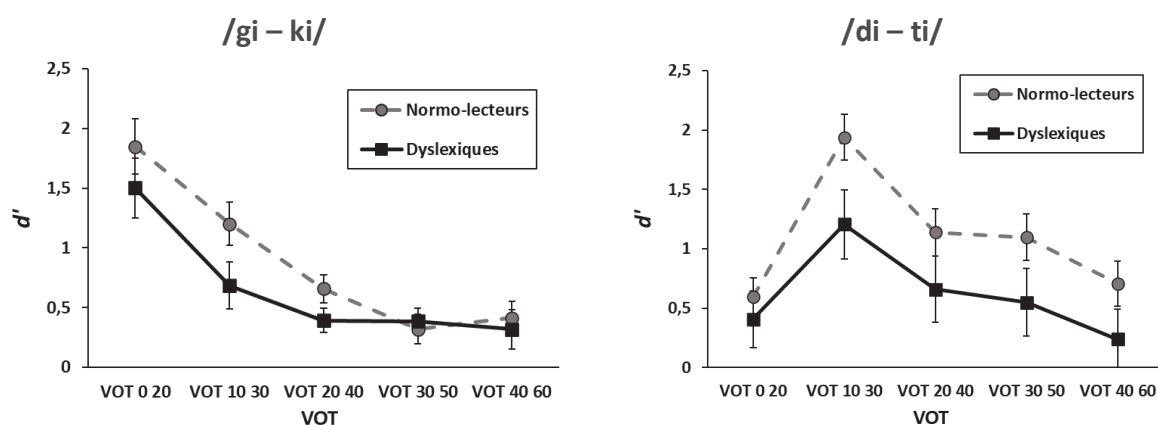


Figure 37. Courbes de discrimination en anglais pour les paires /gi-ki/ (gauche) et /di-ti/ (droite) dans le groupe DYS (trait plein noir) et NL (pointillés gris) à T1.

**Identification.** Les données en termes de valeurs de seuil ou de pente en identification de phonèmes anglais ne différaient pas significativement entre les groupes DYS et NL, et ce, pour les deux paires (Tableau 6).

Tableau 6. Valeurs de  $p$  au test  $t$  de Student pour les comparaisons entre les groupes DYS et NL concernant les valeurs de seuil et de pente pour les paires /gi-ki/ et /di-ti/ à T1.

		Test $t$ de Student			Test $t$ avec permutation	
		$t$	$ddl$	$p$	$Z$	$p$
/gi-ki/	Seuil	0.10	33.07	.920	0.10	.918
	Pente	0.99	38.27	.328	0.99	.324
/di-ti/	Seuil	0.31	37.74	.761	0.31	.755
	Pente	0.93	34.61	.357	0.94	.346

**Comparaison des performances entre L1 et L2, pour chacun des groupes.** Le groupe de NL présentait deux différences de traitement des paires de consonnes selon la langue. Ainsi, leur pic de discrimination était significativement plus haut en anglais qu'en français pour la paire /gi-ki/ ( $t(19) = 2.42, p = .026, d = .60$ ;  $Z = 1.83, p = .068$ ), et une pente d'identification était significativement plus abrupte pour la paire /di-ti/ en anglais qu'en français ( $t(19) = 2.08, p = .052, d = .61$ ;  $Z = 1.87, p = .062$ ).

Pour le groupe dyslexique, seule une différence au niveau du seuil d'identification de la paire /gi-ki/ était significative ( $t(20) = 2.05, p = .054, d = .64$ ;  $Z = 2.00, p = .046$ ) avec un seuil à +12 ms en anglais et à +20 ms en français.

**Etude qualitative des données des adultes dyslexiques en comparaison à un premier groupe normatif.** La figure 38 montre que quel que soit la variable dépendante observée, peu de dyslexiques sont dans la norme. Cependant, deux sous-groupes se distinguent. La moitié



des adultes dyslexiques se trouve en dessous de la norme définie par l'intervalle de confiance pour le pic de discrimination (50%), une majorité a un seuil d'identification plus proche de la frontière catégorielle française qu'anglaise (VOT trop court) (41%) ou un seuil correspondant à un VOT anormalement long (36%), et encore une majorité dont la pente d'identification est moins abrupte que la norme (52%). Il est tout de même intéressant de noter qu'environ un quart des participants ont de meilleures performances que leurs homologues sans pathologie. Les participants dyslexiques avaient été recrutés sans prise en compte de l'origine de leur dyslexie. Cependant, une analyse qualitative par individu nous assure qu'aucun participant du groupe DYS n'était au-dessus de la norme pour les trois mesures.

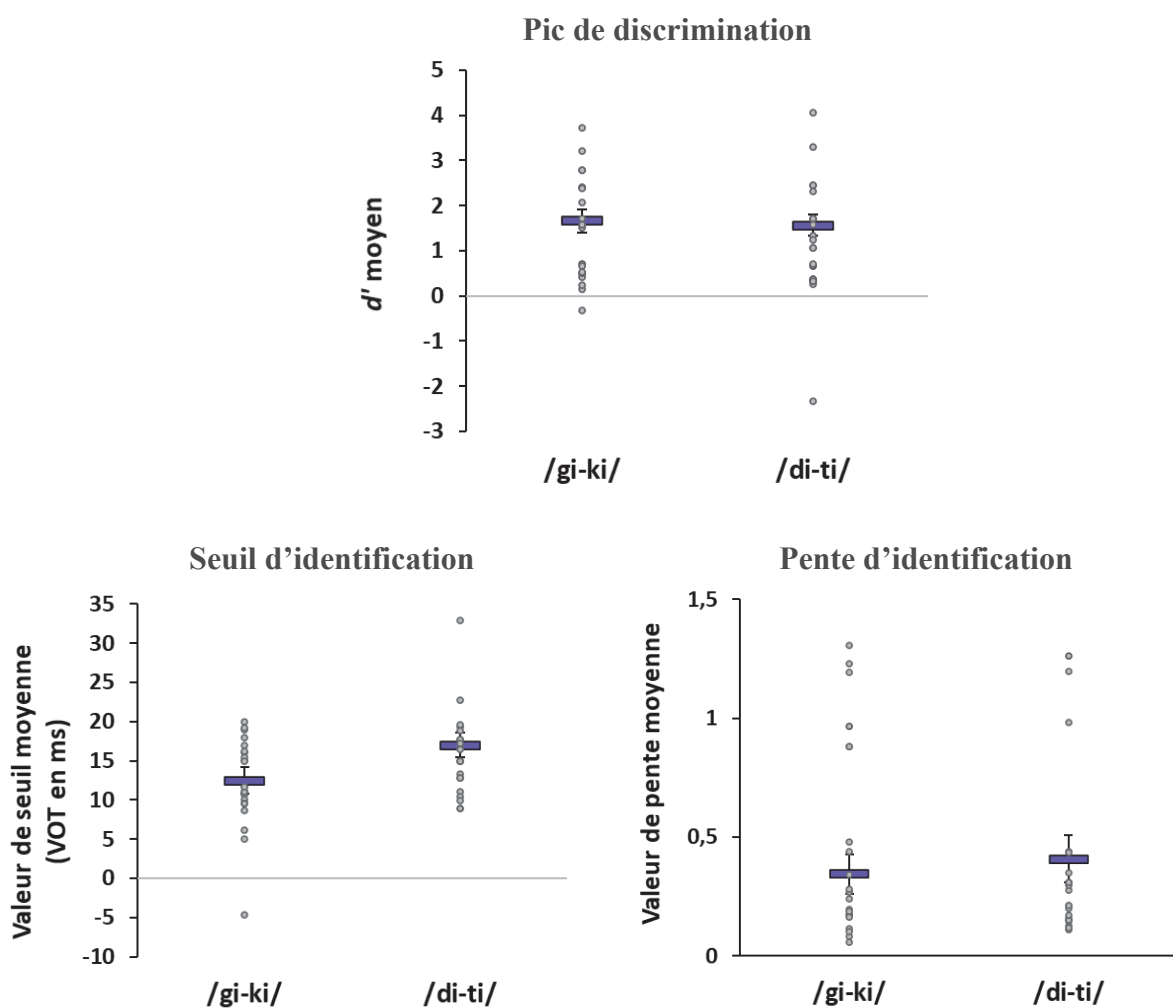


Figure 38. Répartition des participants du groupe DYS (points) par rapport à la moyenne de la norme (barre violette) pour les valeurs de  $d'$  du pic de discrimination (en haut), de seuil (bas, gauche) et de pente (bas, droite) d'identification. Les barres d'erreurs représentent l'intervalle de confiance à 95%.

#### **4.8. Discussion sur les tests de perception catégorielle anglaise (Expériences 4 et 5)**

Ce chapitre avait deux objectifs principaux. Le premier était d'évaluer quelques qualités psychométriques de tests de perception catégorielle (PC) en anglais destinés à mesurer ensuite des effets d'entraînement (phases pré- et post-entraînement) pour l'apprentissage du voisement en anglais L2. Il s'agissait aussi de commencer à créer des normes pour permettre son utilisation dans le milieu académique et permettre une mesure objective du niveau des étudiants en perception de phonèmes anglais et pour situer un apprenant par rapport à ses pairs et évaluer ainsi ses besoins. Le second objectif, était d'évaluer s'il existait une différence de performances en PC de l'anglais entre des adultes normo-lecteurs et des adultes dyslexiques, et s'ils traitaient de la même façon la différence entre le français et l'anglais pour le contraste de voisement. Nous avons tenté d'y répondre à travers deux expériences portant chacune sur un des objectifs.

La discussion de ce chapitre commentera les qualités et les limites des tests de PC élaborés ici en anglais, ainsi que sur leur intérêt dans un contexte de diagnostic ou d'évaluation académique. Nous reviendrons ensuite sur les différences observées entre les adultes atteints d'un trouble neuro-développemental du langage et les adultes sans trouble, lorsqu'il s'agit de traiter le contraste de voisement en français et en anglais, langues où il est source d'ambiguïtés ; des interprétations seront proposées au sujet de ces différences. Enfin, la discussion s'ouvrira sur les limites de ces deux expériences et les nouveaux questionnements qui en découlent.

##### **4.8.1. Qualités diagnostiques des tests**

Le test de PC en anglais construit dans l'**Expérience 5** était basé sur le même principe que celui créé en français dans le Chapitre 2 (Expérience 1). Évaluer ses qualités psychométriques et commencer à constituer des normes était nécessaire avant de l'utiliser pour tester des effets d'entraînement au voisement en anglais auprès de francophones. Ces évaluations ont porté sur deux paires de consonnes occlusives anglaises opposées par le voisement comme dans les épreuves en français : /d-t/ et /g-k/.

Les tests de PC ont su montrer de bonnes qualités psychométriques. Le premier résultat fort est qu'aucun effet de test-retest n'a été observé et ce, sur aucune des deux paires. En effet, il n'y a eu ni amélioration ni dégradation des performances que ce soit pour le *d'* du pic de discrimination ou pour les valeurs de seuil et de pente d'identification lorsque les participants réalisaient deux fois le même test à 5 jours d'intervalle. Cela montre qu'il n'y a pas

d'apprentissage de la tâche et pas d'amélioration dû à la pratique. Ce résultat est important car il va permettre de quantifier l'effet d'un entraînement basé sur l'opposition du voisement en anglais. En effet, si un progrès est observé, il sera imputable à l'entraînement et non simplement à un effet de pratique.

Les résultats sur la fidélité du test sont quant à eux plus hétérogènes. En effet, on observe une différence entre les deux paires, comme en français (Chapitre 2). La paire /d-t/ présente, de manière générale, une meilleure fidélité que la paire /g-k/ pour les deux épreuves (identification et discrimination). Même s'il n'y a pas de fidélité pour la pente concernant cette paire, cette dernière semble tout de même la plus stable pour évaluer la PC anglaise chez des francophones. Le calcul de pente a ses limites, avec l'ajustement à une courbe qui n'est pas toujours très représentative des résultats bruts. La valeur de seuil, en revanche, montre une bonne fidélité, et cela avec une variation inter-individuelle présente entre les valeurs, et suivant globalement bien une droite. La paire /g-k/ montre tout de même des qualités : elle permet un test fidèle en identification pour les valeurs de seuil mais aussi de pente ; la fidélité est par contre médiocre en discrimination.

Ces résultats montrent que s'il y a un choix à faire entre les deux paires pour une évaluation pré-post entraînement, le test avec la paire /d-t/ est meilleur si on veut tester l'effet sur la discrimination, ainsi que l'identification mais en se limitant alors à la valeur de seuil. En revanche, si un effet sur la pente doit être évalué, le choix devrait plutôt se tourner vers /g-k/, qui a son intérêt, avec tout de même une réserve un peu générale sur la pente : il y a peu de variabilité entre les sujets, et seuls de rares individus ont des valeurs extrêmes.

Les deux paires ont leurs points forts, mais montrent des différences. Que faire dans une volonté d'évaluation d'un entraînement ou dans le cadre scolaire ? Il serait possible d'envisager de réaliser l'évaluation sur une seule des deux paires (plutôt /d-t/), ce qui permettrait de diminuer le temps de passation de moitié (15 minutes au lieu de 30). Cette option serait envisageable pour évaluer le niveau d'un apprenant dans le cadre académique. Elle est moins recommandée pour l'évaluation d'un apprentissage à la suite d'un protocole d'entraînement. En effet, cela limiterait la possibilité d'observer un effet de transfert des apprentissages depuis un matériel appris vers un matériel non-appris.

Or, observer un effet de transfert peut être une mesure particulièrement intéressante dans le cadre de l'étude menée ici. L'évaluation de la paire /t-d/ a pour cela tout intérêt à être présente, compte tenu du contenu des entraînements proposés par la suite (entraînement

seulement sur les paires /b-p/ et /g-k/, cf. Chapitre 5). Mesurer l'effet de transfert donne une estimation de la profondeur des apprentissages.

Si l'on s'intéresse d'un peu plus près à l'évaluation de la fidélité des épreuves, particulièrement pour la pente d'identification, trois types de comportements ont été observés pour la paire /d-t/ : un maintien, une amélioration et une détérioration des compétences. Ces variabilités inter-individuelles sont surprenantes. Un premier élément pouvant expliquer ce résultat pourrait être le côté peu habituel de la tâche d'identification de phonèmes pour les participants : elle est peu répandue dans le domaine scolaire. Un effet de surprise aurait pu créer un biais. Une telle surprise n'a cependant pas gêné les participants outre mesure, car aucun effet global de la pratique n'a été trouvé, cet argument n'est donc pas le plus convaincant. Un autre élément est que les sons entendus sont modifiés, présentent peu de différence entre eux, et la présentation isolée de tout contexte sémantique ou lexical accentue sans doute leur ambiguïté. Certains participants peu habitués à ce type de matériel ont pu être pris au dépourvu lors de la première séance, mais moins lors de la deuxième, ce qui expliquerait l'amélioration observée chez un quart d'entre eux. Une autre piste pour expliquer l'amélioration des performances pour cette fraction de l'échantillon serait une amélioration spontanée due à la répétition de la tâche.

La différence entre les deux paires pourrait aussi trouver son origine dans l'ordre des tests. En effet, tous les participants réalisaient d'abord l'évaluation sur la paire /d-t/ puis sur la paire /g-k/. La surprise due à la nature de la tâche d'identification, difficile et peut-être déroutante au début pour certains, pourrait être accentuée pour la paire /d-t/. Notons toutefois que cette différence de performances entre les paires ne s'est produite que dans la tâche d'identification et non en discrimination. Elle se distingue aussi de la tâche d'identification par le fait de ne pas demander de recours explicite aux représentations phonémiques. En effet, en discrimination, le participant répond « identique » ou « différent » sur une base acoustique à la suite de l'écoute de deux sons. La perception d'une petite différence acoustique, quelle qu'elle soit, justifie la réponse, ce qui est simple, et pourrait rendre le participant insensible au fait de réaliser cette tâche pour la première fois (forcément sur /d-t/) ou la deuxième fois : l'atténuation de l'effet d'ordre par la facilité pourrait expliquer l'absence de diminution des performances pour /d-t/ par rapport à /g-k/ en discrimination. A contrario, dans une tâche d'identification, il est demandé de catégoriser le son en choix forcé. Dans ce cas, et selon les modèles d'apprentissage d'une L2, le participant va essayer de rapprocher le son entendu d'une catégorie connue de sa L1. Cependant, ce procédé a ses limites en L2, d'autant plus

avec des sons jouant sur l'ambiguïté. Il est difficile d'en savoir plus sans réaliser le même protocole en évaluant d'abord la paire /g-k/ puis la paire /d-t/.

Si l'on souhaite évaluer la PC en anglais il faudra prendre des précautions. Dans le premier cas, peut-être qu'utiliser la discrimination sur les deux paires serait plus profitable que d'en utiliser une seule : la tâche de discrimination est bien souvent utilisée dans des recherches expérimentales, de manière isolée, pour une estimation de la PC. Dans le cadre scolaire, il faudra aussi être vigilant si l'on s'adresse à une population plus jeune. Cette tâche, qui est difficile et place forcément devant des ambiguïtés, pourrait poser particulièrement problème avec des enfants. Afin de compenser cet effet, il faudrait donner des consignes très précises et rassurantes sur ce qui est attendu, et proposer un retour sur la tâche à la fin de l'évaluation. Cela éviterait que l'enfant reste perplexe sur ce qu'il a réussi à faire par rapport à ce qu'on attendait de lui, perplexité parfois spontanément exprimée par les adultes de l'étude.

En fin de compte, trois pistes peuvent être proposées pour l'utilisation de ces tests. La première serait de baser l'évaluation uniquement sur la paire /d-t/ qui montre un faible effet test-retest ainsi qu'une bonne fidélité générale aussi bien en discrimination qu'en identification. Cependant, cela implique de ne pas pouvoir observer la pente, ni évaluer un possible effet de transfert. La deuxième piste serait d'évaluer uniquement la discrimination mais cette fois en utilisant les deux paires /d-t/ et /g-k/. Cette possibilité permet de tester un effet de transfert mais ne nous apprend rien sur la mobilisation volontaire des catégories phonémiques (i.e. pas d'identification). Ces deux pistes ont l'avantage de réduire le temps de passation, ce qui n'est pas négligeable dans certains contextes, hors expérimentation ou évaluation clinique approfondie. La troisième piste consiste à tout faire (les deux paires en identification et en discrimination) mais dans l'ordre exact qui a servi aux normes. Cela permettrait d'observer différentes mesures, optimales pour chaque épreuve.

#### **4.8.2. Différence de PC entre dyslexiques et normo-lecteurs à l'âge adulte : une réalité ?**

La plainte des francophones quant à leurs difficultés d'apprentissage de l'anglais n'est pas dénuée de fondement et particulièrement en cas de dyslexie. Dans cette étude (Expérience 5), les participants adultes (dyslexiques ou non) avaient été sélectionnés pour avoir étudié l'anglais de façon scolaire sans dépasser le niveau B2 à l'université, ce qui est un faible niveau. Nous avons aussi noté une forte variabilité inter-individuelle dans le groupe dyslexique quant à leurs performances en PC anglaise.

En l'absence de données venant d'un groupe d'anglophone natif, nous avons simplement comparé les performances en anglais, à T1, entre les groupes d'adultes dyslexiques et normo-lecteurs. Il n'y avait pas de différence en identification, mais en discrimination : les adultes dyslexiques se sont montrés moins sensibles à la frontière catégorielle, leur pic de discrimination était moins élevé que celui des normo-lecteurs. Ce résultat est en accord avec l'étude de Crombie (1997) insistant sur les difficultés de discrimination d'un son nouveau dans une L2 chez des enfants dyslexiques. Il s'agit de l'argument le plus fort dans l'expérience 5, en faveur d'une difficulté accrue en apprentissage de l'anglais, en cas de dyslexie, cette fois chez l'adulte, sur le plan phonologique. La tâche de discrimination de nouveaux phonèmes semble plus sensible à la différence entre dyslexiques et normo-lecteurs que la tâche d'identification. Des données objectives sur ces difficultés en apprentissage d'une L2 chez des adultes dyslexiques sont difficiles à apporter (Soroli *et al.*, 2010). Le résultat obtenu dans notre étude (Expérience 5) peut contribuer à légitimer les plaintes des apprenants atteints de cette pathologie (en français, Abadie & Bedoin, 2016 ; Mazur-Palandre *et al.*, 2016), et les aides qui devraient être mises en place pour rétablir une meilleure égalité des chances.

D'autres comparaisons ont permis de révéler des différences dans la façon dont les deux groupes traitent le contraste de voisement selon la langue. Pour la paire /g-k/, les normo-lecteurs montrent seulement une hauteur de pic de discrimination différente entre l'anglais et le français, tandis que les dyslexiques montrent seulement un seuil d'identification différent entre ces langues. Pour la paire /d-t/, la pente d'identification en anglais des normo-lecteurs est plus abrupte qu'en français tandis que les dyslexiques ne montrent aucune différence entre leurs performances en français et en anglais. Les deux groupes réagissent différemment aux variations du contraste du voisement selon la langue. Cela témoigne de différences en perception de la parole. Dans la langue la moins familière pour eux, il apparaît que les normo-lecteurs soit plus sensibles à l'intensité de la frontière catégorielle. Dans cette langue, les participants dyslexiques ont quant à eux, montré une plus grande sensibilité à une différence de position de la frontière, mais ils ne l'ont pas perçue pour ce qu'elle est : ils ont fait l'erreur de percevoir un seuil anglais au niveau d'un VOT plus court qu'en français, et ce devrait être le contraire. Leurs réponses montrent qu'ils sont sensibles à une différence de seuil mais, ils n'en ressortent pas une frontière correcte en L2. Leurs réponses placent la frontière française en moyenne à +20 ms au lieu de 0 - +10 ms et la frontière anglaise en moyenne à +12 ms au lieu de + 25 ms. Ce dernier résultat est tout à fait cohérent avec la

littérature scientifique sur la PC chez des dyslexiques, qui décrit chez eux des frontières catégorielles souvent floues et instables (Serniclaes *et al.*, 2001 ; Noordenbos & Serniclaes, 2015).

Il faut toutefois reconnaître que l'Expérience 5 ne montre pas de différence massive entre les dyslexiques et les normo-lecteurs en anglais à T1. Cela montre que leur trouble du langage n'impacte pas beaucoup plus l'apprentissage de la phonologie anglaise qu'en l'absence de dyslexie. Cela s'explique sans doute en partie par le fait que les participants dyslexiques inclus dans l'étude avaient tous suivi un cursus universitaire et reçu un apprentissage standard de l'anglais dans leur parcours scolaire comme leurs pairs normo-lecteurs. Les dyslexiques réussissant à atteindre les études supérieures ont su développer de fortes compensations par rapport à leur trouble (Colé *et al.*, 2020). La différence entre leurs performances et leur ressenti peut trouver une explication dans la forte conscience qu'ils ont de leurs difficultés avec les sons du langage, que les adultes sans pathologie n'ont pas. Les normo-lecteurs surestiment probablement leurs performances en anglais, là où des étudiants dyslexiques, moins confiants en eux, sous-estiment leur niveau. Enfin, les étudiants dyslexiques sont aussi connus pour avoir un haut niveau d'anxiété (Plaisant, 1989) pouvant contribuer à une perception exagérée de leurs difficultés dans certains domaines, en particulier lorsqu'il s'agit de langage.

## **Chapitre 5. Aide à l'apprentissage de la perception du voisement anglais**

---

À cette étape de la thèse, nous disposons d'épreuves de perception catégorielle (PC) en anglais qui ont montré une insensibilité à la répétition (absence d'effet test-retest) ainsi qu'une bonne fidélité (Expérience 4), et le système phonologique français des participants dyslexiques a pu être stabilisé (Expériences 2 et 3). Comme recommandé par Flege et Bohn (2021), les participants semblent de ce fait dans de bonnes dispositions pour l'apprentissage d'une L2. L'étude présentée dans ce chapitre a alors permis d'évaluer l'effet d'un entraînement de 5 jours (30 min/jour) utilisant la stimulation dichotique pour l'apprentissage du voisement des consonnes anglaises chez des adultes francophones normo-lecteurs et dyslexiques. Cet effet a été mesuré au niveau comportemental par l'utilisation des épreuves de PC précédemment évaluées et au niveau neuronal par l'enregistrement EEG permettant des mesures en potentiels évoqués.

### **5.1. Dispositifs pour l'apprentissage de nouvelles consonnes en L2**

Bien qu'ils soient brefs, certains entraînements dits intensifs en L2 s'avèrent efficaces pour que des adultes apprennent ou maîtrisent mieux le rôle phonologique d'indices acoustiques absents du système phonologique de leur L1 ou dont l'aspect est inhabituel. De nombreuses expériences d'entraînement conduites auprès d'apprenants japonais ont par exemple porté sur l'opposition anglaise /r-l/, absente du système initial du japonais en L1. D'autres expériences portent sur la prise en compte de nouveaux indices de durée pour le voisement des consonnes.

#### **Aide à l'apprentissage de nouveaux phonèmes par des entraînements courts**

Pour apprendre une différence phonémique n'impliquant pas d'indices acoustiques temporels (e.g. contraste /r-l/ anglais), Callan *et al.* (2003) ont par exemple montré que l'entraînement d'adultes japonais pendant 1 mois (45 sessions d'environ 15 min), exerçant l'identification de ces phonèmes avec un feedback auditif immédiat (présentation d'un son plaisant après une réponse correcte et d'un son déplaisant en cas de réponse incorrecte), améliorait les performances en identification perceptive. Ce progrès évalué à partir de la même tâche et des mêmes phonèmes que ceux de l'entraînement n'est pas très surprenant. La réorganisation corticale fonctionnelle des aires temporeles supérieures et médianes observée grâce à l'IRMf après l'entraînement l'est davantage. La plasticité cérébrale fonctionnelle des adultes s'est aussi traduite par l'activation d'un réseau cérébral plus large après l'entraînement



qu'avant celui-ci, avec l'implication bilatérale du gyrus supra-marginal, du planum temporale, de l'aire de Broca et des aires prémotrices et motrices supplémentaires.

Hazan *et al.* (2005) ont travaillé sur la même paire de consonnes auprès de Japonais apprenant l'anglais avec un dispositif encore plus court, mais des séances plus longues : 10 séances de 40 minutes. Ils ont comparé deux types d'entraînement basés sur la haute variabilité phonétique des stimuli. L'un fournissait des informations seulement en modalité auditive, l'autre était audio-visuel et le mot entendu était dit par un agent conversationnel visible (la tête parlante « Baldi », Massaro, 1998, cité par Hazan *et al.*, 2005). Les chercheurs ont montré que, pour ce contraste, une amélioration des performances survenait après chaque entraînement, mais de manière accentuée s'il était audio-visuel. Bon nombre de travaux sur l'acquisition de phonèmes anglais montrent qu'offrir à l'apprenant la possibilité de voir le visage (généralement naturel) parler améliore l'apprentissage (pour une revue, voir Inceoglu, 2015). Par exemple, dans le contexte anglais/espagnol, Ortega-Llebaria *et al.* (2001) ont décrit l'amélioration de la perception des consonnes anglaises (/b, d, g, p, t, k, v, z, ô, f, s, dz, m, n/ etc.) par des étudiants espagnols à la suite d'un entraînement audio-visuel utilisant l'agent conversationnel de la technologie Baldi. Ici, Baldi guidait le participant à travers les tâches et prononçait les mots de l'entraînement en rendant visibles les mouvements du bas de son visage. L'information audio-visuelle favorisait aussi un transfert de compétences vers de nouveaux mots, d'autres locuteurs, et vers la production (Hardison, 2003). Cela s'expliquerait fondamentalement par le fait que 80% des variations dans le tractus vocal du locuteur peuvent être extraites en voyant son visage (Jiang *et al.*, 1998).

### **Aide à l'apprentissage de nouveaux phonèmes basée sur la variété des exercices**

Faire varier les types d'exercices est un autre moyen de garantir la profondeur des effets d'un entraînement purement auditif. Ainsi, Shinohara et Iverson (2018) ont montré que l'entraînement de Japonais à l'opposition anglaise /r-l/ à partir d'exercices d'identification mais aussi de discrimination permettait une amélioration de la précision non seulement en perception mais aussi en production, à la suite de seulement 10 sessions de 30 minutes. Un tel transfert de compétence de la perception à la production témoigne d'un apprentissage relativement profond. Plusieurs recherches ont montré que des entraînements perceptifs peuvent donner lieu à ce transfert pour des consonnes (Alves & Luchini, 2017), des voyelles (Lambacher et al, 2002 ; Thomson, 2011), ou même des tons du mandarin (Wang *et al.*, 2003).

Sur le plan théorique, le transfert d'apprentissage de la perception à la production est un argument en faveur de la théorie motrice du traitement de la parole. Les exercices, en particulier lorsqu'ils offrent une haute variabilité des informations (plusieurs voix, plusieurs contextes constitués par les autres phonèmes), inciteraient les apprenants à réaliser des ajustements perceptifs à des représentations de gestes (*gestural features*). Ces derniers sont des éléments plus stables que l'information auditive à propos des phonèmes. L'activation des représentations de gestes serait bénéfique à la production d'une réponse orale (Kangatharan *et al.*, 2021). Ainsi, selon la Théorie Motrice de la Perception de Parole (Lieberman, 1967, 1985, pour une synthèse voir NGuyen, 2005), perception et production s'appuieraient sur un module de représentation commun lié aux commandes articulatoires (Lieberman & Mattingly, 1989) et au système cérébral moteur (Galantucci *et al.*, 2009). Le transfert perception-production est compatible avec cette théorie et suggère ici un apprentissage profond, basé sur l'élaboration de nouvelles catégories mentales (Grenon *et al.*, 2019). Une réorganisation de l'espace acoustique-phonétique utile à la perception et à la production de parole se serait produite (Kangatharan *et al.*, 2021).

La profondeur de l'apprentissage dépendrait aussi du choix des exercices. Contrairement à l'hypothèse de la supériorité des exercices d'identification sur les exercices de discrimination, des données montrent que ces derniers peuvent permettre l'établissement de représentations phonologiques particulièrement robustes et persistantes (Flege, 1995). L'étude de Shinohara et Iverson (2018) a permis de préciser que le transfert était favorisé si les exercices ne se limitaient pas à d'identification. Des progrès particulièrement élevés survenaient en effet lorsque des exercices d'identification et de discrimination étaient combinés dans l'entraînement.

### **Aide à l'apprentissage de nouveaux phonèmes par la haute variabilité des informations**

La haute variabilité phonétique des stimuli présentés pendant l'apprentissage serait un gage de réussite pour les entraînements perceptifs purement auditifs. Ils se réclament d'une approche réaliste (*direct-realistic approach*). Les entraînements perceptifs à haute variabilité (*High Variability Phonetic Training*, HVPT) constituent un mode d'entraînement à part entière : généralement, plusieurs voix sont utilisées pour les stimuli des exercices, avec de nombreux mots dans des contextes phonétiques eux aussi variés, et un feed-back est donné en réponse aux décisions. Le succès de ce type d'entraînement s'est révélé par exemple pour l'apprentissage de la différence anglaise /r-l/ par des Japonais, avec un transfert vers de nouveaux mots et vers la voix de nouveaux locuteurs (Lively *et al.*, 1993). Même avec un

entraînement très court (45 minutes réparties sur 3 semaines), cette technique s'est montrée efficace pour des oppositions de consonnes anglaises non basées sur des indices de durée /b/-/v/ et /d/-/ð/ chez des adultes bilingues espagnol/catalan (Cebrian & Carlet, 2014). La perception de mots opposés pour ces consonnes s'est améliorée, et un transfert s'est produit vers des mots non entendus pendant l'entraînement, autre critère intéressant pour un apprentissage non superficiel. Toujours avec un programme très court (5 × 15 min par semaine, pendant 4 semaines), Hwang et Lee (2015) ont rapporté une amélioration de la perception de consonnes anglaises absentes en coréen (/p/-/f/) par de jeunes apprenants coréens (11-13 ans) à la suite d'un entraînement en perception où les exemples étaient choisis pour leur haute variabilité. De même, lorsque l'anglais est appris tardivement, Bradlow *et al.* (1997) ont montré que la perception de consonnes nouvelles pour les apprenants de cette L2 s'améliore grâce à un entraînement HVPT : l'identification mais aussi la production des consonnes /r/-/l/ gagne alors en qualité. Au contraire, en cas de faible variabilité, l'entraînement ne garantit pas la généralisation de l'opposition /r-l/ par des Japonais à des mots produits d'une manière un peu différente (Strange & Dittman, 1984). Des comparaisons directes entre des groupes ayant réalisé un entraînement avec une variabilité élevée ou faible confirment des progrès plus élevés et davantage de généralisations pour un contraste entre des consonnes japonaises chez des apprenants néerlandais si la variabilité était forte (Sadakata & McQueen, 2013).

#### **Aide à l'apprentissage de nouveaux phonèmes par accentuation des traits phonétiques**

Une autre technique consiste à présenter au début des exercices des exemplaires avec des caractéristiques phonétiques exagérées, puis des exemplaires de plus en plus naturels pour une paire de phonèmes. Pour l'apprentissage du contraste anglais /r-l/ par des Japonais, elle accroît les compétences en identification et en discrimination (même en l'absence de feedback), davantage qu'une technique présentant uniquement des stimuli difficiles (McClelland *et al.*, 2002), mais sans généralisation à un autre continuum faisant varier le même contraste (McCandliss *et al.*, 2002).

#### **Aide à l'apprentissage de nouveaux phonèmes intégrant la production et le retour visuel**

L'entraînement peut aussi porter directement sur la production. Il exploite alors souvent le retour audio-visuel articulatoire. Des locuteurs chinois qui prononçaient le /r/ français en assimilant son lieu d'articulation à celui du /x/ en mandarin ont ainsi réalisé des exercices de prononciation accompagnés de deux types d'informations (Wu *et al.*, 2015). Un groupe recevait une description explicite de la manière de prononcer le /r/ français, comme dans une

pratique scolaire. Les participants de l'autre groupe recevaient une illustration en temps réel recueillie par ultra-sons : une courte vidéo présentant l'image dynamique de la position de leur propre langue lors de leur prononciation. Ils estimaient ainsi en temps réel si la position de leur langue était comparable à celle d'un locuteur français. Ce retour visuel pour faciliter l'ajustement articulatoire a généré les plus grands progrès en production.

Avec l'électropalatographie – une autre technique utilisant un retour visuel sur l'information articulatoire –, Hacking *et al.* (2017) ont entraîné des étudiants anglophones apprenant le russe. Cette technique visualise le contact de la langue avec le palais pour étudier le déplacement des zones de contact pendant la production de parole. L'entraînement ciblait le contraste entre des consonnes palatales et non-palatales. Les apprenants répétaient des segments, mots ou phrases selon la session, en tentant de produire une articulation proche du modèle visuel (informations sur la position de la langue et le palais d'un locuteur russe). Le dispositif comptait deux périodes de deux semaines, avec  $8 \times 15$  minutes d'entraînement pour chaque période. Les sessions étaient donc brèves (15 min) et l'ensemble court (4 heures cumulées) pour une prise en charge d'1 mois. Les résultats ont révélé des niveaux d'efficacité hétérogènes pour cet entraînement : malgré un effet global significatif, certains participants ne se sont pas améliorés. D'après les auteurs, l'absence d'amélioration significative chez ces derniers pourrait être due à une période d'entraînement trop courte (seulement 4 semaines), ou à la période pendant laquelle il s'est déroulé (fin d'un semestre universitaire, où les participants n'avaient plus de cours et désinvestissaient peut-être leurs apprentissages). Il existe un argument en faveur de la première interprétation, car Schmidt (2012) avait utilisé avant eux cette technique pendant une période plus longue, afin d'aider des Coréens pour l'apprentissage phonologique et articulatoire de contrastes entre des consonnes anglaises proches pour leur lieu d'articulation (i.e. /s/ - /ʃ/, /z/ - /dʒ/, et /l/ - /ɹ/). Une amélioration de la production de ces contrastes avait été observée avec une homogénéité satisfaisante. Les participants réalisaient un entraînement beaucoup plus long que celui d'Hacking *et al.* : 10 heures étalées sur 10 semaines, contre seulement 4 heures sur 1 mois dans le protocole d'Hacking *et al.*. En réduisant drastiquement la durée (moins de la moitié), Hacking *et al.* ont sans doute pénalisé des participants dont le rythme d'apprentissage est globalement lent, d'où une forte hétérogénéité dans l'efficacité de leur très court programme. Une limite vraisemblable des entraînements brefs est sans doute l'inégalité de leurs effets imputable à la rapidité d'apprentissage propre à chacun.

### **Aide à l'apprentissage de nouveaux phonèmes opposés par des indices temporels**

Les entraînements phonologiques brefs ciblent parfois des indices phonologiques temporels. Par exemple, Menning *et al.* (2002) ont proposé à des adultes germanophones apprenant le japonais seulement 10 jours d'entraînement (mais 1h30/jour) à la discrimination de la mora. En magnétoencéphalographie (MEG), ils ont recueilli l'indice d'une meilleure sensibilité (MMF) à la mora (division temporelle particulière de la structure des mots), ainsi que de meilleures performances en détection explicite de différences de durées de consonnes et de voyelles dans des paires minimales.

Un entraînement phonologique encore plus bref (3 jours) a été conduit sur la différence de voisement des fricatives anglaises – largement basée sur le délai de voisement (VOT) – auprès de locuteurs du finnois, langue n'ayant pas ce contraste sur les fricatives (Tamminen *et al.*, 2015). L'entraînement sans feed-back associait cette fois perception et production. Absente avant l'entraînement, la MMN traduisant la sensibilité pré-attentionnelle à ce trait est apparue dès le deuxième jour à l'écoute de fricatives déviantes par leur voisement dans une série de fricatives standards.

Shum *et al.* (2021) offrent un récent exemple d'entraînement particulièrement bref et destiné à mieux faire prendre en compte des indices de durée pour percevoir le voisement de consonnes anglaises par des apprenants chinois, dont la L1 ne requiert pas ce traitement phonologique. Dans l'entraînement, les exemples présentaient une grande variabilité perceptuelle. Les chercheurs ont montré une amélioration de la perception des consonnes occlusives anglaises par les apprenants chinois, les tests portant toutefois seulement sur les phonèmes perçus lors de l'entraînement. Les exercices étaient particulièrement brefs, mais largement répartis dans le temps : 20 séances de 5 minutes de perception auditive de la voix d'anglophones natifs réparties sur 4 à 6 semaines. Cette étude s'est déroulée avec succès auprès de différents groupes d'âge : enfants, adolescents, et jeunes adultes.

Toujours pour l'apprentissage de la différence de voisement, un exercice de production sans modèle audio de seulement 50 minutes a été testé avec un retour visuel représentant la durée du VOT des consonnes (spectrogramme) produites (Olson, 2019). Les apprenants (étudiants anglophones apprenant l'espagnol) pouvaient comparer visuellement leur production à celle de locuteurs natifs. Les consonnes espagnoles et anglaises sont opposées par le voisement selon des VOTs différents. La production des différences de VOT pour les consonnes entraînées s'est améliorée, avec transfert du bénéfice à des consonnes non-entraînées. Le dispositif d'entraînement, assez sophistiqué et combinant production et

observation d'un retour visuel, participait au succès du programme : il compensait sa brièveté, ainsi que l'absence de distribution dans le temps. En 2021, Olson a tenté de répliquer ces effets, mais l'entraînement comportait cette fois deux phases et il était temporellement plus distribué. La première phase durait 25 minutes pendant 1 heure de classe de langue. La seconde se déroulait à domicile pendant 3 jours : le participant refaisait les exercices en vue de s'améliorer. La diminution de la durée du VOT attendue pour les consonnes sourdes s'est avérée significative après l'entraînement, mais seulement pour le matériel étudié, sans transfert de compétence vers d'autres mots non-entraînés. Le succès plus mitigé de ce type de dispositif dans cette étude peut être interprété de deux façons. Cet entraînement est peut-être plus efficace quand il est massé – sous la forme d'une longue séance de 50 mn – plutôt que dispersé en petites séances. Il est aussi possible que l'objet de l'apprentissage soit trop subtil sur le plan temporel dans cette deuxième étude. En effet, une réduction de durée d'un VOT positif était attendue et non une simple différence entre VOTs positif (pour les sourdes) et négatif (pour les sonores). Réduire la durée du VOT pour les consonnes sourdes espagnoles consiste pour les apprenants anglophones à ne plus produire d'aspiration pour de telles consonnes dans cette L2. Cela semble plus difficile à apprendre (ou à réaliser sur le plan articulatoire) que de produire une vibration des plis vocaux avant l'explosion dans une sonore espagnole. S'empêcher de produire un trait familier pourrait davantage résister à un entraînement que s'habituer à produire un trait conçu comme supplémentaire.

## 5.2. Objectif

Il existe dans la littérature scientifique de nombreuses études sur les entraînements perceptifs pour aider à l'apprentissage d'un nouveau contraste de consonnes dans une seconde langue. Certains entraînements perceptifs ont montré leur efficacité malgré leur brièveté. De telles études sont cependant rares au sujet des adultes francophones apprenant l'anglais.

Le cas du couple anglais/français présente pourtant une particularité intéressante concernant les manières dont des différences de durée d'un indice phonétique – le VOT – fondent l'opposition entre les consonnes occlusives sonores et sourdes dans ces langues : alors que des VOTs longs et courts opposent respectivement les occlusives sonores et les occlusives sourdes françaises, c'est l'inverse en anglais. Or, selon le modèle AST (Poeppl, 2003), les indices acoustiques longs et courts sont mieux traités respectivement dans les cortex auditifs droit et gauche. Pour que les francophones traitent mieux le voisement des consonnes anglaises, il semble pertinent de proposer un dispositif d'aide à l'apprentissage les conduisant

à investir prioritairement l'HG dans le traitement des consonnes sonores anglaises et l'HD dans celui des occlusives sourdes anglaises. Cela suppose de se détacher d'un mode d'apprentissage de la phonologie anglaise calqué sur le traitement phonologique français, pour passer à un mode de traitement phonétique où les caractéristiques acoustiques des sons de parole sont prises en compte. Il faut pour cela abandonner l'analogie avec les phonèmes français, largement suscitée par l'écrit. En effet, ce qui commence par la lettre B en anglais est phonétiquement proche du [p] français qui s'écrit avec la lettre P, et ce qui s'écrit avec le P en anglais est une consonne aspirée, phonétiquement éloignée du [p] français comme de tout autre phonème français. L'objectif est d'aider à traiter correctement, et de la manière la plus judicieuse possible en termes d'investissement hémisphérique, la longue aspiration du [p<sup>h</sup>] anglais et le court VOT du [b] anglais pour ce qu'ils véhiculent linguistiquement (des consonnes respectivement sourdes et sonores).

L'écoute dichotique pourrait apporter une réponse à ce besoin. Dans l'Expérience 6, une nouvelle méthode de stimulation basée sur ce principe est proposée. Sans que l'apprenant soit informé de la façon dont les sons lui parviennent dans un casque audio, il est convié à apprendre du vocabulaire anglais. Il entend pour cela des mots prononcés par un anglophone, tout en voyant une image illustrant son signifié et le plus souvent sa forme orthographique. Le dispositif consiste à diffuser le mot anglais seulement dans l'oreille gauche s'il commence par une consonne occlusive anglaise sourde, et seulement dans l'oreille droite s'il commence par une consonne occlusive anglaise sonore. La présentation simultanée d'un bruit neutre (bruit blanc) de même intensité dans l'autre oreille est destinée à créer un phénomène de concurrence pour renforcer le traitement privilégié de la consonne par l'hémisphère controlatéral, toujours choisi pour être le plus efficace dans le traitement d'indices temporels longs ou courts. Si cette nouvelle aide, très ciblée, s'avère efficace chez l'adulte, il conviendra de discuter ses fondements cognitifs et neurolinguistiques. Sur le plan cognitif, ce dispositif mise sur un apprentissage du meilleur moyen de traiter phonétiquement les consonnes anglaises sur le plan neuronal : il s'agit d'une alliance entre approches cognitive et neurolinguistique.

Le premier objectif de cette partie de la recherche (Expérience 6) était de tester si la stimulation dichotique constitue un apport significatif par rapport à la stimulation binaurale classique dans un apprentissage perceptif. Les critères de jugement étaient pour cela, au niveau comportemental, une perception plus catégorielle de consonnes occlusives anglaises opposées en voisement après un entraînement dichotique plutôt que binaural et, au niveau



neuronal, une sensibilité pré-attentionnelle davantage accrue si l'entraînement était dichotique plutôt que binaural. Le second objectif était de tester si cette technique d'aide à l'apprentissage phonologique était efficace chez des adultes atteints de dyslexie (Expérience 8).

Afin de répondre à ces objectifs, quatre expériences ont été conduites chez des adultes francophones apprenant l'anglais. L'**Expérience 6** comparait l'effet de deux entraînements brefs de 5 jours (30 min/jour) centrés sur les consonnes occlusives, l'un utilisant la stimulation dichotique et l'autre la stimulation binaurale, sur la perception catégorielle d'occlusives sonores et sourdes anglaises. L'**Expérience 7** évaluait l'effet de l'entraînement par stimulation dichotique au niveau neuronal, par l'enregistrement de potentiels évoqués en écoute passive (paradigme de *oddball*). L'**Expérience 8** évaluait ensuite l'effet de l'entraînement par stimulation dichotique sur la perception catégorielle anglaise chez des adultes dyslexiques. Enfin, l'**Expérience 9** évaluait l'effet potentiel de l'entraînement par stimulation dichotique en anglais sur la perception catégorielle en français, c'est-à-dire sur la L1 des participants normo-lecteurs et dyslexiques, avant tout pour vérifier que l'entraînement proposé en anglais n'avait pas d'impact négatif sur la phonologie dans la langue maternelle.

### **5.3. Expérience 6 : effets des entraînements par stimulation dichotique vs. binaurale sur la PC anglaise chez des adultes normo-lecteurs**

#### **5.3.1. Principe de l'expérience et hypothèses**

Le but de l'Expérience 6 était de comparer l'effet de deux entraînements en anglais (l'un par stimulation dichotique, l'autre par stimulation binaurale) sur la perception catégorielle (PC) des consonnes occlusives anglaises chez des adultes francophones sans pathologie.

Deux hypothèses opérationnelles ont été formulées :

- 1) Les deux entraînements devraient améliorer la PC (identification et discrimination) des consonnes occlusives ayant été entendues lors de l'entraînement, témoignant d'un effet d'entraînement perceptif, et cette amélioration pourrait être plus élevée après l'entraînement dichotique plutôt que binaural.
- 2) L'entraînement par stimulation dichotique devrait améliorer la perception catégorielle de consonnes occlusives n'ayant pas été entendues pendant l'entraînement, alors que l'entraînement par stimulation binaurale ne devrait pas permettre un tel transfert.



## 5.3.2. Méthode

### 5.3.2.1. Participants

Quarante adultes normo-lecteurs (34 femmes, 6 hommes) âgés de 18 à 29 ans ( $M = 24.1$  ans,  $ET = 1$  ans et 3 mois) ont participé à l'Expérience 6. Tous étaient monolingues français, avec une vue normale ou corrigée, et ils ne présentaient aucun déficit auditif connu. Ils déclaraient n'avoir à leur connaissance aucun trouble neuro-développemental et avaient étudié l'anglais, mais leur niveau ne dépassait pas B2 (Cadre Européen Commun de Références pour les Langues, CECRL) d'après une évaluation universitaire effectuée l'année du testing. Tous les participants étaient droitiers ( $\geq 80\%$  de réponses « droite » dans la version abrégée de l'*Edinburgh Handedness Inventory* (Oldfield, 1971)) et présentait un REA (Bedoin *et al.*, 2010).

Les participants ont été répartis aléatoirement dans deux groupes (20 participants chacun) : dans l'un, ils réalisaient un entraînement perceptif utilisant la stimulation dichotique (Groupe Dichotique, GD), dans l'autre ils réalisaient un entraînement perceptif contrôle où les sons étaient diffusés de façon binaurale (Groupe Binaural, GB).

### 5.3.2.2. Matériel et équipement

La perception catégorielle a été testée à l'aide du matériel décrit dans les Expériences 1 et 4 du Chapitre 2 partie 2.10.2 et du Chapitre 4 partie 4.6.2.3.

Les exercices de l'entraînement en anglais ont été créés en 2016 (Moulay, 2016). Initialement programmés avec E-Prime (Psychology Software Tools, Inc., Sharpsburg, PA), ils ont eux aussi été reprogrammés avec OpenSesame (Mathôt, S., *et al.*, 2012) pour faciliter leur diffusion auprès des participants (licence gratuite, utilisation possible sur PC et Mac). Chacune des 860 images était libre de droits. Les exercices étaient composés de 172 mots monosyllabiques anglais formant des paires minimales (e.g. *gall/call*, 69%) et des paires quasi-minimales variant par la consonne initiale, mais aussi la consonne finale (e.g., *gaze/case*, 31%). Ils commençaient tous par une consonne occlusive sourde ou sonore avec un lieu d'articulation vélaire (/k-g/) ou bilabial (/p-b/). Ils ont été enregistrés dans une salle insonorisée et anéchoïque par un locuteur féminin adulte anglophone natif, à l'aide du logiciel ROCme! (Ferragne *et al.*, 2013). Leur niveau d'intensité a été égalisé. Les entraînements et tests réalisés avant et après l'entraînement étaient réalisés à partir de l'ordinateur personnel du participant, mais un ordinateur PC et un casque audio (Beyerdynamic DT 770 Pro) pouvaient lui être prêtés en cas de besoin.

### 5.3.2.3. Procédure d'évaluation

Afin d'évaluer l'efficacité de l'entraînement, il était demandé aux participants de réaliser à l'étape pré-entraînement (T1) et à l'étape post-entraînement (T2) quatre épreuves de perception catégorielle (i.e. deux en français, Expériences 1 Chapitre 2 partie 2.10.2, et deux en anglais, Expériences 4 dans le Chapitre 4 partie 4.6.2.3) à 5 jours d'intervalles (Fig. 39). Comme pour l'entraînement en français, les paires /g-k/ et /d-t/ étaient utilisées dans chacune des quatre épreuves de test. La paire /g-k/ permettait d'évaluer l'effet de l'entraînement sur la perception des consonnes sur lesquelles portaient les exercices. La paire /d-t/ permettait de mesurer un potentiel transfert de compétences vers une paire de consonnes non présentées pendant l'entraînement. Les tests étaient toujours présentés dans l'ordre suivant : discrimination en français, identification en français, discrimination en anglais et identification en anglais. La consigne insistait toujours sur la langue d'origine des stimuli et la paire /d-t/ était toujours présentée avant la paire /g-k/.



Figure 39. Procédure générale de l'Expérience 6

Les participants effectuaient aussi un test d'écoute dichotique en français (Bedoin *et al.*, 2010, 2011) permettant de mesurer la dominance hémisphérique pour le traitement d'indices acoustiques courts à l'écoute d'un matériel verbal (144 paires de mots monosyllabiques diffusés de manière parfaitement synchronisée, un tiers des couples différant seulement par le lieu d'articulation de leur consonne initiale). Les participants n'étaient pas informés que des mots différents leur étaient présentés dans l'une et l'autre oreilles (il leur était expliqué que le son de la voix du locuteur était volontairement altéré pour les besoins du test). Ils devaient focaliser leur attention sur un point placé en face d'eux, pour minimiser les biais attentionnels vers une partie de l'espace. Ils devaient répéter à haute voix le mot entendu et l'expérimentatrice notait les réponses. Aucun feed-back n'était donné mais des encouragements étaient systématiquement fournis lors des trois pauses. Dans les tests

d'écoute dichotique, il peut arriver que certains participants rappellent systématiquement un mot particulier quelle que soit l'oreille de présentation, ce qui constitue un biais pour une mesure précise de l'asymétrie hémisphérique (Voyer & Techentin, 2009). Les couples pour lesquels cela se produisait étaient systématiquement retirés avant le calcul de l'asymétrie, selon les recommandations de Wexler et Halwes (1983) et Zatorre (1989). L'indice lambda ( $\lambda = \ln\left(\frac{R+1}{L+1}\right)$ ) où  $\ln$  = logarithme naturel, R = nombre de réponses indiquant le mot de l'oreille droite, L = nombre de réponses indiquant le mot de l'oreille gauche) (Bryden & Sprott, 1981) a été calculé pour évaluer l'avantage de l'oreille droite (*Right Ear Advantage*, REA) traduisant un avantage de l'HG. Cet indice est considéré comme particulièrement fiable, car il ne dépend pas du niveau général d'efficacité (de Bode *et al.*, 2007 ; DiStefano *et al.*, 2004 ; Fernandez & Smith, 2000). Lorsqu'il est positif, le REA indique une dominance de l'oreille droite - HG, tandis qu'il indique une dominance de l'oreille gauche - HD (*Left Ear Advantage*, LEA) s'il est négatif.

#### 5.3.2.4. Procédure d'entraînement

Les entraînements étaient réalisés pendant 5 jours (30 min/jour), en autonomie au domicile. Les exercices portaient sur les oppositions /k-g/ et /p-b/. Ils étaient présentés au participant comme des exercices d'apprentissage de vocabulaire. Ils étaient de deux types : exercices d'apprentissage lexical et exercices de jugement portant sur les mots étudiés dans les exercices d'apprentissage lexical. Le jugement plaçait l'apprenant face à un choix forcé portant sur le voisement de la consonne initiale de mots formant des paires minimales. Les jours 1-2-4-5 du protocole, le participant réalisait un exercice d'apprentissage lexical et un exercice de jugement. Le jour 3, il faisait deux exercices de jugement. Les participants devaient toujours porter le casque audio avec l'écouteur droit sur l'oreille droite et l'écouteur gauche sur la gauche (cette consigne lui était rappelée chaque jour par téléphone). Ils n'étaient pas informés qu'ils recevraient des informations différentes dans les deux oreilles, mais les exercices étaient programmés de sorte qu'un mot commençant par une occlusive sonore anglaise soit toujours présenté à l'oreille droite-HG et qu'un mot commençant par une sourde anglaise le soit à l'oreille gauche-HD. A la fin des 5 jours d'entraînement, le participant nous envoyait par mail les fichiers réponses des exercices d'entraînement. La date des fichiers permettait de vérifier que les exercices avaient bien été réalisés les jours convenus. L'examen des temps de réponse et la vérification des touches ayant servi aux réponses permettaient de vérifier que des réponses avaient bien été données et ceci en respectant la consigne pour les

touches. Si aucune réponse n'était donnée au bout de 5000 ms, l'exercice passait automatiquement à l'item suivant, et la mention NA apparaissait sur le fichier-réponse, permettant d'exclure des participants n'ayant pas vraiment réalisé l'entraînement. Aucun participant n'a été exclu en respectant ces critères.

Lors de l'**Exercice d'apprentissage de vocabulaire**, le participant entendait un mot à travers le casque, et voyait à l'écran sa représentation orthographique, sa traduction française ainsi qu'une image sémantiquement liée. Chaque mot était présenté deux fois de suite. Les mots donnés à apprendre (172 au total) changeaient tous les jours. La partie gauche de la figure 40 illustre ce type d'exercices.

Dans l'**Exercice de jugement 1**, l'apprenant entendait un des mots anglais appris le jour même, puis voyait apparaître à l'écran simultanément une image sémantiquement liée au mot entendu et deux mots anglais écrits l'un en dessous de l'autre (paire minimale). Il devait choisir entre les deux mots écrits celui qui correspondait au mot entendu et à l'image en appuyant sur l'une des deux touches de réponse (haut vs. bas). Un feedback visuel le renseignait sur ses erreurs. Chaque mot faisait l'objet de plusieurs jugements, et il était présenté avec l'image utilisée lors de l'apprentissage ou avec d'autres images, pour stimuler l'intérêt et la motivation. La partie centrale de la figure 40 illustre cet exercice.

Dans l'**Exercice de jugement 2**, le participant entendait un mot anglais puis voyait deux images (i.e. cette fois, il n'y avait plus de support écrit, ce qui augmentait progressivement le niveau de difficulté). Il devait choisir la touche du clavier (haut vs. bas) dont la position correspondait à celle de l'image représentant le mot entendu. La partie droite de la figure 40 illustre cet exercice.

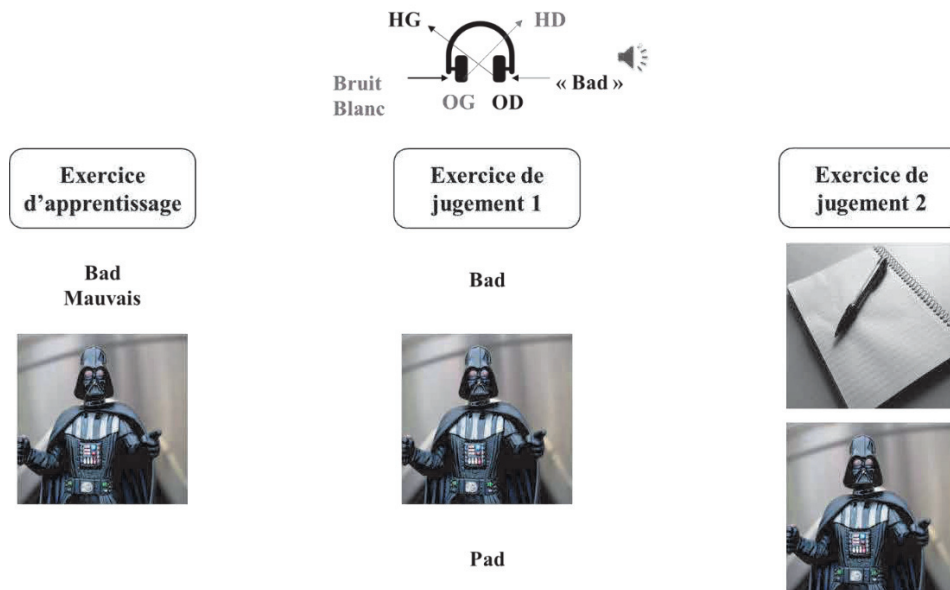


Figure 40. Déroulement des exercices d’entraînement concernant les consonnes anglaises en écoute dichotique, avec les trois types d’exercices proposés. Dans les trois types d’exercices, les stimuli auditifs étaient présentés selon le dispositif illustré en haut de la figure. L’entraînement contrôle binaural était identique excepté la présentation des sons (le mot était alors présenté dans les deux oreilles et il n’y avait pas de bruit blanc).

Pour le **groupe GD**, les exercices informatisés présentaient les mots au moyen de l’écoute dichotique dans un casque audio pour renforcer l’association durée du VOT – hémisphère le plus compétent pour traiter cette durée (i.e. les mots initialisés par une occlusive sonore anglaise présentés dans l’oreille droite – HG avec bruit blanc simultané dans l’oreille gauche ; les mots initialisés par une occlusive sourde anglaise présentés dans l’oreille gauche – HD avec bruit blanc simultané dans l’oreille droite). Pour le **groupe GB**, l’entraînement avait le même contenu que pour le groupe GD, aussi bien pour les mots que pour les images et les exercices (tâches), mais les mots étaient diffusés de manière binaurale, c’est-à-dire simultanément dans les deux oreilles (pas de bruit blanc).

Les fichiers de réponses des entraînements envoyés par les participants étaient examinés pour vérifier que le participant n’avait pas appuyé sur d’autres touches que celles désignées pour les réponses, et qu’il n’avait pas produit plus de quatre temps supérieurs à 3 écart-types calculés à partir de ses données. Il n’y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Dans de cas où les réponses étaient globalement très lentes ou particulièrement rapides, les données du participant n’étaient pas conservées si cette lenteur ou rapidité extrêmement ne concernait que T1 ou que T2, car cela pouvait témoigner d’un manque de sérieux lors de l’une des étapes. Aucun participant n’a été exclu en respectant ces critères.

### 5.3.3. Analyses des données

Les données des épreuves de perception catégorielle (PC) ont été traitées comme décrit pour les Expériences 1 et 4 dans les Chapitre 2, partie 2.10.4 et Chapitre 4 partie 4.6.3. Les mêmes traitements statistiques que ceux décrits dans le Chapitre 3 – Expérience 2, partie 3.3.3 ont été réalisés sur les résultats de cette expérience, avant et après l’entraînement.

### 5.3.4. Résultats

**Discrimination.** Avant l’entraînement (T1), les performances en discrimination (indice de sensibilité  $d'$ ) pour le couple de stimuli bornant le pic (VOT 0-20 ms) n’étaient pas significativement différentes entre les groupes GD et GB, et ce, pour les paires /gi-ki/ ( $t(36.68) = 0.97, p = .336$ ), et /di-ti/ ( $t(33.23) = 1.51, p = .140$ ).

Pour la paire /gi-ki/, qui faisait partie de l’entraînement, les performances en discrimination au niveau du pic (VOT 0-20 ms) s’amélioraient entre T1 et T2 dans le groupe GB ( $t(19) = 2.15, p = .045, d = .49$ ) (Fig. 41 à droite), ainsi que dans le groupe GD (Fig. 41, à gauche), avec un effet de plus grande taille dans ce dernier groupe qui avait reçu l’entraînement dichotique ( $t(19) = 2.42, p = .026, d = .57$ ). Cette progrès pour le groupe GD n’était pas dépendante du niveau initial des participants ( $r(18) = .13, p = .588$ ).

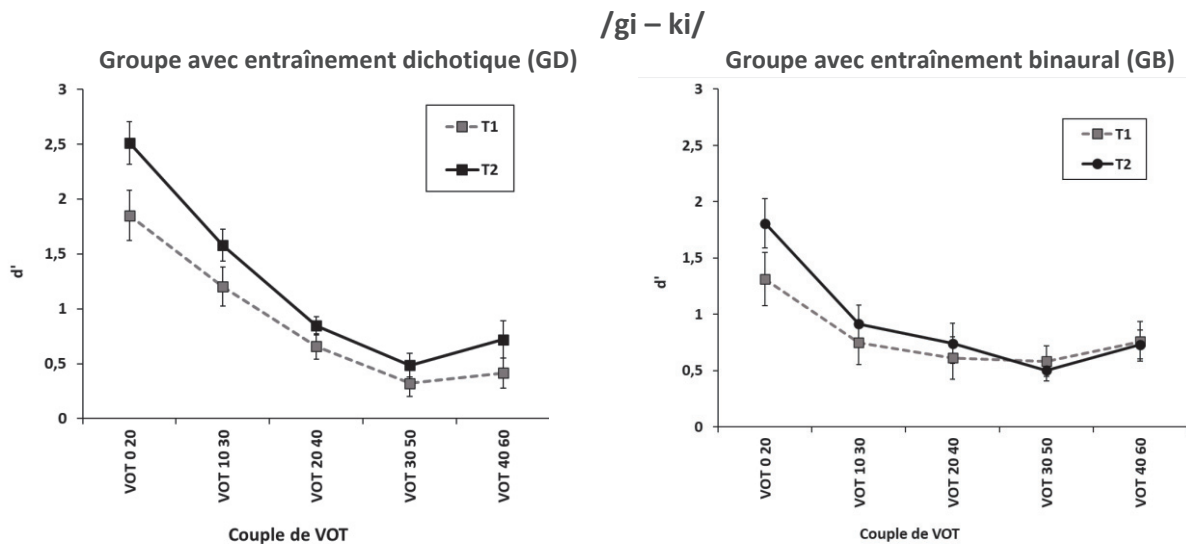


Figure 41. Courbes de discrimination (indice de sensibilité  $d'$ ) en anglais pour la paire /gi-ki/, dans les groupes GD (à gauche) et GB (à droite) entre T1 (pointillés) et T2 (trait plein). Les barres d’erreurs représentent l’erreur-type.

Comme illustré par la figure 41, après l’entraînement en anglais (T2), la pente de discrimination en partant du pic (VOT 0-20 ms) devenait plus abrupte qu’à T1, pour les deux groupes, car la sensibilité à la différence de VOT dans le couple de stimuli n’augmentait entre

T1 et T2 que pour le pic, et en aucun cas de manière significative dans les autres couples de stimuli (Tableau 7), ce qui traduisait un accroissement de la nature catégorielle de la perception avec l'émergence d'une frontière catégorielle plus nette.

Tableau 7. Résultats du test *t* de Student pour comparer les performances en discrimination réalisée entre T1 et T2 sur les couples de stimuli de la paire /gi-ki/, pour les groupes GD et GB. L'astérisque \* indique une différence significative. *ddl* = 19

Groupe	VOT 0-20	VOT 10-30	VOT 20-40	VOT 30-50	VOT 40-60
GD	$p = .045 *$	$p = .079$	$p = .234$	$p = .198$	$p = .138$
GB	$p = .026 *$	$p = .392$	$p = .611$	$p = .571$	$p = .889$

Pour la paire (/di-ti/) (Fig. 42), le pic de discrimination était situé 10 ms plus tard que pour la paire (/di-ti/), dans le sens d'un VOT positif. Pour cette paire de phonèmes, qui ne faisait pas partie de l'entraînement, l'augmentation de l'indice *d'* en discrimination entre T1 et T2 au niveau du pic (VOT 10-30 ms) était tout juste significative dans le groupe avec entraînement dichotique, GD ( $t(19) = 2.06, p = .053, d = .47$ ) (Fig. 42, à gauche), et pas dans le groupe avec entraînement binaural, GB ( $t(19) = 1.27, p = .220, d = .25$ ) (Fig. 42, à droite).

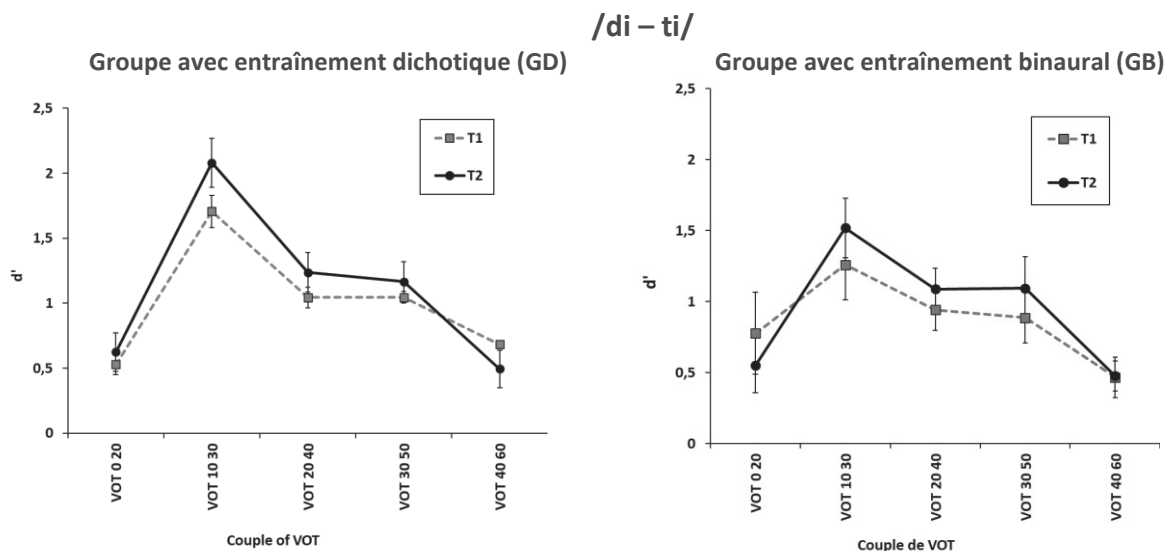


Figure 42. Courbes de discrimination en anglais pour la paire /di-ti/ dans les groupes GD (à gauche) et GB (à droite) entre T1 (pointillés) et T2 (plein). Les barres d'erreurs représentent l'erreur-type.

Aucune différence entre T1 et T2 n'était observée pour les autres couples de VOT, quel que soit le groupe (voir Tableau 8).

Tableau 8. Résultats du test *t* de Student pour comparer les performances en discrimination entre T1 et T2 sur les couples de stimuli de la paire /di-ti/, pour les groupe GD et GB. L'astérisque \* indique une différence significative et l'astérisque entre parenthèses (\*) indique une tendance. *ddl* = 19

Groupe	VOT 0-20	VOT 10-30	VOT 20-40	VOT 30-50	VOT 40-60
--------	----------	-----------	-----------	-----------	-----------

GD	$p = .542$	$p = .053(*)$	$p = .354$	$p = .548$	$p = .366$
GB	$p = .450$	$p = .220$	$p = .471$	$p = .358$	$p = .944$

**Identification.** Que ce soit pour les valeurs de seuil ou de pente, il n’y avait pas de différence de performances entre les deux groupes à T1, avant les entraînements (Tableau 9).

Tableau 9. Valeurs de  $p$  pour les comparaisons réalisées avec les tests  $t$  de Student entre les performances des deux groupes (GD vs. GB) en identification de phonème à l’étape précédant l’entraînement (T1).

	Seuil		Pente	
	<i>ddl</i>	<i>p</i>	<i>ddl</i>	<i>p</i>
/gi-ki/	29.88	.813	37.61	.799
/di-ti/	36.27	.561	33.21	.305

La valeur de la pente d’identification des deux paires de consonnes évaluées n’a pas changé significativement après l’entraînement, aussi lorsqu’il était dichotique (/gi-ki/ :  $t(19) = 0.49$ ,  $p = .629$ ,  $d = .12$  ; /di-ti/ :  $t(19) = 1.53$ ,  $p = .142$ ,  $d = .37$ ) que binaural (/gi-ki/ :  $t(19) = 0.46$ ,  $p = .651$ ,  $d = .12$  ; /di-ti/ :  $t(19) = 0.54$ ,  $p = .599$ ,  $d = .18$ ).

Pour la paire qui avait été entraînée (/gi-ki/), la valeur de seuil s’est quant à elle déplacée significativement entre T1 et T2, que ce soit dans le groupe entraîné de façon dichotique ( $t(19) = 3.53$ ,  $p = .002$ ,  $d = .71$ ) ou binaurale ( $t(19) = 3.50$ ,  $p = .002$ ,  $d = .74$ ) passant de 12.24 ms à 17.81 ms pour le groupe GD et de 11.71 ms à 15.41 ms pour le groupe GB (Fig. 43).

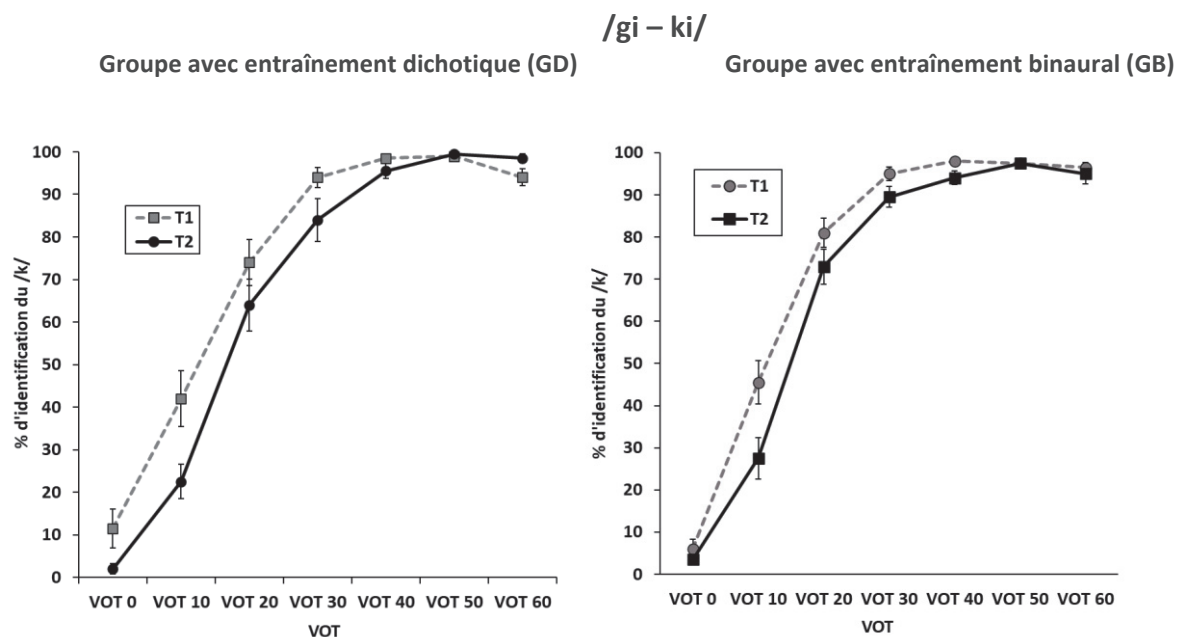


Figure 43. Courbes d’identification pour le groupe GD (à gauche) et GB (à droite) à T1 (pointillés) et T2 (plein) pour la paire /gi-ki/. Les barres d’erreurs représentent l’erreur-type.



L'observation de la figure 43 a révélé entre T1 et T2 une diminution du pourcentage de réponses /k/ pour le phonème anglais /g/ (i.e. diminution des erreurs d'identification du /g/ anglais, matérialisée par une diminution de hauteur de la partie gauche de la courbe, qui pourrait correspondre au début d'un plateau typique d'une perception catégorielle, au niveau du VOT 0 ms ( $t(19) = 2.17, p = .043, d = .63$ ) et VOT 10 ms ( $t(19) = 4.03, p < .001, d = .79$ ) pour le groupe GD, mais seulement au niveau du VOT 10 ms ( $t(19) = 4.41, p < .001, d = .76$ ) pour le groupe GB. Plutôt qu'un début de plateau, ce changement évoque surtout un resserrement des réponses autour d'un prototype pour le /g/ anglais. La figure 43 montre aussi que les réponses du groupe GD (et non celles du groupe GB) pour les VOTs 0 et 10 se caractérisent par une diminution très nette de la variabilité des décisions entre T1 et T2, reflétant des réponses plus unanimes pour ces VOTs, et donc plus caractéristiques de la perception catégorielle.

Pour la paire non-entraînée (/di-ti/), le déplacement du seuil d'identification vers un VOT plus long n'était significatif que dans le groupe ayant réalisé l'entraînement dichotique ( $t(19) = 3.29, p = .004, d = .75$ ), passant de 16.75 à 21.37. Comme l'illustre la figure 44, ce déplacement du seuil s'expliquait surtout par une diminution du pourcentage de réponses /ti/ pour le stimulus /di/ au niveau du VOT 20 ms (i.e. plus de 20% d'erreurs de catégorisation en moins pour les sons [d] anglais s'écartant du prototype du phonème /d/ anglais) ( $t(19) = 5.30, p < .001, d = .76$ ). La figure 44 montre ce resserrement des réponses à T2 autour de la réponse /di/ pour plusieurs VOTs courts.

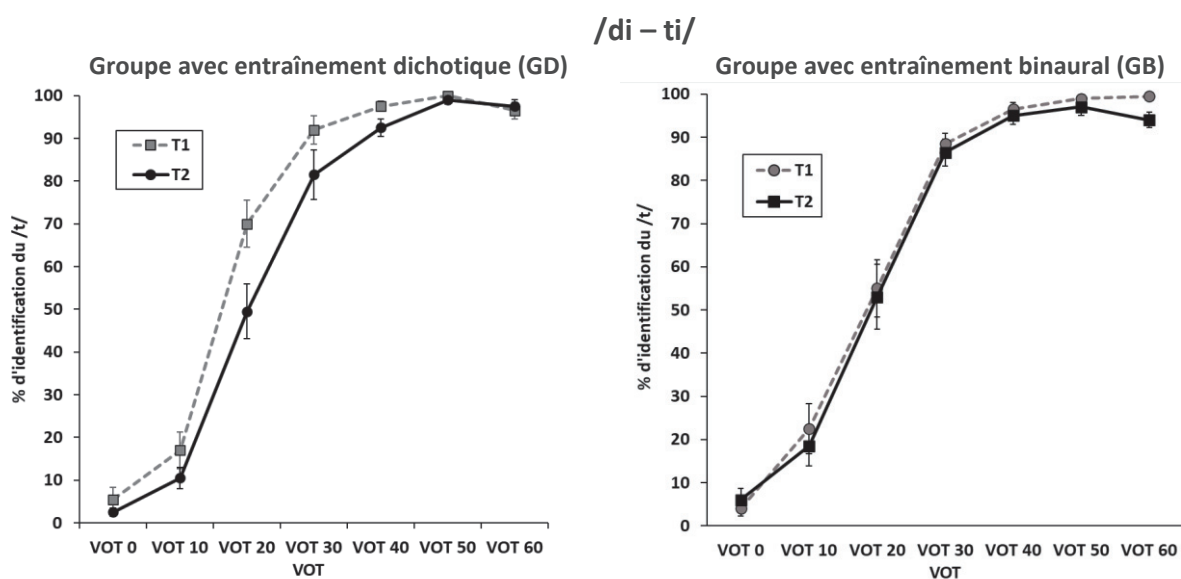


Figure 44. Courbes d'identification pour le groupe GD (à gauche) et GB (à droite) à T1 (pointillés) et T2 (plein) pour la paire /di-ti/. Les barres d'erreurs représentent l'erreur-type.

### 5.3.5. Discussion intermédiaire

L'Expérience 6 visait à comparer les effets d'un entraînement par stimulation dichotique à ceux d'un entraînement par stimulation binaurale pour l'apprentissage de consonnes opposées par le voisement en anglais. Les résultats ont montré que les deux entraînements permettaient d'accentuer le caractère catégoriel de la perception des phonèmes anglais appris lors de l'entraînement. Les deux groupes présentaient une augmentation du pic de discrimination sans modification des autres points de la courbe, ainsi qu'un déplacement du seuil d'identification vers la norme anglaise, et un traitement moins allophonique du /g/. L'ensemble témoigne d'une position plus adéquate de la frontière catégorielle, au niveau d'un VOT typique de cette frontière pour l'anglais. Cette amélioration était de taille plus importante à la suite de l'entraînement par stimulation dichotique. Le dispositif dichotique était par ailleurs le seul à induire une généralisation des compétences vers des phonèmes non-appris lors des exercices. Ce transfert de compétences se traduisait par l'augmentation du pic de discrimination pour la paire /di-ti/, ainsi que par le déplacement du seuil d'identification dans le sens attendu en anglais, et une perception moins allophonique du /d/. Les résultats sont ainsi en faveur de l'intérêt de la stimulation dichotique pour l'apprentissage de consonnes opposées par le voisement en anglais L2 chez des adultes normo-lecteurs.

## 5.4. Expérience 7 : effet de l'entraînement dichotique sur les potentiels évoqués chez des adultes non-dyslexiques

### 5.4.1. Principe de l'expérience et hypothèses

Suite aux résultats comportementaux encourageant l'utilisation de l'écoute dichotique pour l'apprentissage des consonnes occlusives anglaises, nous avons voulu voir si des modifications étaient observables au niveau cérébral. Cette étude avait pour but de mesurer un possible effet de l'entraînement dichotique sur les potentiels évoqués chez des adultes francophones sans pathologies.

Pour cela, les potentiels évoqués ont été analysés dans une épreuve de *oddball* passif, où les apprenants percevaient une série de syllabes standards /gi/ anglaises dans laquelle survenaient de rares syllabes /ki/ anglaises, ou l'inverse dans un autre bloc. Trois ondes (P2, MMN et P3) ont été examinées en tant qu'indicateurs potentiels de la perception d'une différence faisant partie d'un système phonologique, et non en réponse à de simples différences acoustiques.

Les hypothèses suivantes ont été formulées :

- 1) Les participants ayant déjà été exposés à l'anglais, au moins de manière scolaire, une MMN devrait être observée avant l'entraînement pour l'écoute passive d'une consonne occlusive vélaire anglaise déviante sur le plan du voisement dans le contexte d'une série d'occlusives vélaire anglaises standards présentant l'autre valeur de voisement.
- 2) L'amplitude de cette MMN devrait croître à la suite de l'entraînement dichotique.
- 3) La topographie de cette MMN devrait être typique de cette onde, enregistrée davantage avec les électrodes posées sur la partie antérieure du scalp plutôt que sur la partie postérieure.
- 4) Suite à l'entraînement dichotique, la MMN devrait être plus saillante sur les électrodes gauches pour les sonores anglaises et sur les électrodes droites pour les sourdes anglaises, conformément aux caractéristiques de durée des VOTs en anglais et à l'asymétrie hémisphérique supposée par le cadre théorique AST pour la perception d'indices temporels courts et longs.
- 5) À la suite de l'entraînement, l'amplitude de l'onde P3 devrait augmenter, témoignant elle aussi de l'augmentation de la sensibilité à la différence phonologique perçue et une meilleure catégorisation.
- 6) À la suite de l'entraînement, l'onde P2 devrait diminuer pour les stimuli déviants, témoignant d'un filtrage précoce des stimuli qui, bien que surprenants par leur faible fréquence d'écoute, ont trouvé leur place dans le système phonologique en construction et ne doivent donc plus être analysés en profondeur comme des stimuli très surprenants.
- 7) L'entraînement dichotique pour le voisement en anglais ne devrait pas avoir d'impact sur la MMN associée à la perception d'une différence de voisement en français.

#### **5.4.2. Méthode**

##### **5.4.2.1. Participants**

Seize participants ayant réalisé l'entraînement dichotique ont accepté d'effectuer, en plus des épreuves de PC (Expérience 6), une épreuve d'écoute passive dans des conditions d'enregistrement en EEG avant et après l'entraînement (Expérience 7) (Annexe 3).

#### 5.4.2.2. Matériel et équipement

Quatre stimuli ont été utilisés : /gi/ et /ki/ en anglais d'une part, et en français d'autre part. Il s'agissait des syllabes produites avec une voix naturelle par le même locuteurs bilingue natifs anglais/français que pour les épreuves de PC et ayant servi de prototype pour créer les continua des Expérience 1 à 4. En français, le /gi/ avait un VOT négatif de - 60 ms et le /ki/, un VOT positif de +20 ms. En anglais, le /gi/ avait un VOT très bref de + 10 ms et le /ki/ un VOT positif de + 60 ms.

Un total de 1012 items a été présenté dans un paradigme *oddball* passif (85% d'items standards, 15% d'items déviants) par le logiciel Presentation® (v. 14.9, [www.neurobs.com](http://www.neurobs.com)) à travers un casque audio (Beyerdynamics DT 770 Pro 250#). Un bonnet comportant 32 électrodes était posé (Electro-Cap International, INC., selon le système international 10–20), et le système Biosemi ActiveTwo EEG system—Version 5.36 (02-06-2006) fonctionnant à une fréquence d'échantillonnage de 512 Hz et une résolution de 24-bit a été utilisé. Sept électrodes actives externes étaient placées sur le visage : quatre autour des yeux afin d'enregistrer les clignements et mouvements oculaires sur les deux axes (2 Vertical Electro-OculoGraphy (VEOG) pour les électro-oculogrammes verticaux ; 2 HEOG pour les électro-oculogrammes horizontaux), une électrode sur le nez, et deux électrodes sur les mastoïdes droite et gauche.

#### 5.4.2.3. Procédure d'évaluation en EEG

Les mesures EEG et ERP ont été réalisées à T1 et T2 (i.e. avant et après l'entraînement dichotique). Chaque session d'enregistrement comportait deux parties, l'une présentant des syllabes françaises, l'autre des syllabes anglaises. La moitié des participants ont été testés avec la partie française avant la partie anglaise, et l'autre moitié a été testée avec la partie anglaise en premier. Chaque partie était divisée en deux blocs, de sorte que /gi/ était le stimulus standard dans l'un d'eux, et /ki/ le stimulus standard dans l'autre, dans des ordres contrebalancés entre les participants dans la partie française comme dans la partie anglaise. Les stimuli étaient présentés de manière binaurale. Ils ont été pseudo-randomisés pour s'assurer qu'au moins quatre standards étaient présentés avant chaque déviant. Le signal correspondant au standard qui suivait immédiatement un déviant était systématiquement exclu de l'analyse. Le participant était assis dans un fauteuil confortable, dans une salle insonorisée et anéchoïque. Il était seul dans le box, mais un dispositif permettait à l'expérimentateur de veiller visuellement à son bien-être pendant l'expérience. Il regardait un film silencieux non

sous-titré et avait pour instruction de prêter attention au film, et non aux sons diffusés dans le casque. L'expérience était divisée en quatre blocs de 10 minutes – un bloc pour chaque déviant possible –, et un temps de repos était proposé entre les blocs si le participant le souhaitait. L'enregistrement durait au total 45 minutes.

#### **5.4.2.4. Enregistrement électrophysiologique**

L'EEG a été enregistré en continu. Les électrodes ont été ajustées une à une pour obtenir une impédance stable inférieure à 20 mV. Les données ont été analysées hors ligne avec le module ERPLAB (v. 1.0.0.42) de la boîte à outils EEGLAB (v. 9.0.2.3.b) développée pour Matlab (v. 7.0.9.R2009b). Les données ont ensuite été filtrées avec un filtre passe-bande de 0.1 à 30 Hz (12 dB/oct). Elles ont été découpées de - 200 à 800 ms par rapport au début du stimulus. Pour la correction par la ligne de base, la tension moyenne dans l'intervalle de 200 ms avant le stimulus a été soustraite de chaque période. Un rejet automatique des artefacts a été pratiqué, excluant les époques dont le tracé s'écartait de  $\pm 100 \mu\text{V}$ . Pour chaque condition, les signaux ont été moyennés pour chaque participant, et référencés aux deux mastoïdes. Enfin, le signal des réponses aux stimuli standards a été soustrait des réponses aux stimuli déviants dans chacun des 4 blocs expérimentaux pour extraire les réponses MMN.

#### **5.4.2.5. Procédure d'entraînement**

Comme les participants étaient ceux de l'Expérience 6, ils ont réalisé l'entraînement par stimulation dichotique décrit dans l'Expérience 6, partie 5.3.2.4 de ce chapitre.

#### **5.4.3. Analyse de données**

La MMN a été calculée comme la réponse au stimulus déviant moins la réponse au stimulus standard. Une inspection visuelle a permis de choisir les fenêtres d'analyse séparément pour les 4 blocs, sur les électrodes d'intérêt citées ci-après, et selon l'étape (pré- et post-entraînement) à partir des courbes de moyennes calculées sur l'ensemble des participants. Le pic a pu être enregistré pour le déviant /ki/ dans une même fenêtre de 150-250 ms en français et en anglais, alors que pour le déviant /gi/, la fenêtre la plus appropriée était de 100-200 ms en français et de 200-300 ms en anglais. Les électrodes F3, F4, Fz, FC1, FC2, C3, C4, Cz, CP1, CP2, P3, P4, Pz ont été sélectionnées pour les analyses statistiques guidées par les hypothèses. Elles ont été regroupées en cinq positions : frontales, fronto-centrales, centrales, centro-pariétales, pariétales. Les analyses ont porté sur l'amplitude moyenne de l'onde. Un test *t* de Student bilatéral apparié (avec correction de Bonferroni) a d'abord été effectué pour déterminer si la MMN différait significativement de zéro pour chaque déviant à T1. Puis,

l'amplitude moyenne de la MMN a été soumise à deux ANOVA à mesures répétées distinctes :

- l'une portait sur l'ensemble des données, selon le plan 16 (sujet) × 2 (session) × 5 (position) × 2 (voisement), avec Session, Position, Voisement comme facteurs intra-sujet,
- l'autre portait sur le signal recueilli sur les seules électrodes posées à gauche ou à droite, pour évaluer l'effet de latéralité, selon le plan 16 (sujet) × 2 (session) × 5 (position) × 2 (voisement) × 2 (latéralité), avec Session, Position, Voisement et Latéralité comme facteurs intra-sujet.

Le test *t* de Student bilatéral apparié et avec correction de Bonferroni permettait des comparaisons de moyennes pour décomposer des interactions.

La P3 a été enregistrée dans une fenêtre de 200-350 ms et la P2 dans une fenêtre de 100-200 ms pour les déviants /ki/ et /gi/, en français comme en anglais.

#### 5.4.4. Résultats

**Consonnes anglaises.** Les tracés correspondant à l'amplitude moyenne du signal (condition déviant moins condition standard) sont présentés dans la figure 45, séparément lorsque le déviant était /gi/ et /ki/, sur l'électrode Fz à T1 et à T2. Les tests *t* de Student ont révélé que l'amplitude moyenne était significativement différente de zéro et négative avant l'entraînement pour /ki/ ( $t(15) = 3.06, p = .008, d = 1.08$ ), mais pas pour /gi/ ( $t(15) = 1.11, p = .285, d = .39$ ) dans la fenêtre où se produisait la MMN.

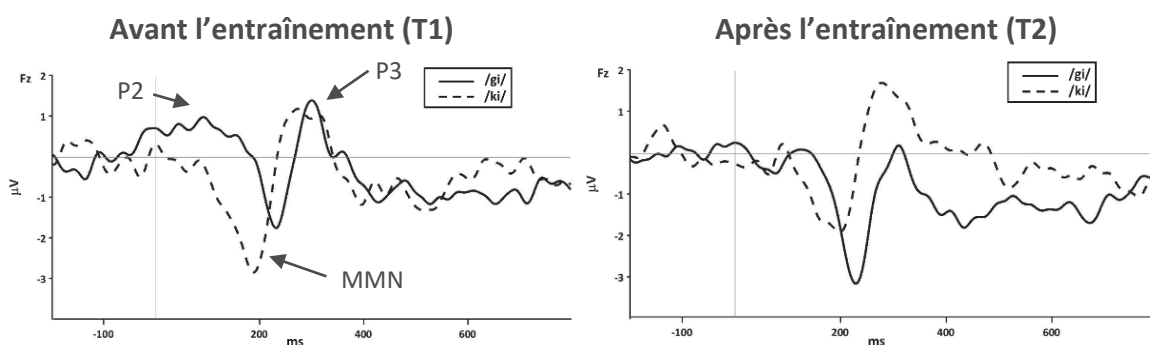


Figure 45. Grande moyenne ( $\mu\text{V}$ ) de la différence Déviant moins Standard, avant l'entraînement (T1, à gauche) et après l'entraînement en anglais (T2, à droite), lorsque le déviant était /gi/ (trait plein) ou /ki/ (pointillés) en anglais.

Les résultats de la première ANOVA ont indiqué un effet principal de position ( $F(4,60) = 6.41, p < .001, \eta_p^2 = .30$ ). Comme l'illustre la figure 46, l'amplitude de la MMN était plus élevée sur les électrodes antérieures (i.e. frontales et frontocentrales) que sur les électrodes plus postérieures (i.e. centrales, centropariétales et pariétales). L'amplitude de la

MMN ne différait pas significativement entre les positions frontale et frontocentrale ( $t(15) = 1.20, p = .249, d = .58$ ), mais la MMN la plus élevée était enregistrée en position frontocentrale. Cette dernière dépassait significativement la MMN recueillie sur les électrodes plus postérieures : centrales ( $t(12) = -2.68, p = .017, d = .53$ ), centropariétales ( $t(12) = -2.25, p = .040, d = .48$ ) et pariétales ( $t(12) = -4.30, p < .001, d = .34$ ).

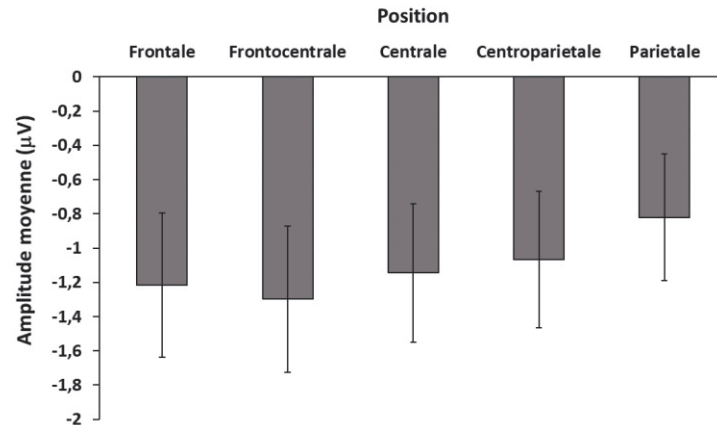


Figure 46. Amplitude moyenne ( $\mu\text{V}$ ) selon les positions des électrodes. Les barres d'erreurs représentent l'erreur-type.

L'ANOVA a aussi montré l'interaction Session  $\times$  Voisement ( $F(1,15) = 6.23, p = .025, \eta_p^2 = .29$ ) illustrée par la figure 47, et reflétant une MMN plus ample après l'entraînement en anglais qu'avant celui-ci pour /gi/ ( $t(15) = 2.40, p = .030, d = .26$ ), mais pas pour /ki/ déjà ample à T1.

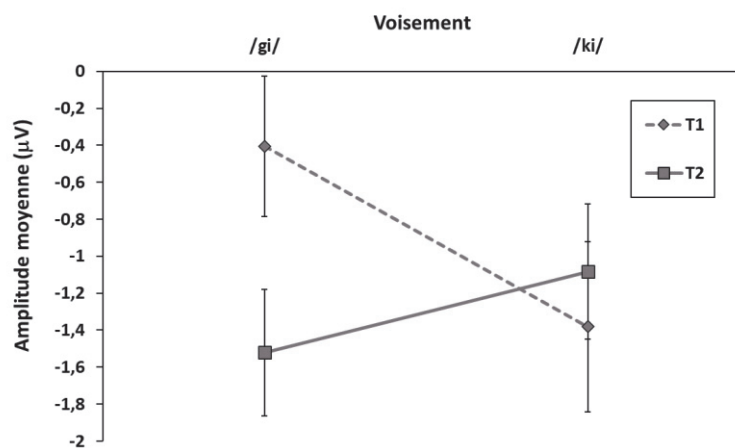


Figure 47. Amplitude moyenne ( $\mu\text{V}$ ) pour /gi/ (gauche) et /ki/ anglais (droite) avant l'entraînement (T1, pointillés) et après l'entraînement (T2, trait plein). Les barres d'erreurs représentent l'erreur-type.

Guidée par nos hypothèses, l'ANOVA restreinte aux électrodes gauches et droites a permis de mieux décrire un changement survenu pour la MMN entre T1 et T2 pour /ki/, changement invisible si la latéralité n'était pas prise en compte. L'interaction Session  $\times$  Voisement  $\times$  Latéralité  $\times$  Position était significative ( $F(4,60) = 4.72, p = .002, \eta_p^2 = .24$ ). Afin

de mieux appréhender les changements entre les sessions T1 et T2, cette interaction a été décomposée pour représenter sur la figure 48 l'interaction moins complexe Session × Voisement × Latéralité. Cette dernière n'atteignait pas le seuil de significativité ( $F(1,15) = 2.33, p = .148, \eta_p^2 = .13$ ), mais elle a été inspectée pour tester le phénomène prédit par les hypothèses, et s'intéresser seulement ensuite à sa modulation par la position des électrodes.

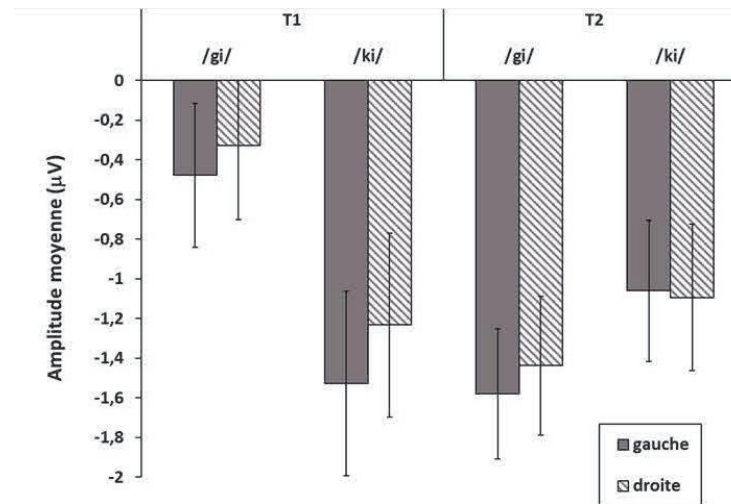


Figure 48. Amplitude moyenne ( $\mu\text{V}$ ) avant l'entraînement (T1, à gauche) et après (T2, à droite) pour les électrodes gauches (gris) et les électrodes droites (rayé) pour /gi/ et /ki/ en anglais. Les barres d'erreurs représentent l'erreur-type.

En ce qui concerne le /ki/ anglais, l'observation de la figure 49 montre que la MMN était plus ample sur les électrodes gauches que sur les électrodes droites à T1, comme cela pourrait se justifier pour le /k/ français dont le VOT est court. Cette différence disparaissait à T2, ce qui suggère que l'implication de l'HG avait perdu sa dominance pour le traitement du /k/ anglais, disparition pertinente compte-tenu de la longueur du VOT du /k/ anglais pour lequel l'HD était censé être peu compétent. L'interaction significative Session × Voisement × Latéralité × Position reflétait une modulation de ce changement entre T1 et T2 par la position de l'électrode. A T1, l'amplitude plus élevée de la MMN sur les électrodes gauches plutôt que droites pour /ki/ était significative seulement en positions centrale ( $t(15) = 2.34, p = .034, d = .85$ ) et centropariétale ( $t(15) = 2.52, p = .024, d = .78$ ) (Fig. 49, gauche). Cet effet de latéralité a disparu pour ces deux positions à T2 (respectivement,  $t(15) = -0.16, p = .873, d = .81$  ;  $t(15) = -0.23, p = .822, d = .72$ ) (Fig. 49, droite).



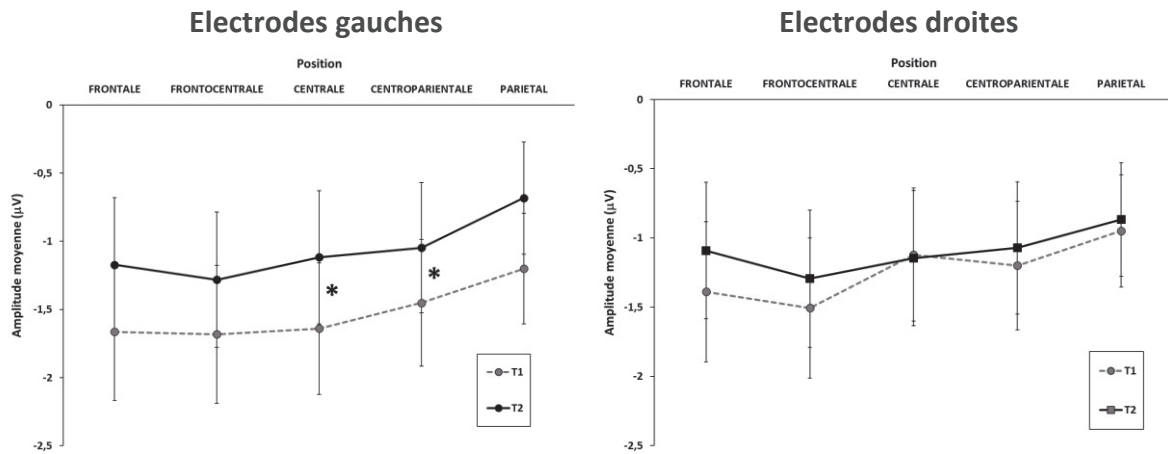


Figure 49. Amplitude Moyenne ( $\mu\text{V}$ ) de la MMN pour les consonnes anglaises en fonction de la position des électrodes avant l'entraînement (T1, pointillés) et après l'entraînement (T2, trait plein) pour les électrodes gauches (à gauche) et les électrodes droites (à droite). Les barres d'erreurs représentent l'erreur-type.

Par ailleurs, l'examen visuel de la figure 45 suggérait, entre T1 et T2, une augmentation de l'amplitude de l'onde P3 pour /ki/, ainsi qu'une diminution de l'onde P2 pour le /gi/. Des analyses complémentaires ont été réalisées pour tester la significativité de ces observations. Entre T1 et T2, l'amplitude de la P3 pour /ki/ passait de 0.32 à 0.49  $\mu\text{V}$ , ce qui n'était pas significatif ( $t(15) = 0.52, p = .612$ ). La diminution de la P2 pour /gi/ était quant à elle marginale, ( $t(15) = 2.02, p = .062, d = .59$ ) avec une taille d'effet modérée, suggérant l'émergence du *sensory-gating* prédit.

**Consonnes françaises.** Les tracés correspondant à l'amplitude moyenne du signal (condition déviant moins condition standard) sont présentés dans la figure 50, séparément lorsque le déviant était /gi/ et /ki/, sur l'électrode Fz à T1 et à T2. Une réponse MMN a été mise en évidence par rapport à zéros ( $t(15) = -3.43, p = .004, d = .70$ ). Cette réponse était présente dès T1 pour /gi/ ( $t(15) = 2.85, p = .012, d = 1.01$ ), mais pas pour /ki/ ( $t(15) = 0.49, p = .632, d = .17$ ).

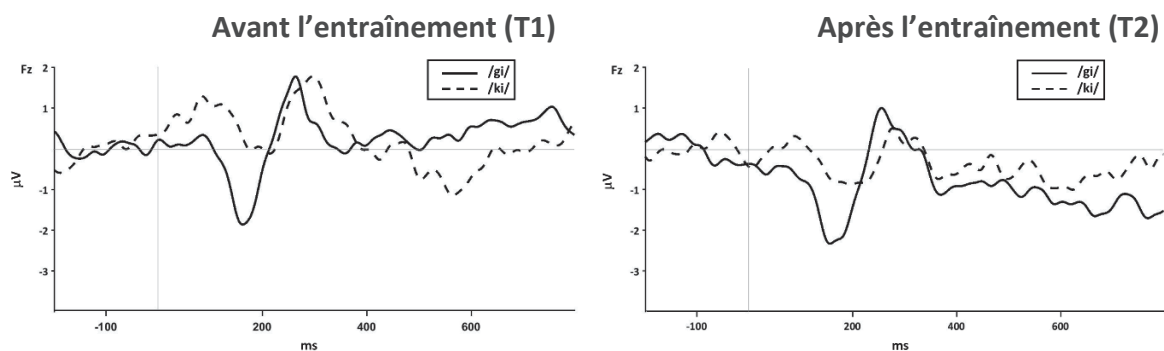


Figure 50. Grande moyenne ( $\mu\text{V}$ ) de la différence Déviant moins Standard, avant l'entraînement (T1, à gauche) et après l'entraînement en anglais (T2, à droite), lorsque le déviant était /gi/ (trait plein) ou /ki/ (pointillés) en français.

Comme pour les consonnes anglaises, deux ANOVA ont été réalisées sur l'amplitude moyenne des MMN pour les consonnes françaises.

L'ANOVA générale a révélé une amplitude de la MMN significativement plus élevée pour /gi/ que pour /ki/ ( $F(1,15) = 5.22, p = .037, \eta_p^2 = .26$ ). La figure 50 suggère une MMN plus ample de façon générale à T2 qu'à T1, et cette différence était statistiquement marginale, mais de grande taille ( $F(1,15) = 3.33, p = .088, \eta_p^2 = .18$ ). L'effet de la Position était significatif ( $F(4,60) = 17.24, p < .001, \eta_p^2 = .53$ ), en raison d'une plus grande amplitude aux emplacements antérieurs plutôt que postérieurs mais, comme l'illustrent la figure 51 et l'interaction Session  $\times$  Position ( $F(4,60) = 4.62, p = .003, \eta_p^2 = .24$ ), cette différence de MMN selon la Position était essentiellement présente à T2. La MMN était au final significativement plus ample à T2 qu'à T1 uniquement en position frontale ( $t(15) = 2.41, p = .029, d = .20$ ). Une tendance similaire dans cette direction était observée en position frontocentrale ( $t(15) = 2.02, p = .062, d = .22$ ), mais l'effet de la session n'était pas significatif sur les positions plus postérieures (Fig. 51). Il n'y avait pas d'autres interactions.

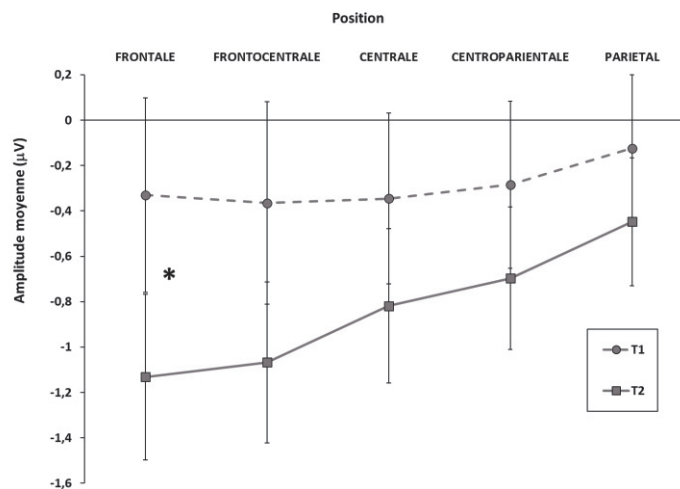


Figure 51. Amplitude moyenne ( $\mu\text{V}$ ) de la MMN pour les consonnes françaises en fonction de la position des électrodes avant l'entraînement en anglais (T1, pointillés) et après l'entraînement (T2, trait plein). Les barres d'erreurs représentent l'erreur-type.

L'ANOVA portant sur les électrodes latéralisées a montré une plus grande amplitude de la MMN recueillie avec les électrodes posées du côté gauche ( $-0,62 \mu\text{V}$ ) qu'avec les électrodes posées à droite ( $-0,48 \mu\text{V}$ ), mais cette différence n'atteignait pas le seuil de significativité ( $F(1,15) = 3.70, p = .074, \eta_p^2 = .18$ ). Aucune autre interaction n'approchait ce seuil.

Il n'y avait pas de changement significatif pour la P2 ou la P3 en français après l'entraînement en anglais.

#### 5.4.5. Discussion intermédiaire

L'Expérience 7 mesurait, en potentiels évoqués, l'effet de l'entraînement par stimulation dichotique auprès d'étudiants normo-lecteurs. Les résultats montrent que l'entraînement par stimulation dichotique permettait l'émergence d'une MMN pour le déviant /gi/ parmi des standards /ki/. Cette MMN n'existait pas avant l'entraînement. La diminution de l'onde P2 pour le déviant /gi/ montre aussi un changement de traitement du /g/ après l'entraînement. Ces résultats apportent des arguments en faveur de la création d'une nouvelle catégorie pour la consonne sonore anglaise /g/ (dont le VOT court est ambigu pour les francophones susceptibles de le percevoir comme un /k/). La MMN est en effet censée être d'autant plus large que le déviant fait partie du système phonologique de l'auditeur (Sams *et al.*, 1985). La diminution de l'onde P2 pourrait traduire la mise en place d'un *sensory-gating*, phénomène montrant que le stimulus n'est plus inhabituel et ne nécessite pas d'être traité de manière plus profonde.

L'occlusive sourde anglaise /k/ était elle aussi mieux traitée après l'entraînement. Même si elle faisait déjà l'objet d'une MMN avant l'entraînement, cette onde n'était alors recueillie que par les électrodes posées à gauche. Après l'entraînement, elle devenait bilatérale et d'amplitude symétrique entre les deux hémisphères, à cause d'une diminution d'amplitude à gauche. Ce résultat pourrait être interprété comme un désengagement de l'HG pour le traitement d'un indice de durée long dans une consonne anglaise. L'augmentation de l'onde P3 pour le déviant /ki/ après l'entraînement suggère aussi que le traitement de la consonne /k/ s'est modifié, y compris pour une étape d'analyse plus tardive que le mécanisme pré-attentionnel associé à la MMN.

Enfin, l'entraînement sur les consonnes anglaises par stimulation dichotique semble avoir renforcé des capacités assez générales de traitement de la durée. Des effets de cet entraînement en anglais se sont transférés au traitement des consonnes françaises pour lesquelles la MMN provoquée par les déviants devenait plus frontale.

## **5.5. Expérience 8 : effet de l'entraînement dichotique en anglais sur la PC en anglais chez des adultes dyslexiques**

### **5.5.1. Principe de l'expérience et hypothèses**

Le but de cette expérience était d'évaluer l'effet de l'entraînement dichotique des deux précédentes expériences chez des adultes francophones dyslexiques sur la perception catégorielle (PC) des occlusives anglaises.

Les hypothèses suivantes ont été formulées :

- 1) Suite à l'entraînement dichotique, les performances des participants dyslexiques en identification et en discrimination de phonèmes anglais devraient davantage refléter une PC qu'avant cet entraînement.
- 2) L'entraînement par stimulation dichotique devrait améliorer la perception catégorielle de consonnes occlusives n'ayant pas été entendues pendant l'entraînement.

### **5.5.2. Méthode**

#### **5.5.2.1. Participants**

Dans cette expérience, les 21 participants (13 femmes, M = 21,9 ans, ET = 1 ans et 7 mois) faisaient partie du groupe de 27 dyslexiques ayant réalisé l'entraînement dichotique en français (voir Chapitre 3 – Expérience 2 et 3, partie 3.3.2.1.). En effet, 21 d'entre eux avaient accepté d'effectuer l'entraînement en anglais après l'entraînement en français. Les performances en PC en anglais à T1 ont été recueillies avant le premier entraînement proposé aux dyslexiques (en français) (Annexe 3). L'entraînement en français présenté dans le Chapitre 3 était en partie réalisé afin de protéger le système phonologique français des adultes dyslexiques d'une potentielle contamination par un entraînement dans une L2.

#### **5.5.2.2. Matériel et équipement**

Le matériel et l'équipement étaient les mêmes que ceux décrits dans l'Expérience 6 partie 5.3.2.4 de ce chapitre.

#### **5.5.2.3. Procédure d'évaluation**

L'évaluation de la perception catégorielle de phonèmes anglais opposés par le voisement était la même que celle décrite dans le Chapitre 4 – Expérience 4, partie 4.6.2.3., avec les épreuves de discrimination et d'identification.

#### 5.5.2.4. Procédure d'entraînement

L'entraînement réalisé par les participants était le même que celui décrit dans l'Expérience 7 partie 5.3.2.4 de ce chapitre utilisant la stimulation dichotique. Comme dans l'Expérience 6, les fichiers de réponses étaient examinés pour vérifier que le participant n'avait pas appuyé sur d'autres touches que celles désignées pour les réponses, et qu'il n'avait pas produit plus de quatre temps supérieurs à 3 écart-types calculés à partir de ses données. Il n'y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Dans de cas où les réponses étaient globalement très lentes ou particulièrement rapides, les données du participant n'étaient pas conservées si cette lenteur ou rapidité extrêmement ne concernait que T1 ou que T2, car cela pouvait témoigner d'un manque de sérieux lors de l'une des étapes. Aucun participant n'a été exclu en respectant ces critères.

#### 5.5.3. Analyses de données

Les performances du groupe de participants dyslexiques aux épreuves de PC ont été traitées et représentées comme décrit dans les Expériences 1 et 4 dans les Chapitre 2, partie 2.10.4 et Chapitre 4 partie 4.6.3. Puis, les mêmes traitements statistiques que ceux décrits dans le Chapitre 3 – Expérience 2 dans la partie 3.3.3 ont été réalisés pour évaluer l'effet de l'entraînement en anglais sur la PC en anglais dans le groupe dyslexique.

#### 5.5.4. Résultats

**Discrimination.** Comme le montre la Figure 52 pour la paire /gi-ki/, il n'y avait pas d'augmentation significative du  $d'$  entre T1 et T2 au niveau du pic de discrimination (VOT 0-20 ms) ( $t(20) = 1.45, p = .163, d = .27$ ), mais une augmentation du  $d'$  au niveau du VOT 10-30 ms à la fois significative et de grande taille ( $t(20) = 3.19, p = .005, d = .81$ ). Cela traduisait un déplacement du pic pour se rapprocher de la frontière anglaise, avec néanmoins le maintien du pic (VOT 0-20 ms).

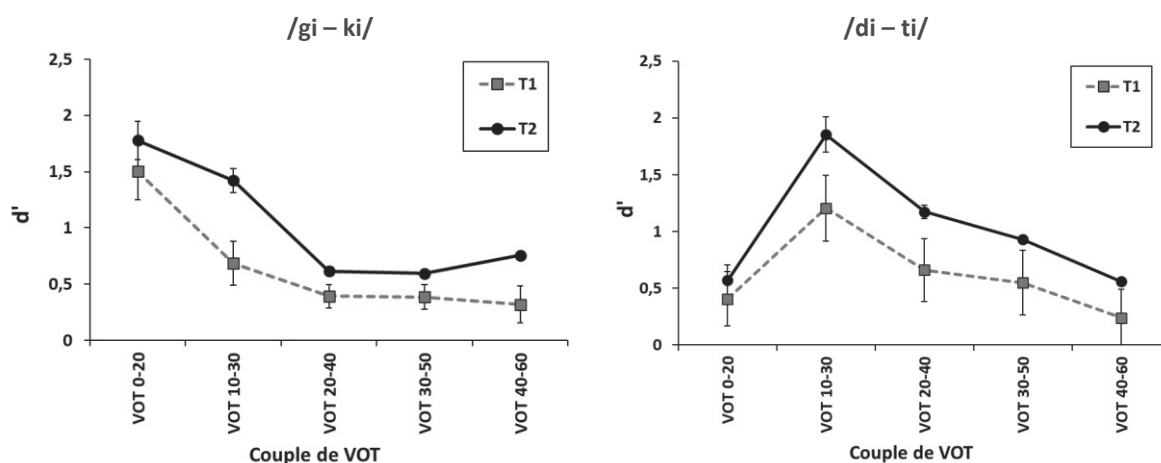


Figure 52. Courbes de discrimination en anglais pour les paires anglaises /gi-ki/ (à gauche) et /di-ti/ (à droite) entre T1 (pointillés) et T2 (trait plein) chez les adultes dyslexiques. Les barres d'erreurs représentent l'erreur-type.

Pour la paire /di-ti/, les analyses ont montré une amélioration significative de taille moyenne de la sensibilité à la différence au niveau du pic typique de la discrimination de voisement en anglais (VOT 10-30 ms) après l'entraînement ( $t(20) = 3.63, p = .002, d = .53$ ). Au niveau du VOT 20-40 ms, l'augmentation apparente du  $d'$  n'était pas significative (Tableau 10). La configuration des résultats est proche de celle des normo-lecteurs à T1 pour cette paire (voir figure 37, Chapitre 4 – Expérience 5), à ceci près que seul le groupe dyslexique accroissait significativement l'amplitude du pic entre T1 et T2, se rapprochant ainsi de celui des normo-lecteurs. Enfin, l'examen de la figure 52 faisait ressortir une diminution frappante des barres d'erreurs à T2 par rapport à T1 pour l'ensemble des paires de VOT, pour /g-k/ et pour /d-t/, soulignant l'émergence d'une plus grande homogénéité des réponses entre les participants.

Tableau 10. Résultats du test  $t$  de Student pour comparer les performances en discrimination entre T1 et T2 pour les couples de stimuli des paires /gi-ki/ et /di-ti/. L'astérisque \* indique une différence significative.  $ddl = 20$ .

Paire de phonèmes	VOT 0-20	VOT 10-30	VOT 20-40	VOT 30-50	VOT 40-60
/gi-ki/	$p = .163$	$p = .005^*$	$p = .273$	$p = .271$	$p = .040^*$
/di-ti/	$p = .490$	$p = .002^*$	$p = .096$	$p = .103$	$p = .230$

**Identification.** Entre T1 et T2 (Fig. 53), seul le seuil de la paire non entraînée (/di-ti/) s'est significativement déplacé, vers une durée de VOT plus longue et plus conforme à la frontière catégorielle anglaise ( $t(20) = 2.29, p = .033, d = .50$ ). Cette modification du seuil a été permise par un abaissement de la courbe au niveau des VOT 0 ms ( $t(20) = 2.02, p = .057, d =$

.56) et 10 ms ( $t(20) = 3.33, p = .003, d = .51$ ). Entre ces deux sessions, la pente n'a pas significativement évolué, que ce soit pour /gi-ki/ ou /di-ti/.

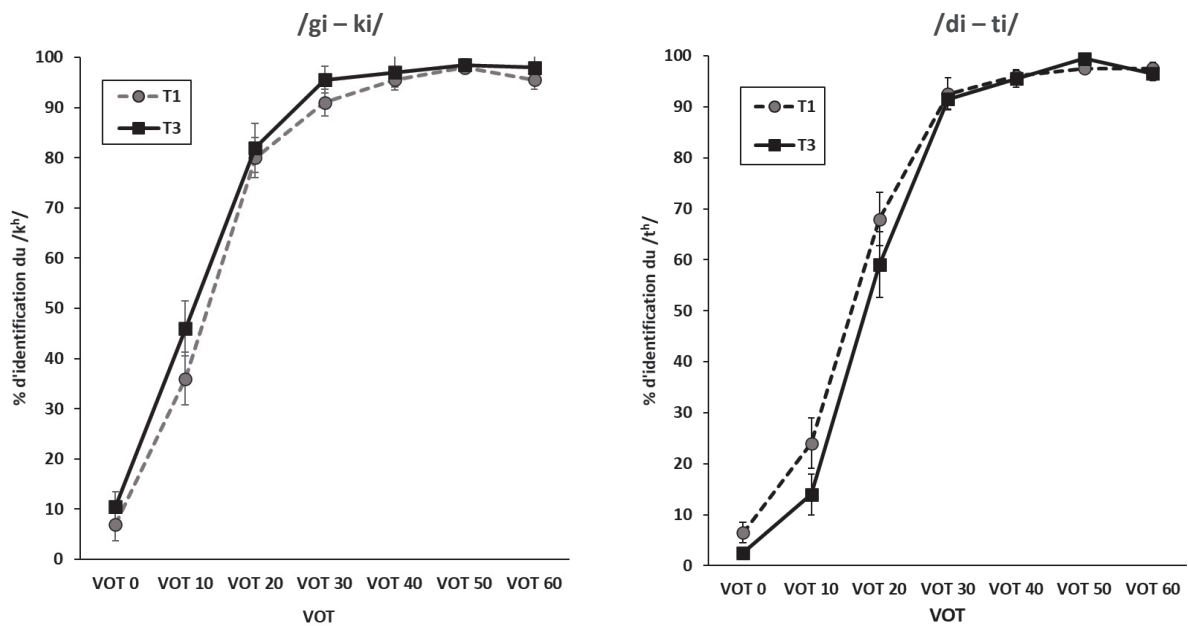


Figure 53. Courbes d'identification pour les paires anglaises /gi-ki/ (à gauche) et /di-ti/ (à droite) entre T1 (pointillés) et T2 (trait plein). Les barres d'erreurs représentent l'erreur-type.

### 5.5.5. Discussion intermédiaire

L'entraînement en anglais par stimulation dichotique proposé à des étudiants dyslexiques dans l'Expérience 8 a apporté des résultats encourageants. Il leur a permis de présenter une perception plus catégorielle pour la paire de consonnes entraînée, comme le montre l'augmentation du pic de discrimination. De plus, cet entraînement a permis un transfert de compétences vers une paire de consonnes opposées en voisement mais non entraînée. Le pic de discrimination augmentait en effet pour la paire /d-t/. Le seuil d'identification s'était aussi déplacé vers la norme anglaise et le /d/ était mieux catégorisé après l'entraînement. Ces résultats nous apprennent que, même en cas de dyslexie à l'âge adulte, il est possible d'apprendre un nouveau contraste de consonnes opposées par le voisement, particulièrement si un entraînement ciblé sur ce contraste est associé à une stimulation dichotique.

## 5.6. Expérience 9 : effet de l'entraînement en anglais sur la PC en français chez des adultes normo-lecteurs et dyslexiques

### 5.6.1. Principe de l'expérience et hypothèses

La dernière expérience de cette partie de l'étude visait à évaluer l'impact de l'entraînement en anglais sur la perception catégorielle en français, ceci pour les trois groupes de participants des protocoles précédant (i.e., Groupe non-dyslexique dichotique, Groupe non-dyslexique binaural, Groupe dyslexique dichotique).

L'hypothèse formulée pour cette expérience était que les performances en PC française des différents groupes ne seraient pas détériorées à T2 par rapport à T1, sous l'influence de l'entraînement réalisé sur le voisement anglais.

### 5.6.2. Méthode

#### 5.6.2.1. Participants

Les 40 participants normo-lecteurs (NL) de l'Expérience 6 ayant réalisé les entraînements dichotique ( $n = 20$ ) et binaural ( $n = 20$ ), ainsi que les 21 participants dyslexiques (DYS) de l'Expérience 8 ont été inclus dans ce protocole, leur point commun étant d'avoir réalisé l'entraînement en anglais, et d'avoir réalisé – avant et après – les tests de PC en français. Pour le groupe DYS, dans cette expérience, les performances à T1 ont été recueillies après leur entraînement en français (visant à renforcer leur système phonologique français), étape précédant leur entraînement en anglais (Annexe 3).

### 5.6.3. Analyse de données

Les données en perception catégorielle (PC) en français ont été traitées et analysées comme décrit dans le Chapitre 2 – Expérience 1, partie 2.10.4.

### 5.6.4. Résultats

**Participants normo-lecteurs.** Le seul changement significatif était une modification de la pente d'identification en français après l'entraînement dichotique en anglais, pour les deux paires françaises /gi-ki/ ( $t(19) = 2.20, p = .041, d = .70$ ) et /di-ti/ ( $t(19) = 2.95, p = .008, d = .87$ ) : la pente devenait plus abrupte pour chacune des paires (Fig. 54). Ce changement de pente résultait de la disparition de deux petits défauts présents à T1 comme le montre la figure 54 : pour /gi-ki/, les réponses devenaient plus homogènes pour /gi/, avec une tendance à moins d'hésitation sur l'identité de cette consonne voisée au VOT 0 ms ( $t(19) = 1.92, p =$



.071,  $d = .60$ ) et cela de manière significative au VOT 10 ms ( $t(19) = 3.11, p = .006, d = .64$ ). Pour /di-ti/, cette amélioration pour la consonne voisée était beaucoup plus discrète sur la figure 54, et l'accentuation de la pente à T2 s'expliquait surtout par un plateau plus net pour l'identification de /ti/, avec notamment des réponses tendant à plus homogénéité pour /ti/ au VOT 40 ms ( $t(19) = 1.79, p = .090, d = .54$ ).

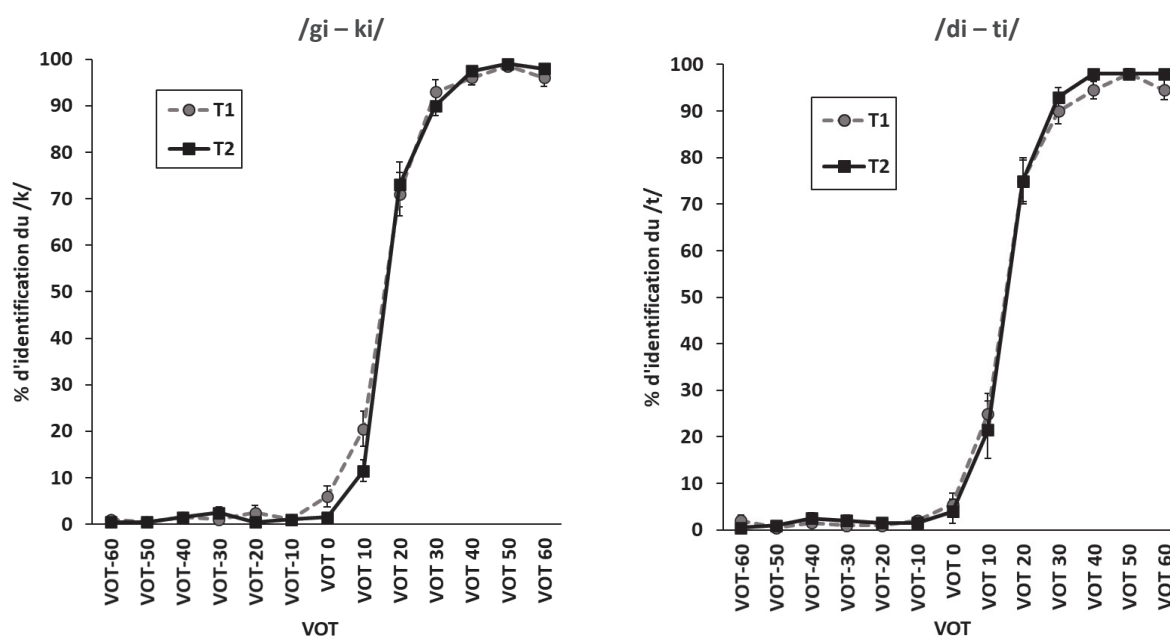


Figure 54. Courbes d'identification pour les paires françaises /gi-ki/ (à gauche) et /di-ti/ (à droite) entre T1 (pointillés) et T2 (trait plein) pour le groupe de normo-lecteurs avec entraînement dichotique. Les barres d'erreurs représentent l'erreur-type.

Aucun autre effet significatif ne ressortait des analyses (Tableau 11).

Tableau 11. Résultats aux tests  $t$  de Student pour comparer avant et après l'entraînement en anglais les valeurs  $d'$  au pic de discrimination, le seuil et la pente d'identification pour les paires françaises /gi-ki/ et /di-ti/ après l'entraînement en anglais pour les groupes GD et GB.  $ddl = 19$

		Pic	Seuil	Pente
GD	/gi-ki/	$p = .253$	$p = .371$	$p = .041^*$
	/di-ti/	$p = .073$	$p = .837$	$p = .008^*$
GB	/gi-ki/	$p = .623$	$p = .496$	$p = .195$
	/di-ti/	$p = .200$	$p = .906$	$p = .340$

**Participants dyslexiques.** Il n'y a eu aucun changement significatif en PC des consonnes françaises par le groupe DYS suite à l'entraînement en anglais (Tableau 12).

Tableau 12. Résultats aux tests *t* de Student pour les valeurs *d'* au pic de discrimination, le seuil et la pente d'identification pour les paires françaises /gi-ki/ et /di-ti/ après l'entraînement dichotique en anglais pour le groupe DYS. *ddl* = 20

	Pic	Seuil	Pente
/gi-ki/	$p = .443$	$p = .504$	$p = .486$
/di-ti/	$p = .062$	$p = .199$	$p = .612$

### 5.6.5. Discussion intermédiaire

L'Expérience 9 permettait de vérifier si les entraînements en anglais L2 avaient un impact sur le traitement des consonnes occlusives françaises de la L1. Les résultats montrent que l'entraînement en anglais n'a pas eu d'impact négatif sur le traitement de ces consonnes françaises, que ce soit chez les adultes normo-lecteurs ou dyslexiques. Le seul changement notable concernait les normo-lecteurs, dont la pente d'identification devenait plus précise pour les deux paires de consonnes françaises (i.e. /g-k/ et /d-t/) après l'entraînement en anglais. Cette amélioration est due à des réponses devenues plus homogènes et moins hésitantes pour identifier le /g/ et le /t/ français.

### 5.7. Discussion sur l'aide à l'apprentissage de la perception du voisement en anglais (Expériences 6 à 9)

Ce chapitre avait deux objectifs principaux. Le premier était d'évaluer l'apport de la stimulation dichotique pour l'apprentissage des consonnes occlusives anglaises en L2 chez des adultes sans pathologie par rapport à la stimulation binaurale, classiquement utilisée dans les protocoles d'entraînement perceptifs. Le second objectif était d'évaluer l'impact de la stimulation dichotique pour l'apprentissage des consonnes occlusives anglaises en L2 chez des adultes atteints d'une dyslexie développementale. Nous avons tenté d'y répondre à travers quatre expériences.

La discussion de ce chapitre reviendra sur les difficultés des adultes francophones en apprentissage tardif de l'anglais L2, mais aussi sur leurs possibles progrès en phonologie pour cette langue, en évoquant les modèles d'apprentissage d'une L2. L'apport de l'entraînement proposé sera ensuite synthétisé, en lien avec ses caractéristiques qui sont 1) de faire l'économie d'un enseignement explicite et 2) d'utiliser la stimulation dichotique dans des exercices d'entraînement perceptif au voisement en anglais. Il s'agira aussi d'interpréter ces

effets, en tentant d'expliquer quels mécanismes cognitifs et neuronaux sous-jacents peuvent être impliqués dans un apprentissage réalisé en stimulant préférentiellement l'un ou l'autre hémisphère grâce à l'écoute dichotique. Les liens seront alors établis avec le cadre théorique de l'échantillonnage asymétrique (AST) issu des neurosciences. Enfin, la discussion s'ouvrira sur les limites de l'étude menée à travers ces quatre expériences et les nouveaux questionnements qui en découlent.

### **5.7.1. Difficulté d'apprentissage tardif de l'opposition de voisement en anglais : un argument supplémentaire en ERP**

Les données comportementales (Chapitre 4), ont montré que les francophones sans pathologie traitaient plutôt bien les sons anglais. Cependant, il semblerait que dans la langue la moins familière pour eux (en L2), les normo-lecteurs soit plus sensibles à l'intensité de la frontière catégorielle. Les données recueillies en potentiels évoqués dans l'Expérience 7, notamment concernant la MMN, ont apporté des éléments objectifs sur certaines faiblesses des représentations phonologiques du voisement des consonnes anglaises chez des apprenants tardifs, avant l'entraînement. La MMN indique une sensibilité à une différence perceptive considérée comme pertinente par le système cérébral et faisant l'objet de représentations mentales de catégories. La MMN enregistrée présentait sa localisation classique : une distribution fronto-centrale conforme à la littérature sur ce sujet (Cheour *et al.*, 1996 ; Näätänen *et al.*, 1989, 2007). Le fait le plus marquant pour cette onde, avant l'entraînement, était une configuration de résultats présentant une étonnante asymétrie selon le voisement (sour/sonore) du déviant : en anglais, les participants ne présentaient pas de MMN à l'écoute du déviant /gi/ dans une série de standards /ki/, alors qu'ils présentaient une MMN à l'écoute du déviant /ki/ dans une série de standards /gi/.

Cette absence de cohérence dans l'émergence de la MMN témoigne de lacunes dans la représentation du contraste phonologique de voisement entre ces deux consonnes anglaises. Elle permet même de dire que les deux sons présentés ne renvoyaient alors pas clairement à des phonèmes anglais pour les apprenants. Cette présence partielle et asymétrique de la MMN entre les deux conditions montre que la sensibilité pré-attentionnelle et automatisée à une différence entre les occlusives anglaises /g/ et /k/ n'était pas encore opérationnelle, et que le système phonologique mobilisé pour traiter la parole anglaise était inadapté. L'observation d'une MMN dans une seule des deux conditions montre que les participants n'ont pas construit de représentations phonologiques des phonèmes /g/ et /k/ sur la base des fondements phonétiques propres au contraste de voisement en anglais.

Une interprétation peut être proposée en se basant sur le cadre théorique offert par les modèles SLM-r (Flege & Bohn, 2021) et PAM (Best, 1995 ; Best *et al.*, 2001). Une hypothèse commune aux deux modèles est qu'un fort recouvrement phonétique entre un phonème de la L2 et un phonème de la L1 peut conduire à assimiler le phonème de la L2 à cette catégorie connue de la L1. Cette perception – dite allophonique – pourrait conduire à traiter le /g/ anglais comme un /k/ français, sur la base phonétique de son court VOT positif. Au niveau neuronal, le /g/ anglais activerait ainsi la représentation du /k/ français, si l'on suit l'hypothèse d'assimilation. Ce phénomène pourrait être renforcé par la procédure et le matériel de l'expérience de oddball passif, où l'information sonore était délivrée comme une succession de syllabes isolées, en-dehors d'un contexte sémantique ou lexical qui aurait peut-être permis plus de nuances. Ici, selon le modèle SLM, des indices autres que le court VOT, qui pourraient éventuellement permettre de ne pas considérer le [g] anglais comme un /k/, seraient filtrés. Le mécanisme d'assimilation conduirait en effet à un traitement superficiel et limité à l'extraction des seules propriétés acoustiques permettant une classification rapide (Flege, 1995). En d'autres termes, le phonème /k/ de la L1 agirait comme un fort attracteur et forcerait la catégorisation.

Cette assimilation du son [g] anglais au phonème /k/ français conduit à une erreur, car il pourrait amener à croire que le graphème associé est la lettre C (ou K). Cependant, ce rapprochement excessif du [g] anglais avec le phonème /k/ français présente – s'il reste transitoire – l'intérêt de constituer un contexte de standards perçus comme exagérément familiers, propre à faire percevoir le déviant anglais /k/ comme surprenant. Ce dernier présente en effet un indice sensoriel saillant qui le rend étrange par rapport au système phonétique-phonologique de la L1 : sa longue aspiration est à la fois un indice marquant par sa durée et inédit pour sa position si l'on se réfère au Français.

Ici, le modèle PAM permet d'ajouter une hypothèse. Il décrit trois situations de discrimination de phonèmes en L2, et même s'il s'agit ici de discrimination passive, c'est la situation qui s'en rapproche le plus. La situation 1 est celle où les deux sons de la L2 sont perçus comme des exemplaires aussi représentatifs d'une catégorie de la L1, et la discrimination serait alors très difficile ; d'après les résultats de l'Expérience 7, ce n'est pas ce qui se produit lorsque le déviant /ki/ intervient dans le contexte de standard /gi/. La situation 2 est celle où deux sons sont assimilés à des exemplaires d'une catégorie de la L1 mais différents dans la capacité à bien la représenter et la discrimination serait alors plus facile ; c'est ce qui se produit probablement dans l'expérience lorsque le déviant /ki/ intervient dans le contexte

de standards /gi/. Dans la situation 3, les deux sons de la L2 sont assimilés à des catégories distinctes dans la L1 et leur discrimination est excellente, ce qui n'est pas le cas dans notre expérience, car le changement de contexte (de rares /gi/ anglais parmi des /ki/ anglais) suffit à supprimer la sensibilité pré-attentionnelle à la différence. Ainsi, dans les termes du modèle PAM, la MMN enregistrée à l'écoute du déviant /k/ anglais en contexte de standards /g/ anglais serait le signe d'un début de différenciation, probablement plus proche d'une différence de proximité au prototype du /k/ français, que d'une franche séparation phonologique entre les /g/ et /k/ anglais dans la L2. Il ne s'agit pas encore d'un nouveau phonème /k/ dans son opposition au /g/ anglais. C'est ce que suggère l'asymétrie de la MMN entre les conditions avec déviant /k/ et avec déviant /g/.

L'asymétrie observée dans les résultats pour la MMN montre que, lorsque le déviant /g/ anglais survient dans un contexte de standards /k/ anglais, il ne provoque pas de MMN. Pour rester cohérents avec l'interprétation qui vient d'être proposée, on peut penser que les standards /k/ anglais sont assimilés à la catégorie /k/ dont ils sont une forme atypique : le VOT est positif comme dans un /k/ français, mais avec une durée exagérée. Dans le modèle PAM, il s'agit toujours bien de la situation 2, où les deux phonèmes entendus en L2 sont des exemplaires ayant des degrés de typicalité différents pour la catégorie phonologique générale /k/, pour l'instant issue de la L1. L'écoute d'un contexte formé de /k/ anglais activerait alors la représentation du /k/ français, et lorsqu'un /g/ anglais survient, il ne surprend pas car il est un bon exemplaire de la catégorie /k/ évoquée par le contexte des standards. Cela expliquerait l'absence de MMN pour le déviant /g/ anglais.

Inclure dans une même catégorie mentale (celle du /k/) les /g/ et /k/ anglais constitue une erreur de catégorisation phonologique pour l'anglais L2 : il faudrait considérer /g/ et /k/ comme deux phonèmes distincts en anglais, mais cela n'est apparemment pas encore possible avant l'entraînement. Cependant, dès cette étape pré-entraînement, l'aspect atypique du /k/ anglais semblait déjà faire reconnaître celui-ci comme très atypique et orienter vers la création d'une nouvelle représentation, et donc d'un nouveau phonème. C'est ce que prévoit la théorie SLM-r (Flege & Bohn, 2021). Cette dernière admet en effet que l'apprentissage tardif d'une L2 puisse générer des modifications de l'espace des représentations phonologiques. L'objectif de notre entraînement était justement d'inciter une telle modification à se produire, plus en profondeur, et de façon plus appropriée à l'anglais. Un début de traitement approprié du contraste de voisement des occlusives anglaises se dessinait déjà avant l'entraînement, avec la marginalisation du /k/ anglais perçu comme atypique du système phonologique de la L1. Le

changement de représentation phonologique n'était cependant pas encore complet, puisque le /g/ anglais n'était pas traité automatiquement comme déviant dans un contexte répétant le standard /k/ anglais.

Ce début de changement de représentation avant l'entraînement risque de rester incomplet, si la situation reste bloquée à cet état. De plus, il est inadapté, car il se fait au prix de l'assimilation erronée du [g] anglais au /k/, ce qui ne sera pas adapté à l'orthographe des mots. Si une des causes du blocage est bien l'erreur d'avoir assimilé le [g] anglais au /k/ français, auquel il ressemble sur le plan acoustique et articulatoire, un entraînement spécifique pour permettre de rétablir une différence de perception et de catégorisation claire des /k/ et /g/ anglais ne pourrait ici se faire qu'en levant l'ambiguïté. Une option (1) serait d'expliquer et faire intégrer aux apprenants qu'il convient, en anglais, de considérer comme un /g/ ce que l'on a l'habitude d'entendre comme un /k/ français et de considérer le /k/ aspiré comme le seul /k/ dans cette L2. Une deuxième option (2) serait d'aider implicitement à percevoir la différence entre les /g/ et /k/ anglais par la mise en évidence répétée de la nécessité de ce contraste pour des distinctions lexicales. Une troisième option (3) serait de contredire l'assimilation du son [g] au phonème /k/ en incitant à associer les graphèmes G et C (ou K) des mots écrits à des indices de durée (VOTs) opposés à ceux de la L1. Il faudrait pour cela insister sur ces aspects phonétiques temporels en anglais au moyen d'une stimulation dissociée des cortex auditifs gauche et droit selon leurs compétences temporelles. L'entraînement proposé dans notre programme combinait les options (2) et (3).

L'assimilation du /g/ anglais au /k/ français est une erreur dont l'apprenant devrait se rendre compte lors des exercices du programme proposé, incluant les représentations orthographiques des mots. Pour aider les apprenants à dépasser cette erreur, des informations sur l'écrit peuvent être d'une aide précieuse : grâce à la lettre G au début d'un mot anglais commençant par un son indûment perçu comme un /k/ français, alors que les lettres C et K sont réservées au /k/ anglais, l'émergence d'une opposition de voisement entre deux consonnes vélaires pourrait s'imposer. L'apport systématique d'informations orthographiques sur les mots anglais pendant l'entraînement était destiné à orienter vers l'apprentissage de cet aspect de la phonologie anglaise. L'entraînement proposé dans l'étude était par ailleurs perceptif. Si l'objectif affiché était d'aider à acquérir du vocabulaire en anglais, l'objectif expérimental était de stimuler de façon répétitive les apprenants avec des mots anglais commençant par des consonnes opposées en voisement. Au-delà d'un entraînement perceptif

classique, les supports cérébraux du traitement des traits phonétiques particuliers au voisement anglais étaient ciblés, sur le plan temporel.

### **5.7.2. Efficacité d'un entraînement perceptif bref (binaural) pour la phonologie chez des apprenants tardifs de l'anglais L2**

Les données comportementales ont montré des progrès en perception catégorielle après l'entraînement perceptif de durée limitée qui a été réalisé, et ce, même en dehors de l'utilisation de stimulations dichotiques.

Dans sa version classique, c'est-à-dire binaurale, l'entraînement a présenté une efficacité significative malgré sa durée globale courte (2h30 d'entraînement, réparti sur 5 jours). L'étude ajoute ainsi un argument supplémentaire à un ensemble de travaux montrant l'intérêt de tels dispositifs brefs (Menning *et al.* 2002 ; Olson, 2019 ; Tamminen *et al.*, 2015 ; Shum *et al.*, 2021). En effet, plusieurs indices montrent, dans les résultats, qu'après l'entraînement binaural les participants ont manifesté une perception plus catégorielle pour la paire de phonèmes entendue durant l'entraînement. Notamment, le pic de discrimination s'est élevé, sans que cette augmentation s'étende au reste de la courbe. De plus, lorsqu'il fallait catégoriser les phonèmes, un déplacement du seuil d'identification s'est produit dans le sens de la norme anglaise, essentiellement grâce à une meilleure catégorisation du /g/ anglais. Ce dernier point est remarquable, car ce phonème est par ailleurs celui qui, lorsqu'il intervenait comme déviant dans une série de /ki/ anglais ne provoquait pas d'indice électrophysiologique d'une sensibilité pré-attentionnelle avant l'entraînement. Le programme avec stimulation binaurale présente des qualités, par sa capacité à rendre plus catégorielle, et donc phonémique, la perception des occlusives sonores anglaises, très ambiguës pour les francophones, et assimilables très facilement au /k/ de leur L1. Il induit aussi un réajustement de la position de la frontière catégorielle qui devient plus conforme à ce qui est attendu en anglais.

L'entraînement binaural a toutefois présenté des limites importantes. En particulier, il n'a pas conduit les apprenants à généraliser leurs nouvelles compétences en perception catégorielle pour la paire de consonne entraînée (/k-g/) vers une paire elle aussi opposée par le voisement mais avec un autre lieu d'articulation (/t-d/), non présentée pendant les exercices. Cette absence de généralisation des compétences, en discrimination comme en identification, peut être interprétée comme le reflet d'un apprentissage assez superficiel de la différence de voisement. En l'absence d'une telle généralisation, il est seulement possible de dire qu'une



nouvelle paire de phonèmes, opposés par le voisement, vient d'être apprise en anglais, mais le voisement lui-même n'est pas acquis pour l'articulation particulière entre les plans phonologique et phonétiques propres à l'anglais.

L'intérêt de dispositifs d'aide à l'apprentissage de nature aussi ponctuelle est d'être faciles à intégrer dans des séances d'enseignement classiques, notamment à l'université, mais aussi, nous l'espérons à plus long terme, au collège et au lycée. L'aide apportée est séduisante, mais nous pouvons aussi anticiper sa dilution si elle n'est pas renforcée de manière à optimiser les bienfaits. Par exemple, il reste à imaginer et recommander des manières d'expliquer et certainement d'illustrer les différences acquises de manière implicite. Il se peut que les bénéfiques soient plus durables grâce à l'explicitation des différences données à entendre ; cela reste à vérifier. Relayer cet apprentissage par des images de spectrogrammes (comme réalisé par Olson en 2019), simplifiés ou non, pourrait aider les apprenants en mobilisant cette fois leur attention sur les indices pertinents et sur les pièges dont ils ont pu être victimes lors d'assimilations trop rapides entre L1 et L2. Ces prolongements ne dispensent pas d'imaginer aussi un renforcement régulier de l'apprentissage du voisement anglais, par la reprise de nouvelles séances avec des exercices comme ceux de nos expériences, par exemple sur un semestre. En vue de permettre une meilleure maîtrise du voisement pour différentes paires minimales, une solution pourrait aussi être d'entraîner avec toutes les paires d'occlusives. Toutefois, l'étude a montré qu'il existe une autre façon d'étendre l'apprentissage à d'autres paires de consonnes : utiliser l'écoute dichotique pour stimuler différemment les hémisphères cérébraux lors de l'apprentissage.

Une limite classique pour ce type de résultat est de ne pas s'assurer de la pérennité de l'effet. Il faudrait pour cela tester les participants encore une fois, un ou plusieurs mois après la fin de l'entraînement. Pour des raisons pratiques, il était difficile de prolonger à ce point l'engagement des participants. De plus, l'interprétation de données recueillies longtemps après un entraînement est toujours délicate, en particulier chez de jeunes adultes pour qui beaucoup d'évènements peuvent interférer avec ce qui est mesuré, pendant la période sans entraînement. De manière très variable entre les individus, sans doute plus encore que chez les enfants, leur accès à l'information ou à la stimulation en anglais peut changer, avec par exemple de nouveaux enseignements en anglais pour les étudiants (et tous ne l'étaient pas), des changements dans les pratiques culturelles (suivre des séries en anglais), communiquer avec des nouveaux collègues ou amis en anglais... ce qui peut introduire de multiples biais. Plutôt que d'ajouter une mesure décalée dans le temps, nous avons choisi de privilégier le



recueil de données supplémentaires avec une autre technique (l'EEG), pour répondre à des questions sur la nature pré-attentionnelle de la nouvelle sensibilité au voisement en L2. Le recueil de ces données électrophysiologiques a été réalisé uniquement pour l'entraînement dichotique.

### **5.7.3. Efficacité de la stimulation dichotique pour optimiser l'effet d'entraînement perceptif au voisement en anglais L2**

La comparaison entre les effets des deux types d'entraînement est possible car, en dehors de la stimulation différente des deux oreilles dans le casque, l'entraînement dichotique était identique à l'entraînement binaural. De plus, il n'y avait pas de différence de performance dans la phase pré-entraînement entre les groupes entraînés de manière binaurale ou dichotique.

L'objectif de la stimulation dichotique dans les exercices était de favoriser le fait que la consonne initiale d'un mot anglais soit envoyée dans l'oreille permettant à l'hémisphère controlatéral de la prendre en charge prioritairement, en faisant en sorte que cet hémisphère soit le plus compétent pour la durée du VOT dans cette consonne. Selon le modèle AST (Poeppel, 2003), il s'agit de l'HG (préférentiellement stimulé par l'information adressée à l'oreille droite) et de l'HD (préférentiellement stimulé par l'information adressée à l'oreille gauche), croisement renforcé par la concurrence dans un dispositif de véritable écoute dichotique, puisqu'un bruit blanc était toujours adressé à l'autre oreille. Les occlusives sourdes anglaises (long VOT) étaient ainsi jouées dans l'oreille gauche pour être traitées prioritairement par l'HD et les sonores anglaises (court VOT) dans l'oreille droite - HG. Ce procédé de stimulation auditive, innovant dans une démarche d'aide à l'apprentissage, tente de travailler sur les associations hémisphère – durée des traits acoustiques les plus adaptées au traitement phonologique du voisement anglais, à partir de ses particularités phonétiques, opposées à celles du français.

#### **5.7.3.1. Arguments comportementaux en faveur de la stimulation dichotique : effet de transfert.**

Plusieurs aspects des résultats comportementaux en perception catégorielle sont encourageants quant à l'intérêt de cette stimulation dichotique pour renforcer l'effet de l'entraînement perceptif au voisement des occlusives anglaises, puisque les effets de l'entraînement dichotique répliquaient ceux qui se produisaient après l'entraînement binaural, mais de façon plus nette, et aussi sur des indicateurs supplémentaires.

L'entraînement dichotique a ainsi permis, comme l'autre entraînement, une amélioration des compétences pour la paire de consonnes apprise lors de l'entraînement – /g-k/, avec une augmentation du pic de discrimination sans augmentation au niveau des autres VOTs, ce qui va dans le sens d'une perception plus catégorielle, avec une frontière plus fermement perçue. La taille de cet effet était un peu plus grande après l'entraînement s'il était dichotique. De plus, d'une manière générale, les tailles d'effet entre les étapes pré- et post-entraînement étaient plus importantes si la stimulation dichotique avait été employée, et la variabilité autour des réponses diminuait aussi plus nettement aux VOT 0 et 10, autour du prototype du /g/ anglais.

Comme l'entraînement binaural, l'entraînement dichotique était par ailleurs suivi d'un déplacement du seuil d'identification vers la norme anglaise, mais cela s'accompagnait cette fois d'une augmentation des réponses correctes pour catégoriser le phonème /g/ anglais aux deux VOTs 0 et 10, alors que cela ne se produisait que pour le VOT 10 après l'entraînement binaural. Le phonème anglais /g/ semble se comporter comme un attracteur plus fort après l'entraînement dichotique, et produit un resserrement des réponses autour de lui. Ce phénomène était moins fort après l'entraînement binaural, et il était présent pour la paire non-entraînée – /d-t/ – seulement après l'entraînement dichotique. Il s'agit là d'un premier effet de transfert, gage d'un apprentissage qui se décontextualise : la généralisation suggère que quelque chose a été appris sur le voisement des occlusives anglaises et pas seulement sur le voisement des occlusives vélaires.

Un autre argument en faveur d'un apprentissage plus profond grâce à la stimulation dichotique est un deuxième effet de transfert seulement après ce type d'entraînement. Un des résultats forts en faveur de ce transfert est le déplacement du seuil d'identification pour la paire non entraînée /d-t/, seulement après l'entraînement dichotique. Ce déplacement était en partie dû à une réduction de l'identification erronée du /d/ anglais comme un /t/ au niveau de la frontière. Cet infléchissement de la courbe a eu lieu au niveau du VOT 20 ms et a rapproché ce VOT du seuil d'identification. De ce fait, après l'entraînement dichotique, la frontière catégorielle attendue en anglais correspondait cette fois au niveau d'ambiguïté maximale perçue par les participants, ce qui les rapprochait d'une perception catégorielle anglaise. Un troisième argument en faveur d'un effet de transfert seulement après l'entraînement dichotique a été relevé. Le pic de discrimination pour la paire /d-t/ augmentait seulement après cet entraînement. Un tel transfert de compétence, ou généralisation, plaide en faveur de la profondeur de l'apprentissage induit par l'entraînement dichotique.

Le dispositif dichotique optimise l'effet de l'entraînement perceptif proposé dans l'étude, et pourrait efficacement participer l'aide à l'apprentissage de nouveaux phonèmes en anglais L2 auprès d'adultes francophones. Ce résultat apporte aussi un argument supplémentaire au modèle AST (Poeppel, 2003 ; Poeppel *et al.*, 2008) qui constituait le fondement théorique de l'expérience. L'asymétrie et la complémentarité hémisphérique pour les traitements temporels opérés sur le signal de parole semblent conformes à ce que proposent aussi l'ensemble des modèles d'échantillonnage à échelle multiples proposés en neurosciences (Hickok & Poeppel, 2007) pour les étapes précoces du traitement de la parole dans le cortex auditif.

### **5.7.3.2. Arguments électrophysiologiques en faveur de la stimulation dichotique**

Sur le plan de la sensibilité pré-attentionnelle au voisement anglais, l'analyse des données en potentiels évoqués a fait ressortir quatre principaux résultats témoignant de modifications des réactions cérébrales au voisement des occlusives anglaises après l'entraînement dichotique.

Le premier et le plus important est un accroissement de la MMN pour l'occlusive sonore anglaise après l'entraînement, seulement si celui-ci était dichotique. En effet, avant l'entraînement, le déviant /g/ anglais parmi les /k/ anglais ne provoquait pas de MMN. L'apparition de la MMN pour ce déviant à la suite de l'entraînement dichotique semble montrer que les participants ont créé une représentation mentale de cette catégorie phonémique qui était auparavant probablement assimilée aux sourdes françaises. Les occlusives sonores anglaises constituent la catégorie la plus difficile à interpréter sur le plan phonologique pour un francophone, à cause de leur proximité avec les occlusives sourdes de la L1. La stimulation dichotique semble participer à dépasser cette ambiguïté. L'émergence de la MMN pour l'occlusive sonore anglaise et la persistance de la MMN pour l'occlusive sourde anglaise constituent des arguments sur le plan neuronal pour la création de nouvelles connaissances phonologiques propres au voisement anglais, indépendamment du système phonétique-phonologique français.

Cela va dans le sens du modèle SLM-r (Flege & Bohn, 2021), concernant le point qui l'oppose au modèle PAM : l'apprentissage tardif de nouveaux phonèmes serait possible. Le maintien de dispositions cérébrales pour l'analyse élémentaire de l'information acoustique pourrait participer à cet apprentissage tardif. L'inhibition qui les réduit pour restreindre la sensibilité aux seuls aspects phonétiques compatibles avec le système phonologique de la L1

n'apparaît pas insurmontable. Des dispositifs d'aide à l'apprentissage tenant compte des biais liés aux associations phonétiques-phonémiques de la L1 doivent pour cela être mis en place. La stimulation dissociée des cortex auditifs gauche et droit en fonction de leurs compétences complémentaires pour l'analyse temporelle, ajoutée à l'apport d'informations orthographiques, à l'entraînement perceptif court mais ciblé, dans un contexte d'apprentissage lexical rendant le contraste de voisement incontournable, semble une combinaison prometteuse.

Un autre résultat en faveur d'un changement positif pour le traitement de l'occlusive sonore anglaise après l'entraînement dichotique est l'observation d'une diminution de l'amplitude de l'onde P2 pour ce phonème. Les modifications d'amplitude de la P2 sont moins traditionnellement explorées pour évaluer l'apprentissage dans le domaine phonologique. La diminution d'amplitude de cette onde est pourtant un phénomène intéressant, interprété comme le reflet d'un filtrage précoce et non-conscient – *sensory-gating* –, qui consisterait à éviter à des régions cérébrales impliquées dans des analyses de plus haut niveau d'être submergées par des informations de moindre intérêt, peu surprenantes (Venables, 1964). Ici, la diminution d'amplitude de l'onde P2 pour le déviant /gi/ seulement après l'entraînement dichotique suggère que le son n'est plus perçu comme étrange au point de nécessiter une exploration plus fine, puisqu'il fait déjà partie d'un système de représentation. En bloquant l'approfondissement du traitement de telles informations, ce mécanisme de filtrage est adaptatif (Lijffijt *et al.*, 2009a). L'interprétation de la diminution de l'onde P2 dans certaines conditions expérimentales par rapport à d'autres est encore l'objet de débats. Cependant, des données suggèrent qu'elle est associée à un message adressé par des mécanismes précoces en perception vers des mécanismes plus élaborés, les incitant à ne pas s'engager fortement, pour ne pas dépenser des ressources inutilement (Benikos *et al.*, 2013). Dans notre problématique, la diminution d'amplitude de l'onde P2 à l'écoute du déviant /gi/ après l'entraînement dichotique serait ainsi un témoin supplémentaire de modifications des réactions cérébrales à une différence phonologique. Plus précisément, elle pourrait traduire le fait que ce qui est perçu est devenu suffisamment familier pour ne pas nécessiter une grande dépense de ressources d'analyse : l'accès aux représentations phonologiques pour l'identification suffirait, ce qui est bien manifesté pour le déviant /gi/ anglais, par le couple formé d'une onde P2 de faible amplitude et d'une ample MMN, après l'entraînement dichotique.

Le troisième changement est observé au sujet des réactions à l'occlusive sourde anglaise /k/, au niveau de la MMN. Après le programme, la réaction à la déviance pour le /k/ anglais parmi des /g/ anglais est demeurée présente, et ce, avec une MMN d'amplitude aussi grande qu'avant l'entraînement. Cependant, sa localisation s'est modifiée. Principalement latéralisée sur les électrodes de l'HG avant l'entraînement, elle était enregistrée bilatéralement après ce dernier. Ce changement peut être interprété comme la disparition de la spécificité de l'HG (voire un certain désengagement de celui-ci) pour traiter le /k/ anglais et son long VOT. Avant l'entraînement, le cerveau réagissait au /k/ anglais comme s'il s'agissait d'un /k/ français contenant un VOT court facile à traiter par l'HG. L'apprenant tentait peut-être de le traiter phonologiquement comme un /k/ français, bien qu'il soit phonétiquement atypique de cette catégorie. Il est intéressant de remarquer qu'après l'entraînement dichotique l'HD participe davantage au traitement du /k/ anglais, car cela se justifie par son efficacité pour traiter la durée du long VOT du /k/ anglais. Ce changement de l'asymétrie hémisphérique pour le traitement du déviant /k/ anglais montre que ce n'est pas seulement le traitement du /g/ anglais qui se modifie après l'entraînement dichotique : la paire /g-k/ est traitée autrement. La disparition de la dominance gauche pour le traitement de l'occlusive sourde anglaise /k/ pourrait traduire une réorganisation fonctionnelle inter-hémisphérique tout à fait favorable à un traitement phonétique plus adapté en anglais, et donc à la diminution des risques d'erreurs de communication. Il serait intéressant de savoir si cette réorganisation fonctionnelle est due à la stimulation dichotique ou simplement à l'entraînement perceptif. Pour tester cette question, il conviendra de conduire une expérience d'entraînement équivalente, mais avec stimulations binaurales, avec enregistrement des potentiels évoqués avant et après. Une telle expérience a bien été réalisée (Expérience 6), mais avec seulement le recueil de données comportementales et pas de données en EEG avant et après le programme. Il s'agit d'un prolongement envisageable pour l'étude présentée ici.

Le dernier changement observé est une petite augmentation de l'amplitude de l'onde P3 pour l'occlusive sourde anglaise après l'entraînement dichotique. Cette dernière est un indicateur supplémentaire au sujet d'un changement neuro-fonctionnel pour le traitement du /k/ anglais après l'entraînement dichotique. Il accompagne vraisemblablement un traitement plus adapté à la langue anglaise, puisque les données comportementales montrent que les occlusives sourdes sont mieux traitées en PC après l'entraînement dichotique. L'augmentation d'amplitude de la P3 pourrait refléter le traitement plus approfondi de cette consonne, après son identification, soit parce qu'elle est devenue un nouveau phonème – le /k/ anglais perçu

dans sa singularité encore un peu étrange –, soit parce qu'elle est perçue dans un large espace phonétique-phonologique qui s'adapte, tel que le propose Flege (1995), et où elle est un /k/ atypique.

#### **5.7.4. Aspects cognitifs et neuronaux de l'effet du dispositif dichotique sur l'apprentissage du voisement anglais**

La stimulation dichotique comme aide à l'apprentissage du voisement en anglais L2 semble avoir un intérêt pour de jeunes adultes francophones. Étant donné l'inversion entre les indices de durée pour le voisement des occlusives entre l'anglais et le français, l'apprenant francophone doit apprendre à utiliser autrement les compétences de régions cérébrales spécialisées pour le traitement des indices acoustiques courts et longs lorsqu'il passe de la L1 à la L2. Il est confronté à des phonèmes orthographiés de la même façon dans ces langues, alors que leurs caractéristiques temporelles phonétiques ne sont pas adaptées à une même utilisation des compétences d'analyse temporelle des cortex auditifs gauche et droit. Inciter les apprenants francophones à engager fortement l'HG dans le traitement des occlusives sourdes et l'HD dans le traitement des occlusives sonores lorsqu'il sait qu'ils doivent traiter de la parole anglaise est apparu comme une méthode efficace pour qu'ils apprennent les particularités phonétiques temporelles du voisement en anglais.

Il s'agissait d'un apprentissage implicite, puisque rien n'est expliqué au participant, dont l'attention est même attirée vers l'enrichissement de son lexique. Il apparaît qu'en seulement 2h30 d'exercices étalés sur 5 jours, des connaissances phonologiques peuvent se mettre en place. Elles sont de bonne qualité, et non superficielles, puisque les caractéristiques temporelles du voisement anglais sont mieux traitées après l'entraînement dichotique pour des consonnes qui, dans les exercices, n'ont pas fait directement l'objet de ces associations occlusives sourdes - HD d'une part, et occlusives sonores - HG d'autre part. Le voisement anglais et son support phonétique si différent de celui du voisement français est mieux maîtrisé, de manière abstraite, c'est-à-dire phonologique. Ce résultat va dans le sens de l'élaboration d'une règle phonologique nouvelle et de catégories phonémiques nouvelles, c'est-à-dire l'élaboration d'un système phonologique pour la L2, indépendant de celui de la L1. C'est ce qui est attendu pour l'avancée dans la maîtrise d'une langue étrangère.

Sur le plan neurologique, lorsque des phonèmes sont extraits correctement et que leurs traits phonétiques sont bien encodés dans le signal de parole, comme cela devient le cas dans nos expériences après l'entraînement, la région-clef alors activée serait le gyrus temporal

supérieur (STG). Dans cette aire cérébrale, une réponse sélective aux divers traits phonétiques a été mise en évidence à partir d'enregistrements de surface corticale directe à haute densité (Mesgarani *et al.*, 2014). Par exemple, pour des anglophones, une électrode relève une réaction cérébrale à l'écoute de toutes les occlusives dont le VOT est inférieur à 36 ms (sonores), alors qu'une autre enregistre une réaction seulement aux occlusives dont le VOT dépasse 36 ms (sourdes) : il s'agit bien d'une réaction non-linéaire (i.e. catégorielle) au trait de voisement, à travers une sensibilité fine et catégorielle à un indice de durée, et cela pour toutes les occlusives. Cela correspond à un plus grand niveau d'abstraction qu'une réaction spécifique à chaque phonème. L'encodage d'un trait est assuré par la réponse d'une population de neurones distribués dans cette région spécialisée pour les représentations acoustiques-phonétiques de la parole chez l'Homme. Les changements survenus dans le traitement du voisement anglais chez nos participants après l'entraînement pourraient s'appuyer sur ces capacités de réaction spécifiques au niveau du STG. La nature de nos données ne nous permet pas d'assurer qu'il en est ainsi, mais il s'agit d'un mécanisme probable qui, chez des apprenants tardifs, reposerait sur une bonne plasticité cérébrale sur le plan fonctionnel.

L'amélioration des connaissances phonologiques en L2, et le fait d'échapper à une assimilation approximative et inadaptée au système phonologique de la L1, n'est certes qu'un aspect de l'apprentissage d'une nouvelle langue, mais le fait que ce type d'apprentissage soit possible chez des apprenants tardifs est très encourageant. Quelques indices électro-cérébraux recueillis dans cette étude suggèrent que de nouvelles représentations mentales se sont créées pour les phonèmes de la L2 : c'est ce dont témoigne notamment l'accentuation de la MMN en situation de *oddball* passif. Cette accentuation après l'entraînement est d'autant plus marquante que la MMN provoquée par des déviants qui se démarquent par leur durée est peu sensible au simple effet de répétition du testing (i.e. peu d'effet test-retest, Chobert *et al.*, 2012). De plus, l'analyse des potentiels évoqués selon la latéralité des électrodes a montré que la plasticité fonctionnelle reste forte chez l'apprenant adulte. Non seulement la sensibilité cérébrale peut s'accroître pour des différences de sons nouveaux par rapport au registre phonémique de la langue maternelle, mais, sous l'effet d'une courte série d'exercices ciblant l'asymétrie hémisphérique fonctionnelle, de nouvelles habitudes peuvent se créer rapidement et conduire à utiliser, autrement que par assimilation à la L1, les compétences d'analyse temporelle de l'HG et de l'HD d'une manière adaptée à la L2. Cet apprentissage est d'autant plus intéressant qu'il ne se fait pas au prix de l'abandon d'une utilisation toute différente de



ces mêmes compétences d'analyse temporelle pour la phonologie en L1, puisque le voisement en français reste toujours aussi bien traité de façon catégorielle.

#### **5.7.5. La stimulation dichotique comme aide à l'apprentissage du voisement en anglais L2 chez des adultes dyslexiques**

Les participants dyslexiques ont tous réalisé l'entraînement de façon dichotique. Compte tenu de leur pathologie d'apprentissage, il semblait important de leur proposer en priorité le dispositif le plus susceptible de les aider. A la suite de cet entraînement, plusieurs indices dans l'analyse des performances montrent une amélioration dans le sens d'une perception plus catégorielle en anglais.

Ainsi, le pic de discrimination augmentait après l'entraînement pour la paire de consonnes /g-k/ utilisée pour l'entraînement. La significativité de ce progrès est intéressante, car un entraînement perceptif – sans dispositif dichotique – ciblant lui aussi la perception phonémique (mais dans la langue maternelle des participants dyslexiques) tendait à avoir le même effet sur le pic de discrimination, mais de façon marginale sur le plan statistique (Zoubrinetzky *et al.*, 2019). L'analyse des données montre aussi que la position du pic de discrimination en anglais s'est modifiée, avec une augmentation de la sensibilité à la différence au niveau du couple de VOT 10-30 ms. Cette modification au niveau du pic pourrait traduire d'un début de modification de la frontière catégorielle vers la norme anglaise.

Il faut cependant noter la persistance du pic pour le couple VOT 0-20. Ce maintien est étonnant, et pourrait traduire un mode d'apprentissage singulier chez les dyslexiques. Ils sont connus pour laisser coexister des frontières catégorielles multiples pour un seul contraste phonologique. Cela correspond à leur tendance à percevoir la parole de façon allophonique, avec la prise en compte de frontières qui ne sont pas utiles dans une langue (Serniclaes *et al.*, 2001, 2004, 2021 ; Serniclaes & Seck, 2018), et souvent une zone de pic élargie dans la courbe de discrimination (Maassen *et al.*, 2001), voire deux pics pour un même contraste (Bogliotti *et al.*, 2002). Cette anomalie est censée disparaître des données comportementales lorsque les dyslexiques deviennent adultes (il en reste des traces dans les enregistrements neurophysiologiques) (Noordenbos *et al.*, 2012), mais nos données montrent que la situation d'apprentissage de nouveaux phonèmes en L2 permet de retrouver des indices de cette tendance allophonique dans les données comportementales. Certaines recherches associent la perception de type allophonique à une instabilité des réponses des oscillations cérébrales aux



caractéristiques acoustiques du signal, réduisant les compétences neuronales à s'adapter pour un apprentissage statistique (Hornickel & Kraus, 2013), ce qui donnerait lieu à des représentations de phonèmes qui, d'après nos données, peuvent se modifier, mais avec un certain manque de robustesse. Cette modification est néanmoins encourageante, car elle va globalement dans le bon sens. Nous pouvons donc émettre l'hypothèse que certains dispositifs d'aide puissent permettre à certains dyslexiques de dépasser les anomalies de connectivité entre les cortex auditifs gauche et droit et le gyrus frontal inférieur gauche (Boets *et al.*, 2013) ou l'ensemble de la voie dorsale (Dufor *et al.*, 2007), à la source de difficultés à récupérer des connaissances sur les phonèmes, mais sans doute aussi à l'origine de difficultés à créer ici des liens entre les traitements temporels de bas niveau et les traitements phonologiques de plus haut niveau.

Ce progrès au niveau du pic de discrimination est aussi observé chez les adultes dyslexiques pour la paire de consonnes qui n'a pas été entraînée. Il s'agit d'un transfert de compétences d'autant plus remarquable qu'il n'est pas observé chez tous les normo-lecteurs. Il semble qu'une généralisation des compétences acquises sur les supports phonétiques temporels spécifiques au voisement anglais soit possible, même en cas de dyslexie. De plus, pour cette paire, le seuil d'identification s'est déplacé dans le sens attendu en anglais. Ces résultats forment un ensemble cohérent pour argumenter un progrès pour une perception catégorielle plus adaptée à l'anglais. Le fait que les progrès sur /d-t/ s'observent sur la hauteur du pic et sur le seuil est également intéressant, car ce sont deux variables dépendantes qui se sont montrées fidèles pour cette paire dans l'évaluation des effets test-restest (Chapitre 4).

Les adultes dyslexiques inclus dans l'étude se sont montrés de bons répondants au dispositif proposé pour apprendre le contraste de voisement, de manière généralisée, en anglais. Cela se confirme aussi avec la diminution de la variabilité de leurs réponses à T2 en discrimination. Malgré leur pathologie, ils peuvent tirer un certain bénéfice, mesurable, d'un entraînement perceptif par stimulation dichotique de seulement 2h30 réparti sur 5 jours.

L'observation des courbes d'identification chez les participants dyslexiques soulève un point sur la pente de la courbe, en tant que variable dépendante. Même si visuellement la pente d'identification semble s'améliorer après l'entraînement, cette mesure est peu sensible dans notre étude (cf. Chapitre 2 – Expérience 1). Il est possible que les faibles effets observés sur la pente d'identification viennent de la faible adaptation du mode de calcul de celle-ci au cas des dyslexiques. Ce calcul consiste à ajuster les données du participant à une fonction. Or, les performances des dyslexiques en identification sont parfois si faibles qu'elles ne se

distribuent pas du tout sous la forme d'une courbe ayant l'allure d'une sigmoïde. Un biais peut alors survenir, car l'ajustement est tout de même fait, et la valeur de pente qui en ressort, bien qu'un peu faible, est surestimée par rapport à la faiblesse des performances. Comme des courbes atypiques peuvent apparaître particulièrement à T1, la surévaluation de la pente avant l'entraînement a pu empêcher d'observer une amélioration après l'entraînement.

Au niveau comportemental, l'entraînement dichotique semble avoir eu un effet bénéfique sur la perception phonologique du voisement anglais pour notre groupe d'adultes dyslexiques. Nous n'avons pas d'information de nature électrophysiologique, et un prolongement de cette recherche pourrait consister à réaliser le test de *oddball* passif déjà proposé aux adultes non-dyslexiques, avant et après l'entraînement, afin de pouvoir observer un éventuel changement au niveau cérébral. Cela n'a pas été une priorité ici car, en cas de dyslexie, les ondes habituellement observées en présence d'une déviance phonologique (i.e. MMN) ne sont pas toujours présentes chez les dyslexiques (Shulte-Körne *et al.*, 1998, 1999, 2001) même si nous en avons mesuré dans leur langue maternelle (Chapitre 3 – Expérience 3).

#### **5.7.6. La stimulation dichotique : un entraînement favorisant le traitement des sons courts anglais**

Après l'entraînement réalisé avec stimulation dichotique, les progrès sont parfois asymétriques entre les consonnes sonores et sourdes. Il est apparu que, dans tous ces cas, c'était au profit de des occlusives anglaises contenant un VOT court. Ici, l'entraînement dichotique semble donc avant tout renforcer le traitement approprié des sons courts. Après les exercices réalisés dans ces conditions, les normo-lecteurs ont en effet parfois fait des progrès en PC pour une paire, mais surtout pour celle ayant un VOT bref : l'occlusive sonore anglaise. Par exemple, les progrès en identification correspondent surtout au resserrement des réponses autour des prototypes du /g/ et du /d/ anglais, qui ont de courts VOT ; dans le cas des dyslexiques ce resserrement, indiquant une perception moins allophonique, est présent seulement pour la sonore /g/. Chez les normo-lecteurs, d'après les potentiels évoqués, c'est aussi pour le phonème porteur du VOT court (/g/) que la MMN augmente après l'entraînement.

Cela peut en partie s'expliquer par le fait que les occlusives ayant un VOT positif court sont celles qui créent le plus d'ambiguïté au départ pour les francophones, car un tel indice de durée, positionné après le bruit d'explosion, doit être interprété comme une sourde dans leur

L1, mais comme une sonore en anglais L2. L'ambiguïté est moindre pour les sourdes anglaises, qui sont surtout atypiques. Une plus grande marge de progression serait donc possible pour les sonores anglaises. Le problème peut aussi s'envisager sous un autre angle : les durées de VOT qui posent problème au francophone apprenant le voisement anglais sont les durées courtes, après le bruit d'explosion. Une zone d'ambiguïté se formerait ici, dans l'espace phonétique.

L'origine de l'amélioration de la perception des sons courts anglais après l'entraînement dichotique pourrait être interprétée de différentes manières.

Une hypothèse serait que la stimulation répétée du cortex auditif gauche par des stimuli contenant un VOT court renforce la puissance des oscillations rapides dans cette aire cérébrale. Une autre interprétation, sans doute plus crédible en cas d'entraînement bref, est que la stimulation répétée du cortex auditif gauche avec des mots initialisés par une consonne contenant un VOT bref renforcerait les capacités de synchronisation des oscillations rapides dominantes dans cette région de l'HG. Cette synchronisation des oscillations cérébrales au rythme des informations entrantes est supposée déficitaire en cas de dyslexie (Cutini *et al.*, 2016 ; Di Liberto *et al.*, 2018 ; Goswami, 2011 ; Lizarazu *et al.*, 2015), et l'échantillonnage de la parole en fines unités, sur la base de l'*entrainement* des oscillations gamma, a été décrit comme défaillant dans cette pathologie (Dushanova *et al.*, 2020; Lehongre *et al.*, 2013 ). Le dispositif d'entraînement dichotique pourrait améliorer cette compétence, au moins transitoirement, y compris chez des adultes où elle est anormalement peu développée. Même si elle est transitoire, une telle amélioration pourrait faire davantage prendre en considération la véritable durée des VOT des phonèmes anglais, pour échapper à une assimilation en l'occurrence trompeuse. Cette prise en compte est finalement une orientation de l'attention vers un aspect déterminant pour apprendre de nouvelles règles d'association phonétiques-phonologiques.

Une autre possibilité est que la stimulation spécifique du cortex auditif gauche par des informations contenant des informations cruciales de courte durée renforce l'automatisation de la prise en charge de tout indice acoustique court par l'HG, ce qui serait favorable au traitement du court VOT du /g/ anglais, même si un indice long est attendu pour cette consonne en L1. Rien ne permet encore de trancher entre ces hypothèses sur l'origine détaillée des progrès.

Par ailleurs, les exercices faisaient alterner fréquemment entre la présentation d'un mot à gauche ou à droite, sollicitant la complémentarité inter-hémisphérique pour le traitement du langage de la manière la plus appropriée. Si le cortex auditif gauche, par exemple, est capable de traiter des indices courts et longs, l'alternance entre des mots commençant par des VOTs courts et longs dans les exercices dichotiques incite probablement le cortex auditif droit à prendre en charge plus systématiquement les VOTs longs et donc à identifier les occlusives sourdes anglaises, pour répartir les efforts. Aider à mieux répartir ces fonctions pourrait être source d'amélioration en perception de la parole.

Enfin, une interprétation en termes attentionnels est aussi possible. L'entraînement perceptif dichotique permet peut-être, en distribuant judicieusement les informations sur les durées, d'attirer l'attention des apprenants sur ce qui est particulier à la phonétique du voisement anglais : la brièveté du VOT des occlusives sonores anglaises et/ou leur ressemblance avec les sourdes françaises, ainsi que la longue aspiration des occlusives sourdes dans cette L2. Sans aller jusqu'à parler d'une prise de conscience de ces caractéristiques, l'orientation sélective de l'attention temporelle dans l'analyse des indices acoustiques de durée est organisée, avec l'écoute dichotique, de sorte que cette attention puisse s'associer de manière répétitive à des unités infra-phonémiques adaptées au traitement optimal du voisement en anglais. Une attention mieux distribuée pendant un apprentissage garantit sans doute le succès de celui-ci. Cette interprétation attentionnelle est bien sûr tout à fait compatible avec celles qui évoquent l'ajustement de rythmes cérébraux distincts dans l'HG et dans l'HD.

#### **5.7.7. Processus cognitifs sous-jacents : mécanismes *top-down* et *bottom-up* ?**

Les résultats ont apporté des arguments sur l'intérêt de la stimulation dichotique par rapport à la stimulation binaurale dans le dispositif d'aide à l'apprentissage du contraste de voisement en anglais L2 par des francophones. Cependant, la nature des mécanismes impliqués dans les changements produits par la stimulation dichotique reste à mieux comprendre. Les conséquences de l'entraînement dichotique pourraient être doubles : l'un de nature *top-down*, l'autre *bottom-up*.

Tout d'abord, il a permis la création de représentations phonologiques plus adaptées pour les couples de phonèmes anglais opposés en voisement. Cela dote l'apprenant de connaissances nouvelles et appropriées qui peuvent rapidement l'aider sur le versant *top-down* des mécanismes d'identification de phonèmes. Des données montrent que des représentations

phonémiques (i.e. spécifiques au langage) participent à la survenue de la MMN dans les paradigmes de *oddball* sur la sensibilité aux traits phonologiques. En effet, des stimuli artificiels contenant les mêmes caractéristiques que les phonèmes mais sous une forme simplifiée ne permettent pas de reproduire une MMN comme à l'écoute d'un déviant issu de la parole naturelle (Christmann *et al.*, 2014). Des activations spécifiques à la perception de la parole, non relevées pour des sons non-langagiers de complexité pourtant très proche, dans la partie antérieure et médiane du cortex temporal supérieur gauche en témoignent également (Narain *et al.*, 2003 ; Scott *et al.*, 2000), même en l'absence de contenu sémantique, par exemple pour des syllabes CV (Liebenthal *et al.*, 2005 ; Obleser *et al.*, 2007), surtout si les phonèmes font partie de l'inventaire phonémique du participant (Näätänen *et al.*, 1997 ; Chladkova *et al.*, 2013). Tout cela renvoie au rôle de représentations linguistiques dans la perception phonologique appropriée à une langue, telle qu'elle semble émerger après l'entraînement dichotique.

Par ailleurs, les modifications du tracé EEG pour des ondes très précoces comme la P2 suggèrent en plus un effet de l'entraînement sur la prise en charge cérébrale des indices acoustiques pertinents dans le signal de parole anglais à un niveau de traitement élémentaire (*bottom-up*). En effet, après l'entraînement dichotique, le déviant /gi/ anglais produit une MMN plus ample, mais aussi une onde P2 dont l'amplitude diminue. Or, la réduction d'amplitude de la P2 est interprétée comme l'indice d'un *sensory-gating*, mécanisme attentionnel très élémentaire empêchant les fonctions cérébrales de plus haut niveau de s'impliquer dans des analyses profondes. Ce mécanisme permet d'éviter d'être submergé par des informations peu pertinentes et de dépenser inutilement les ressources de traitement. Les sons courts semblent devenir ici moins étranges après l'entraînement, puisque le *sensory-gating* se met en place.

#### **5.7.8. Processus cognitifs sous-jacents : une amélioration des traitements temporels ?**

L'entraînement dichotique avec du matériel anglais a eu des retentissements sur le traitement phonologique du français chez les adultes dyslexiques. Ils vont dans le sens d'un meilleur traitement général des durées.

Après l'entraînement en français, les dyslexiques ont réalisé des progrès en PC de consonnes françaises (cf. Chapitre 3). L'entraînement dichotique qu'ils ont ensuite reçu en anglais n'a pas fait disparaître ces progrès. Pourtant, les participants ont subi, pour ces deux

entraînements, des stimulations hémisphériques opposées en référence aux phonèmes... mais cohérentes en référence à la durée des VOTs. La persistance des progrès en français est rassurante, et suggère que c'est bien le traitement approprié des durées qui est favorisé par le dispositif dichotique tel que proposé, dans son ensemble.

L'entraînement en L1 proposé était avant tout une précaution pour protéger le système phonologique français des dyslexiques, réputé pour sa fragilité, avant l'entraînement en L2. Il a cependant fait plus que protéger le système de la L1, il l'a renforcé. Cela pourrait avoir contribué au succès de l'entraînement ultérieur en anglais dans ce groupe. Le modèle SLM-r insiste sur le fait que plus le système phonologique de la L1 est stable avec des catégories bien délimitées, plus l'apprentissage de la L2 se fera correctement. Il est possible qu'avoir renforcé, par la remédiation dichotique en français, le système phonologique de nos participants dyslexiques en L1, les ait aidés à créer de nouvelles représentations mentales des catégories anglaises pour le voisement. Cet enchaînement d'entraînements est sans doute un aspect qui devrait rester dans les recommandations pour de futures applications.

Un autre argument en faveur de l'effet de l'entraînement dichotique sur le traitement des durées en général est le renforcement du traitement des sons français au niveau des zones antérieures (et non plus centrales) du cerveau chez des participants sans pathologies. Sans changer d'amplitude ni de latence, la MMN était en effet dominante dans une position plus typique (antérieur) de sa manifestation pour des traitements temporels.

Un autre changement cohérent avec l'hypothèse d'un bon potentiel de la stimulation dichotique pour améliorer le traitement de plusieurs types de durées est une amélioration de la pente d'identification en français uniquement après l'entraînement par stimulation dichotique, toujours chez les normo-lecteurs. Les pentes deviennent alors plus abruptes à la fois parce que l'occlusive sonore /g/ française (avec un long VOT) est mieux identifiée, mais aussi parce que l'occlusive sourde /k/ (avec un court VOT) est perçue de manière moins allophonique. Ainsi, l'entraînement dichotique en anglais a exercé un transfert de compétences pour traiter les durées dans une autre langue : la L1. Même si nous avons vu que certaines données suggèrent un effet surtout sur les consonnes dotées d'un court VOT, le dispositif de stimulation dissocié des cortex auditifs gauche et droit semble capable d'améliorer les deux aspects des traitements temporels, d'une manière dissociée selon les durées (i.e. pour les courtes et les longues).

La stimulation dichotique semble plus efficace que la stimulation binaurale pour l'apprentissage des règles phonétiques-phonologique du voisement en anglais L2 pour des

francophones, qu'ils présentent un trouble du langage ou non. Il serait intéressant de réaliser ce même type d'entraînement auprès d'anglophones souhaitant apprendre le français en L2. En effet, il est possible d'imaginer un entraînement de même durée, travaillant sur l'opposition du voisement en français avec les mots commençant par des occlusives sonores françaises adressées à l'oreille gauche – HD, et des occlusives sourdes françaises à l'oreille droite – HG. Tout comme en anglais, il existe en français de nombreuses paires minimales variant uniquement par la première lettre et opposées par le voisement comme « bain » et « pain », qu'il est impératif de bien identifier pour éviter les malentendus.

Dans la suite de cette thèse, l'effet d'un entraînement dichotique a été testé sur une autre opposition basée sur la durée et importante dans la langue anglaise : les voyelles courtes/longues (Chapitre 6).

### **5.7.9. Limites et résultats inattendus**

Les données en PC de l'anglais pour des paires d'occlusives opposées en voisement ont fait ressortir une différence de position du pic de discrimination selon le lieu d'articulation de la paire avant comme après l'entraînement. Le pic était au VOT +10 ms pour les vélares /k-g/ anglaises et au VOT +20 ms pour les alvéolaires /t-d/ anglaises. L'observation d'une différence n'est pas étonnante, puisque les durées des VOTs sont variables et très sensibles aux lieux d'articulation (Cho & Ladefoged, 1999). Cependant, le sens de cette différence ne correspond pas à ce qui pourrait être prédit en se basant sur la description du VOT moyen en anglais de Lisker et Abramson (1964) (Tableau 13). La position du pic relevé dans notre étude pour les vélares /k-g/ situe la frontière catégorielle à +10 ms, et donc un peu avant la durée moyenne du court VOT positif pour /g/ selon Lisker et Abramson (21 ms). Par contre, la durée moyenne du VOT anglais pour l'alvéolaire /d/ serait de seulement +5 ms, et les réponses de nos participants situent la frontière catégorielle à +20 ms, c'est-à-dire plus tard que pour les vélares. Cela suggère que les participants francophones de notre étude ont un niveau trop faible en anglais, avant comme après l'apprentissage, pour connaître la façon subtile dont le VOT varie en fonction des lieux d'articulation dans cette L2. L'interaction subtile entre VOT et lieu d'articulation n'était pas la cible de l'entraînement et il n'est pas surprenant que les positions des frontières catégorielles pour l'anglais soient peu typiques et le restent après les exercices.

Notons que cette anomalie ne s'explique pas directement par une assimilation au français. En effet, en français, avant la voyelle /i/, le court VOT positif du phonème /t/ est



légèrement plus bref que celui du phonème /k/, et le long VOT négatif du /d/ est nettement plus long que celui du /g/ (Wallet, 2015). Cette configuration ferait attendre, comme en anglais, une frontière catégorielle plus proche de 0 pour les alvéolaires /t-d/ que pour les vélaires /k-g/, ce qui n'est pas le cas dans les réponses en discrimination chez nos participants en anglais. La position variable de la frontière catégorielle selon le lieu d'articulation des consonnes anglaises ne s'explique pas simplement par une assimilation à la L1.

Ce résultat peut être discuté par rapport au modèle PAM (Best, 1995 ; Best *et al.*, 2001), initialement proposé pour le bilinguisme puis étendu à l'apprentissage d'une L2 (PAM-L2, Best & Tyler, 2007). Une de ses hypothèses dit que, lorsque des consonnes se ressemblent phonétiquement, comme le /g/ anglais et le /k/ français, la consonne de la L2 est traitée comme un exemplaire de la catégorie abstraite qui lui ressemble le plus en L1 (ici, le /k/ français). La méthode d'évaluation quantitative « cross-language assimilation overlap » a confirmé l'intérêt d'étendre le modèle PAM à l'apprentissage d'une L2, car la similarité perçue entre la langue maternelle et la L2 prédit globalement la précision avec laquelle les phonèmes de la L2 sont discriminés, pour les voyelles (Levy, 2009) et les consonnes (Guyon *et al.*, 2000). Mais nos résultats montrent que l'assimilation n'explique pas toutes les données.

Le modèle SLM-r (Flege & Bohn, 2021) considère qu'un phonème en L2 peut être traité comme « identique », « similaire » ou « nouveau », par rapport à l'inventaire phonologique natif. A T1, en l'absence d'information orthographique, le /g/ anglais devrait alors être perçu comme « identique » au /k/, ce qui conduit à une erreur. Nous avons vu que cette assimilation explique en partie les résultats. Elle n'explique cependant pas les différences de frontières de voisement entre les occlusives anglaises selon leur lieu d'articulation, ni les progrès après l'entraînement dichotique. Comme le soutient par ailleurs Flege (1995), il est possible pour un apprenant tardif d'aller au-delà, et de profiter de la plasticité de l'espace phonétique phonologique pour apprendre, certes avec difficulté ou avec des aides adaptées, de nouveaux phonèmes.

Tableau 13. Durée moyenne du VOT (ms) en français (Serniclaes, 1987) et en anglais (Lisker & Abramson, 1964) pour les occlusives vélaires et alvéolaires

	/d/	/g/	/t/	/k/
Français		-110		20
Anglais	5	21	70	80



Pour finir, le cas du voisement anglais pour un francophone, est si particulier qu'il ne correspond directement à aucune des trois situations prévues par le modèle PAM, et l'interprétation des difficultés posées a nécessité de s'écarter parfois des modèles classiques. En effet, apprendre l'opposition de voisement en anglais ne consiste pas simplement pour le francophone 1) à apprendre des paires de phonèmes quasi-équivalentes à celles de la L1 (ce qui est très facile), 2) ni à apprendre une différence entre un exemplaire typique et un exemplaire moins typique d'un phonème français (un peu moins facile), 3) ni à apprendre deux nouveaux phonèmes constituant des exemplaires aussi typiques l'un que l'autre d'une seule catégorie phonémique de la L1 (ce qui est particulièrement difficile). Ce qui est étudié dans cette thèse est encore différent, puisqu'il s'agit de quelque chose de proche du cas 2, mais avec une difficulté en plus : l'un des nouveaux segments (l'occlusive sonore anglaise) a un VOT typique d'un phonème français, mais qui est sourd. Cela crée une forte ambiguïté puisqu'il s'écrit avec une lettre qui ne tient pas compte de cette inversion. Seule la prise en compte du trait inexistant en L1 – le long VOT positif des sourdes anglaises – peut constituer un nouveau repère fiable pour se représenter la nouvelle opposition phonémique anglaise. Cette interprétation est cohérente avec le *Ontogeny Phylogeny Model* (Major, 2001) pour qui les phénomènes linguistiques qui se ressemblent sont appris plus lentement que les autres. L'entraînement par stimulation dichotique s'appuie en grande partie sur le repère que peut constituer la longue aspiration des sourdes anglaises, en incitant à bien traiter ce trait avec les régions cérébrales les plus adaptées à l'analyse d'indices longs, et à traiter les sonores anglaises pour ce qu'elles sont en termes temporels : des consonnes dotées d'un VOT court.

## **Chapitre 6. Aide à l'apprentissage de la perception des durées des voyelles anglaises**

---

La stimulation dichotique semble vectrice d'une aide potentielle pour l'apprentissage du voisement en anglais chez des adultes francophones. En effet, l'opposition de voisement des consonnes occlusives, basée principalement sur un indice acoustique de durée (VOT), a permis d'exploiter la complémentarité hémisphérique pour le traitement de VOTs courts et longs en stimulant de manière différenciée, et optimale pour les durées, les oreilles et les hémisphères cérébraux. Il existe en anglais un autre type de phonèmes s'opposant par la durée : les voyelles. Dans cette langue, il existe des voyelles courtes et longues qui ne se distinguent parfois que par ce trait phonologique. Ce trait n'a pas de valeur phonologique pour les voyelles dans les langues non quantitatives comme le français. En anglais, la différence entre voyelles courtes et longues, repose principalement sur deux indices : la durée, mais aussi les formants. Cette nouvelle étude (Expérience 10, 11 et 12) a permis d'évaluer l'effet d'un entraînement perceptif par stimulation dichotique pour l'apprentissage d'un couple de voyelles courtes et longues anglaises chez des adultes normo-lecteurs et des adultes dyslexiques francophones, avec des mesures comportementales.

### **6.1. La différence entre voyelles courtes et longues en anglais**

Produites par la vibration des cordes vocales, les voyelles sont modifiées par la forme de la cavité buccale et de la langue. Leur description articulatoire porte sur la position de la langue et de la mâchoire (Crothers, 1978 ; Maddieson, 1984) qui déterminent les propriétés spectrales (hauteur de formants) sur le plan acoustique. Contrairement aux consonnes, leur perception n'est pas catégorielle mais plus graduelle.

En anglais, certaines voyelles s'opposent phonologiquement par leur durée tandis qu'en français elles ne sont pas phonologiquement discriminables sur cette base. De nombreuses paires minimales anglaises existent à partir de cette différence (e.g. fill [fɪl] et feel [fi:l]). Cette différence de durée s'accompagne d'autres différences acoustiques, quant à elles spectrales et subtiles et comme des différences de hauteur des formants. Ainsi, les formants F1 et F2 du /ɪ/ anglais sont plus proches de ceux du /e/ français que du /i/ français. Des expériences faisant varier indépendamment la durée et la hauteur du deuxième formant (F2) des voyelles /ɪ/ et /i:/ ont montré que les anglophones accordent plus de poids au critère spectral statique (hauteur du F2) qu'à la durée pour identifier ces voyelles. L'inverse est

observé en cas de déficience auditive (Van Summers & Leek, 1992). Dans les Expériences 11 et 12 présentées dans la thèse, il s'agit en un premier temps de mettre au point une technique d'aide à l'apprentissage pour des francophones, exploitant la différence entre les indices de durée opposant les voyelles.

En production, pour apprendre une L2 dotée de différences de durées des voyelles (i.e. L2 quantitative), l'apprenant dont la L1 est également quantitative peut apprendre à produire les différences spectrales plus fines qui participent à leur opposition. En revanche, si sa L1 est non quantitative (e.g. le français), il tendrait à se limiter à faire varier la durée (Bogacka, 2004 ; Flege, *et al.*, 1997 ; Kondaurova & Francis, 2008). Toujours en production, même si l'apprenant tardif devient bilingue, une influence forte de la L1 se produit vers les voyelles de la L2 qui resteront « colorées » par les propriétés acoustiques de la L1 (Baker & Trofimovich, 2005). Capliez (2011) a pu montrer chez des francophones apprenant l'anglais l'utilisation systématique du son français [i] comme substitut des deux voyelles /ɪ/ (court) et /i:/ (long) sur lesquelles repose par exemple la distinction de la paire minimale *live* - *leave*. Bien qu'ils s'appuient plus sur la durée que sur les formants, les apprenants francophones produisent, à cause de cette assimilation, un contraste de durée en L2 trop peu marqué, particulièrement pour les voyelles /ɪ/ et /i:/ (Krzonowski *et al.*, 2018).

En perception, les voyelles d'une L2 quantitative comme l'anglais sont généralement identifiées comme appartenant à la même catégorie : un phonème de la L1 non quantitative (Baker *et al.*, 2008). Des recherches ont montré que cette assimilation perceptuelle des voyelles de la L2 aux catégories phonémiques de la L1 persiste chez les adultes, même des années après l'apprentissage et même s'ils pratiquent beaucoup la L2 (Grimaldi *et al.*, 2014 ; Levy, 2009). D'autres travaux comme ceux de Chladkova *et al.* (2013) vont dans le sens d'une telle assimilation excessive. Evans et Iverson (2007) relèvent que plusieurs voyelles de l'anglais sont assimilées en perception à la même catégorie de voyelle française, en particulier /ɪ/ et /i:/ assimilées à la voyelle /i/ française. Les apprenants d'une L2 quantitative sont davantage sensibles à la différence de durée entre deux voyelles de cette L2 si ces voyelles n'ont pas leur équivalent dans leur L1, c'est-à-dire si l'assimilation directe à leur phonologie native n'est pas possible ; l'apprenant francophone de l'anglais L2 n'est pas dans ce cas. D'une manière plus générale, chez des natifs d'une langue non-quantitative, la sensibilité pré-attentionnelle à la différence de durée des voyelles de la L2 est plus faible que celle des natifs d'une langue quantitative. En effet, à l'écoute d'une voyelle longue de L2 dans une série répétant la voyelle courte qui s'y oppose, l'onde négative associée à la perception même non-

consciente d'un déviant (i.e. l'onde MMN), observée chez les locuteurs natifs, est anormalement faible chez les apprenants (Nenonen, *et al.*, 2005). Une partie des difficultés de compréhension orale des francophones en anglais relève d'une prise en compte inadaptée des indices phonétiques de durée, non seulement pour le voisement des consonnes étudié dans les précédents chapitres, mais aussi pour les voyelles.

Cette difficulté persistante d'apprentissage en L2 entre dans le cadre posé par Flege *et al.* (1997) et dans le modèle SLM (Flege, 1995 ; SLM-r, Flege & Bohn, 2021) qui s'applique aussi bien à l'apprentissage de nouvelles consonnes que de nouvelles voyelles. Pour rappel, le modèle SLM-r souhaite rendre compte de la façon dont les systèmes phonétiques se réorganisent au cours de la vie en réponse à l'exposition à une L2. La particularité de ce modèle, notamment par rapport à un autre comme PAM (Best, 1995 ; Best *et al.*, 2001), est d'envisager une telle réorganisation comme possible, même en cas d'apprentissage tardif d'une L2. D'après le modèle SLM-r, la création de nouvelles catégories phonémiques et l'assimilation de phonèmes de la L2 à des catégories de la L1 (ici pour les voyelles) dépendrait de la précision du système de représentations phonétique / phonologique de la L1, de la qualité de l'input de la L2 et de la ressemblance du son de la L2 avec des sons de la L1. Evans et Iverson (2007) ont constaté, par exemple, que la précision des locuteurs non-natifs en identification de voyelles anglaises produites naturellement variait en fonction du degré de ressemblance entre les voyelles anglaises entendues et celles existant dans le système phonologique des locuteurs natifs.

Flege et Bohn (2021) avancent aussi que les mécanismes de formation des catégories de la L1 restent disponibles pour la L2. Escudero *et al.* (2014) ont proposé que, pour percevoir les voyelles de la L2 avec plus de précision, les apprenants peuvent dans certains cas "déplacer la frontière" entre les catégories de la L1 pour construire celles de la L2, ce qui est un simple aménagement. Les adultes qui identifient de manière très précise les voyelles de leur L1, identifieraient aussi plus précisément les voyelles d'une L2 (Lengeris & Hazan, 2010), car les mécanismes leur ayant permis d'élaborer efficacement leur système phonologique initial resteraient mobilisables et performants. La disponibilité des mécanismes de formation de catégorie de la L1 permet de soutenir l'hypothèse qu'un apprentissage tardif de bonne qualité est possible. Dans le cadre de l'étude menée ici pour l'apprentissage des voyelles /ɪ/ et /i:/ anglaises par des francophones, d'après le modèle SLM-r, il est possible d'anticiper deux comportements possibles. Le premier serait que les francophones identifient les deux voyelles anglaises (/ɪ/ et /i:/) comme faisant partie de la catégorie du /i/ de leur L1.

Cela est possible par l'assimilation du phonème perçu à la catégorie la plus proche de leur L1. Cela se traduirait en perception par l'acceptation de deux exemplaires atypiques du /i/ français, et en production par deux exemplaires identiques (inspirés du /i/ français) ou alors par un /i:/ produit légèrement plus long que le /i/ (peut-être par analogie avec les graphèmes *ee* ou *ea* associés au /i:/ en anglais, si l'apprentissage de cette L2 comporte des éléments orthographiques) sans différence formantique entre les deux. Une deuxième possibilité serait que le /i/ court soit assimilé à la catégorie du /e/ français, car cela concorderait au niveau temporel et au niveau formantique (le /i/ anglais est phonétiquement proche du /e/ français, même si les francophones ont du mal à en prendre conscience par eux-mêmes). Le /i:/ quant à lui pourrait être perçu comme un exemplaire atypique (un peu trop long) du /i/ français, car il est similaire à cette voyelle au niveau formantique. Cette dernière hypothèse trouve un appui dans le modèle ASP (Strange, 2011) selon lequel des routines de sélection perceptives (SPRs) ont été mises en place chez les adultes. Cela impliquerait qu'ils privilégient certains indices plus pertinents dans leur L1 pour catégoriser les nouveaux phonèmes de la L2. En français, les différences de longueur n'étant pas pertinentes pour les voyelles, les caractéristiques formantiques seraient l'objet de ces routines pour les voyelles (d'où le rapprochement /i/-/e/). Les SPRs des francophones apprenant l'anglais (surtout si c'est beaucoup à travers l'écoute) pourraient ainsi les conduire à catégoriser le /i/ anglais comme un /e/ français, et le /i:/ anglais comme un /i:/, et ainsi dans des catégories différentes.

## **6.2. Entraînements disponibles sur la durée des voyelles**

Il existe de nombreuses études sur des entraînements (perceptifs ou incluant la production de mots) pour l'aide aux apprenants de voyelles en L2. La plupart des recherches présentées ci-dessous portent sur les différences de longueur de voyelles en anglais, sujet au cœur de ce chapitre de la thèse.

En 2010, Kondaurova et Francis ont utilisé des programmes de 4 jours (30 min/jour) auprès de 61 jeunes adultes espagnols (M = 23.1 ans) pour les aider à percevoir la différence entre /i/ et /i:/ en anglais. Les anglophones natifs s'appuient pour cela surtout sur des indices spectraux, mais les hispanophones apprenant l'anglais L2 prennent plus en compte la durée. L'étude comparait l'efficacité de programmes exerçant à inhiber la prise en compte de la durée, ou à augmenter l'attention portée aux caractéristiques spectrales, ou les deux. Tous les programmes ont réduit l'attention portée à la durée, mais l'acquisition du contraste /i/-/i:/ de la L2 était plus précise après un entraînement ciblé (sur la focalisation ou sur l'inhibition)

qu'après l'entraînement combinant les deux mécanismes. Un entraînement centré sur un mécanisme cognitif précis serait particulièrement efficace pour l'acquisition phonétique en L2 chez l'adulte.

Pour ce qui est de la persistance des acquis, Hu *et al.* (2016) ont réalisé un entraînement perceptif chez 34 adultes chinois apprenant l'anglais (M = 22.9 ans). Le programme ciblait des voyelles anglaises rendues artificiellement équivalentes pour leur durée, et il durait 6 jours (1h/jour). Les chercheurs ont comparé les performances des apprenants entraînés à celles d'un groupe contrôle qui regardait des vidéos en anglais. Les deux groupes se sont améliorés en identification et en discrimination des formants des voyelles anglaises, avec toutefois un plus grand progrès pour le groupe ayant reçu l'entraînement perceptif ciblé. Les améliorations étaient encore mesurables 3 mois plus tard. L'entraînement perceptif a aussi permis aux participants d'apprendre à moins se focaliser sur la durée pour percevoir la différence entre des voyelles comme /ɪ/ et /i:/. Après l'entraînement qui orientait l'attention sur les différences de formant en ayant supprimé les différences de durée, les apprenants accordaient en effet plus de poids aux indices spectraux lorsque les indices temporels et spectraux étaient présents (et redondants) dans des stimuli naturels, ce qui n'était pas le cas dans le groupe contrôle.

Cinq sessions de 60 min sur 2-3 semaines utilisant un entraînement perceptif phonologique ont amélioré la discrimination, mais aussi cette fois la production de voyelles anglaises par 28 locuteurs de l'azéri (Azerbaïdjan) adultes (M = 25.3 ans) (Mokari & Werner, 2017). Cet entraînement, plus long, confirme qu'il est possible d'améliorer la perception et la production de voyelles avec un entraînement perceptif.

Certaines recherches se sont aussi intéressées à l'utilisation de la haute variabilité phonétique (HVPT) comme technique d'entraînement pour l'aide à l'apprentissage de voyelles. Lengeris et Hazan (2010) ont utilisé ce procédé pour l'apprentissage de voyelles anglaises auprès de 18 apprenants grecs de 18 à 35 ans. Les participants ont réalisé 5 sessions de 45 min réparties sur 2 semaines. Les auteurs ont montré une amélioration de la perception et de la production des voyelles entraînées. D'autres données ont montré qu'un programme utilisant la HVPT amenait un petit groupe de 8 apprenants grecs à se focaliser davantage sur les indices spectraux que sur les indices temporels pour identifier les voyelles anglaises (Kangatharan *et al.*, 2021), ce qui constitue un changement qualitatif sur le plan phonétique. L'utilisation de la HVPT a aussi amélioré la production de mots contenant la voyelle cible en lecture de phrases et en production de parole spontanée (Lengeris, 2018). L'avantage de la haute variabilité dans les exemplaires d'apprentissage s'observe aussi à travers la profondeur

de l'apprentissage, par exemple avec un contraste de voyelles françaises (L2) mieux généralisé à la voix d'autres locuteurs en perception, et une prononciation plus stable chez des apprenants espagnols (Kartushina & Martin, 2019).

L'entraînement ne se restreint pas toujours à des exercices perceptifs. Il passe parfois par la répétition de modèles. Saloranta *et al.* (2015) ont par exemple entraîné neuf Finlandais âgés de 21 à 30 ans pendant deux jours (2 sessions par jour) à l'imitation de modèles de voyelles. Le succès du programme est apparu à travers des progrès mesurables dès le premier jour. Les auteurs ont ainsi montré qu'un entraînement explicite en production pouvait améliorer la production du contraste /y/-/u/. La voyelle /u/ n'existe pas en finois mais peut être rapprochée du /u/ ou du /y/. Cette étude est intéressante dans notre recherche, bien que l'entraînement soit en production et le nôtre en perception, car elle montre que de jeunes adultes sont capables de mieux produire un phonème n'existant pas dans leur L1 en l'assimilant à une autre catégorie de leur L1, comme cela pourrait être le cas pour le /ɪ/ de la paire /ɪ/-/i:/ avec le /e/ français.

Plusieurs études ont exploité l'aide apportée par un retour visuel articulatoire. Par exemple, Kartushina *et al.* (2015) ont montré qu'un entraînement avec retour visuel articulatoire immédiat améliorait la production de voyelles danoises chez des francophones ne connaissant pas du tout le danois, après 1h d'entraînement par voyelle. Avec une seule séance d'exercices ajoutant un feedback visuel basé sur les données articulatoires de l'apprenant, Suemitsu et Dang (2015) ont aussi amélioré la prononciation de la voyelle anglaise /æ/ par des Japonais adultes.

Dans le cas de francophones apprenant l'anglais, Krzonowski *et al.* (2016) ont montré qu'un entraînement sur les voyelles anglaises était plus efficace pour des apprenants francophones si les exercices étaient en perception (identification et discrimination avec feedback) plutôt qu'en production (imitation avec feedback visuel informant sur l'adéquation entre la durée et les indices spectraux de la voyelle produite d'une part, et ceux de la moyenne des locuteurs natifs d'autre part). Seul l'entraînement basé sur la perception améliorait vraiment la perception et la production (effet de transfert) des voyelles anglaises. Carlet et Cerbian (2019) ont apporté une information complémentaire en montrant qu'un entraînement à l'identification de voyelles offrait de meilleurs résultats qu'un entraînement à la discrimination.



Peu de recherches se sont intéressées aux mesures en potentiels évoquées pour l'évaluation d'entraînement perceptifs de nouveaux contrastes de voyelles en L2. Reinke *et al.* (2003) ont évalué l'effet de 4 sessions d'entraînement sur 4 jours (35 min/session) à la discrimination de cinq voyelles anglaises américaines (/i/, /a/, /æ/, /u/ et /ɜ/) chez 16 apprenants anglophones canadiens. Les auteurs ont montré que l'amélioration en discrimination des voyelles présentées par couple s'accompagnait d'une diminution de la latence du complexe N1-P2 sur les électrodes fronto-centrales et temporales et d'une augmentation de l'amplitude de la P2 dans le groupe entraîné comparé à un groupe non-entraîné. Ce résultat renforce l'idée selon laquelle les apprentissages perceptifs peuvent être suivis de changements au niveau du fonctionnement neuronal. Ylinen *et al.* (2010) ont montré qu'un entraînement perceptif à l'identification de mots anglais permettait une augmentation de la MMN à partir des électrodes correspondant à l'HG en position centrale, davantage si le déviant était /i:/ parmi de /ɪ/ que dans le cas inverse chez 12 apprenants finois. La moitié des mots utilisés dans l'entraînement avaient été modifiés pour inciter les apprenants à utiliser les indices spectraux des voyelles plutôt que les indices temporels. Ces résultats montrent qu'une asymétrie peut apparaître dans les progrès au sein d'une paire de voyelles en L2 pour le traitement pré-attentionnel de celles-ci. Krzonowski (2020), proposait deux types d'entraînement à des adultes francophones apprenants l'anglais, portant sur la durée des voyelles : un entraînement HVPT (avec des tâches d'identification de voyelles contenu dans les mots entendu en choix forcé et de discrimination de mots) et un entraînement à la production de mots anglais. Elle a montré qu'un entraînement perceptif améliorait d'autant plus les performances en perception et en production qu'un entraînement à la production. Des mesures électrophysiologiques étaient aussi menées dans le cadre de ce travail de thèse. Cependant, elle n'a pu montrer aucuns effets avec ce type de mesures pour les entraînements proposés. Cela montre qu'il est difficile de faire émerger des effets avec des mesures en potentiels évoqués pour les voyelles, particulièrement dû à la perception graduelle de celles-ci et non catégorielle comme dans le cas de consonnes.

### **6.3. Objectifs**

Tout comme pour les consonnes, des équipes de recherches se sont intéressées aux entraînements perceptifs brefs pour aider à l'apprentissage de la différence de durée des voyelles dans une L2. Bon nombre de ces entraînements s'appliquent aux voyelles anglaises et particulièrement à la différence /ɪ/-/i:/. Ce couple présente un avantage par rapport aux



autres couples de voyelles opposées par la durée en anglais, car il offre la possibilité de travailler sur un grand nombre de paires minimales qui peuvent l'illustrer.

Les voyelles anglaises sont un nouvel exemple de différences temporelles entre des phonèmes. Selon le modèle AST (Poeppel 2003) ainsi que les différents arguments sur les spécialisations des hémisphères pour les traitements temporels courts (Fiez *et al.*, 1995 ; Spironelli & Angrilli, 2010) et longs (Boemio *et al.* 2005 ; Poeppel *et al.*, 2008), les événements acoustiques court (e.g. la voyelle /ɪ/) sont mieux traités dans le cortex auditif gauche et les longs (e.g. la voyelle /i:/) dans le droit. Les Expériences 10, 11 et 12 présentées dans ce chapitre ont été élaborées en se basant sur la même logique que celles ayant inspiré les expériences sur le VOT des consonnes dans le Chapitre 5. En effet, la durée moyenne du /ɪ/ court anglais est de 30-75 ms et celle du /i:/ long est de 125-250 ms. Ces durées sont conformes aux fenêtres temporelles de durées différentes décrites dans le modèle AST (Poeppel, 2003). C'est pourquoi, ici aussi, il semblait pertinent de proposer un dispositif d'aide conduisant les apprenants à investir prioritairement l'HG dans le traitement de la voyelle anglaise courte /ɪ/ et l'HD dans celui de la voyelle anglaise longue /i:/. Cela suppose pour un francophone d'apprendre à traiter de manière nouvelle (et distincte) des sons proches d'un phonème français, et de se départir de leurs habitudes en L1. Le plus souhaitable serait qu'un tel apprenant réorganise ses connaissances phonologiques pour créer deux nouveaux phonèmes anglais L2, plutôt que de se limiter à assimiler les deux à une seule catégorie phonémique du français. L'idée sous-tendant cette étude est d'aider à traiter correctement, et de la manière la plus correcte en termes d'investissement hémisphérique, la paire de voyelle /ɪ-/i:/ afin de percevoir plus précisément leurs particularités (que ce soit leur durée ou leurs formants) et de produire un contraste de meilleure qualité.

C'est pourquoi, l'écoute dichotique qui avait déjà montré quelques avantages pour l'apprentissage de l'opposition de voisement pour des consonnes occlusives anglaises (Chapitre 5) semblait une piste à explorer pour les voyelles courtes et longues. Ce procédé devait favoriser l'association plus directe du /ɪ/ à l'HG et du /i:/ à l'HD, par la mise en concurrence avec un bruit blanc dans l'oreille controlatérale. Ce dispositif dichotique était destiné à permettre, une fois de plus, de proposer un entraînement bref (3 jours, 20 min/jour), sans passer par un enseignement explicite sur la nature des différences entre les phonèmes ciblés.

Cette étude avait **trois objectifs** principaux.

Le premier était de valider des outils de mesures comportementales pouvant par la suite évaluer l'effet de divers entraînements par des tests relevant des performances en perception et en production de paires minimales s'opposant par des voyelles courtes vs. Longues.

Le deuxième était d'évaluer l'apport de l'écoute dichotique par rapport à la stimulation binaurale dans le cadre d'un entraînement perceptif portant sur l'opposition de durée des voyelles, à l'aide de mesures comportementales.

Le troisième objectif était de mesurer au niveau comportemental l'apport du dispositif par stimulation dichotique auprès d'adultes dyslexiques.

Afin de répondre à ces objectifs, trois expériences ont été conduites chez des adultes francophones apprenant l'anglais. L'**Expérience 10** évaluait certaines qualités psychométriques des tests de perception et de production construits pour l'évaluation de l'entraînement. L'effet test-retest ainsi que la fidélité ont été mesurés. Un début de mise en place de normes a été engagé en vue d'un futur usage des épreuves pour l'évaluation dans le cadre académique ou dans la recherche sur les apprentissages. L'**Expérience 11** comparait l'effet de deux entraînements brefs de 3 jours (20 min/jour) centrés sur l'opposition /ɪ/ - /i:/, l'un utilisant la stimulation dichotique et l'autre la stimulation binaurale, à l'aide des outils évalués dans l'Expérience 10 et utilisés cette fois aux étapes pré- et post-entraînement. Cette fois, le contenu de l'entraînement tenait compte des recherches montrant l'efficacité de la haute variabilité des stimuli dans les exercices. Pour cela, les stimuli dans l'entraînement avaient été produits par deux hommes et une femme, pour ajouter la variabilité des voix à la diversité des contextes consonantiques dans lesquels les voyelles étaient entendues. Enfin, l'**Expérience 12** évaluait l'effet de l'entraînement par stimulation dichotique sur les performances en perception et en production des voyelles anglaises /ɪ/ - /i:/ chez des adultes dyslexiques. Seul le programme avec écoute dichotique a été testé auprès de ce public, car il semblait prioritaire d'évaluer l'effet du dispositif potentiellement le plus efficace dans le cas d'une aide à apporter à une population atteinte d'une pathologie. L'objectif de l'Expérience 12 était restreint à l'évaluation des effets de l'entraînement perceptif dichotique, sans avoir pour l'instant l'ambition de mesurer si la stimulation dichotique constituait un apport dans le cas d'apprenants dyslexiques.

## **6.4. Expérience 10 : Evaluation de qualités psychométriques des épreuves de perception et de production de mots en anglais**

### **6.4.1. Principe de l'expérience et hypothèses**

L'Expérience 10 visait à évaluer si les épreuves de perception et de production des voyelles /ɪ/ et /i:/ en anglais élaborés pour l'étude étaient de bons outils pour tester l'effet d'entraînements en anglais pour des apprenants adultes francophones. L'**hypothèse** était qu'il y aurait un faible effet de répétition (faible effet test-retest), ainsi qu'une bonne fidélité pour les quatre épreuves de perception et de production proposées. Nous avons également fourni des normes pour ce test en termes de moyennes des performances, de médiane, d'erreur-type et d'intervalle de confiance, à partir d'un groupe de 60 adultes. La taille de l'échantillon est encore faible, mais il pourrait être étoffé par la suite.

### **6.4.2. Méthode**

#### **6.4.2.1. Participants**

Un total de 60 participants normo-lecteurs, tous étudiants, a réalisé ces épreuves de perception et de production en anglais (Annexe 3). Vingt participants ont réalisé les tests deux fois à 3 jours d'intervalle pour estimer les effets test-retest et la fidélité (18 femmes, M = 21 ans 9 mois ; ET = 3 ans 6 mois). Par ailleurs, les données de 40 autres participants (36 femmes, M = 21 ans 7 mois, ET = 2 ans 7 mois) ayant réalisé ces mêmes épreuves à T1 avant un autre entraînement en anglais (Expérience 11) se sont ajoutées à celles obtenues à T1 par les 20 participants du groupe ayant permis l'évaluation de la sensibilité à la répétition. Cela a permis d'étoffer la taille du groupe pour débiter la constitution de normes. Tous les participants étaient monolingues français, avec une vue normale ou corrigée et ils ne présentaient aucun déficit auditif connu. Ils déclaraient n'avoir à leur connaissance aucun trouble neuro-développemental. Ils avaient étudié l'anglais, mais leur niveau ne dépassait pas B2 (Cadre Européen Commun de Références pour les Langues, CECRL) selon l'évaluation réalisée à l'université l'année du testing.

#### **6.4.2.2. Matériel et équipement**

Les 68 mots monosyllabiques utilisés dans les épreuves de perception et de production de mots ont été enregistrés dans une salle anéchoïque par trois locuteurs anglophones natifs (1 femme, 2 hommes) à l'aide du logiciel ROCme! (Ferragne *et al.*, 2013). Les membres d'une paire s'opposaient par la longueur de leur voyelle /ɪ/ ou /i:/. Les tests pré- et post-entraînement

comptaient quatre tâches. Les épreuves de perception (tâche de vocabulaire et tâche de perception auditive) contenaient 12 paires minimales travaillées pendant l'entraînement (mots appris) : chaque mot était associé à une image (libre de droit) illustrant son contenu. Les deux tâches de production contenaient 6 paires de mots appris et 6 paires de mots non-appris, et ces 12 paires minimales étaient les mêmes dans les deux épreuves de production. Les épreuves ont été programmées avec OpenSesame.

#### 6.4.2.3. Procédure d'évaluation

Tous les participants réalisaient les épreuves dans le même ordre : la perception (tâche de vocabulaire et tâche de perception auditive) puis la production (avec et sans modèle audio). La tâche de vocabulaire était toujours réalisée avant la tâche de perception auditive et la tâche de lecture sans modèle audio toujours avant celle avec modèle audio.

**En perception** deux tâches de jugement perceptif étaient proposées (Fig. 55).

Dans la **tâche de vocabulaire**, uniquement en modalité visuelle, à chaque item deux mots écrits apparaissaient simultanément à l'écran de part et d'autre d'une image. Le participant devait appuyer sur la touche de réponse (gauche vs. droite) pour désigner le mot représenté par l'image dès qu'il pensait avoir la réponse exacte. Etant donné que l'entraînement proposait des informations sur la forme phonologique et sur l'orthographe des mots, il était possible de mieux réussir cette tâche après l'entraînement sur la base de l'acquisition de connaissances sur l'orthographe précise des mots, sans que les aspects phonologiques comme la durée des voyelles jouent obligatoirement un rôle. C'est pourquoi une deuxième tâche de perception a été proposée (tâche de perception auditive) où la réussite ne pouvait pas s'expliquer par des connaissances orthographiques, puisque les seules indications sur la forme des mots étaient données en faisant écouter les mots.

Dans la **tâche de perception auditive**, deux mots se succédaient dans le casque, et il fallait décider lequel était représenté par l'image affichée en pressant, à la fin du signal sonore, la touche gauche pour le premier mot ou la droite pour le second, dès qu'il pensait avoir la réponse exacte. Ici, il n'y avait pas d'indication sur l'orthographe des mots. La compréhension de la consigne, présentée oralement et à l'écrit avant chaque tâche, était vérifiée par l'expérimentatrice. Il fallait répondre le plus vite possible. L'épreuve totale (les deux tâches de perception) durait 10 minutes.

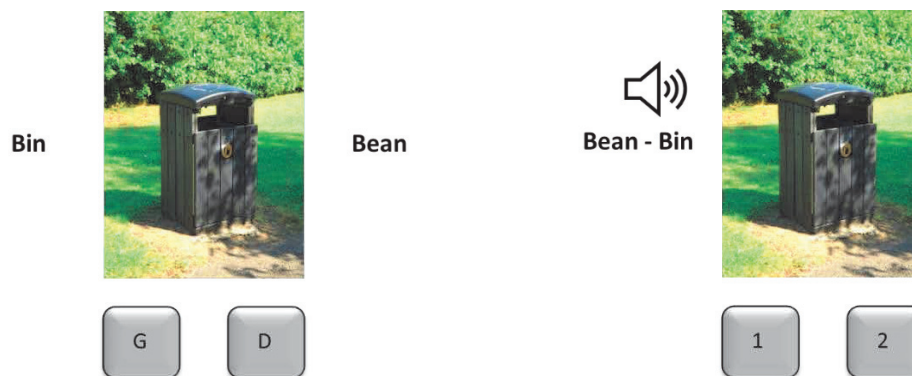


Figure 55. Procédure d'évaluation pour les tâches de vocabulaire (gauche) et de perception auditive (droite).

L'épreuve de production comportait deux tâches (Fig. 56).

Dans la **tâche de lecture sans modèle audio**, uniquement visuelle, un mot écrit apparaissait à l'écran. Le participant enregistrait sa voix lisant ce mot avec le meilleur accent anglais possible.

Dans la **tâche de lecture avec modèle audio**, le participant entendait un mot dans le casque et le voyait en plus simultanément écrit à l'écran. Il avait pour consigne d'imiter au mieux ce qu'il avait entendu en s'enregistrant. La consigne, orale et écrite, était présentée avant chaque tâche. Sa bonne compréhension était vérifiée par l'expérimentatrice. L'épreuve totale (les deux tâches de production) durait 10 minutes.

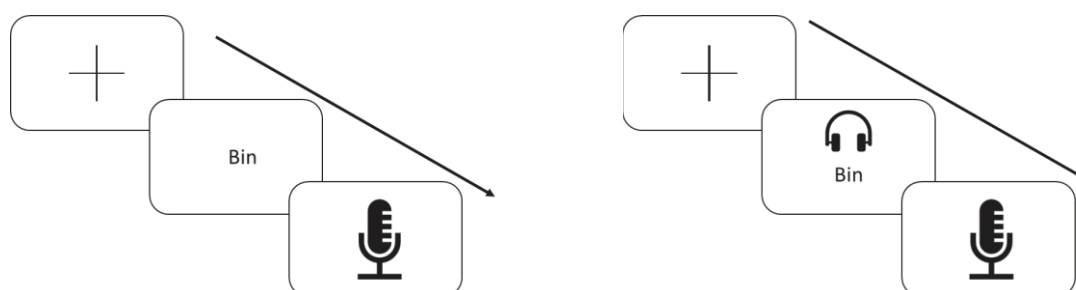


Figure 56. Procédure d'évaluation pour les tâches de lecture sans modèle audio (gauche) et avec modèle audio (droite).

Les participants de ce protocole ont réalisé les épreuves à distance. Ils étaient suivis par téléphone et avait reçu, avant le début de l'expérimentation, une fiche explicitant toutes les étapes à réaliser depuis l'installation du logiciel Opensesame sur leur ordinateur jusqu'à l'envoi des fichiers de réponse par mail ainsi que les consignes de chaque tâche. Au téléphone, les instructions insistaient sur les conditions dans lesquelles les épreuves devaient être réalisées : dans un environnement silencieux, à un moment de la journée où le participant se sentait intellectuellement disponible et concentré. Comme dans toute épreuve réalisée à

distance, il n'est pas exclu que certains participants aient effectué les tâches dans de mauvaises conditions. La réalisation correcte des épreuves a été vérifiée lors de la réception des fichiers de réponses, en utilisant les critères suivants pour exclure les données du participant : s'il avait appuyé sur d'autres touches de réponse que celles désignées pour cela, ou si elles comportaient plus de quatre temps supérieurs à 3 écart-types calculés à partir des données du participant. Il n'y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Aucun participant n'a été exclu en respectant ces critères.

### 6.4.3. Analyse des données

Les réponses dans les deux tâches de l'épreuve de perception ont été transformées en taux de réponses correctes, ensuite moyennées pour chaque participant dans chacune des tâches. Les temps de réponses pour les décisions exactes ont aussi été moyennés.

Pour les deux tâches de production, la durée de la partie stable de la voyelle produite a été mesurée à l'aide du logiciel Praat. Deux juges ont effectué les mesures séparément. En cas de décalage de plus de 10 ms, les mesures étaient refaites et l'avis d'un troisième juge était sollicité.

L'effet test-retest a été évalué à l'aide de tests  $t$  avec permutations entre T1 et T2. Des tests  $t$  de Student bilatéraux appariés avec correction de Bonferroni ont aussi été appliqués, assortis du calcul des tailles d'effet. La fidélité des épreuves a été évaluée par l'étude des corrélations (test de Pearson) avec une estimation de leur taille d'effet (i.e.  $r$ ) afin d'évaluer le lien entre les performances à T1 et T2. Les tailles des effets ( $d$ ,  $r$ ) étaient interprétées selon Cohen (1988) comme faible autour de .20, moyenne si  $>.50$  ou grande si  $>.80$ .

### 6.4.4. Résultats

#### 6.4.4.1. Evaluation de l'effet test-retest

**Epreuves de perception.** Comme l'illustre la figure 57 (à gauche), il n'y avait pas différence entre les taux de réponses exactes à T1 et à T2 dans la tâche de vocabulaire ( $t(19) = 0.24$ ,  $p = .815$ ,  $d = .04$  ;  $Z = 0.24$ ,  $p = .808$ ) ni dans la tâche de perception auditive ( $t(19) = 0.55$ ,  $p = .589$ ,  $d = .11$  ;  $Z = 0.56$ ,  $p = .575$ ). Une accélération des réponses se produisait entre T1 et T2 dans les tâches de vocabulaire ( $t(19) = 2.96$ ,  $p = .008$ ,  $d = .63$  ;  $Z = 2.51$ ,  $p = .012$ ) et de perception auditive ( $t(19) = 2.78$ ,  $p = .012$ ,  $d = .48$  ;  $Z = 2.41$ ,  $p = .016$ ) (Fig. 57, à droite).

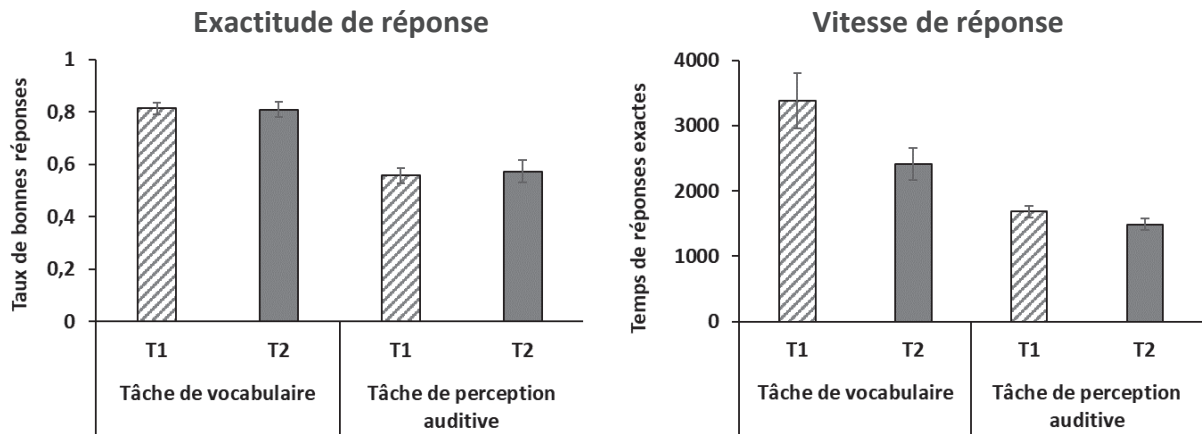


Figure 57. Moyennes des taux de réponses exactes (gauche) et des temps de réponses (droite) entre T1 (rayures) et T2 (plein), pour les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite). Les barres d'erreurs représentent l'erreur-type.

**Epreuves de production.** Dans la tâche de lecture sans modèle audio, seul la durée du /i:/ s'allongeait entre T1 et T2 ( $t(19) = 5.27, p < .001, d = .83 ; Z = 3.45, p < .001$ ). La durée du /ɪ/ ne changeait pas entre T1 et T2 ( $t(19) = 1.43, p = .170, d = .35 ; Z = 1.39, p = .164$ ) (Figure 58).

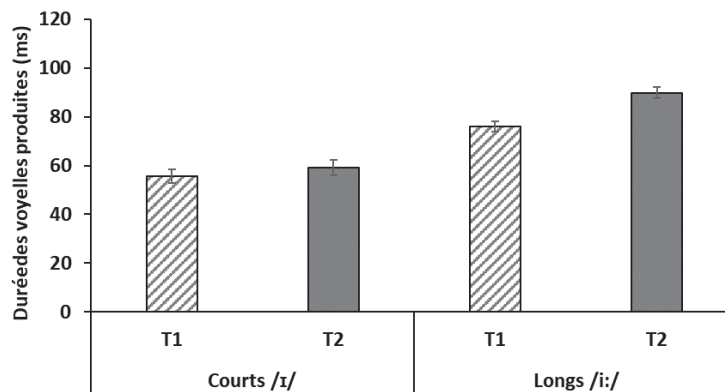


Figure 58. Moyennes des longueurs des voyelles produites entre T1 (rayures) et T2 (plein) pour la tâche de lecture sans modèle audio pour le /ɪ/ (voyelle courte) et le /i:/ (voyelle longue) séparément. Les barres d'erreurs représentent l'erreur-type.

Dans la tâche de lecture avec modèle audio (Fig. 59), il n'y avait aucune modification dans la durée des /ɪ/ entre T1 et T2. La durée des /i:/ produits n'augmentait pas significativement entre T1 et T2 ( $t(19) = 1.73, p = .100, d = .19 ; Z = 1.65, p = .100$ ) passant de 98 ms à 102 ms (Tableau 14).

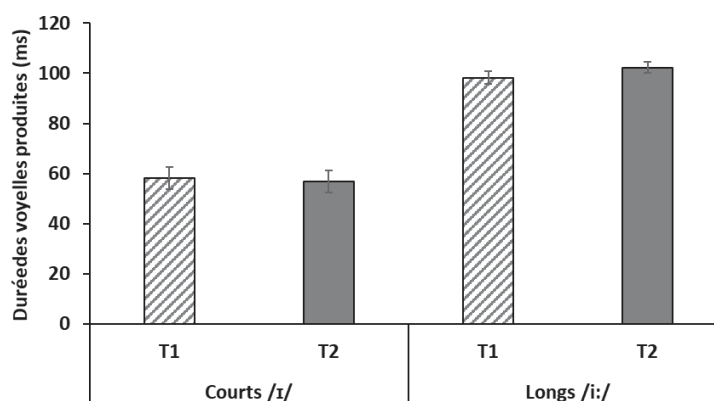


Figure 59. Moyennes des longueurs des voyelles produites entre T1 (rayures) et T2 (plein) pour la tâche de lecture avec modèle audio, pour le /ɪ/ (voyelle courte) et le /i:/ (voyelle longue). Les barres d'erreurs représentent l'erreur-type.

Tableau 14. Résultats aux tests *t* de Student et *t* avec permutations pour les durées des voyelles produites entre T1 et T2, toutes longueurs confondues, et pour la voyelle courte /ɪ/ et la voyelle longue /i:/ séparément, dans la tâche de production avec modèle audio. *ddl* = 19

	Test <i>t</i> de Student		Test <i>t</i> avec permutation	
	<i>t</i>	<i>p</i>	<i>Z</i>	<i>p</i>
Courtes /ɪ/	0.76	.458	0.77	.443
Longue /i:/	1.73	.100	1.65	.100

Par ailleurs, à T1, le fait que la production se fasse avec ou sans modèle audio jouait un rôle pour la durée de la voyelle longue /i:/. Pour le /ɪ/, la durée n'était pas différente entre les deux épreuves ( $t(19) = 1.20, p = .246, d = .24$ ) (lecture sans modèle : 56 ms ; lecture avec modèle : 58 ms), mais pour le /i:/ il y avait une différence significative, car à T1 la présence du modèle audio incitait à produire un /i:/ plus long qu'en absence de modèle ( $t(19) = 5.12, p < .001, d = 1.08$ ) (lecture sans modèle : 76 ms ; lecture avec modèle : 98 ms).

#### 6.4.4.2. Evaluation de la fidélité

**Epreuves de perception.** Dans la tâche de vocabulaire, les taux de réponses exactes étaient corrélés positivement entre T1 et T2 ( $r(18) = .80, p < .001$ ) (Figure 60, haut gauche) et la corrélation positive était aussi significative pour les temps de réponses ( $r(18) = .64, p = .003$ ) (figure 60, haut droite). Dans la tâche de perception auditive, les taux de réponses exactes étaient positivement corrélés entre T1 et T2 ( $r(18) = .68, p < .001$ ) (figure 60, bas gauche) et la corrélation positive était également significative pour les temps de réponses ( $r(18) = .71, p = .001$ ).



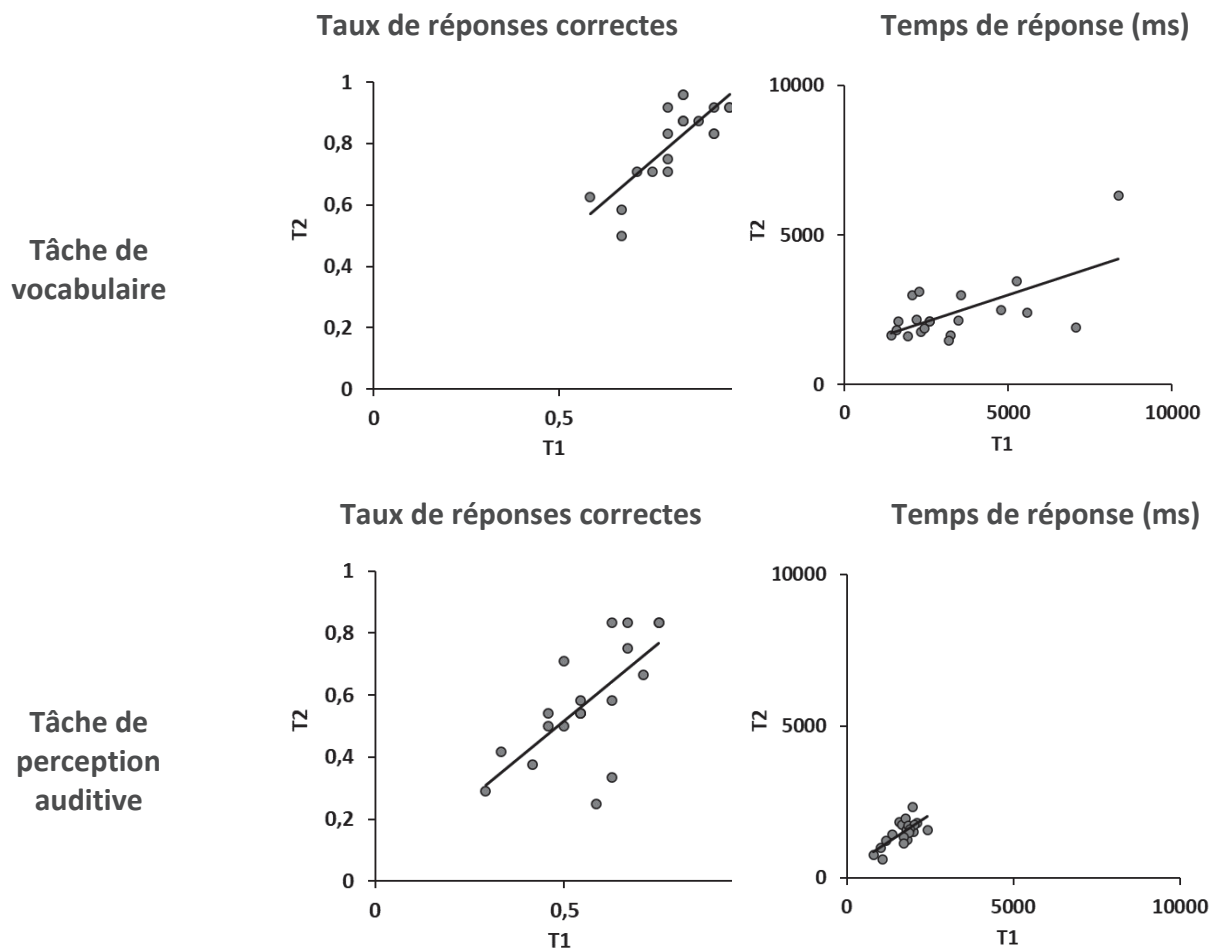


Figure 60. Nuages de points représentant la corrélation entre T1 et T2 pour les taux de réponses (à gauche) et les temps de réponses (à droite), pour la tâche de vocabulaire (haut) et de perception auditive (bas). La barre noire représente la droite de régression.

**Epreuves de production.** Les résultats de l'étude des corrélations entre la longueur des voyelles produites à T1 et à T2 dans les deux épreuves de production, avec et sans modèle, sont détaillés dans le tableau 15.

Tableau 15. Résultats de l'étude des corrélations pour la longueur des voyelles produites entre T1 et T2 pour les tâches de lecture sans modèle audio et avec modèle audio, sur toutes les longueurs confondues, pour la voyelle courte /ɪ/ et la voyelle longue /i:/. ddl = 18

		<i>r</i>	<i>p</i>
<b>Lecture sans modèle audio</b>	Courtes /ɪ/	.39	.089
	Longues /i:/	.76	<.001
<b>Lecture avec modèle audio</b>	Courtes /ɪ/	.73	<.001
	Longues /i:/	.89	<.001

Ces corrélations étaient significatives et correspondaient à des nuages de points bien resserrés autour de la droite de régression (Fig. 61). La seule corrélation non significative était marginale et relevée pour la voyelle courte /ɪ/ (Fig. 61, en haut à gauche) ( $r(18) = .39, p = .089$ ).

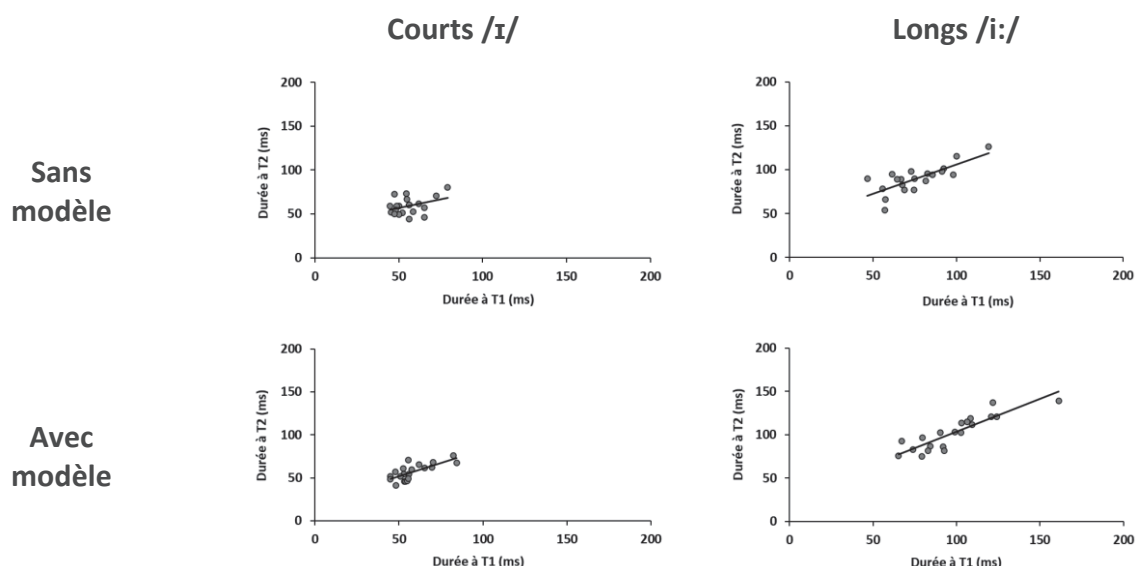


Figure 61. Nuages de points représentant la corrélation entre T1 et T2 dans les tâches de production avec et sans modèle audio, pour les voyelles /ɪ/ (à gauche) et /i:/ (à droite). La barre noire représente la droite de régression.

#### 6.4.4.3. Vers la création de normes

**Epreuves de perception.** Les résultats des calculs pour de premières normes (ici, sur 60 participants) sur les taux de réponses exactes et les temps de réponses aux décisions exactes en tâches de vocabulaire et de perception auditive sont indiqués dans le Tableau 16. Il s’agit de la moyenne, de la médiane, de l’erreur-type, des intervalles de confiance à 95% et à 99% avec indications des bornes formées par les intervalles de confiance autour de la moyenne, pour les taux de réponse exactes (Fig. 62, gauche) et les temps de réponse (Fig. 62, droite).

Tableau 16. Résultats de normes (en construction) des épreuves de perception en anglais sur les taux de réponses exactes et les temps de réponses dans les tâches de vocabulaire et de perception auditive à T1. (M = Moyenne, Méd. = Médiane, Err-T = Erreur-type, IC = Intervalle de Confiance). Les bornes sont indiquées par la moyenne + ou - l’IC.

		M	Méd.	Err-T	IC (95%)	M + IC (95%)	M - IC (95%)	IC (99%)	M + IC (99%)	M - IC (99%)
<b>Vocabulaire</b>	Taux	0.77	0.79	0.02	0.03	0.80	0.74	0.04	0.81	0.73
	Temps	2915	2579	174	341	3256	2574	448	3363	2467
<b>Perception Auditive</b>	Taux	0.56	0.54	0.02	0.04	0.60	0.52	0.05	0.61	0.51
	Temps	1691	1721	59	116	1807	1575	152	1843	1539

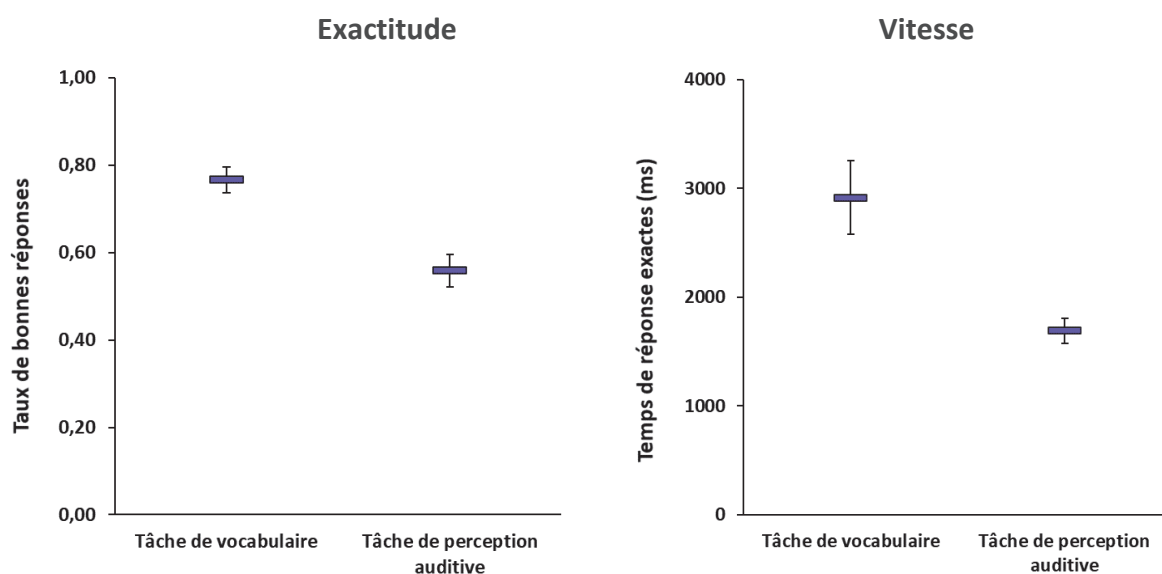


Figure 62. Moyennes des taux de réponses exactes (gauche) et des temps de réponses aux décisions exactes (droite) pour les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite). Les barres d'erreurs représentent l'intervalle de confiance à 95% autour de la moyenne.

**Epreuves de production.** Les résultats des calculs pour de premières normes (ici, sur 60 participants) sur les durées des voyelles /ɪ/ et les /i:/ produites dans les tâches de lecture sans (Fig. 63, gauche) et avec modèle audio (Fig. 63, droite) sont indiqués dans le tableau 17. Il s'agit de la moyenne, de la médiane, de l'erreur-type, des intervalles de confiance à 95% et à 99% avec indications des bornes formées par les intervalles de confiance autour de la moyenne.

Tableau 17. Résultats de normes des épreuves de production en anglais sur la durée des voyelles courtes /ɪ/ et longues /i:/ produites en tâche de lecture sans et avec modèle audio à T1. (M = Moyenne, Méd. = Médiane, Err-T = Erreur-type, IC = Intervalle de Confiance)

		M	Méd.	Err-T	IC (95%)	M + IC (95%)	M - IC (95%)	IC (99%)	M + IC (99%)	M - IC (99%)
Sans modèle	/ɪ/	71	101	5	11	82	60	14	85	57
	/i:/	111	133	5	10	121	101	13	124	98
Avec modèle	/ɪ/	70	71	2	5	75	65	6	76	64
	/i:/	137	133	2	5	141	133	6	142	131

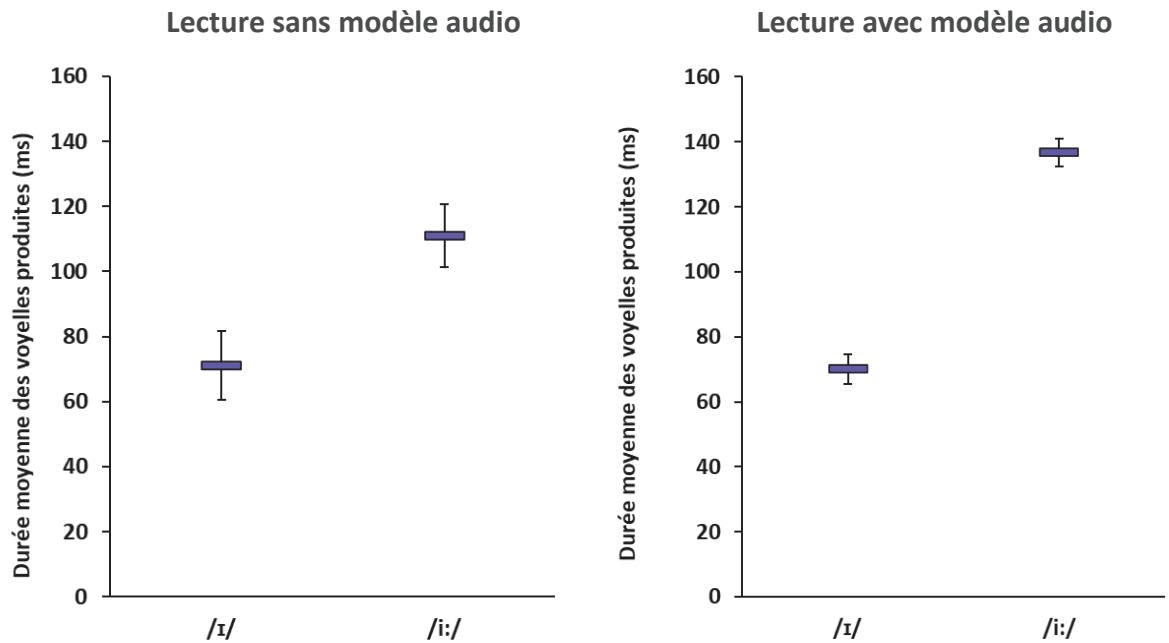


Figure 63. Moyennes des durées des voyelles courtes /ɪ/ (partie gauche) et longues /i:/ (partie droite) produites en tâche de lecture sans modèle audio (gauche) et avec modèle audio (droite). Les barres d'erreurs représentent l'intervalle de confiance à 95% autour de la moyenne.

#### 6.4.5. Discussion intermédiaire

Les résultats de l'Expérience 10 montrent, de manière générale, d'assez bonnes qualités psychométriques pour les nouvelles épreuves, ce qui apporte des éléments favorables à l'hypothèse formulée pour cette expérience. L'hypothèse est validée pour les deux épreuves de perception, qui deviennent ainsi particulièrement intéressantes pour de futures évaluations d'effets de programmes d'entraînement aux étapes pré- et post-entraînement. En effet, dans ces deux épreuves de perception, l'exactitude des réponses est insensible à l'effet de répétition (pas d'effet test-retest) et cette mesure est aussi fidèle.

En production, l'hypothèse est partiellement validée. L'épreuve de répétition avec modèle audio est particulièrement intéressante en tant qu'outil d'évaluation avant et après une prise en charge, car la durée des voyelles /ɪ/ et /i:/ produites est insensible à l'effet test-retest et la fidélité est significative. L'épreuve de répétition sans modèle audio présente des qualités un peu moins bonnes. Les mesures pour la voyelle /ɪ/ sont certes insensibles à l'effet de répétition et la fidélité est proche du seuil de significativité. Par contre, pour /i:/ la fidélité est significative, mais cette voyelle a le désavantage de montrer un effet de répétition avec un allongement à T2. Dans cette tâche, cette voyelle est donc moins intéressante pour l'évaluation d'effets de prises en charge. Il en sera tenu compte dans l'utilisation des quatre épreuves pour tester les effets d'entraînement dans les Expériences 11 et 12.

## **6.5. Expérience 11 : effets des entraînements par stimulation dichotique vs. binaurale sur la perception et la production de voyelles anglaises chez des adultes non-dyslexiques**

### **6.5.1. Principe de l'expérience et hypothèses**

Le but de l'Expérience 11 était de comparer l'effet de deux entraînements en anglais (l'un par stimulation dichotique, l'autre par stimulation binaurale) sur la perception et la production des voyelles anglaises /ɪ/ et /i:/ chez des adultes francophones normo-lecteurs.

Deux hypothèses opérationnelles ont été formulées :

- 1) Après les deux entraînements perceptifs, par rapport à l'étape précédant l'entraînement, les mots entendus lors de l'entraînement devraient susciter de meilleures performances qu'avant l'entraînement, avec davantage d'exactitude en perception (épreuves de vocabulaire et de perception auditive) sans que ce soit au prix d'un ralentissement des réponses, et un transfert vers la production qui devrait faire produire des /ɪ/ plus courts et des /i:/ plus longs en lecture sans et avec modèle audio.
- 2) L'entraînement par stimulation dichotique devrait davantage améliorer les performances que l'entraînement par stimulation binaurale.
- 3) L'entraînement par stimulation dichotique devrait améliorer la perception (vocabulaire et perception) et la production (lecture sans et avec modèle audio) de mots contenant les voyelles /ɪ/ et /i:/ n'ayant pas été entendus pendant l'entraînement, alors que l'entraînement par stimulation binaurale ne devrait pas permettre un tel transfert.

### **6.5.2. Méthode**

#### **6.5.2.1. Participants**

Quarante adultes normo-lecteurs (36 femmes, M = 21 ans 7 mois, ET = 2 ans 7 mois), tous étudiants, ont participé à cette étude. Tous étaient monolingues français, avec une vue normale ou corrigée et ils ne présentaient pas de déficit auditif connu. Ils déclaraient n'avoir à leur connaissance aucun trouble neuro-développemental et avaient étudié l'anglais, mais leur niveau ne dépassait pas B2 (Cadre Européen Commun de Références pour les Langues. CECRL) selon une évaluation dans l'année à l'université. Tous les participants étaient droitiers ( $\geq 80\%$  de réponses « droite » dans la version abrégée de l'*Edinburgh Handedness Inventory* (Oldfield. 1971)).

Les participants ont été répartis aléatoirement dans deux groupes (20 participants chacun) : dans l'un, ils réalisaient un entraînement perceptif utilisant la stimulation dichotique (Groupe Dichotique, GD), dans l'autre ils réalisaient un entraînement perceptif contrôle où les sons étaient diffusés de façon binaurale (Groupe Binaural, GB).

### **6.5.2.2. Matériel et équipement**

La perception et la production ont été testées à l'aide du matériel décrit dans l'Expérience 10 (partie 6.4.2.2 de ce chapitre).

Les exercices de l'entraînement en anglais ont été créés en 2018 (Bouhon, 2018). Ils ont été programmés avec OpenSesame (Mathôt, *et al.*, 2012) pour faciliter leur diffusion auprès des participants (licence gratuite, utilisation possible sur PC et Mac). Chacune des 204 images était libre de droits. Les exercices étaient composés de 28 paires de mots monosyllabiques anglais formant des paires minimales (e.g. *bin/bean*). Ils contenaient tous la voyelle courte /ɪ/ (M = 52 ms, ET = 19 ms, Min = 28 ms, Max = 112 ms) ou longue /i:/ (M = 113 ms, ET = 259 ms, Min = 62 ms, Max = 160 ms). Les durées moyennes sont en accord avec les fenêtres d'analyses temporelles du modèle AST, la fenêtre courte dominant dans le cortex auditif gauche et la fenêtre longue dans le cortex auditif droit (Poeppel, 2003). La comparaison des durées maximales et minimales montre toutefois que quelques exemplaires ne respectaient pas cette règle. Les mots ont été enregistrés dans une salle anéchoïque par trois locuteurs anglophones natifs (1 femme, 2 hommes), à l'aide du logiciel ROCme! (Ferragne *et al.*, 2013). Leur niveau d'intensité a été égalisé. Les entraînements et tests effectués avant et après l'entraînement étaient réalisés à partir de l'ordinateur personnel du participant, mais un ordinateur PC et un casque audio (Beyerdynamic DT 770 Pro) pouvaient lui être prêtés en cas de besoin. Cette expérience a été réalisée entièrement à distance pour le groupe GB, les données du groupe GD ayant été acquises avant la pandémie. Le suivi à distance a été réalisé comme décrit dans l'Expérience 10, dans la partie 6.4.3.2 de ce chapitre.

### **6.5.2.3. Procédure d'évaluation**

Afin d'évaluer l'efficacité de l'entraînement, il était demandé aux participants de réaliser à l'étape pré-entraînement (T1) et à l'étape post-entraînement (T2) les deux tâches de perception de mots et les deux tâches de production de mots décrites dans l'Expérience 10, dans la partie 6.4.2.3 de ce chapitre à 3 jours d'intervalles (Fig. 64). Tous les mots des tests étaient des paires minimales opposées par la durée de la voyelle. Les trois quarts ont été appris durant l'entraînement et l'autre quart non-appris pour évaluer l'effet de l'entraînement

et un potentiel transfert de compétences vers des mots non-appris. Les tests ont été réalisés dans le même ordre que celui utilisé pour l'évaluation de l'effet test-retest et de la fidélité dans l'Expérience 10, dans la partie 6.4.2.3 de ce chapitre (voir Fig. 64).

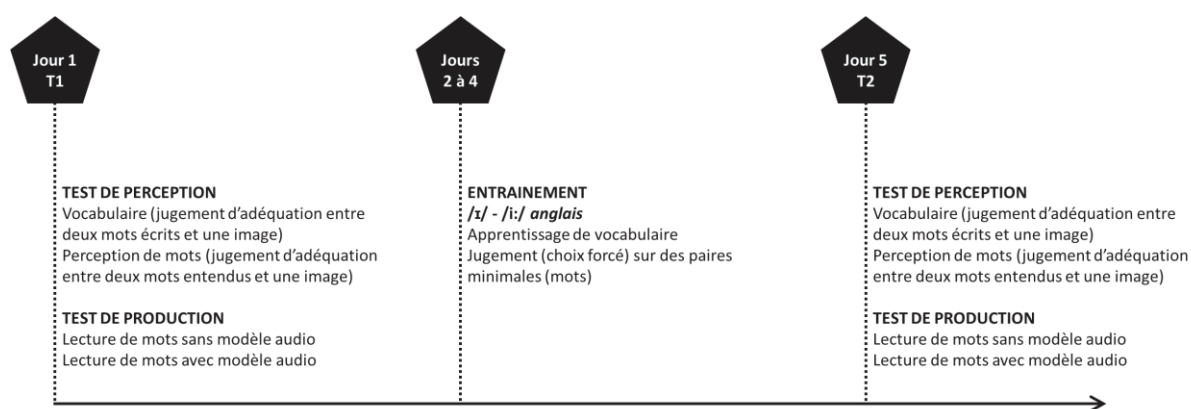


Figure 64. Procédure générale de l'Expérience 6.

#### 6.5.2.4. Procédure d'entraînement

Les entraînements étaient réalisés pendant 3 jours (20 min/jour), en autonomie au domicile. Ils se composaient de deux types d'exercices perceptifs : des exercices d'apprentissage lexical et des exercices de jugement portant sur les mots des exercices d'apprentissage traités auparavant. Les deux types d'exercices sont détaillés dans le paragraphe suivant. Pour réaliser les jugements, l'apprenant était face à un choix basé sur la durée de la voyelle (/ɪ/ ou /i:/) contenue dans les mots formant des paires minimales les jours 1-2. Les jours 1 et 2, l'apprenant réalisait ainsi chaque jour un exercice d'apprentissage et les deux exercices de jugement qui lui correspondaient. Le jour 3, il n'y avait plus d'exercice d'apprentissage, mais seulement deux exercices de jugement qui portaient sur tous les mots appris les deux jours précédents. Les exercices étaient tous présentés au participant comme visant l'apprentissage de vocabulaire. Les participants devaient toujours porter le casque audio avec l'écouteur droit sur l'oreille droite et l'écouteur gauche sur l'oreille gauche. Ils n'étaient pas informés qu'ils recevraient des informations différentes dans les deux oreilles, mais les exercices étaient programmés de sorte qu'un mot contenant une voyelle courte anglaise soit toujours présenté à l'oreille droite-HG et qu'un mot contenant une voyelle longue anglaise le soit à l'oreille gauche-HD.

Dans l'**Exercice d'apprentissage de vocabulaire**, le participant entendait un mot à travers le casque, voyait sa représentation orthographique à l'écran, sa traduction française écrite ainsi qu'une image sémantiquement liée. Chaque mot était présenté deux fois de suite. Le participant avait pour consigne d'être attentif et de tenter d'apprendre ce vocabulaire

anglais. Les mots donnés à apprendre (56 au total) changeaient tous les jours. La partie gauche de la figure 65 illustre cet exercice.

Dans l'**Exercice de jugement 1**, l'apprenant entendait un des mots anglais appris le jour même, puis voyait apparaître simultanément une image sémantiquement liée au mot entendu et deux mots anglais écrits l'un en dessous de l'autre (paire minimale). Il devait choisir entre les deux mots écrits celui qui correspondait au mot entendu et à l'image en appuyant sur l'une des deux touches de réponse (haut vs. bas). Il avait pour consigne d'appuyer dès qu'il pensait avoir la réponse exacte. Un feedback visuel le renseignait sur ses erreurs. Chaque mot avait été appris dans l'exercice d'apprentissage de vocabulaire et faisait l'objet de plusieurs jugements. Il était présenté avec l'image utilisée lors de l'apprentissage ou avec d'autres images, pour stimuler l'intérêt et la motivation. La partie centrale de la figure 65 illustre cet exercice.

Dans l'**Exercice de jugement 2**, le participant entendait un mot anglais puis voyait deux images (i.e. cette fois, il n'y avait plus de support écrit, ce qui augmentait progressivement le niveau de difficulté). Il s'agissait des mêmes mots que pour l'exercice de jugement 1. Cette fois, il n'y avait pas d'information orthographique. Le participant devait choisir la touche du clavier (haut vs. bas) dont la position correspondait à celle de l'image représentant le mot entendu, dès qu'il pensait avoir la réponse exacte. La partie droite de la figure 65 illustre cet exercice.



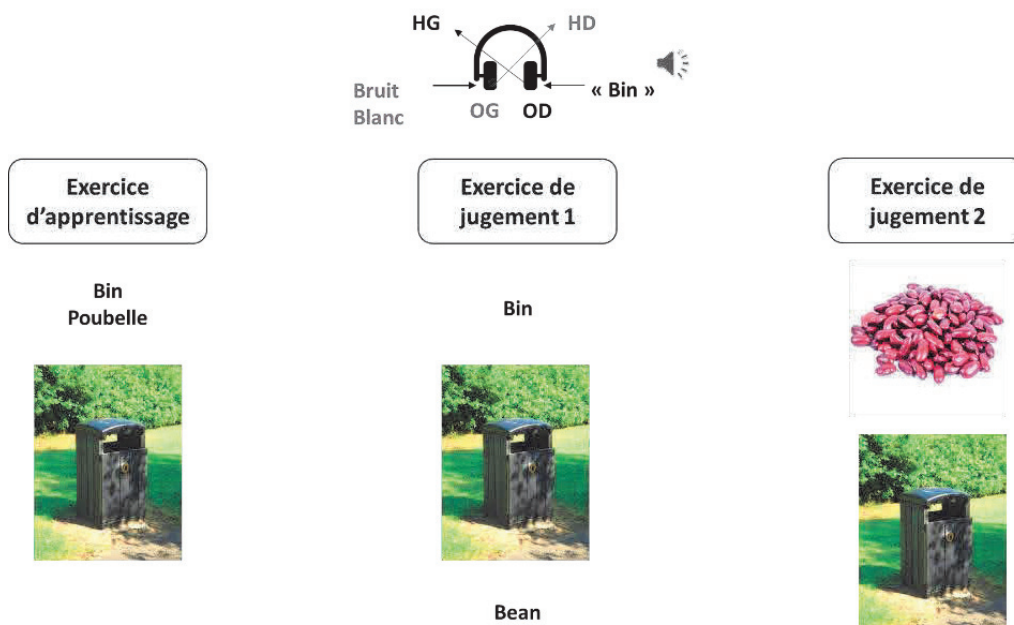


Figure 65. Déroulement de l’entraînement concernant les voyelles anglaises en écoute dichotique, avec les trois types d’exercices proposés. Dans les trois types d’exercices, les stimuli auditifs étaient présentés selon le dispositif illustré en haut de la figure. L’entraînement contrôle binaural était identique excepté la présentation des sons (le mot était alors présenté dans les deux oreilles et il n’y avait pas de bruit blanc).

Pour le **groupe GD**, les exercices informatisés présentaient les mots au moyen de l’écoute dichotique dans un casque audio pour renforcer l’association durée de la voyelle – hémisphère le plus compétent pour traiter cette durée (i.e. les mots contenant une voyelle courte présentés dans l’oreille droite – HG avec bruit blanc simultané dans l’oreille gauche ; les mots contenant une voyelle longue présentés dans l’oreille gauche – HD avec bruit blanc simultané dans l’oreille droite). Pour le **groupe GB**, l’entraînement avait le même contenu que pour le groupe GD, aussi bien pour les mots que pour les images et les exercices (tâches), mais les mots étaient diffusés de manière binaurale, c’est-à-dire simultanément dans les deux oreilles (pas de bruit blanc).

Au téléphone, les instructions insistaient sur les conditions (silence, calme, concentration) dans lesquelles les épreuves devaient être réalisées, comme dans l’Expérience 10. La réalisation correcte des épreuves et des exercices a été vérifiée lors de la réception des fichiers de réponses, en inspectant les dates d’enregistrement des fichiers. Comme dans l’Expérience 10, les fichiers de réponses étaient examinés pour vérifier que le participant n’avait pas appuyé sur d’autres touches que celles désignées pour les réponses, et qu’il n’avait pas produit plus de quatre temps supérieurs à 3 écart-types calculés à partir de ses données. Il n’y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Dans de cas où les réponses étaient globalement très lentes ou particulièrement rapides, les données du

participant n'étaient pas conservées si cette lenteur ou rapidité extrêmement ne concernait que T1 ou que T2, car cela pouvait témoigner d'un manque de sérieux lors de l'une des étapes. Aucun participant n'a été exclu en respectant ces critères.

### 6.5.3. Analyses des données

Les données ont été traitées comme décrit dans l'Expérience 10 (partie 6.4.3 de ce chapitre). Les analyses statistiques ont été réalisées en deux temps. Les groupes ont d'abord été comparés entre eux à T1 à l'aide de tests  $t$  de Student non-appariés bilatéraux avec ajustement de Bonferroni. Puis, l'effet de chaque entraînement a été évalué séparément pour chaque groupe sur les performances entre T1 et T2 à l'aide de tests  $t$  de Student appariés avec correction de Bonferroni.

### 6.5.4. Résultats : comparaison des groupes GD et GB à T1

#### Epreuves de perception.

Avant l'entraînement (T1), il n'y avait pas de différence entre les deux groupes pour le taux de réponses exactes de la **tâche de vocabulaire** ( $t(37.30) = 0.49, p = .625, d = .16$ ), mais le groupe GD répondait significativement plus vite que le groupe GB ( $t(37.14) = 2.10, p = .042, d = .67$ ). Selon l'intervalle de confiance à 95%, les deux groupes étaient dans la norme pour les taux de réponses exactes (GD : 0.74 ; GB : 0.75), mais seul le groupe GB était dans la norme pour les temps de réponse (2945 ms), le groupe GD répondant plus vite que la norme (2381 ms) (Fig. 66, gauche).

Dans la **tâche de perception auditive**, le groupe GD répondait de façon plus exacte ( $t(37.92) = 2.02, p = .050, d = .64$ ) et plus rapide ( $t(37.86) = 2.09, p = .044, d = .66$ ) que le groupe GB. A T1, le groupe GB répondait au hasard (Binomiale = .55,  $p = .50$ ), alors que le groupe GD réussissait davantage que ne le prédisait le hasard (Binomiale < .001,  $p = .50$ ). Le groupe GD avait un taux de réponses exactes légèrement plus élevé que la norme (0.61) et répondait plus vite que la norme (1546 ms), alors que le groupe GB avait un taux de réponses exactes tout juste dans la norme (0.52) et répondait plus lentement que la norme (1851 ms) (Fig. 66, droite).

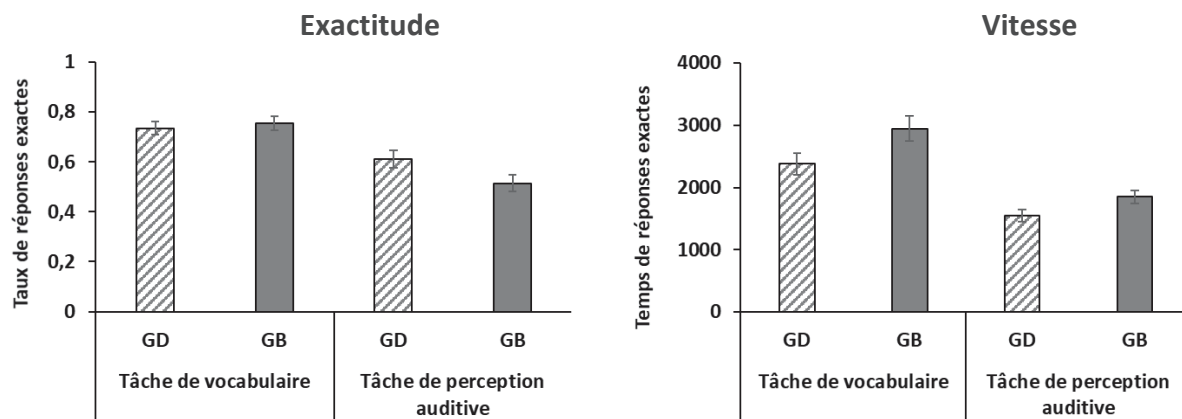


Figure 66. Taux de réponses exactes (gauche) et temps de réponses (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite) à T1 pour les groupes GD (rayures) et GB (plein). Les barres d'erreurs représentent l'erreur-type.

### Epreuves de production.

Dans la **tâche de production sans modèle audio**, le groupe GB produisait des /ɪ/ plus longs que le groupe GD à T1 ( $t(37.37) = 3.83, p = .001, d = 1.21$ ). Cette durée était anormalement longue dans le groupe GB, car au-dessus de la norme définie par l'intervalle de confiance à 95% (87 ms), alors que la durée des /ɪ/ du groupe GD était dans la norme (70 ms). Pour les /i:/, les deux groupes ne différaient pas entre eux, malgré une durée dans la norme pour le groupe GD (105 ms) et plus longue que la norme dans le groupe GB (125 ms) (Fig. 67, gauche).

Dans la **tâche de production avec modèle audio**, le groupe GB produisait des /ɪ/ plus longs que le groupe GD à T1 ( $t(34.33) = 4.40, p < .001, d = 1.39$ ). Le groupe GB était pour cela hors-norme, car la durée de ses /ɪ/ était au-dessus de la norme définie par l'intervalle de confiance à 95% (85 ms), tandis que la durée des /ɪ/ du groupe GD était dans la norme (67 ms). Il n'y avait pas de différence de longueur du /i:/ entre les groupes qui produisaient chacun une durée moyenne un peu au-dessus de la norme (GD : 155 ms ; GB : 157 ms) (Fig. 67, droite).

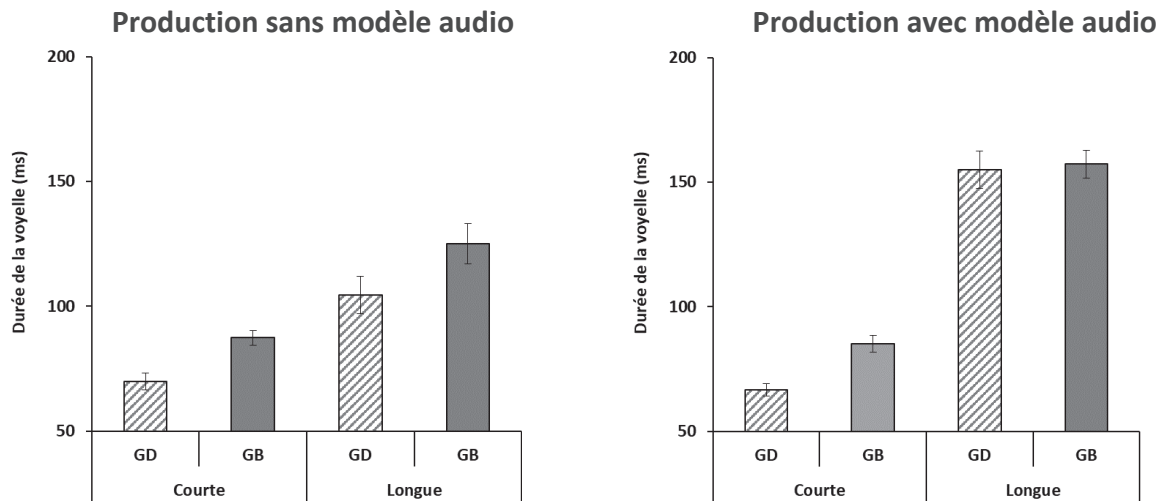


Figure 67. Durée des voyelles produites entre les groupes GD (rayure) et GB (plein) pour la voyelle courte /ɪ/ (partie gauche) et la voyelle longue /i:/ (partie droite) en tâche de production sans modèle (gauche) et avec modèle (droite). Les barres d'erreurs représentent l'erreur-type.

Les groupes n'étant pas comparable directement à cause de différences entre eux dès T1, nous avons réalisé des analyses séparées pour chacun des groupes.

#### 6.5.5. Résultats : effet de l'entraînement par stimulation dichotique

**Epreuves de perception.** Dans la **tâche de vocabulaire** (Fig. 68), les participants du groupe GD répondaient de façon plus exacte ( $t(19) = 6.05, p < .001, d = 1.36$ ) et plus rapide ( $t(19) = 5.19, p < .001, d = 1.38$ ) à T2 qu'à T1. Dans la **tâche de perception auditive**, les réponses du groupe GD devenaient plus exactes à T2 qu'à T1 ( $t(19) = 2.66, p = .016, d = .67$ ), sans altération de leur vitesse de réponse ( $t(19) = 0.42, p = .680, d = .10$ ).

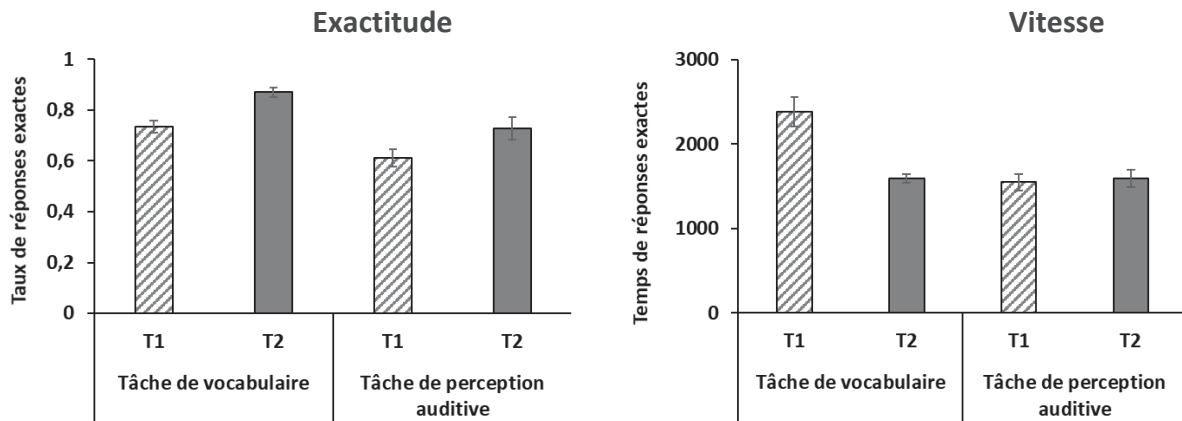


Figure 68. Taux de réponses exactes (gauche) et temps de réponses (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite) entre T1 (rayures) et T2 (plein) pour le groupe GD. Les barres d'erreurs représentent l'erreur-type.

**Epreuves de production.** Dans la tâche de **production sans modèle audio**, après l'entraînement (T2), les participants du groupe GD produisaient des /ɪ/ plus courts ( $t(19) = 2.52, p = .021, d = .41$ ) et des /i:/ plus longs ( $t(19) = 4.90, p < .001, d = .65$ ) qu'à T1 (Fig. 69, gauche). Dans la tâche de **production avec modèle audio** (Fig. 69, droite), il n'y avait pas de changements significatifs entre les deux étapes.

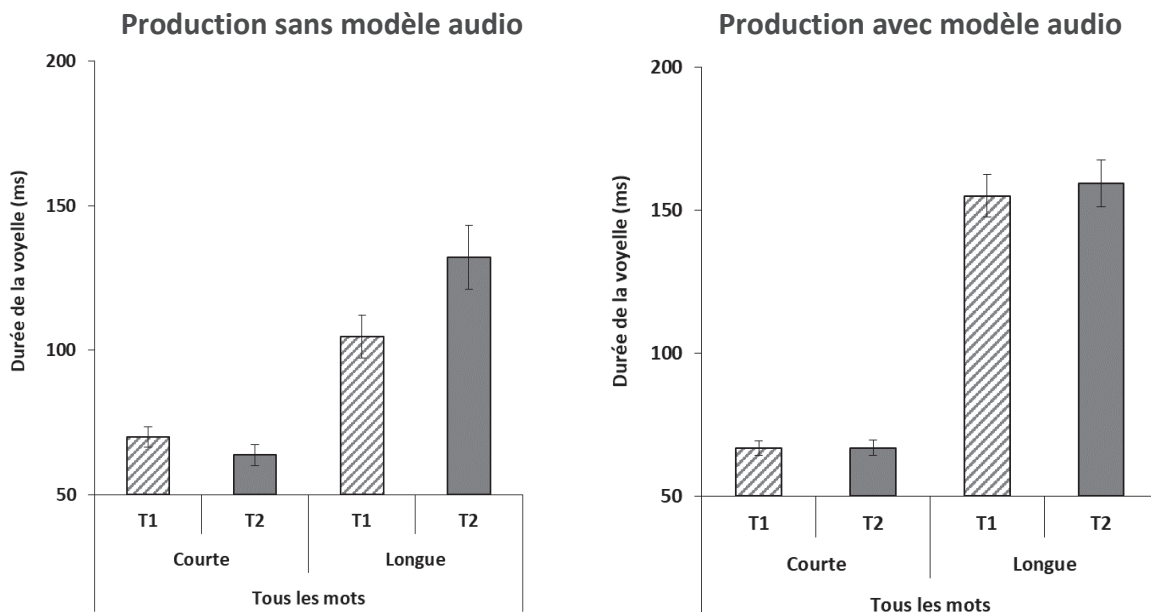


Figure 69. Durée des voyelles produites entre T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (partie gauche) et la voyelle longue /i:/ (partie droite). Les barres d'erreurs représentent l'erreur-type.

Afin de répondre à l'hypothèse sur un possible effet de transfert de compétences, des analyses complémentaires ont été réalisées séparément sur les mots ayant été étudiés pendant l'entraînement et ceux ne l'ayant pas été. Dans la tâche de **production sans modèle audio** (Fig. 70), les participants allongeaient la durée du /i:/ que les mots aient été étudiés pendant

l'entraînement ( $t(19) = 4.39, p < .001, d = .65$ ) ou non ( $t(19) = 3.64, p = .002, d = .59$ ). Cependant, seuls les /ɪ/ contenus dans des mots appris durant l'entraînement étaient significativement raccourcis à T2 ( $t(19) = 2.78, p = .012, d = .44$ ). Il n'y avait pas de changement significatif dans la tâche de **production avec modèle audio**.

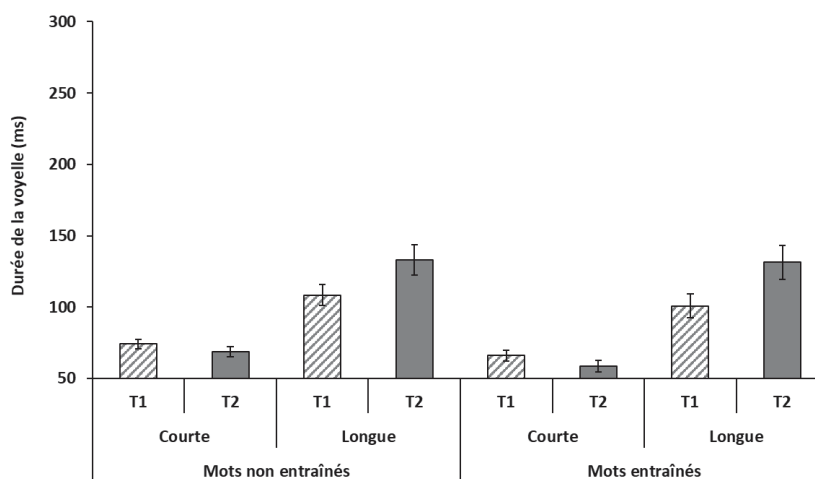


Figure 70. Durée des voyelles produites dans la tâche de production sans modèle audio entre T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (gris) et la voyelle longue /i:/ (violet). Les barres d'erreurs représentent l'erreur-type.

#### 6.5.6. Résultats : effet de l'entraînement par stimulation binaurale

**Epreuves de perception.** Dans la **tâche de vocabulaire**, les participants du groupe GB répondaient de façon plus exacte ( $t(19) = 4.63, p < .001, d = 1.07$ ) et plus rapide ( $t(19) = 5.77, p < .001, d = 1.59$ ) à T2 qu'à T1 (Fig. 71). C'était aussi le cas pour la tâche de **perception auditive** avec une amélioration de l'exactitude ( $t(19) = 5.55, p < .001, d = 1.17$ ) et de la vitesse ( $t(19) = 2.51, p = .021, d = .55$ ) entre T1 et T2. A T2, les réponses devenaient significativement différentes du hasard (Binomiale  $< .001$  ;  $p = .50$ ) alors qu'elles ne l'étaient pas dans ce groupe à T1.

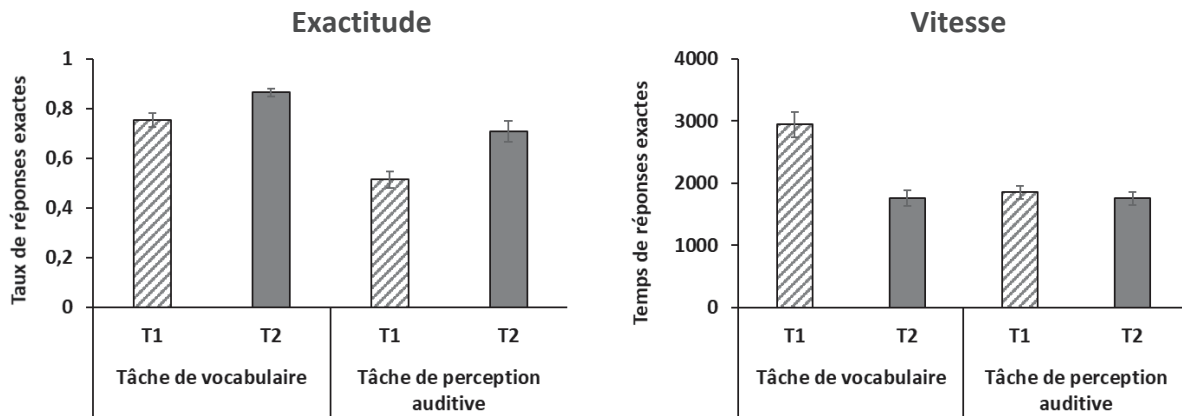


Figure 71. Taux de réponses exactes (gauche) et temps de réponses (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite) à T1 (rayures) et T2 (plein) pour le groupe GB. Les barres d'erreurs représentent l'erreur-type.

**Epreuves de production.** Dans la tâche de **production sans modèle**, après l'entraînement (T2), les participants du groupe GB produisaient des /ɪ/ plus courts ( $t(19) = 3.14, p = .005, d = .43$ ) et des /i:/ plus longs ( $t(19) = 3.23, p = .004, d = .83$ ) qu'à T1 (Fig. 72, gauche). L'allongement du /i:/ dans la tâche de **production avec modèle** était tout juste significatif ( $t(19) = 2.07, p = .053, d = .50$ ) entre T1 et T2 (Fig. 72, droite).

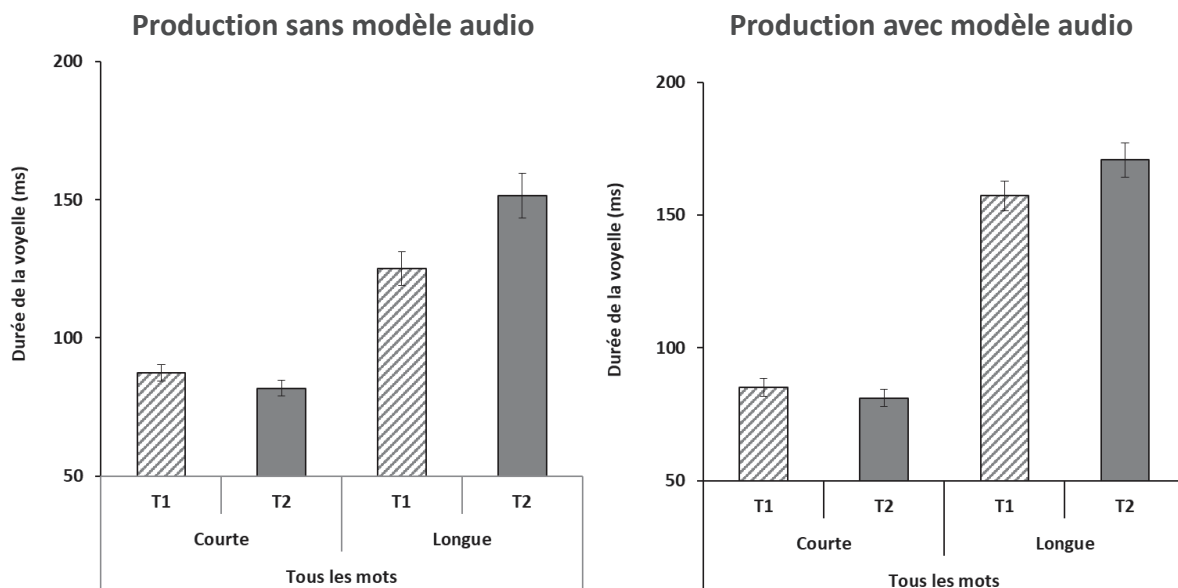


Figure 72. Durée des voyelles produites entre T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (gris) et la voyelle longue /i:/ (violet), dans les tâches de répétition sans modèle audio (gauche) et avec modèle audio (droite). Les barres d'erreurs représentent l'erreur-type.

Concernant les analyses complémentaires réalisées séparément sur les mots ayant été étudiés pendant l'entraînement et ceux ne l'ayant pas été, dans la tâche de **production sans modèle** (Fig. 73, haut) les participants allongeaient la durée du /i:/, que les mots aient été vus pendant l'entraînement ( $t(19) = 3.54, p = .002, d = .95$ ) ou non ( $t(19) = 2.42, p = .026, d =$

.64). Seuls les /ɪ/ contenus dans des mots non-appris durant l'entraînement étaient significativement raccourcis ( $t(19) = 2.21, p = .040, d = .43$ ), et non dans ceux ayant été appris ( $t(19) = 1.87, p = .077, d = .36$ ). Dans la tâche de **production avec modèle audio** (Fig. 73, bas), seul les /i:/ contenus dans des mots ayant été entraînés s'allongeaient significativement entre T1 et T2 ( $t(19) = 2.41, p = .026, d = .52$ ).

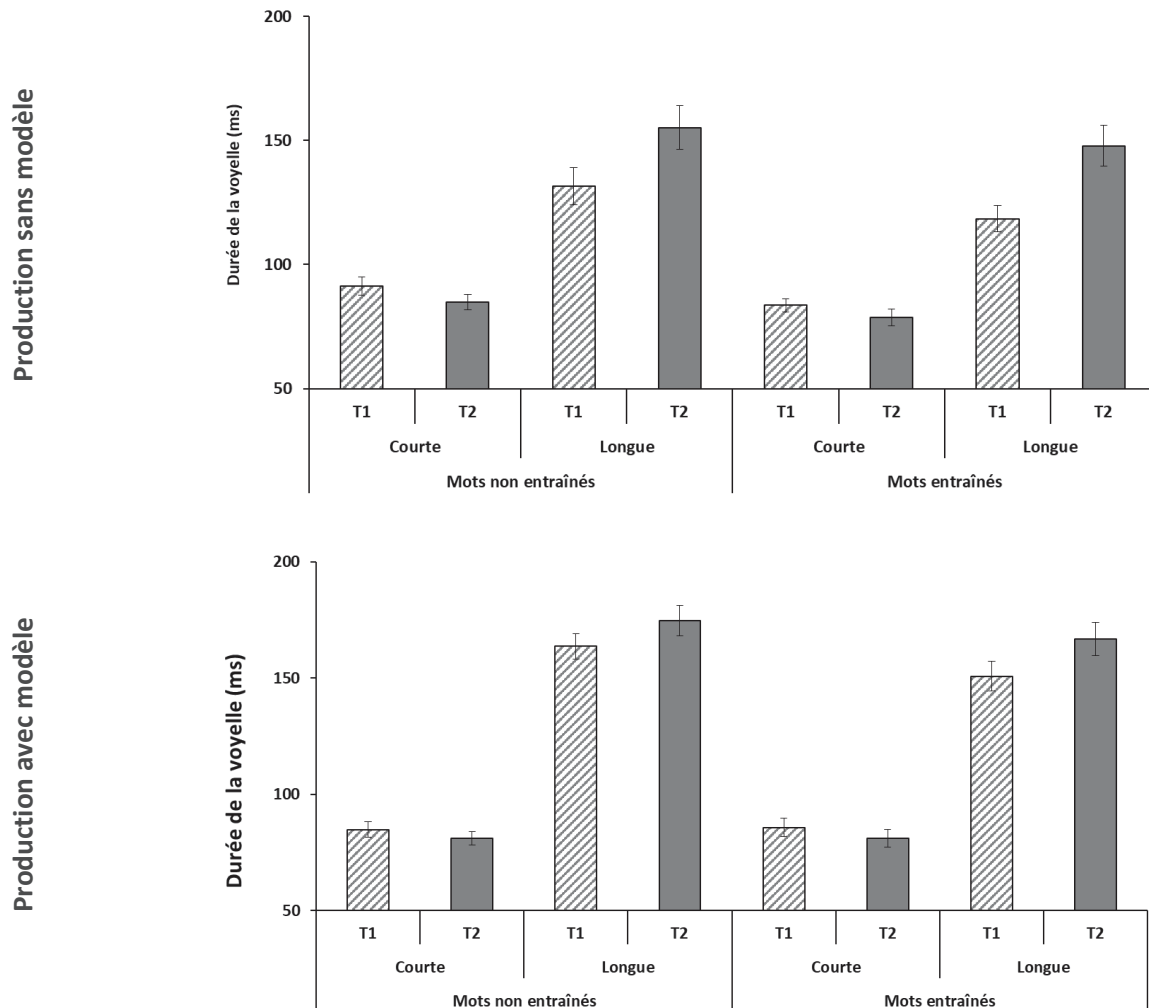


Figure 73. Durée des voyelles produites dans les tâches de production sans modèle audio (haut) et avec modèle audio (bas) entre T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (gris) et la voyelle longue /i:/ (violet). Les barres d'erreurs représentent l'erreur-type.

### 6.5.7. Discussion intermédiaire

Les deux épreuves de perception n'étant pas sensibles à l'effet de répétition sur l'exactitude (Expérience 10), et les groupes GD et GB ne présentant pas de différence d'exactitude à T1 dans l'une d'elles (épreuve de vocabulaire), le progrès de grande taille en vocabulaire après les entraînements dichotique ( $d = 1.36$ ) et binaural ( $d = 1.07$ ) montre un effet bénéfique de l'entraînement perceptif, de taille un peu plus grande s'il est dichotique. Dans l'épreuve de



perception auditive, l'absence d'effet de répétition (Expérience 10) tranchait avec le progrès de taille moyenne après l'entraînement dichotique ( $d = .67$ ) et de grande taille après l'entraînement binaural ( $d = 1.17$ ). Pour cette épreuve, le groupe GD avait de meilleures performances à T1 que le groupe GB, ce qui laissait moins de marge d'amélioration possible et pourrait expliquer la moindre taille de l'amélioration après l'entraînement dichotique.

En production, il n'y avait pas d'effet de répétition dans l'épreuve sans modèle audio pour /ɪ/ (Expérience 10), ce qui constituait une bonne base pour tester l'effet d'entraînement. L'entraînement a permis de raccourcir le /ɪ/ avec des tailles d'effet similaires pour les entraînements dichotique ( $d = .41$ ) et binaural ( $d = .43$ ). Comme le groupe GD produisait des /ɪ/ plus courts que le groupe GB à T1, le progrès du groupe GD peut être valorisé, car il y avait pour lui moins de marge de modification possible et le raccourcissement est une modification qui a ses limites. Par ailleurs, il est difficile de discuter l'hypothèse sur l'effet de transfert vers des mots non-appris, car le progrès se limitait aux mots appris pour le groupe GD et aux mots non-appris pour le groupe GB (difficile de parler de transfert en l'absence de progrès sur les mots appris). Pour le /i:/, il avait un effet de répétition avec un allongement de cette voyelle en l'absence d'entraînement entre T1 et T2 ( $d = .83$ ) (Expérience 10), et les performances des groupes GD et GB étaient comparables entre elles à T1. A partir de cette même base à T1, l'entraînement dichotique est suivi d'un progrès de taille plus petite que l'effet de répétition ( $d = .65$ ) et l'entraînement binaural d'un progrès de même taille celle de l'effet de répétition ( $d = .83$ ). Ainsi, pour /i:/ il n'y a pas de progrès dû à l'entraînement perceptif. Enfin, la production en présence d'un modèle audio ne subissait pas d'effet de répétition (Expérience 10) et les groupes GD et GB étaient équivalents à T1, mais seul le groupe GB a allongé la voyelle longue /i:/ entre T1 et T2, de manière cependant tout juste significative et seulement sur les mots appris.

L'ensemble des données soutiennent l'hypothèse de l'efficacité de l'entraînement perceptif sur les deux tests perceptifs, et sur la production surtout pour la voyelle courte /ɪ/. Si quelques aspects des données suggèrent un effet légèrement plus grand de l'entraînement dichotique pour le contraste /ɪ-/i:/ en perception et sur la brièveté du /ɪ/ en production, l'entraînement binaural est suivi d'un allongement plus net du /i:/. L'hypothèse d'une optimisation de l'effet de l'entraînement perceptif par la stimulation dichotique n'a pas d'argument fort en sa faveur dans cette expérience sur les voyelles. L'hypothèse sur le transfert à des mots nouveaux peut difficilement être discutée, mais un transfert perception-production s'est produit à partir de l'entraînement perceptif de manière générale.

## **6.6. Expérience 12 : effets de l'entraînement dichotique sur la perception et la production de voyelles anglaises chez des adultes dyslexiques**

### **6.6.1. Principe de l'expérience et hypothèses**

Le but de l'Expérience 12 était d'évaluer l'effet de l'entraînement dichotique utilisé dans l'Expérience 11 chez des adultes francophones dyslexiques, sur la perception et la production des voyelles anglaises /ɪ/ et /i:/.

Les hypothèses suivantes ont été formulées :

- 1) Avant l'entraînement, les performances en perception des voyelles anglaises /ɪ/ et /i:/ devraient être plus élevées chez les adultes normo-lecteurs (NL) que chez les adultes dyslexiques (DYS), et il devrait en être de même en production avec un /ɪ/ plus court et un /i:/ plus long dans le groupe NL que dans le groupe DYS.
- 2) Suite à l'entraînement par stimulation dichotique la perception du contraste /ɪ-/i:/ (tâches de vocabulaire et de perception) pourrait s'améliorer chez les adultes dyslexiques, par une diminution de la durée du /ɪ/ et/ou un allongement de la durée du /i:/. Pour les potentiels effets de l'entraînement, un intérêt particulier sera porté à l'exactitude car cette mesure est insensible à l'effet de répétition selon l'Expérience 10.
- 3) L'effet de l'entraînement pourrait être suffisamment profond pour qu'un transfert se fasse vers la production de mots contenant les voyelles /ɪ/ et /i:/ (lecture sans et avec modèle audio), avec éventuellement aussi un transfert vers des mots n'ayant pas été entendus pendant l'entraînement. Pour les potentiels effets de l'entraînement, un intérêt particulier sera porté au /ɪ/ dans les deux tâches et au /i:/ dans la tâche avec modèle audio, car ces mesures sont insensibles à la répétition d'après l'Expérience 10.

### **6.6.2. Méthode**

#### **6.6.2.1. Participants**

Les 20 participants (DYS) (16 femmes) de 18.7 ans à 34.7 ans (M = 22 ans 8 mois ; ET = 3 ans 6 mois) étaient des étudiants monolingues francophones ayant reçu un diagnostic de dyslexie développementale par un orthophoniste durant l'enfance. Ils n'avaient pas d'autre trouble neurodéveloppemental (notamment pas de TDAH), ni de trouble articulaire ou bégaiement. Ils déclaraient avoir une bonne audition et une vision normale ou corrigée. Tous

les participants étaient droitiers (80% de réponses « droite » dans la version abrégée de l'*Edinburgh Handedness Inventory* (Oldfield, 1971)).

Les participants normo-lecteurs (NL) (20) formant le groupe contrôle pour l'Expérience 12 étaient les participants testés avant et après l'entraînement par stimulation dichotique de l'Expérience 11 (Annexe 3).

#### **6.6.2.2. Matériel et équipement**

Le même matériel que celui de l'Expérience 11 décrit dans la partie 6.5.2.2 de ce chapitre était utilisé dans l'Expérience 12.

#### **6.6.2.3. Procédure d'évaluation**

Les participants réalisaient la même procédure que celle de l'Expérience 11 décrite dans la partie 6.5.2.3 de ce chapitre.

#### **6.6.2.4. Procédure d'entraînement**

Les participants dyslexiques de cette expérience ont réalisé le même entraînement dichotique que dans l'Expérience 11 décrit dans la partie 6.5.2.4 de ce chapitre. Comme dans l'Expérience 11, les fichiers de réponses étaient examinés pour vérifier que le participant n'avait pas appuyé sur d'autres touches que celles désignées pour les réponses, et qu'il n'avait pas produit plus de quatre temps supérieurs à 3 écart-types calculés à partir de ses données. Il n'y a pas eu de cas de participants avec des temps de réponse étonnamment rapides. Dans de cas où les réponses étaient globalement très lentes ou particulièrement rapides, les données du participant n'étaient pas conservées si cette lenteur ou rapidité extrêmement ne concernait que T1 ou que T2, car cela pouvait témoigner d'un manque de sérieux lors de l'une des étapes. Aucun participant n'a été exclu en respectant ces critères.

#### **6.6.3. Analyses des données**

Les données ont été traitées comme décrit dans l'Expérience 10 (partie 6.4.3 de ce chapitre). Les premières analyses portaient sur les performances à T1 des participants dyslexiques (DYS) par rapport à leurs pairs normo-lecteurs (NL) pour mesurer une possible différence entre les deux groupes à l'aide de tests *t* de Student non-appariés bilatéraux avec correction de Bonferroni. Les performances du groupe DYS ont aussi été confrontées aux normes (encore partielles) présentées dans l'Expérience 10. Ensuite, l'effet de l'entraînement chez les dyslexiques a été mesuré comme pour l'Expérience 11 (partie 6.5.3 de ce chapitre) entre T1 et T2.

#### 6.6.4. Résultats : Comparaison des performances entre les groupes NL et DYS avant l'entraînement (à T1)

**Epreuves de perception.** Dans la **tâche de vocabulaire**, à T1, il n'y avait pas de différence entre les deux groupes pour le taux de réponses exactes ( $t(37.98) = 1.23, p = .228, d = .39$ ) (Fig. 74, à gauche), mais le groupe NL répondait significativement plus vite que le groupe DYS ( $t(22.82) = 3.41, p = .002, d = 1.08$ ) (Fig. 74, à droite). Dans la tâche de **perception auditive**, le groupe NL répondait de manière plus exacte ( $t(30.76) = 2.20, p = .035, d = .70$ ) et plus rapide ( $t(33.97) = 2.88, p = .007, d = .91$ ) que le groupe DYS.

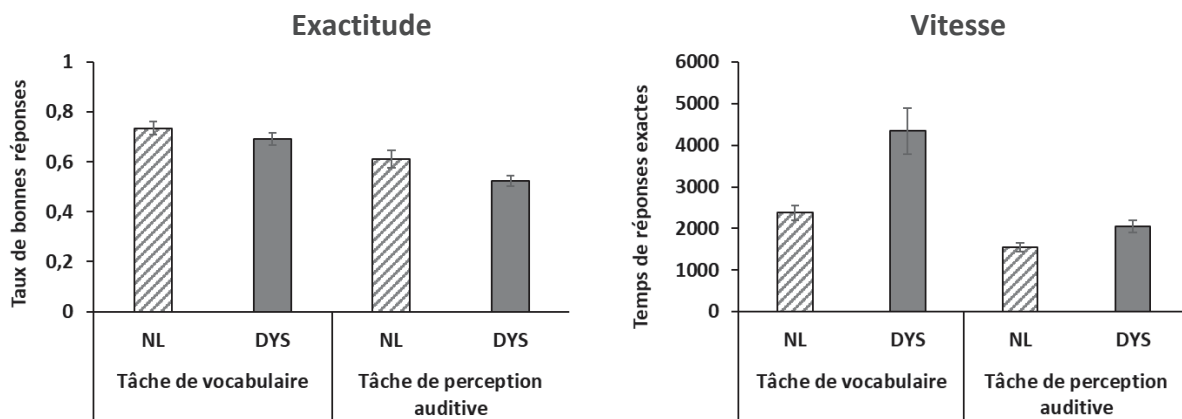


Figure 74. Taux de réponses exactes (gauche) et temps de réponses (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite) à T1 pour les groupes NL (rayures) et DYS (plein). Les barres d'erreurs représentent l'erreur-type.

**Epreuves de production.** Avant l'entraînement (T1), pour la production du /i/, il n'y avait pas de différence entre les deux groupes que ce soit sans modèle ( $t(38.00) = 0.18, p = .855, d = .05$ ) ou avec modèle ( $t(37.39) = 1.40, p = .169, d = .44$ ). La voyelle longue /i:/ était produite avec une durée plus longue par le groupe NL que par le groupe DYS dans les tâches sans modèle audio ( $t(26.82) = 2.74, p = .011, d = .87$ ) et avec modèle audio ( $t(36.65) = 2.73, p = .010, d = .86$ ) (Fig. 75).

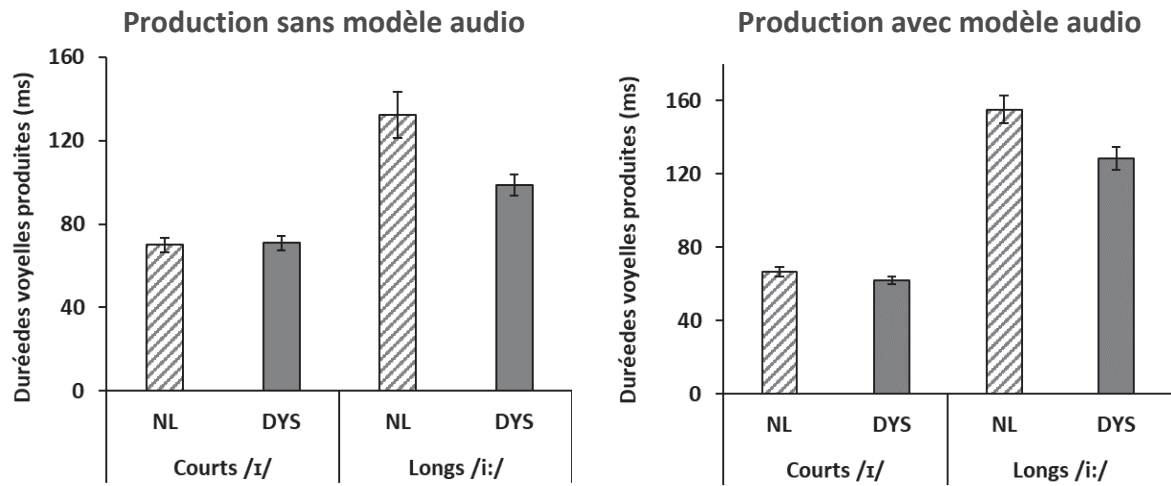


Figure 75. Durée des voyelles produites dans les tâches de lecture sans modèle audio (gauche) et avec modèle audio (droite) par les groupes NL (rayure) et DYS (plein) pour la voyelle courte /ɪ/ (partie gauche) et la voyelle longue /i:/ (partie droite). Les barres d'erreurs représentent l'erreur-type.

**Etude qualitative des données des adultes dyslexiques par rapport au groupe normatif.**

Concernant l'épreuve de perception, la figure 76 permet d'observer les performances de chaque participant dyslexique comparées à la norme (n = 60) proposée dans l'Expérience 10. Dans la tâche de vocabulaire, la majorité des participants dyslexiques répondaient moins correctement (65%) et plus lentement (55%) que la norme définie par l'intervalle de confiance à 95%. Beaucoup répondaient aussi moins correctement (45%) et plus lentement (70%) que la norme dans la tâche de perception auditive. Cela montre aussi que quelques participants dyslexiques de l'échantillon étaient aussi performants voire meilleurs que la norme dans les deux épreuves de perception.

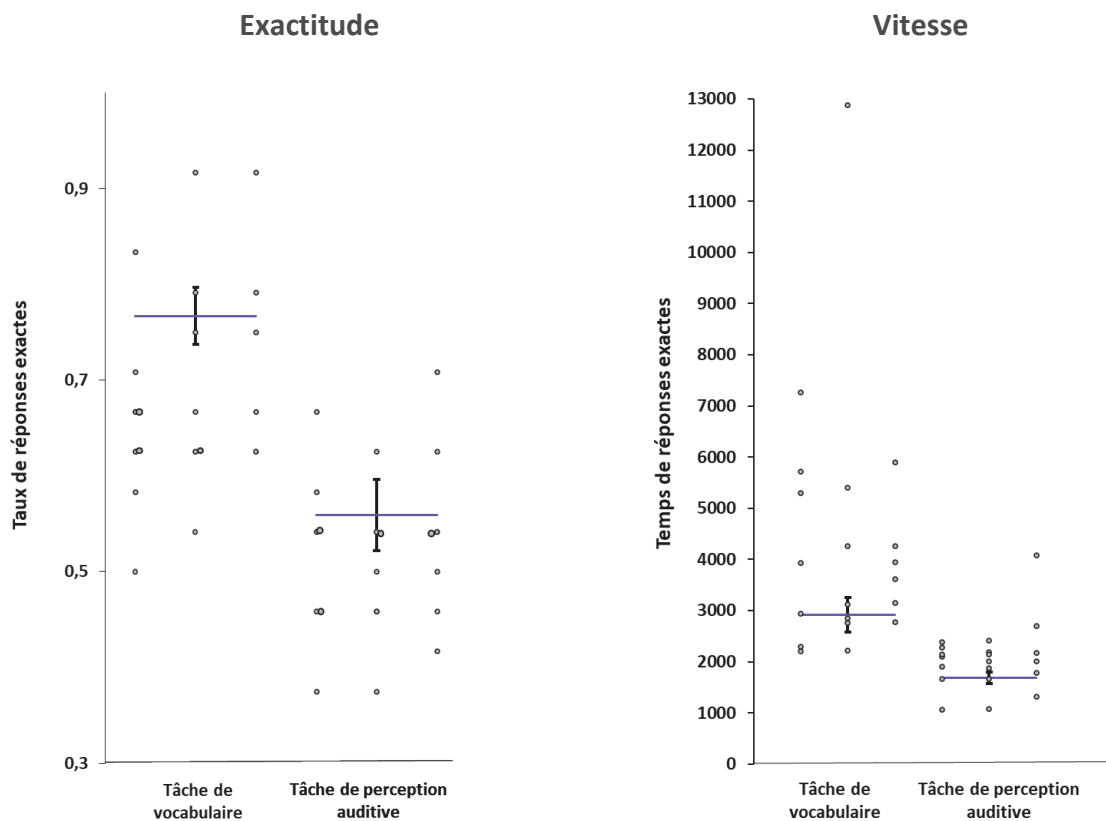


Figure 76. Répartition des 20 participants du groupe DYS (points) par rapport à la moyenne des 60 adultes normo-lecteurs ayant passé les épreuves (Expérience 10) (barre horizontale violette) et aux bornes définies par l'intervalle de confiance à 95%, pour les taux de réponses exactes (gauche) et les temps de réponses aux décisions exactes (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite).

Dans l'épreuve de production (Fig. 77), la position individuelle des participants DYS par rapport à la norme indique qu'une forte proportion d'entre eux (65%) produisaient des voyelles /i/ d'une durée située dans la norme pour la tâche de production sans modèle audio, et une proportion aussi grande (65%) la produisaient cette fois de façon plus courte que la norme s'il y avait un modèle audio. La figure indique aussi que la durée du /i:/ était anormalement courte pour une proportion élevée de participants DYS dans la tâche sans modèle audio (65%). Dans la tâche avec modèle audio, la durée du /i:/ était plus courte que la norme pour 55% du groupe DYS, mais on remarque surtout que les durées pour cette voyelle /i:/ étaient particulièrement hétérogènes entre les participants DYS (la moitié produisaient un /i:/ plus long que la moyenne, et l'autre moitié un /i:/ plus court que la moyenne).

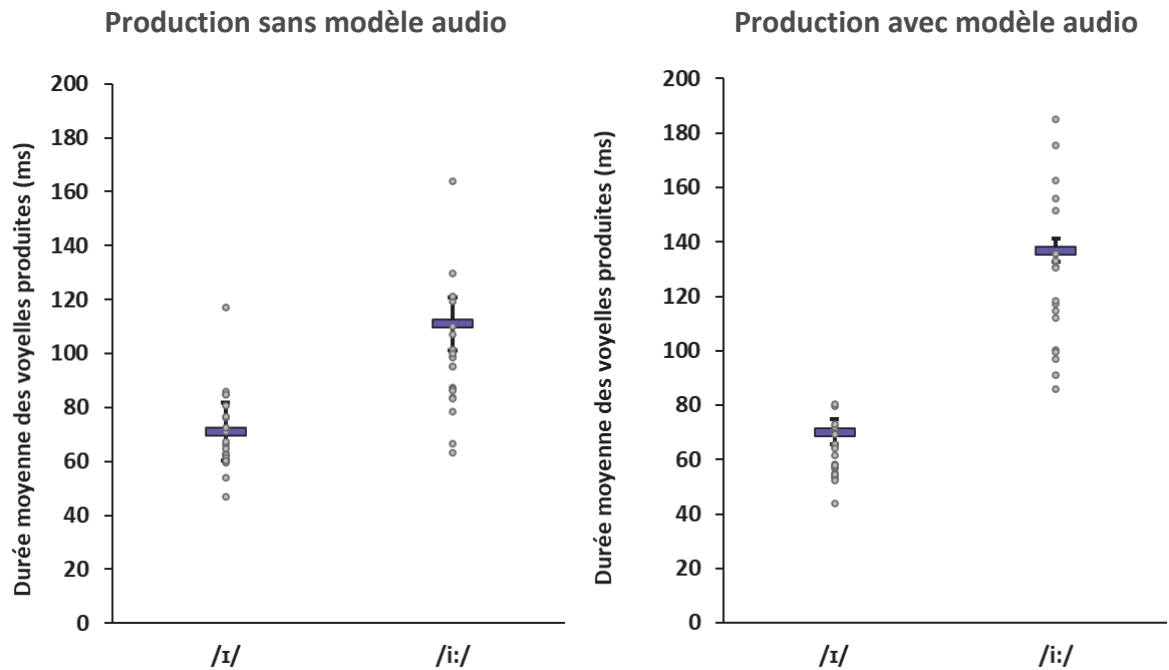


Figure 77. Répartition des participants du groupe DYS (points) par rapport à la moyenne de la norme (barre violette) pour les durées des voyelles /ɪ/ (partie gauche) et /i:/ (partie droite) produites dans les tâches de production sans modèle (gauche) et avec modèle (droite). Les barres d'erreurs représentent l'intervalle de confiance à 95%.

#### 6.6.5. Résultats : Effet de l'entraînement dichotique pour le groupe DYS

**Epreuves de perception.** Dans la **tâche de vocabulaire**, les participants dyslexiques répondaient de façon plus exacte ( $t(19) = 7.03, p < .001, d = 1.36$ ) et plus rapide ( $t(19) = 4.03, p = .001, d = 1.19$ ) à T2 qu'à T1 (Fig. 78). Dans la tâche de **perception auditive**, leurs réponses tendaient à être plus exactes ( $t(19) = 1.87, p = .077, d = .58$ ), sans altération de leur temps de réponses ( $t(19) = 1.21, p = .241, d = .15$ ). Cette amélioration marginale entre T1 et T2 leur permettait de passer de réponses non différentes du hasard à T1 (Binomiale = .34,  $p = .50$ ) à une exactitude significativement supérieure au hasard à T2 (Binomiale  $< .001, p = .50$ ).

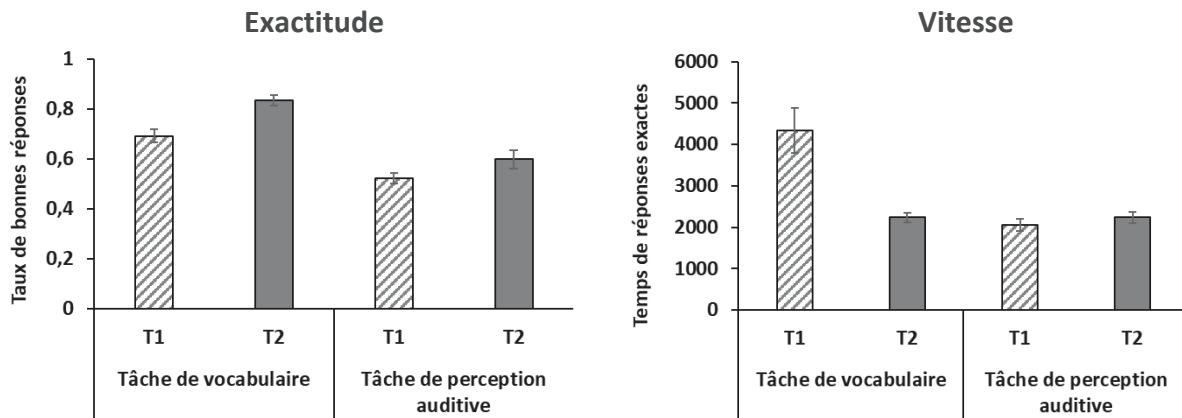


Figure 78. Taux de réponses exactes (gauche) et temps de réponses (droite) dans les tâches de vocabulaire (partie gauche) et de perception auditive (partie droite) à T1 (rayures) et T2 (plein). Les barres d'erreurs représentent l'erreur-type.

**Epreuves de production.** Dans la tâche de **production sans modèle audio** (Fig. 79, gauche), les participants dyslexiques produisaient des /ɪ/ plus courts ( $t(19) = 3.25, p = .004, d = .58$ ) et des /i:/ plus longs ( $t(19) = 2.66, p = .015, d = .56$ ) à T2 qu'à T1. Dans la tâche de **production avec modèle** (Fig. 79, droite), de T1 à T2 la réduction de durée du /ɪ/ était de petite taille et n'atteignait pas le seuil de significativité ( $t(19) = 1.70, p = .105, d = .26$ ), alors que la durée des /i:/ s'allongeait significativement ( $t(19) = 2.41, p = .026, d = .34$ ).

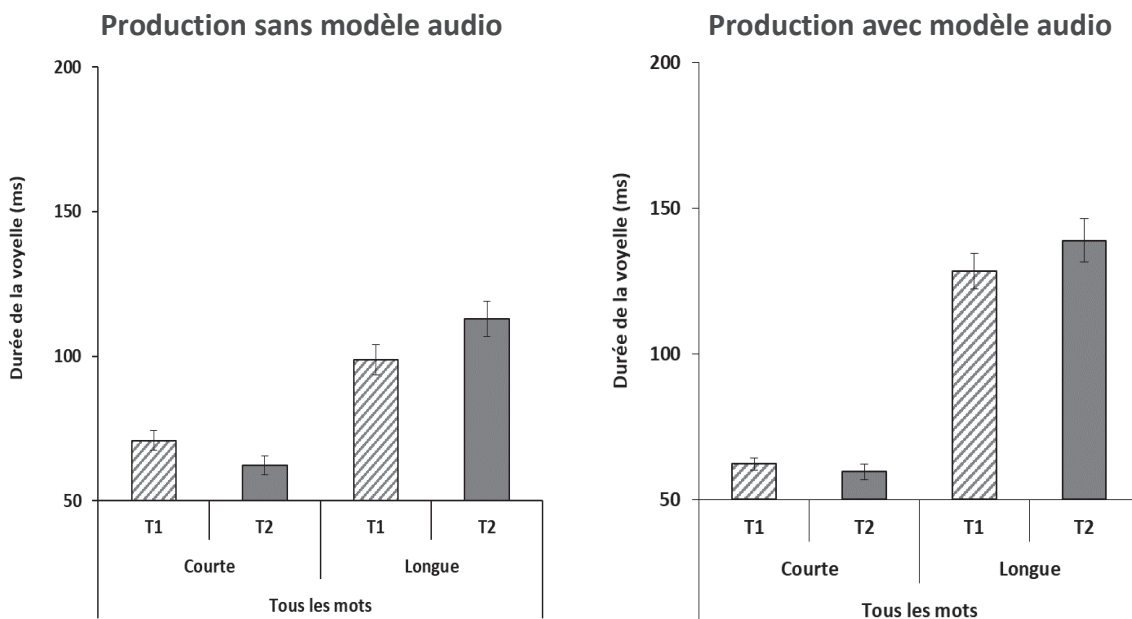


Figure 79. Durée des voyelles produites à T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (partie gauche) et la voyelle longue /i:/ (partie droite). Les barres d'erreurs représentent l'erreur-type.

Afin de tester l'hypothèse d'un potentiel effet de transfert de compétences, des analyses complémentaires ont été réalisées séparément sur les mots qui avaient été étudiés pendant l'entraînement et sur ceux ne l'avaient pas été. Dans la tâche de **production sans modèle**



**audio** (Fig. 80, haut), les participants DYS raccourcissaient la durée du /ɪ/ entre T1 et T2 que les mots aient été vus pendant l'entraînement ( $t(19) = 2.81, p = .011, d = .54$ ) ou non ( $t(19) = 2.80, p = .011, d = .55$ ). C'était aussi le cas pour l'allongement de la durée du /i:/ (respectivement,  $t(19) = 2.05, p = .054, d = .47$  ;  $t(19) = 2.95, p = .008, d = .59$ ). Dans la tâche de **lecture avec modèle audio** (Fig. 80, bas), seuls les /i:/ contenus dans des mots appris lors de l'entraînement étaient allongés à T2 ( $t(19) = 2.91, p = .009, d = .46$ ). Il n'y avait pas de changement sur le /ɪ/, que les mots aient été appris ou non.

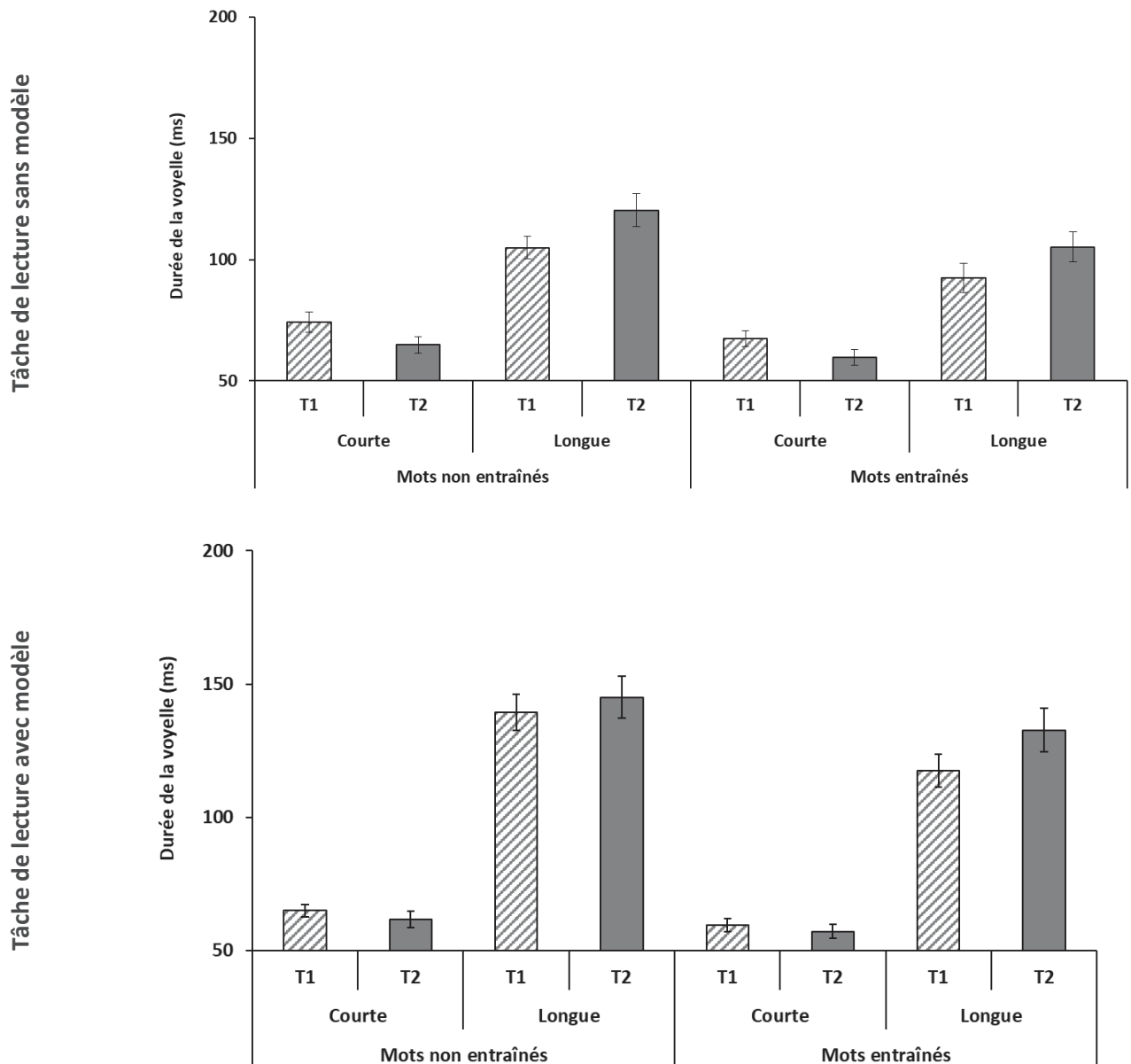


Figure 80. Durée des voyelles produites dans les tâches de lecture sans modèle audio (haut) et avec modèle audio (bas) à T1 (rayure) et T2 (plein) pour la voyelle courte /ɪ/ (gris) et la voyelle longue /i:/ (violet). Les barres d'erreurs représentent l'erreur-type.

### 6.6.6. Discussion intermédiaire

Plusieurs aspects des résultats montrent que les adultes dyslexiques ont progressé après l'entraînement. Dans la tâche de vocabulaire, le taux de réponses exactes était insensible à la répétition (Expérience 10) et ce taux a augmenté après l'entraînement chez les adultes dyslexiques. Dans cette tâche, il y avait un effet de répétition sur les temps de réponse car l'effet test-retest accélérail la vitesse (Expérience 10), mais les dyslexiques ont aussi accéléré leurs réponses après l'entraînement avec un effet de plus grande taille que l'effet test-retest. L'amélioration de l'exactitude ne s'est pas faite au prix de la rapidité, suggérant qu'ils ont appris et qu'ils sont devenus plus sûrs de leurs connaissances. Le progrès de grande taille en rapidité est remarquable compte tenu du fait qu'ils présentaient dans cette tâche un déficit par rapport aux normo-lecteurs avant l'entraînement (à T1) seulement pour leur lenteur.

Pour la tâche de perception auditive où l'exactitude était insensible à la répétition (Expérience 10), les performances des dyslexiques étaient déficitaires par rapport à celles des normo-lecteurs à T1 en exactitude comme en rapidité. Après l'entraînement, les dyslexiques ont réalisé un progrès de taille modérée et qui n'atteignait pas le seuil de significativité. Ce progrès non-significatif et de taille modérée leur permettait tout de même de produire, après l'entraînement, des réponses plus exactes que si elles étaient données au hasard, alors qu'elles ne se distinguaient pas de réponses données au hasard à T1.

Des progrès à la suite de l'entraînement sont aussi apparus en production, ce qui témoigne d'un transfert de compétences développées à partir d'exercices perceptifs et étendues à la production de mots anglais. Dans la tâche sans modèle audio, la durée du /ɪ/ était insensible à la répétition (Expérience 10), et les adultes dyslexiques ont pourtant raccourci la durée de cette voyelle courte à T2. Il s'agit non seulement d'un transfert de la perception vers la production, mais aussi d'un transfert vers des mots non-appris, pour lesquels le /ɪ/ était autant raccourci que pour les mots appris. Parallèlement, les adultes dyslexiques ont augmenté la durée du /i:/ à T2. Cet effet est moins remarquable car un allongement du /i:/ se produisait aussi en effet test-retest (Expérience 10), sous la forme d'un effet de grande taille alors que l'allongement réalisé par les dyslexiques après l'entraînement était de taille modérée. Celui-ci est cependant opéré en faisant un transfert vers des mots non-appris. Il s'agissait aussi d'un vrai progrès, car les dyslexiques produisaient le /i:/ de manière anormalement courte par rapport aux normo-lecteurs à T1. Le raccourcissement du /ɪ/ et l'allongement du /i:/ par les adultes dyslexiques après l'entraînement sont deux effets significatifs qui concouraient à accroître le contraste de durée entre les deux voyelles anglaises après l'entraînement.

En production avec modèle audio, il n’y avait pas d’effet de répétition sur la durée du /i:/ (Expérience 10) et les dyslexiques ont pourtant accru la durée de cette voyelle à T2, véritable progrès là aussi, car ils produisaient cette voyelle longue de manière anormalement courte par rapport aux NL avant l’entraînement.

L’ensemble des données contribue ainsi à montrer qu’en perception comme en production les adultes dyslexiques ont progressé après l’entraînement perceptif qui leur avait été proposé ici, dans sa version avec dispositif dichotique, pour le traitement des voyelles /ɪ/ et /i:/, au-delà de ce que permet un simple effet test-retest.

L’Expérience 12 a été conduite auprès d’étudiants dyslexiques dont nous savions peu de chose. Leur diagnostic avait été posé pendant l’enfance, mais nous n’avions aucune indication sur le degré de sévérité et l’origine de leur pathologie. Le fait qu’ils soient étudiants suggère qu’ils avaient réussi à compenser certaines de leurs difficultés, suffisamment pour obtenir au moins le baccalauréat. Ils ne présentaient sans doute pas, pour la majorité d’entre eux, une forme très sévère de dyslexie, ce qui doit rendre prudent sur la généralisation des résultats dans l’interprétation.

Les déficits cognitifs sous-jacents à leur pathologie neurodéveloppementale variaient sans doute entre les participants, mais les troubles phonologiques sont particulièrement répandus chez les personnes dyslexiques (Bedoin & Krifi, 2009 ; Schulte-Körne *et al.*, 1999 ; Serniclaes *et al.*, 2001, 2004 ; Snowling, 1981, 1987 ; Sprenger-Charolles & Serniclaes, 2004), ils se maintiennent à l’âge adulte (Abadie & Bedoin, 2016) et il s’agirait même des lacunes les plus persistantes ((Dufor *et al.*, 2007 ; Fosker & Thierry, 200 ; Hanley, 1997 ; Noordenbos *et al.*, 2013 ; Svensson & Jacobson, 2006 ; Van Beinum *et al.*, 2005 ; Vandermosten *et al.*, 2010 ; Witton *et al.*, 1998 ; Zabell & Everatt, 2002). D’autres troubles, de nature visuo-attentionnelle, participent aussi parfois à l’émergence de la dyslexie et perturbent surtout l’encodage et la mémorisation de l’orthographe précise des mots (Bedoin, 2016, 2017 ; Bosse *et al.*, 2007 ; Bosse & Valdois, 2009 ; Cassim *et al.*, 2014 ; Lewis & Frick, 1999 ; Moll & Jones, 2013 ; Valdois *et al.*, 2003), difficultés qui caractérisent certains adultes dyslexiques (Abadie & Bedoin, 2016 ; Araujo *et al.*, 2015 ; Mazur-Palandre *et al.*, 2016 ; Milne *et al.*, 2003 ; Moores *et al.*, 2011, 2015 ; Romani *et al.*, 2008 ; Zabell & Everatt, 2002). Dans les données de l’Expérience 12, quelques éléments confirment que les participants dyslexiques présentaient des difficultés persistantes en traitement du langage (ici dans le contexte de l’apprentissage d’une L2), et que ces déficits concernaient le traitement phonologique. Les tests évaluaient en effet des connaissances et compétences au sujet de

phonèmes – voyelles /ɪ/ et /i:/ – et le groupe d’adultes dyslexiques était significativement en échec par rapport au groupe de normo-lecteurs. Ils étaient significativement moins rapides dans les deux tâches de perception et de surcroît moins exacts dans la tâche de perception auditive où, contrairement à la tâche de vocabulaire, aucune indication orthographique n’était fournie. Avant l’entraînement, quand ils devaient s’appuyer ainsi sur leurs seules compétences et connaissances phonologiques, les adultes dyslexiques présentaient une exactitude déficitaire au point qu’ils répondaient d’une manière non différente du hasard, en plus d’être moins performants que les normo-lecteurs.

Par contre, dans la tâche de vocabulaire, la présence de la forme orthographique des mots leur permettait d’exploiter une analogie entre la taille des graphèmes et la durée des voyelles. En anglais, la voyelle courte /ɪ/ s’orthographie généralement avec la lettre I, alors que la voyelle longue /i:/ s’orthographie avec un graphème contenant deux lettres. Les adultes dyslexiques de l’étude étaient davantage en difficulté en anglais lorsqu’ils devaient s’appuyer sur la phonologie en l’absence de telles informations orthographiques. Ils présentaient ainsi sans doute, dans l’ensemble, davantage de difficultés d’ordre phonologique que de nature visuo-attentionnelle. Cela montre aussi que des adultes dyslexiques sont capables d’exploiter efficacement des informations orthographiques dans le cadre de l’apprentissage de l’anglais L2. Sur le plan pédagogique, cela conduit à conseiller pour eux des méthodes d’enseignement qui leur permettent de s’appuyer sur la présentation écrite des mots et des phrases.

Avant l’entraînement, les adultes dyslexiques de l’étude présentaient aussi une particularité en production qui confirme cette fois leurs difficultés dans le domaine des traitements temporels, décrits comme une cause possible des déficits phonologiques (Goswami, 2011 ; Goswami *et al.*, 2002 ; Habib, 2021 ; Habib *et al.*, 2002 ; Leong & Goswami, 2014b ; Soltész *et al.*, 2013 ; Tallal *et al.*, 1996 ; Van Hirtum *et al.*, 2019b). Ils produisaient la voyelle longue /i:/ avec une durée plus courte que les adultes normo-lecteurs dans les deux tâches de production. Même la voyelle courte était produite de façon trop brève, ce qui rendait une bonne partie d’entre eux atypiques par rapport à la norme pour la prononciation de cette voyelle aussi. Cela suggère que leur pathologie et leurs difficultés de traitements temporels, les plaçaient d’emblée en plus grande difficulté que les autres étudiants pour apprendre une L2 dont la phonologie recèle des subtilités temporelles qu’ils ne peuvent pas facilement assimiler à la phonologie de leur L1. C’est le cas de l’apprentissage de l’anglais L2 pour des francophones, tant pour les voyelles (Expérience 12) que pour le voisement des consonnes étudié dans l’Expérience 8.

L'aide proposée pour l'apprentissage de l'anglais dans l'Expérience 12 ciblait les connaissances phonologiques sur les phonèmes de cette L2, mais les moyens utilisés dans l'entraînement mobilisaient à la fois des compétences phonologiques et des compétences orthographiques. Les mots étaient en effet donnés à entendre, associés à une indication sur leur signification et sur leur orthographe. L'analogie entre la durée des voyelles et des graphèmes a probablement été exploitée pendant l'entraînement pour apprendre l'orthographe précise de mots anglais, d'où les progrès entre T1 et T2 dans la tâche de vocabulaire, où ces connaissances orthographiques pouvaient être mises à profit, et l'absence de progrès significatif dans la tâche basée seulement sur la perception auditive.

### **6.7. Discussion sur l'aide à l'apprentissage de perception des durées des voyelles anglaises (Expériences 10, 11 et 12)**

Ce chapitre portait sur l'apprentissage tardif d'un nouveau contraste phonologique en anglais L2 pour des francophones : le contraste de durée entre deux voyelles, à travers l'exemple des phonèmes anglais /ɪ/ et /i:/. Un entraînement perceptif bref (3 jours, 20 min/jour) était proposé et son apport évalué par des épreuves comportementales de perception et de production de mots anglais auprès d'étudiants normo-lecteurs et dyslexiques. Ce chapitre avait trois objectifs principaux. Le premier était de valider la pertinence de quatre épreuves spécialement conçues pour mesurer des performances en perception et en production de paires minimales s'opposant par les voyelles /ɪ/ et /i:/. Tester en production devait permettre d'évaluer un potentiel transfert de compétences entre perception et production de parole. Il s'agissait aussi de commencer à créer des normes pour des tests souhaités dans le milieu académique, susceptibles de fournir des mesures objectives du niveau des étudiants en phonologie anglaise, de préciser leurs besoins et de disposer d'outils d'évaluation objective de l'efficacité de programmes d'entraînement en anglais L2. Le deuxième objectif était d'évaluer l'apport d'un dispositif de stimulation dichotique plutôt que binaural dans le cadre d'un entraînement perceptif sur l'opposition de durée des voyelles /ɪ/ et /i:/. Le troisième objectif était de tester si des adultes dyslexiques pouvaient progresser en perception et en production de ces voyelles après l'entraînement perceptif avec écoute dichotique.

Une synthèse des principaux résultats de chaque expérience étant déjà présentée dans les discussions intermédiaires, la discussion générale pour ce chapitre tente d'aller plus loin en faisant des rapprochements entre les résultats des trois expériences et en proposant des interprétations. Un regard critique sera porté sur les nouvelles épreuves de perception et de

production. Il s'agira ensuite d'une réflexion sur les apports respectifs de l'entraînement perceptif et du dispositif de présentation dissociée des stimuli pour stimuler plus sélectivement les cortex auditifs gauche et droit en fonction de leur compétences supposées complémentaires pour le traitement d'informations de durées différentes. L'efficacité potentielle de ce dispositif de stimulation dichotique sera discutée en tant qu'argument pour soutenir les modèles théoriques qui ont guidé sa conception : le modèle de l'attention dynamique (DAT, Jones, 1976, 2008 ; Jones & Boltz, 1989 ; Jones *et al.*, 2002), les modèles d'analyse du signal de parole à échelles multiples (AST, Poeppel, 2003 ; Poeppel *et al.*, 2008) et, plus indirectement, l'hypothèse selon laquelle la dynamique des oscillations cérébrales guide fondamentalement le fonctionnement cognitif (Benítez-Burraco & Murphy, 2019 ; Giraud & Poeppel, 2012). Enfin, la question du niveau en anglais des étudiants dyslexiques par rapport à celui des étudiants normo-lecteurs sera reprise. La discussion finira sur les aides auxquelles ils sont potentiellement réceptifs pour progresser dans la maîtrise de nouvelles voyelles en anglais.

#### **6.7.1. De nouveaux tests d'évaluation de la perception et de la production des voyelles anglaises**

L'évaluation des quatre tâches de perception et de production a révélé des qualités psychométriques satisfaisantes, et a pointé certaines limites. La principale qualité est la résistance à l'effet de répétition et une fidélité souvent élevée. Seules trois mesures ne satisfont pas au critère d'insensibilité à l'effet test-retest : la vitesse de réponse dans les deux épreuves de perception et la longueur du /i:/ dans la tâche de production sans modèle audio. Testés à 3 jours d'intervalle sans entraînement entre les deux, les adultes étaient ainsi plus confiants dans leurs réponses lorsqu'ils faisaient ces tests pour la deuxième fois, puisqu'ils répondaient plus vite, sans pour autant améliorer leur exactitude.

La vitesse et l'exactitude restent deux données comportementales intéressantes si l'objectif est de situer les performances perceptives d'un adulte par rapport à des normes. Par contre, la mesure de la vitesse ne peut pas être considérée comme très informative pour l'évaluation des effets d'un entraînement car présentait un effet test-retest. Pour ce dernier usage, il est plus indiqué de mesurer l'exactitude.

Dans les épreuves de production, la durée moyenne de la voyelle courte /ɪ/ et de la voyelle longue /i:/ s'est avérée insensible à la répétition, ce qui en fait des mesures pertinentes pour comparer des étapes pré- et post-entraînement. Seul le /i:/ en tâche de production sans

modèle audio ne présentait pas cette qualité : les participants allongeaient la durée de cette voyelle lorsqu'ils faisaient le test pour la deuxième fois. Une interprétation peut être proposée pour expliquer ce progrès en l'absence d'entraînement, spécifiquement dans cette tâche. La consigne est de « lire avec le meilleur accent anglais possible » les mots écrits à l'écran. Les participants ne pouvaient s'appuyer que sur une représentation orthographique. Leur attention n'était pas détournée vers l'imitation de la voix, comme dans la tâche de production avec modèle audio, qui peut les attirer sur d'autres dimensions que la durée (la hauteur de la voix, son timbre...). Etant donné que les étudiants avaient tous pratiqué l'anglais depuis plusieurs années, et que tous les mots des épreuves comportent la voyelle /ɪ/ ou /i:/, ils ont pu inférer au fil des tests (dont certains, perceptifs, leur avaient déjà permis d'entendre la différence) que l'évaluation portait sur la durée des voyelles. Cela a pu les inciter à produire des /i:/ exagérément longs en deuxième session, en suivant l'indication graphémique qui attirait particulièrement leur attention sur ce point dans la tâche de production sans modèle audio.

Comme cet allongement du /i:/ en deuxième session ne se produisait pas lorsque les participants disposaient d'un modèle audio en plus de la représentation orthographique, cela nous apprend qu'un bon moyen d'attirer l'attention sur la différence de durée des voyelles anglaises est de mettre en évidence la différence orthographique associée à ces différences de durée. Faire en sorte que seule cette information orthographique soit disponible conduit l'apprenant à amplifier la différence temporelle. Sur le plan pédagogique, une fois que l'apprenant a une connaissance minimale de la différence de durée entre deux voyelles anglaises, comme c'est le cas ici avec les épreuves préalablement effectuées en perception, il peut être judicieux d'inciter l'apprenant à lire régulièrement à haute voix des listes de mots ou de phrases contenant de telles voyelles. L'amélioration due au simple effet test-retest suggère que cela pourrait renforcer des habitudes de prononciation adaptées. De plus, la production engage le système cérébral moteur en l'associant ici correctement à un phonème particulier. La théorie motrice de la perception de la parole suppose l'activation de ce même système moteur en perception (Lieberman & Mattingly, 1989), ce que confirment des données en neurosciences (Galantucci *et al.*, 2009). L'exercice répété de lecture à haute voix pourrait permettre à ce type d'influence de haut niveau (top-down) en perception de tirer parti de connaissances acquises par apprentissage statistique et adaptées aux mécanismes de perception des voyelles courtes et longues. L'implication d'oscillations endogènes delta, issues d'aires cérébrales dans la voie dorsale impliquant les aires motrices, est supposée renforcer la synchronisation d'une cascade d'oscillations hiérarchisées (Arnal *et al.*, 2015 ;



Morillon *et al.*, 2010). Celles-ci guideraient alors la perception pour un traitement plus précis et temporellement plus adapté des voyelles entendues.

Au final, sur le plan de leurs qualités pour des évaluations, les quatre épreuves élaborées offrent chacune au moins une mesure (deux pour la tâche de production avec modèle audio) pertinente pour évaluer les effets d'un entraînement. Proposer des épreuves en perception mais aussi en production présente aussi l'avantage de tester de potentiels transferts de compétences entre capacités en réception et en production. Des limites ont été pointées et peuvent faire paraître ces épreuves comme imparfaites. Cependant, en prenant les précautions indiquées, elles semblent plus valides que les tests souvent utilisés dans d'autres études qui n'évaluent pas toujours la part de progrès due à l'effet de répétition entre les étapes pré- et post-entraînement. Les évaluations objectives sur quelques points en phonologie permises par ce type de tests pourraient par ailleurs répondre aux besoins des enseignants en anglais pour compléter des échelles d'évaluation plus globales. Elles ont délibérément été conçues pour être très brèves (l'ensemble dure 20 min), afin d'être adaptées à de telles applications. L'ensemble des mesures peut ainsi être utilisé s'il s'agit de tester un participant une seule fois, pour contribuer à situer son niveau dans le milieu académique par exemple.

Il pourrait aussi être intéressant de compléter les normes avec plus de participants, mais aussi en intégrant les données de jeunes adultes normo-lecteurs ayant des niveaux d'études moins élevés, peut-être en dissociant les deux échantillons. Cela permettrait d'utiliser ces outils ainsi que les entraînements et l'évaluation de leurs effets dans le contexte de la formation continue. Des normes seraient aussi utiles pour un public plus jeune, au collège et au lycée. L'augmentation de la taille des échantillons et l'intégration de participants dyslexiques étudiants ou non-étudiants pourrait permettre l'évaluation de la sensibilité et de la spécificité de ces épreuves en tant qu'outils de détection de difficultés majeures en anglais, telles qu'elles sont vécues en cas de dyslexie. Les données obtenues auprès du petit groupe d'adultes dyslexiques montrent que le niveau de difficulté des tâches est adapté à ce public. Il est aujourd'hui nécessaire de disposer de ce type de mesures, car de plus en plus d'étudiants atteints de dyslexie accèdent aux études supérieures (pour les résultats d'une enquête à l'Université de Lyon, voir Abadie & Bedoin, 2016 et Mazur-Palandre *et al.*, 2016). Ils expriment leurs difficultés en apprentissage de l'anglais L2 et les enseignants ont peu d'outils objectifs pour confirmer la sévérité et le caractère atypique de ces difficultés, car les autres étudiants en ont aussi. Les épreuves proposées ici montrent que, comparés à la norme de leurs pairs, les étudiants dyslexiques se démarquent par des réponses moins exactes dans la tâche de



perception, des décisions moins rapides dans les deux épreuves de perception et des voyelles longues d'une durée anormalement courte dans les deux épreuves de production. Cela suggère qu'il devrait être possible, avec des échantillons plus conséquents, d'évaluer la sensibilité/spécificité des épreuves et de les utiliser pour convaincre les enseignants des difficultés atypiques des apprenants dyslexiques.

### **6.7.2. Apports des entraînements perceptifs pour l'apprentissage du contraste de durée des voyelles anglaises chez les adultes normo-lecteurs**

L'Expérience 11 permettait de comparer les effets d'un entraînement perceptif selon une présentation binaurale (groupe GB) ou dichotique (groupe GD) des stimuli. Cette comparaison s'est faite de façon inter-individuelle. Or, il se trouve que, avant l'entraînement, les deux groupes présentaient des différences pour plusieurs aspects des performances. Avant l'entraînement, leurs performances étaient équilibrées seulement pour l'exactitude dans la tâche de vocabulaire et la durée du /ɪ/ dans les deux tâches de production. En dehors de ces critères, les performances du groupe GB étaient moins élevées que celles du groupe GD. Cela réduisait beaucoup les possibilités de comparaisons entre l'efficacité des deux dispositifs de stimulation (binaural vs. dichotique) utilisés pour l'entraînement. C'est pourquoi la discussion porte ici avant tout sur l'effet de l'entraînement perceptif bref et ciblé sur les voyelles /ɪ/ et /i:/, sans pour l'instant prendre en compte la stimulation dichotique ou binaurale.

L'entraînement perceptif a montré une certaine efficacité pour l'apprentissage de la longueur de voyelles anglaises auprès de francophones adultes normo-lecteurs, lorsque son effet est comparé à l'effet test-retest sans entraînement entre les deux sessions (Expérience 10).

Dans la tâche perceptive de vocabulaire, leur exactitude a augmenté après l'entraînement sous la forme d'un effet de grande taille alors qu'il n'y avait pas d'effet test-retest. Leurs réponses se sont accélérées dans cette tâche sous la forme d'un effet de plus grande taille que celui produit par le test-retest. Dans la tâche de perception auditive, leur exactitude a aussi augmenté après l'entraînement, avec des tailles d'effet modérée à grande, alors qu'elle était insensible à l'effet test-retest. Malgré leur brièveté, les entraînements perceptifs ont ainsi amélioré les compétences perceptives, après seulement 1 heure d'entraînement. D'autres chercheurs ont rapporté des bénéfices significatifs pour le même contraste /ɪ/-/i:/ en L2 après un entraînement perceptif légèrement plus long, d'une durée totale de 2h (4 jours, 30 min par jour) (Kondaurova & Francis, 2010). Pour d'autres phonèmes

nouveaux en L2, des entraînements perceptifs brefs se sont aussi avérés efficaces (Cebrian & Carlet, 2014 ; Menning *et al.*, 2002 ; Shum *et al.*, 2021 ; Tamminen *et al.*, 2015 ; Tremblay & Kraus, 2002). Hu *et al.* (2016) ont montré qu'un entraînement bref pouvait produire des améliorations persistantes, pendant 3 mois ; leur entraînement perceptif s'étalait cependant sur 6 jours (1h / jour). Nous n'avons pas vérifié la persistance des effets à long terme dans notre étude, car il n'aurait pas été possible de contrôler les contacts que les participants auraient eus avec l'anglais plusieurs semaines après la fin de l'entraînement. L'objectif était surtout de créer un outil permettant aux enseignants de déclencher un progrès chez les élèves et les étudiants, sur lequel ils puissent s'appuyer sans attendre pour renforcer les acquis, en vue d'assurer leur maintien.

Il n'y avait pas d'effet de transfert de compétences convaincant vers un matériel non-appris, mais un transfert depuis la perception vers la production de mots anglais contenant les voyelles /ɪ/ ou /i:/ s'est produit. En effet, que la diffusion des stimuli soit binaurale ou dichotique, les deux entraînements ont permis aux participants de raccourcir la durée du /ɪ/ vers une durée plus conforme à celle produite par des anglophones, dans la tâche de production sans modèle audio. Cet effet peut être attribué à l'entraînement perceptif, car la durée du /ɪ/ est insensible à l'effet test-retest (Expérience 10). Cela suggère que cet entraînement perceptif favorise le traitement de voyelle courte, car il y avait moins d'effet sur la production de la voyelle longue (hormis pour un groupe (GB) dans une tâche, avec modèle audio). De plus, le raccourcissement n'était pas général mais s'opérait bien sélectivement au /ɪ/. Le /i:/ était quant à lui allongé dans des proportions ne dépassant pas celles observées pour l'effet test-retest.

L'effet particulièrement fort de l'entraînement sur le traitement d'unités courtes fait écho à un résultat analogue obtenu après l'entraînement perceptif par stimulation dichotique pour l'identification des consonnes occlusives sonores anglaises à partir de leur VOT court. L'effet de l'entraînement était moins fort pour les consonnes sonores anglaises et leur long VOT (Expériences 6 et 7, Chapitre 5). Le raccourcissement du /ɪ/ se produisait dans l'Expérience 11 avec des tailles d'effet comparables (modérées) dans les deux groupes GD et GB. Ce raccourcissement était toutefois plus remarquable dans le groupe GD, car celui-ci produisait avant l'entraînement un /ɪ/ plus court que celui du groupe GB. Avec une marge de raccourcissement possible moins grande, le groupe entraîné avec stimulation dichotique a ainsi réalisé une amélioration majeure. L'une des interprétations de l'effet bénéfique de l'entraînement dichotique sur le traitement des indices courts dans l'Expérience 11 était que la

stimulation répétée du cortex auditif gauche par des stimuli contenant un indice court pouvait renforcer la puissance des oscillations rapides (gamma) dans cette aire cérébrale et leurs capacités de synchronisation aux unités les plus fines du signal. Il se peut aussi que la stimulation du cortex auditif gauche par des informations de courte durée (courts VOTs ou voyelles courtes) renforce l'automatisation de la prise en charge de tels indices par l'HG particulièrement compétent pour les traiter.

La différence de durée entre les deux voyelles était déjà présente dans les productions avant l'entraînement : les participants l'avaient peut-être déjà en partie acquise à travers leurs années de pratique de l'anglais, ou bien l'écoute des mots pendant les tâches de perception, toujours réalisées avant, les avait préparés à mieux produire ce contraste.

Les progrès constatés malgré cela après les entraînements perceptifs sont cohérents avec les prédictions du modèle SLM-r sur l'apprentissage tardif d'une L2 (Flege & Bohn, 2021). Selon ce modèle, contrairement au modèle PAM (Best, 1995 ; Best *et al.*, 2001), les mécanismes d'acquisition qui ont permis d'acquérir la L1 restent mobilisables et performants pour l'acquisition de la L2 (Lengeris & Hazan, 2010). C'est aussi ce que prédit le modèle ASP (Strange, 2011) pour qui des routines d'apprentissage restent présentes. Ce dernier modèle fait d'ailleurs des propositions imaginant deux évolutions possibles pour de nouveaux phonèmes en L2. En les appliquant au cas des voyelles étudiées ici, 1) le /ɪ/ et le /i:/ anglais pourraient être simplement assimilés à la même catégorie du /i/ français, considérant le /i:/ comme moins typique de cette catégorie que le /i/ à cause de sa longue durée, ou bien 2) le /ɪ/ pourrait être rapproché du /e/ français car leurs formants sont proches et seul le /i:/ serait assimilé au /i/ français, dont il serait un exemplaire peu typique. L'augmentation substantielle de la distinction entre les deux voyelles anglaises après l'entraînement en perception et production, particulièrement s'il était dichotique, suggère que le deuxième type de modification, évoquant deux catégories phonémiques, s'est opéré. L'analyse des formants des voyelles produites pourrait prochainement permettre de discuter davantage cette hypothèse d'interprétation, selon laquelle les formants F1 et F2 du /ɪ/ devraient se rapprocher de ceux du /e/ après l'entraînement. En attendant, les changements asymétriques opérés sur la production des voyelles /ɪ/ et /i:/ après l'entraînement suggèrent déjà que les deux voyelles n'étaient plus perçues, après l'entraînement, comme des exemplaires d'un seul phonème.

Les progrès réalisés après l'entraînement par les participants pourraient avoir été favorisés par la variabilité des locuteurs dans les entraînements. En effet, un entraînement perceptif HVPT a permis d'améliorer la perception et la production après 5 séances de 45 min

dans une autre étude (Lengeris & Hazan, 2010 ; Lengeris, 2018). Les exercices de notre programme d'entraînement veillaient avec faire varier les contextes consonantiques des voyelles ciblées, et le choix de trois locuteurs au lieu d'un seul dans les Expériences 6, 7 et 8 sur les consonnes, s'appuyait sur le succès des entraînements basés sur la haute variabilité (Bradlow *et al.*, 1997 ; Cebrian & Carlet, 2014 ; Hazan *et al.*, 2005 ; Kangatharan *et al.*, 2021 ; Lively *et al.*, 1993 ; Sadakata & McQueen, 2013). L'entraînement perceptif proposé dans ici visait aussi l'identification de mots, tâche décrite comme plus efficace que la discrimination de sons selon certains travaux (Carlet & Cebrian, 2019).

### **6.7.3. Effet de l'entraînement dichotique chez les adultes normo-lecteurs**

Contrairement à nos hypothèses, les indices d'une supériorité de l'entraînement par stimulation dichotique plutôt que binaurale sont minces : un effet de plus grande taille sur l'exactitude dans la tâche de vocabulaire et un raccourcissement extrême du /ɪ/ en production. Les effets bénéfiques du dispositif dichotique sont plus forts sur l'apprentissage de la différence de voisement des consonnes anglaises (Expériences 6, Chapitre 5). Ce constat décevant peut s'expliquer par la différence de performance avant l'entraînement mais aussi en partie par des caractéristiques temporelles du matériel. En effet, la durée des voyelles courtes contenues dans les mots de l'entraînement était de 52 ms (ET = 19 ms) et celle des voyelles longues de 113 ms (ET = 25). Même si ces moyennes sont en accord avec les fenêtres temporelles respectivement dominantes dans les cortex auditifs gauche et droit selon le modèle AST (Poeppel, 2003 ; Poeppel *et al.*, 2008), les écarts-types sont grands. Les fenêtres d'analyse décrites par Poeppel (2003) sont de 20-50 ms dans l'HG et de 150- 250 ms dans l'HD. Sur le plan de la dynamique neuronale, le découpage d'indices de courte durée impliquerait des oscillations gamma dans l'HG (Fiez *et al.*, 1995 ; Spironeli & Angrilli, 2010) et le découpage d'indice plus longs impliquerait des oscillations thêta dans l'HD (Boemio *et al.*, 2005 ; Poeppel *et al.*, 2008). Ces fenêtres correspondent théoriquement aux durées moyennes des voyelles /ɪ-/i:/, qui seraient de 30-75 ms pour /ɪ/ et de 125- 250 ms pour /i:/. Cependant, certaines voyelles courtes dans les exercices ont une durée qui est celle de la moyenne des voyelles longues (autour de 112 ms). Un /ɪ/ aussi long ne gagne rien à être présenté en priorité à l'HG par l'écoute dichotique. En faisant le choix d'utiliser des stimuli naturels, ce biais s'est produit. Nous aurions pu contrôler la durée des voyelles contenues dans chaque mot utilisé dans l'entraînement mais le faible nombre d'exemplaire de paires minimale s'opposant par la durée du /i:/, rendait cette tâche compliquée comme nous souhaitions avoir le plus grand nombre d'exemplaires possible. La stimulation dichotique n'a

alors sans doute pas permis à certains mots de l'entraînement d'être pris en charge par les fenêtres attentionnelles les plus adaptées, ni de profiter d'une synchronisation aux rythmes cérébraux sous-jacents. Ce biais pourrait être corrigé ultérieurement pour tenter d'optimiser l'effet de la stimulation dichotique en sélectionnant pour les exercices des mots dont les voyelles correspondent mieux aux fenêtres attentionnelles supposées adaptées à /ɪ/ et à /i:/ par le modèle AST.

#### **6.7.4. Effet de l'entraînement dichotique chez les adultes dyslexiques**

Les participants dyslexiques présentaient à T1 des déficiences par rapport aux normo-lecteurs pour le traitement des deux voyelles. Leurs réponses étaient moins exactes que les normo-lecteurs dans la tâche de perception auditive, ils étaient plus lents que les normo-lecteurs dans les deux tâches de perception, et ils produisaient des /i:/ plus courts que les normo-lecteurs dans les deux tâches de production.

Cependant, ces déficits ne les ont pas empêchés de progresser après l'entraînement, proposé pour eux uniquement dans sa version dichotique. L'amélioration de leur exactitude dans les deux tâches de perception, le raccourcissement du /ɪ/ dans les deux tâches de production et l'allongement du /i:/ en production avec modèle audio étaient non seulement significatives, mais survenaient alors que ces mesures sont insensibles à la répétition (pas d'effet test-retest, Expérience 10).

Malgré leurs difficultés initiales en phonologie de l'anglais, les adultes dyslexiques ont ainsi pu progresser de manière équivalente aux adultes normo-lecteurs en perception, où les progrès pour l'exactitude dans la tâche de perception de vocabulaire étaient de même taille que chez les normo-lecteurs. Dans ces deux groupes, un progrès de grande taille était aussi observé pour la rapidité dans cette tâche. Un résultat surprenant est que certains progrès des adultes dyslexiques après l'entraînement dichotique dépassaient ceux des normo-lecteurs. C'était le cas en production, avec un transfert de compétences perception-production particulièrement fort en cas de dyslexie. En production sans modèle audio, l'entraînement était aussi suivi d'un raccourcissement des /ɪ/ sous la forme d'un effet de plus grande taille chez les dyslexiques que chez les normo-lecteurs. De plus, seuls les adultes dyslexiques transféraient cette compétence à des mots non-appris, au-delà du progrès pour les mots appris. En production avec modèle audio, l'entraînement était suivi d'un allongement significatif de la durée du /i:/ chez les adultes dyslexiques, mais pas chez les normo-lecteurs.

Malgré les difficultés de traitements temporels décrites en cas de dyslexie (Habib, 2021 ; Tallal *et al.*, 1996), l'apprentissage d'un contraste de longueur pour deux voyelles courtes/longues en anglais L2 est possible par des adultes dyslexiques. L'entraînement perceptif d'1h a été suivi d'un apprentissage suffisamment profond de cette compétence phonologique pour qu'un transfert se produise en production, et à des mots non étudiés pendant l'entraînement. Ces résultats rejoignent l'efficacité particulière constatée pour un entraînement avec des exercices de perception plutôt que de production observée dans une autre étude sur des voyelles (Krzonowski *et al.*, 2016). Les exercices de perception avaient eu des retentissements positifs en perception et en production de phonèmes anglais chez des adultes francophones sans déficit particulier. Ici, les apprenants étaient dyslexiques, et leurs progrès encouragent à étendre l'utilisation de programmes d'aide à l'apprentissage de l'anglais à ce type de public. Comme ils progressaient autant voire plus que les normo-lecteurs, le dispositif d'entraînement perceptif dichotique testé ici est sans doute une bonne piste dans une démarche d'aide à leur égard. La comparaison des performances avant l'entraînement a mis en évidence de véritables déficits en perception et en production du contraste de durée entre les voyelles anglaises, par rapport aux étudiants normo-lecteurs. Il paraît utile d'encourager les étudiants dyslexiques et les enseignants à utiliser de tels dispositifs pour contribuer à rétablir l'égalité des chances.

Même si l'entraînement par stimulation dichotique montre de bons résultats chez les adultes atteints de dyslexie, il n'est pas encore possible de mesurer si une part du progrès vient de la stimulation dichotique, au-delà de l'entraînement perceptif. Le rôle joué par la stimulation dichotique pourrait être évaluée par la réalisation d'un entraînement avec stimulation binaurale auprès d'un autre groupe d'adultes dyslexiques. Le choix avait été fait de proposer en priorité le dispositif dichotique aux participants dyslexiques, par égard pour leur pathologie, avec le souhait de leur proposer la combinaison potentiellement optimale pour des progrès. Il serait intéressant de poursuivre en comparant chez eux l'effet des stimulations dichotique et binaurale. Pour l'instant, malgré cette limite, il apparaît déjà possible pour un francophone adulte, même atteint de dyslexie, d'apprendre un nouveau contraste phonétique après seulement 1 heure d'entraînement perceptif, de façon suffisamment profonde pour des répercussions en production de parole.

## Chapitre 7. Aide à l'apprentissage de la prosodie anglaise

---

La stimulation dichotique a su montrer ses potentiels apports à l'apprentissage de nouveaux phonèmes anglais auprès de francophones. Cependant, son utilisation est limitée à des petites unités phonémiques, et ne s'applique pas à de plus grandes unités comme les segments prosodiques. Aujourd'hui, de nombreuses études s'intéressent au rôle de la musique et plus particulièrement de l'amorçage rythmique dans la rééducation du langage, et cela notamment pour les traitements grammaticaux qui impliquent justement de repérer de longues unités suprasegmentales (pour une revue, voir Fiveash *et al.*, 2021). Il s'agit ici d'étendre ce principe d'amorçage pour tenter d'attirer l'attention de l'auditeur sur des changements de hauteur de voix survenant sur des sons assez longs : la fin de phrases interrogatives. Cette dernière étude est encore au stade exploratoire. Nous avons souhaité évaluer l'effet d'une amorce rythmique combinée ou non à un retour visuel ayant déjà fait ses preuves dans l'apprentissage et la rééducation du langage (pour ce dispositif précis, Guyot-Talbot *et al.* (2016) ; pour une revue sur le retour visuel, voir Fabre *et al.*, 2016, ou Girod-Roux, 2020). Le domaine d'application est celui de l'aide à l'apprentissage de la prosodie anglaise. Il s'agira plus précisément de l'intonation différente des questions ouvertes et fermées. Cette étude a été menée auprès d'adultes normo-lecteurs et d'adultes dyslexiques.

### 7.1. La prosodie : une nouvelle entrave à l'apprentissage de l'anglais L2

La prosodie est constituée d'unités suprasegmentales, de la taille de mots ou de groupes de mots. Les compétences dans ce domaine se développent tôt. Les suites syllabiques du babillage d'enfants de 8-9 mois ont déjà un contour intonatif propre à leur langue maternelle (De Boysson-Bardies, 2007). Dans le signal de parole, la prosodie permet de véhiculer deux types d'informations : 1) l'état émotionnel du locuteur (prosodie émotionnelle) ; 2) la structure linguistique et le mode d'expression (prosodie intentionnelle). Elles définissent deux domaines de recherche, l'une autour de la prosodie émotionnelle, l'autre autour de la prosodie intentionnelle. Nous nous intéresserons ici à cette dernière.

La notion de rythme est inhérente à la prosodie. Le rythme est défini comme un pattern temporel créé depuis le début d'une séquence et se poursuivant tout le long de l'évènement acoustique (London, 2012 ; Patel, 2008). Dans le langage, ce rythme se manifeste à travers une alternance de moments forts et de moments faibles (i.e. perçus comme moins saillants), mais cette alternance est moins régulière que celle que l'on peut percevoir dans la musique.



Cette différence tient principalement de la nature du langage avec ses caractéristiques lexicales qui ne permettent pas une pulsation rythmique stricte. Une source de difficultés pour le traitement du rythme est la variabilité de ses règles selon les langues, ce qui peut complexifier l'apprentissage tardif d'une L2. L'alternance ne se fait pas à la même échelle selon la langue. Le français est par exemple une langue à rythme syllabique, où toutes les syllabes sont bien prononcées et les alternances concernent la vitesse (e.g. la dernière syllabe des mots est généralement un peu plus forte, car plus longue). La langue anglaise est, quant à elle, une langue accentuelle et fait alterner des syllabes accentuées et non-accentuées en jouant par exemple sur leur intensité. De plus, pour l'anglais, la structure rythmique d'un mot, est présente à deux niveaux hiérarchisés. Par exemple, dans le mot « *Mississippi* », il y a une première structure au niveau syllabique (alternance de quatre syllabes, dont deux fortes), mais aussi une seconde au niveau prosodique avec un « *stress feet* » rendant dans cet exemple la fin du mot plus saillante sur les deux dernières syllabes (Fig. 81). Les indices acoustiques permettant de réaliser la prosodie sont des modifications d'amplitude (perçue comme l'intensité), de durée, et de fréquence fondamentale (F0, perçue comme la hauteur) (Hirst, 2006). La perception du rythme et de sa structure dépend de l'analyse de ces trois dimensions (i.e. amplitude, durée, fréquence) (Allen *et al.*, 2017 ; Besson *et al.*, 2011 ; Patel, 2008).

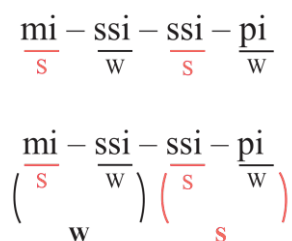


Figure 81. Structure des temps faible (w) et fort (s) du mot « *mississippi* » au niveau syllabique (haut) et prosodique (bas) avec la partie finale du mot accentuée (« *stress feet* » : s) en anglais.

Il existe différents mécanismes d'acquisition de la prosodie en L2 (Li & Post, 2014) et cet apprentissage est différent selon le lien entre la langue maternelle et celle en cours d'apprentissage. En effet, il sera plus facile de segmenter une langue avec une prosodie très différente de la L1 (i.e. français/anglais) qu'une langue trop proche (i.e. français/coréen) (Tremblay *et al.*, 2016), les similitudes faisant oublier les nuances. La réussite de l'apprentissage de la prosodie d'une L2 dépend aussi de l'âge. Ce facteur influe sur la qualité de la prosodie notamment sur la persistance d'un accent étranger (Huang & Jun, 2011). Sur le plan réceptif, la capacité à percevoir si une phrase dite dans une L2 a été produite par un locuteur natif ou par un locuteur ayant une autre L1 est très dépendante à la fois des capacités



de perception des durées et des capacités de perception de la fréquence fondamentale, d'après une étude portant sur des Japonais apprenant le français (Kamiyama, 2004). Interpréter un accent comme étranger repose ainsi sur la perception de caractéristiques rythmiques passant tant par des composantes tant spectrales que temporelles. Cela signifie aussi que l'aptitude à tenir compte de ces différents paramètres du rythme dans ses productions détermine le fait que l'apprenant d'une L2 soit perçu comme un locuteur étranger ou pas.

Jackson et O'Brien (2011) ont montré l'importance de la prosodie dans la compréhension à l'oral et à l'écrit pour une L2. L'exemple des questions, en tant que mode d'expression, résume bien cet aspect. En français, l'intonation est généralement montante à la fin des questions, tandis qu'elle descend à la fin d'un ordre. En anglais, une règle détermine la prosodie de deux types de questions. Pour une question ouverte, à laquelle une réponse développée est attendue, l'intonation finale de la question est descendante. Pour une question fermée, à laquelle une réponse par « oui » ou par « non » est attendue, l'intonation finale est montante. En français, ces variations existent pour les questions, mais sont moins systématiques qu'en anglais, et l'intonation descendante en fin de phrase correspond davantage au mode d'expression impératif : celui des ordres. De ce fait, une méconnaissance de cette particularité de la prosodie anglaise présente le risque de conduire un francophone à mal interpréter certains indices intonatifs en anglais L2. Il pourrait mal comprendre l'intention d'un locuteur anglophone qui pose des questions. Une question motivée par une sincère curiosité peut être par exemple interprétée comme un ordre, une accusation, ou un soupçon déguisé. Cela peut entraîner des malentendus.

Des recherches ont montré l'étendue des difficultés en prosodie pour une L2 chez des apprenants tardifs, plaidant pour la nécessité d'améliorer cet aspect du traitement de la L2, ainsi que pour la faisabilité d'une telle aide. C'est le cas par exemple pour l'apprentissage de l'anglais par des apprenants chinois. Zhang *et al.* (2010) ont réalisé une première étude qui a confirmé les lacunes de ces apprenants dans leurs connaissances et leur perception des indices prosodiques anglais. Les apprenants ont rencontré des difficultés particulières avec les configurations d'accentuation des mots polysyllabiques, l'intonation des questions Wh- (i.e. questions ouvertes), les phrases de continuation, ainsi qu'avec la désambiguïsation prosodique pour l'interprétation sémantique. L'échec des apprenants à utiliser les indices prosodiques pertinents dans une L2, mais pas dans la L1, peut être attribué à un manque de conscience de l'ambiguïté et à une difficulté d'intégration de l'information prosodique, plutôt qu'à une faible sensibilité générale aux indices acoustiques sous-jacents à la prosodie (Zhang & Ding, 2020).

Les auteurs ont conclu à la nécessité d'enrichir l'entraînement à la prononciation en accordant plus d'importance aux connaissances et à la production des traits prosodiques anglais (Zhang *et al.*, 2010). Cela incite à développer des technologies favorisant la compréhension et l'expression à l'oral en prenant plus de soin pour les caractéristiques prosodiques de l'anglais L2. Cette même équipe (Li *et al.*, 2010) a développé un détecteur prosodique capable de repérer les erreurs pour aider à l'apprentissage de l'anglais par des apprenants chinois. Les études de cette équipe de chercheurs sont intéressantes car la difficulté rencontrée par les francophones ressemble à celle des sinophones pour l'intonation des questions et du rythme accentuel en anglais L2. Jusqu'ici les entraînements que nous avons construits étaient uniquement perceptifs. Sur la base des travaux de Li *et al.*, il semble que pour l'apprentissage de la prosodie et des règles impliquant de longues unités, un entraînement à la production orale soit recommandé.

Sur le plan des structures cérébrales sous-tendant le traitement de la prosodie, Gandour *et al.* (2007) ont montré que le traitement de cette dimension de la parole impliquait les lobes frontal, temporal et pariétal. Les activations sont larges et recrutent pratiquement tout le réseau nécessaire au langage. Toutefois, lorsque les participants doivent se focaliser sur l'intonation générale de la phrase, des asymétries hémisphériques subtiles apparaissent. La nécessité de concentrer les ressources d'analyse et sans doute de séquenciation sur de larges unités pour l'analyse prosodique implique ainsi une dominance droite conforme à ce que prédit le modèle AST pour les unités de longue durée. En comparant cette fois les activations cérébrales accompagnant l'écoute en L1 et en L2 chez des bilingues tardifs, ces chercheurs ont montré des modulations similaires dans un système neuronal unitaire malgré l'âge tardif d'acquisition. Là encore, des mécanismes cérébraux comparables apparaissent possibles pour la perception de la prosodie d'une même langue lorsqu'elle est native comme lorsqu'elle est suffisamment bien apprise (cas des bilingues tardifs). Cependant, des ressources neuronales supplémentaires peuvent être nécessaires dans des circonstances spécifiques à la tâche pour des personnes dont le bilinguisme n'est pas équilibré, et a fortiori pour des apprenants tardifs en difficulté comme le sont les participants à nos études. Pour le traitement de la prosodie en L2, des études en EEG ont rapporté des composantes en ERP très proches entre les apprenants de L2 et les locuteurs natifs, même en cas d'apprentissage tardif. Cela suggère qu'il n'est pas nécessaire d'être en immersion totale dans l'environnement de la L2 pour apprendre dans ce domaine (Nickels *et al.*, 2013).

En résumé, la création d'outils utilisables dans le cadre scolaire ou universitaire est souhaitable pour favoriser l'apprentissage de quelques aspects de la prosodie anglaise, puisque celui-ci reste imparfait. Celui-ci ne semble pas nécessiter de changement radical sur le plan des aires et de la dynamique cérébrale impliquées. Le dispositif que nous proposons combinera pour cela le retour visuel sur l'information acoustique issue d'une répétition et un amorçage rythmique.

## **7.2. Musique, oscillations cérébrales et prosodie**

### **7.2.1. Synchronisation des oscillations cérébrales aux rythmes de la musique et de la parole**

Des analogies entre les repères rythmiques de la musique et de la prosodie dans la parole (temps et syllabes ; pulsation et syllabes accentuées) (Hausen *et al.*, 2013), de même que l'organisation hiérarchique du rythme musical et de la prosodie (Lerdahl & Jackendoff, 1983), fondent des liens entre la perception de la musique et du langage. Par ailleurs, sur le plan du traitement de l'information, l'organisation hiérarchique de la distribution temporelle de l'attention et de la dynamique oscillatoire qui la sous-tend offrent des structures dynamiques cognitives et neuronales adaptées aux caractéristiques rythmiques et hiérarchiques de la musique et du langage.

Les oscillations cérébrales joueraient un rôle dans le traitement de la musique et du signal de parole (Jones, 2019). Elles pisteraient les rythmes dans l'information acoustique qui sous-tend la perception de la musique (Fujioka *et al.*, 2012 ; Nozaradan *et al.*, 2011, 2012, 2015), du langage (Giraud & Poeppel, 2012 ; Kösem & Wassenhove, 2017). Leur fonctionnement serait similaire dans les deux domaines (Harding *et al.*, 2019). Par leurs liens avec l'attention temporelle (Jones, 2019), les oscillations permettraient la séquenciation, les prédictions (Arnal & Giraud, 2012), la synchronisation avec l'environnement (Calderone *et al.*, 2014) et sa structure hiérarchique (Jones, 2016 ; Poeppel & Assaneo, 2020). Les oscillations sont ainsi autant de ressources et de mécanismes indépendants des domaines, utiles pour traiter la parole comme la musique (Tillmann, 2012).

Les oscillations cérébrales et leur hiérarchie sont aussi responsables de la communication entre certaines régions perceptives élémentaires (le cortex auditif primaire) et la voie dorsale, notamment le cortex auditif et le cortex moteur (Assaneo & Poeppel, 2018). Un couplage sensori-moteur se crée ainsi, important pour traiter la musique mais aussi le

langage. En effet, la perception du signal de parole active elle aussi les aires motrices par la boucle de perception-production (Wilson *et al.*, 2004). Nous avons vu que cela permet l'exercice d'influences top-down dans la dynamique oscillatoire responsable de l'analyse du signal. Ces couplages sensori-moteurs et les rythmes cérébraux endogènes qui proviennent d'aires cérébrales situées dans la voie dorsale s'ajoutent aux influences bottom-up liées à la synchronisation des oscillations dans les aires auditives primaires aux rythmes externes. L'ensemble contribue à une perception auditive précise du temps (Morillon & Baillet, 2017 ; Peelle & Davis, 2012 ; Van Wijk *et al.*, 2012 ; Yang *et al.*, 2018), pour la musique et le langage.

Dans le cas plus précis de la prosodie dans le langage, les unités porteuses d'information sont les enveloppes de la parole (Ghitza, 2011 ; Leong *et al.*, 2014). L'analyse cognitive de ces enveloppes implique la synchronisation d'oscillations cérébrales lentes. Cet ajustement neuronal au rythme de la parole nécessite de disposer de rythmes endogènes variés (Buzsáki & Draguhn, 2004) et couplés hiérarchiquement pour le traitement et l'intégration d'informations à différentes échelles (Jones, 2016). Parmi les rythmes endogènes lents, les oscillations *delta* permettraient la perception de la prosodie et son découpage (Calderone *et al.*, 2014 ; Keitel *et al.*, 2017). Elles agiraient sur des rythmes cérébraux particulièrement dans l'HD (Kershner, 2019). Elles s'appuieraient sur la récurrence quasi-régulière de modulations saillantes dans l'enveloppe de parole (l'équivalent des temps forts en musique), événements sonores prédominants attirant l'attention dans la parole (Cutler & Foss, 1977 ; Gow & Gordon, 1993 ; Pitt & Samuel, 1990). Elles prédiraient aussi leur survenue. Le traitement des modulations lentes dans la prosodie bénéficie des couplages oscillatoires entre les bandes delta (1-3 Hz) et thêta (4-8 Hz), qui agissent sur les phases et l'amplitude des ondes bêta (13-30 Hz) et gamma (>30 Hz) selon un processus top-down. Celui-ci permet à la dynamique oscillatoire d'être cohérente avec le traitement d'événements brefs mais eux aussi déterminants pour percevoir la structure rythmique de la phrase : certaines syllabes, certains phonèmes, et surtout les attaques d'enveloppe d'amplitude, parfois courtes. En bref, le traitement du rythme musical et de la prosodie passerait directement par l'implication des oscillations lentes delta-thêta qui offrent de larges fenêtres d'analyse pour les unités déterminantes à ce niveau d'analyse. Il passerait aussi par le biais de l'influence descendante exercée par ces oscillations lentes sur des oscillations suffisamment rapides dans des aires élémentaires du cortex auditif pour détecter le début (ou des détails importants) des enveloppes de parole.

Les modulations temporelles lentes de la parole et de la musique montrent des pics larges, mais bien séparés autour de 5 et 2 Hz. Ces échelles de temps acoustiquement dominantes seraient des caractéristiques intrinsèques de la parole et de la musique (Ding *et al.*, 2017). Ce point commun a généré l'hypothèse selon laquelle une stimulation musicale avec un rythme à 2 Hz pourrait pré-ajuster les oscillations delta, par ailleurs bien adaptées à la durée des enveloppes d'amplitude en prosodie. Cela suppose un transfert musique-parole. Stimuler les oscillations lentes engagerait de plus la cascade de couplages et les relations de nichage entre les oscillations delta, thêta, bêta, jusqu'aux oscillations les plus rapides (gamma). C'est aussi pourquoi il semble particulièrement judicieux d'agir sur les rythmes lents. Il s'agit de permettre, indirectement, de mieux repérer des indices acoustiques courts, tels que l'attaque des enveloppes d'amplitude, utiles pour l'analyse prosodique.

Dans le cas des questions en anglais pouvant être proposées à des francophones lors de l'apprentissage en L2, l'intonation des questions peut correspondre à une modulation de la hauteur (F0) en fin de phrase, sur une unité longue compatible avec la fenêtre attentionnelle offerte par le rythme delta. Cependant, cette modulation peut aussi concerner seulement la dernière syllabe de la question, alors compatible avec la fenêtre du rythme thêta, déterminante pour les syllabes (Gross *et al.*, 2013). Un traitement précis du changement de hauteur permettrait de discriminer le type de question – ouverte ou fermée – et de répondre de manière cohérente par rapport à l'attente de l'interlocuteur.

L'hypothèse qui guide notre démarche est que, avant le traitement prosodique du signal de parole, la présentation d'une séquence rythmique musicale avec une pulsation à 2 Hz pourrait pré-activer et/ou pré-ajuster des oscillations lentes. Pré-stimuler ainsi les oscillations delta ou thêta orienterait particulièrement l'attention dans le temps pour extraire de larges unités utiles à l'analyse prosodique. Cela permettrait aussi de réguler en amont les cycles d'attention temporelle nécessaires au traitement du langage à plusieurs échelles (Hickok *et al.*, 2015). Comme les enveloppes prosodiques sont souvent cohérentes avec des unités grammaticales (Cumming *et al.*, 2015 ; Gordon *et al.*, 2015), des essais de remédiation ont été réalisés en vue d'améliorer le traitement syntaxique ou morphosyntaxique (i.e. grammatical) de phrases à partir d'amorces musicales avec une pulsation à 2 Hz. Un bénéfice a été apporté en grammaire par de telles amorces rythmique dans plusieurs langues, tant chez des adultes que des enfants, atteints de dyslexie, de TDL ou sans pathologie (Bedoin *et al.*, 2016b ; Canette *et al.*, 2019, 2020a, b ; Chern *et al.*, 2018 ; Fiveash *et al.*, 2020a,b ; Lananyi *et al.*, 2021 ; Przybylski *et al.*, 2013). Cet ensemble de résultats encourage à utiliser ce type

d'amorces pour améliorer, plus directement, l'analyse d'enveloppes de parole et ainsi l'analyse prosodique.

### **7.2.2. Régularité du rythme musical et aides aux traitements langagiers**

Les humains se synchronisent, s'ajustent, aux structures régulières. La musique ayant un rythme plus régulier et prédictible que celui de la parole, elle est un moyen idéal pour stimuler la synchronisation neuronale. La musique existe dans toutes les cultures, peut être apprise dès le plus jeune âge (Trehub, 2003) et pour certains de ses aspects possiblement in utero (Hepper, 1991). Durant l'enfance, au niveau cérébral, un pattern similaire d'activations se produit pour la mélodie et le rythme. Toutefois, lorsque l'on observe plus particulièrement le gyrus temporal supérieur, on observe dans l'HD une plus grande activation pour les traitements mélodiques que pour les traitements rythmiques (Overy *et al.*, 2004 ; Zatorre *et al.*, 1994). Cela renforce la théorie selon laquelle le cortex auditif droit serait spécialisé dans la résolution spectrale pour des stimuli non langagiers (Zatorre *et al.*, 2002). A l'âge adulte, une prédominance droite pour le traitement de la mélodie s'accroît (Zatorre, 2001), parallèlement à une spécialisation hémisphérique gauche au niveau du gyrus temporal supérieur pour le traitement du rythme (Samson *et al.*, 2001).

Comme il est facile de percevoir le rythme musical, particulièrement régulier, une hypothèse est que l'écoute d'un tel rythme pourrait stimuler la synchronisation des oscillations neuronales dans des bandes de fréquences propices au traitement des enveloppes dans le signal de parole. D'une manière plus générale, l'hypothèse est celle d'un transfert de compétences rythmiques au traitement de la parole (Kraus & Chandrasekaran, 2010). Cette hypothèse a incité des chercheurs à recommander la mise en place et l'évaluation des effets de prises en charge par la musique insistant sur le rythme pour améliorer le traitement du langage (Flaugnacco *et al.*, 2015 ; Goswami, 2012 ; Overy, 2008). Même de courtes amorces rythmiques conformes à la structure syllabique du mot ou de la phrase qui suit, améliorent la détection de phonèmes, la répétition de phrases (Cason *et al.*, 2015 ; Cason & Shön, 2012) et modifient les réponses neuronales (Falk *et al.*, 2017). Un lien prédictif a même été montré entre la capacité de traitement rythmique et la qualité de la prosodie en L2, y compris pour une langue peu familière (Cason *et al.*, 2019).

Le principe d'expérience peut aussi être légèrement différent. Il peut ne pas être basé sur une adéquation terme à terme entre la structure rythmique de l'amorce musicale et la structure rythmique du mot ou de la phrase présentée ensuite. Un principe d'amorçage plus

global consiste en effet à présenter une amorce rythmique musicale d'une trentaine de secondes. Son rythme est adapté à l'analyse qui sera demandée dans une série de phrases présentées ensuite. Par exemple, une amorce avec un rythme très régulier à 2 Hz est supposée stimuler un rythme de diffusion de l'attention à un rythme adapté à l'analyse grammaticale des phrases. Le mécanisme neuronal sous-jacent supposé est un renforcement et/ou un pré-ajustement des oscillations lentes. Elles pourraient alors mieux guider l'attention temporelle pour une séquenciation du signal de parole en grandes unités (enveloppes de parole) et des anticipations adaptées à l'analyse grammaticale. Ce principe d'amorçage rythmique musical s'est avéré efficace chez des enfants avec ou sans pathologies du langage, aussi bien en anglais (Chern *et al.*, 2018), qu'en français (Bedoin *et al.*, 2016b ; Canette *et al.*, 2020b, 2020c, Fiveash *et al.*, 2020a, Przybylski *et al.*, 2013) ou en hongrois (Ladanyi *et al.*, 2020). L'efficacité des amorces a été vérifiée par comparaison à différentes conditions contrôles, comme des bruits de l'environnement non rythmés (Bedoin *et al.*, 2016b), des textures musicales sans rythme (Canette *et al.*, 2020b), ou le silence (Ladanyi *et al.*, 2020). Intercalées entre des exercices de grammaire, de telles amorces musicales avec une pulsation à 2 Hz amplifient aussi l'effet positif de ces exercices pour l'apprentissage implicite de règles grammaticales chez des enfants sourds porteurs d'un implant cochléaire (Bedoin *et al.*, 2016a, 2018). Ces amorces permettent enfin d'accroître l'amplitude de l'onde P600, corrélat électrophysiologique de la sensibilité aux erreurs de grammaire, chez des adultes dyslexiques ou non-dyslexiques (Canette *et al.*, 2020b). Ce principe d'amorçage est celui qui a été retenu pour la dernière étude de la thèse, en combinaison avec un retour visuel sur une information acoustique, en vue d'aider des apprenants francophones à mieux traiter la différence prosodique entre les questions ouvertes et fermées en anglais.

L'utilisation de l'amorçage musical pour tenter d'améliorer le traitement du langage repose sur l'hypothèse d'un partage de mécanismes cognitifs par le traitement de la musique et le traitement du langage. Le modèle *Processing Rhythm In Speech and Music* (PRISM, Fig. 82) de Fiveash *et al.* (2021) met en lumière les trois piliers du traitement du langage et de la musique : un traitement auditif précis, une synchronisation des oscillations neuronales aux stimuli externes et un couplage sensorimoteur. Ce modèle, issu de plusieurs domaines de recherche, regroupe les principes de base de différentes théories : OPERA (Patel, 2011, 2012), ASAP (Patel & Iversen, 2014), Active Sensing (Morillon *et al.*, 2015 ; Schroeder *et al.*, 2010), Dynamic Attending Theory (Jones, 1976, 2016, 2019; Large & Jones, 1999) et Predictive Coding (Friston, 2005, 2010). Son originalité est de mettre en avant une relation



tripartite et de permettre son application directe au traitement du rythme langagier et musical. Ce modèle met en relation ces trois mécanismes, et précise le rôle de chacun dans le domaine temporel. Ainsi, la précision du traitement auditif est cruciale pour la discrimination de courts indices temporels et une perception auditive exacte. La synchronisation des oscillations neuronales aux stimuli externes permet la prédiction des éléments à venir et le suivi de la structure hiérarchique à plusieurs niveaux. Le couplage sensorimoteur permet une connexion étroite entre la perception et la production dans le cerveau, ainsi que l'implication du système moteur engagé dans le réseau cérébral des traitements temporels. Tout cela contribue à la qualité des synchronisations et des prédictions pour traiter la musique ou la parole. L'innovation de ce modèle est la mise en relation possible des trois dimensions. Chaque mécanisme pourrait en effet s'associer aux autres, avec l'idée d'une relation potentiellement tripartite. Ce modèle offre un cadre intéressant pour notre étude, car il permet, par une réflexion sur les liens entre traitements cognitifs et neuronaux, d'envisager des pistes de rééducation et d'aide à l'apprentissage. Il s'agit pour cela de coupler musique et langage. Ce modèle servira ici d'appui pour notre hypothèse au sujet d'une aide potentielle apportée par la musique à l'apprentissage d'un aspect de la prosodie anglaise chez des adultes francophones.

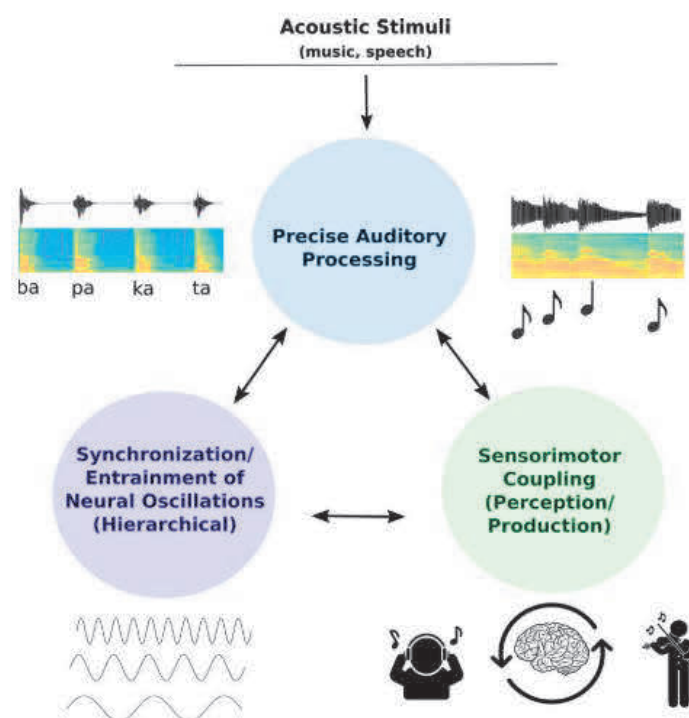


Figure 82. Modèle PRISM. Reproduit à partir de "Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders.", par Fiveash, A. *et al.*, 2021, *Neuropsychology*, 35(8), 771-791.



### 7.3. Dyslexie, prosodie et grammaire

Une faible sensibilité à la perception de la prosodie – rythme de la parole – est corrélée à de faibles capacités en phonologie et en lecture (Wade-Wooley & Wood, 2006). En cas de dyslexie, un trouble du rythme s’observe en production de parole, par exemple avec un débit plus faible de production de syllabes par seconde (De Bree *et al.*, 2006 ; Wolff, 2002), ce qui altère la prosodie. Avec son modèle TSF, Goswami (2011) propose qu’un déficit de perception de la prosodie en cas de dyslexie soit en grande partie à l’origine de la dyslexie. Le déficit fondamental dans la dyslexie serait, selon ce cadre théorique, une déficience du réseau oscillatoire delta qui se synchroniserait très mal aux enveloppes temporelles d’unités prosodiques (Van Hirtum *et al.*, 2019). L’ajustement des oscillations cérébrales à des enchaînements de sons non verbaux diffusés à une fréquence compatible avec le rythme delta est en effet déficitaire dans la dyslexie (Soltész *et al.*, 2013). Il s’agirait fondamentalement d’un déficit non-spécifiquement verbal : le traitement de la récurrence des temps forts chez des enfants dyslexiques, même en-dehors d’un matériel verbal, est un prédicteur longitudinal de la conscience prosodique (Goswami *et al.*, 2013). Chez les adultes dyslexiques le traitement du rythme reste imparfait, avec par exemple des déficits d’anticipation temporelle (Pagliarini *et al.*, 2020) : ils ont des difficultés persistantes à s’appuyer sur les régularités temporelles pour effectuer des prédictions sur les prochains événements. Les difficultés dans le traitement du langage à l’origine de la dyslexie seraient ainsi au moins partiellement liées à un défaut d’ajustement des rythmes cérébraux aux enveloppes d’amplitude (Falk *et al.*, 2015 ; Falter & Noreika, 2014 ; Peter & Stoel-Gammon, 2008). Cet échec entraînerait une cascade de difficultés pour l’extraction, l’analyse et l’apprentissage d’unités de parole plus fines comme les syllabes et surtout les phonèmes.

Il existe des arguments objectifs au sujet de difficultés prosodiques chez les personnes dyslexiques. Des recherches montrent que la prosodie est un domaine mal maîtrisé dans cette pathologie. Dans une étude évaluant la lecture à haute voix, Mendonça Alves *et al.* (2014) ont par exemple montré plusieurs troubles au niveau prosodique chez des enfants dyslexiques parlant le portugais du Brésil. Leur vitesse de lecture et d’articulation étaient réduites. Ils présentaient aussi une altération du nombre et de la durée des pauses, ainsi que des variations atypiques de la fréquence fondamentale. Il leur était également difficile de produire des schémas d’accentuation attendus dans leur langue et de marquer les syllabes accentuées. D’après une autre étude, les enfants dyslexiques obtiennent des scores inférieurs à ceux des normo-lecteurs en perception phonologique et prosodique, mais pas dans les tâches

perceptives de perception d'augmentation d'amplitude (Cuetos *et al.*, 2017). Les auteurs ont conclu que les enfants dyslexiques hispanophones, présentent une déficience prosodique qui peut avoir pour origine de subtils déficits de traitement en audition centrale (i.e. non périphérique). De plus, ils présentent un déficit fondamental de la phonologie supra-segmentale, au niveau des mots et des phrases. Ces difficultés relèveraient de traitements auditifs et temporels inadaptés puisqu'ils se manifestent aussi lorsqu'un rythme équivalent à la prosodie linguistique est proposé dans un matériel non verbal (Calet *et al.*, 2019). Cependant, même si l'interaction entre la prosodie et d'autres composantes du langage telles que la syntaxe et la pragmatique est problématique pour les enfants atteints de Troubles Du Langage (TDL) et/ou de dyslexie, la prosodie elle-même n'est pas au premier plan du tableau clinique des enfants dyslexiques (Marshall *et al.*, 2009). C'est pourquoi ses dysfonctionnements ont été peu étudiés dans la dyslexie.

Le TDL est un autre cas de trouble neurodéveloppemental où les patients ont des difficultés à synchroniser leurs gestes au rythme d'un métronome (Corriveau & Goswami, 2009) ou d'une musique (Cumming *et al.*, 2015), ainsi que des difficultés en prosodie et en traitement de parole rapide (Guiraud *et al.*, 2018). Ces troubles sont encore plus évidents que chez les personnes dyslexiques. Or, ces enfants ont justement des déficits majeurs en grammaire (syntaxe et morphosyntaxe). Cela n'est sans doute pas étranger à des difficultés à traiter temporellement les enveloppes d'amplitude qui sous-tendent les traitements prosodiques, et qui empêcheraient d'organiser les groupes de mots et de réaliser des prédictions (Corriveau *et al.*, 2007). Leurs difficultés à traiter les attaques d'enveloppe d'amplitude (*rise times*) ainsi que les enveloppes elles-mêmes rendent compte d'une part importante de la variance pour leur échec dans des épreuves de langage, notamment phonologiques (Fraser *et al.*, 2010). Leurs performances sont faibles pour discriminer des attaques d'amplitude ou des durées, et elles prédisent leurs difficultés à percevoir l'accentuation dans les mots (Richards & Goswami, 2015). Cet ensemble de données souligne encore le lien entre rythme, prosodie et grammaire. Pour ce qui est de l'apprentissage d'une L2, les adultes TDL rencontrent aussi des difficultés importantes lorsqu'il faut apprendre dans une nouvelle langue des règles prosodiques impliquant des principes hiérarchiques ou même de nouvelles règles d'accentuation simples (Bahl *et al.*, 2009 ; Plante *et al.*, 2010).

Alors que l'état de l'art de McCann et Peppé (2003) montrait un manque de recherche sur la prosodie en cas d'autisme, de nombreuses études ont été réalisées à ce sujet depuis. Hesling *et al.* (2010) ont mis en évidence les difficultés prosodiques de ces patients. Des

adolescents relevant du spectre autistique sont par exemple en difficulté lorsqu'il est nécessaire d'utiliser la prosodie afin de désambiguïser la syntaxe d'une phrase (Diehl *et al.*, 2008). Cette difficulté va là encore de pair avec un déficit de traitement temporel (pour une revue, Jurek *et al.*, 2019). Cependant, en cas d'autisme le niveau de langage est un bon prédicteur pour les capacités perceptives de la prosodie, et cette dernière n'est pas systématiquement atteinte (Lyons *et al.*, 2014). Pour finir avec une dernière pathologie : chez les enfants sourds implantés cochléaires, Torppa *et al.* (2020) ont montré des troubles de la prosodie. Cette étude montre que si ces enfants sont musiciens, leurs déficits dans ce domaine du langage sont atténués. Le rythme de la musique et celui de la parole sont ainsi des domaines pour lesquels certains mécanismes cognitifs communs sont mobilisés.

#### **7.4. Entraînements rythmiques et utilisation d'un retour visuel comme aides au langage**

##### **7.4.1. Effets de ces entraînements sur les traitements dans la langue maternelle**

Récemment, quelques études scientifiques se sont concentrées sur l'utilisation du rythme dans des séances d'entraînement musical pour améliorer les capacités langagières chez l'enfant sain ou atteint d'un trouble du langage, mais peu s'intéressent à l'adulte dyslexique.

Chez des enfants sans pathologie d'abord, un entraînement au rythme peut avoir des répercussions dès le plus jeune âge (9 mois). Dès lors, et après 12 séances d'entraînements, les enfants ont produit une MMN à la suite d'une déviance de la structure temporelle de la musique mais aussi du langage, signe de l'établissement d'une sensibilité pré-attentionnelle et automatisée à la structure rythmique (Zhao & Kuhl, 2016). Après 1 ou 2 ans d'entraînement musical, des enfants de 8 ans segmentent mieux la parole (François *et al.*, 2013). L'amplitude de leur MMN augmente aussi à la suite d'une déviance au niveau syllabique et phonémique (pour des voyelles) (Chobert *et al.*, 2014). Un entraînement musical a aussi modifié les réactions cérébrales à la prosodie inappropriée de phrases chez des enfants de 8 ans (Magne *et al.*, 2006). Notons que la sensibilisation à la musique peut aussi être utile au langage à partir d'autres dimensions communes que le rythme. Une étude a par exemple montré chez des enfants de 8 ans une amélioration de la lecture et de la discrimination de la hauteur dans un signal de parole après 6 mois d'entraînement musical (Moreno *et al.*, 2009). De même, la détection de petits changements de hauteur s'est améliorée à la suite de 6 mois d'entraînement musical basé sur la différence de hauteur en musique (Besson *et al.*, 2007).

Chez des enfants atteints de dyslexie, une revue de question a montré que la formation musicale peut servir d'outil de remédiation pour améliorer les capacités de lecture et d'écriture, bien que le type d'intervention spécifiquement efficace doive faire l'objet de recherches plus approfondies (Rolka & Silverman, 2015). Par exemple, un entraînement musical de 7 mois basé sur le rythme s'est avéré bénéfique pour améliorer la conscience phonologique et la lecture (Flaunacco *et al.*, 2015). Plusieurs études ont utilisé le programme d'entraînement Rythmique Reading Training (RTT) qui propose des exercices de lecture combinée à du rythme (Bonacina *et al.*, 2015). Le but de l'entraînement est de lire des syllabes, mots ou phrases au rythme de la musique jouée en fond sonore. Au cours de l'entraînement, le rythme de la musique s'accélère, augmentant ainsi la difficulté de la tâche. Utilisé chez des collégiens dyslexiques pendant 9 séances individuelles bihebdomadaires de 30 minutes, le programme a été suivi d'une amélioration de la lecture, tant en rapidité qu'en exactitude. Cancer *et al.* (2020) ont comparé l'effet du RTT à deux autres entraînements : la stimulation visuelle hémisphérique-spécifique de Bakker et des jeux vidéo d'action. Chaque groupe réalisait 13h d'entraînement sur 9 jours. Les trois protocoles ont permis d'améliorer les performances des enfants dyslexiques, mais avec des effets différents. Le RTT a augmenté la rapidité de lecture des pseudo-mots, tandis que les deux autres dispositifs ont accru l'exactitude en lecture. Cette étude a mis en lumière le fait que les trois entraînements exercent et améliorent des mécanismes cognitifs différents.

La musique peut aussi améliorer le traitement d'indices phonétiques de durée. Frey *et al.* (2019) ont ainsi montré une normalisation de la sensibilité pré-attentionnelle à des changements de VOT chez des enfants dyslexiques après un entraînement musical de 6 mois, et pas après un entraînement de même durée en peinture. Bhide *et al.* (2013) ont comparé les bénéfices apportés par un autre entraînement basé sur le rythme et un programme informatisé exerçant la traduction graphèmes-phonèmes chez des enfants dyslexiques. Ils ont observé des améliorations des capacités phonologiques et de lecture de même intensité à la suite des deux entraînements. Exercer les compétences rythmiques peut ainsi autant apporter à la lecture qu'un entraînement ciblé sur une compétence plus directement liée à la lecture (conversion graphème-phonème). De même Thomson *et al.* (2013) ont montré que 6 semaines d'entraînement rythmique amélioreraient autant la conscience phonologique que 6 semaines d'exercices phonologiques informatisés, ce qui les a conduits à conseiller de combiner les deux. Les progrès observés par Flaunacco *et al.* (2015) en conscience phonologique et en lecture après 30 semaines d'entraînement musical (1h deux fois par semaine), et non après un

entraînement tout aussi artistique mais en peinture, sont également convaincants. Enfin, l'étude présentée par Habib *et al.* (2016) et réalisée auprès de 24 enfants dyslexiques montre elle aussi les effets positifs d'une prise en charge de troubles du langage par la musique. Elle précise aussi les différences d'effets d'un même entraînement selon sa répartition dans le temps. Ainsi, un entraînement musical de 18h réparti sur 3 jours permet d'améliorer la perception catégorielle et la perception auditive d'unités temporelles du signal de parole. Le même entraînement réparti sur 6 semaines apporte les mêmes bénéfices avec, en plus, une amélioration de l'attention auditive, de la conscience phonologique, des capacités de lecture, de la répétition de pseudo-mots, et un maintien des bénéfices à plus long terme que si l'entraînement est sur 3 jours.

L'utilisation de la stimulation rythmique a aussi été étudiée dans d'autres pathologies du langage comme, par exemple, le bégaiement. Des données montrent qu'une stimulation auditive externe peut diminuer la sévérité du déficit langagier, notamment via un mécanisme de compensation du déficit temporel par un indexage rythmique externe (Toyomura *et al.*, 2011). L'utilisation du chant aide aussi à améliorer la fluence de la parole, en particulier en régulant la structure temporelle des mots en cas de bégaiement mais aussi dans le cadre de troubles du spectre autistique, de la maladie de parkinson ou encore en cas de lésions cérébrales (Wan *et al.*, 2010).

L'utilisation d'un retour visuel sur l'expression orale a aussi fait ses preuves chez des enfants présentant des troubles du langage. Des bénéfices du retour visuel lingual lors d'entraînements à la production de parole sont par exemple décrits pour la prononciation de consonnes chez des enfants sourds (Gallagher, 2013) ou atteints de troubles articulatoires (Cleland, *et al.*, 2015). L'image articulatoire est parfois reconstruite à partir d'une large base de données (e.g. logiciel UltraSpeech-player) (Hueber, 2013) et apparaît sur une représentation de tête (e.g. tête parlante articulatoire) (Badin, *et al.*, 2010). Même son utilisation en simple illustration visuelle (i.e. non personnalisée) améliore la production et l'identification de /t/ et du /k/ chez des enfants qui les confondent (Fabre *et al.*, 2016). Dans ces exemples, les retours visuels portent sur des aspects articulatoires et non acoustiques.

#### **7.4.2. Effets de ces entraînements sur les traitements dans une L2**

Même si, un nombre croissant d'études ont montré les bienfaits de la musique sur différents aspects du langage, peu concernent son impact sur l'apprentissage de la L2 et sa prosodie. Ramirez Verdugo (2006) a utilisé un entraînement en production de parole sous la forme d'un

enseignement explicite avec retour visuel (comparaison de courbes d'intonation d'un anglophone par rapport à celle du participant). Ciblé sur la prosodie et les particularités du système d'intonation anglais (10 sessions de 50 min réparties sur 10 semaines), ce programme a permis d'améliorer la sensibilité des participants à l'intonation. Les productions gagnaient aussi en qualité sur le plan prosodique dans ce groupe d'Espagnols apprenant l'anglais L2.

Le retour visuel en cours d'apprentissage a été utilisé dans quelques recherches sur l'apprentissage de la prosodie. Hardison (2005) a comparé l'effet de deux entraînements à la prosodie chez des Chinois ayant un bon niveau en anglais L2. Il a utilisé *ANVIL*, un outil d'annotation basé sur le web. Cet outil intègre la vidéo d'un discours avec des affichages visuels du contour des variations de la hauteur de la voix de ce modèle. Il ajoute un retour immédiat sur les variations de la hauteur de la voix produite par le participant qui répète. Un groupe s'entraînait avec des discours complets et l'autre avec des phrases isolées. Des groupes contrôles s'entraînaient aussi sur l'un ou l'autre de ces deux matériels, mais sans retour visuel. Les deux groupes entraînés avec *ANVIL* se sont améliorés par rapport aux groupes contrôles. Cependant, le groupe entraîné avec le discours complet (i.e. un ensemble cohérent) a mieux appris puisque ses progrès se sont davantage transférés vers un autre discours. Les effets positifs se sont maintenus une semaine après les entraînements.

Enfin, Guyot-Talbot *et al.* (2016) ont proposé à des francophones apprenant l'anglais un entraînement à la prosodie en production de parole, pendant seulement 40 min. L'entraînement portait sur les intonations des questions anglaises ouvertes et fermées. Les participants recevaient un retour visuel direct sur le contour intonatif du modèle superposé à celui de leur propre production, ce qui permettait une comparaison aisée des tracés par l'apprenant. Les chercheurs ont montré une amélioration marginale mais proche du seuil de significativité des performances des participants pour les questions dont l'intonation était descendante, particularité intonative moins présente dans les questions en français. Ce type de dispositif est encourageant pour l'aide à l'apprentissage de particularités prosodiques dans une langue apprise tardivement. Il présente l'avantage d'être bref et intégrable dans des séances classiques de cours. Les apprenants intégrés dans l'étude de Guyot-Talbot *et al.* (2016) avaient au départ un niveau d'anglais correct (étudiants dans un département universitaire d'anglais). On peut se demander si ce type d'outil est efficace pour des apprenants moins avancés. La combinaison entre ce retour visuel sur des aspects acoustiques des productions et un amorçage rythmique pourrait être envisagée. C'est l'objet des deux dernières expériences de cette thèse.

## 7.5. Objectifs

La prosodie est un nouvel exemple de difficulté rencontrée par les francophones lors de l'apprentissage de l'anglais en tant que seconde langue. La différence de rythmicité entre les deux langues peut être source de malentendus dans la communication. L'exemple des questions illustre bien cette problématique. Alors qu'en français une intonation montante à la fin d'une phrase accompagne le plus souvent le questionnement, en anglais l'intonation finale peut être montante comme descendante pour le même objectif d'interrogation, avec simplement une nuance sur le contenu attendu : respectivement, une réponse oui/non ou une réponse développée. La descente de l'intonation finale en français est plutôt associée à un ordre, ou à une question avec un sous-entendu, et si l'interprétation des questions anglaises est directement assimilée au français, une méprise est possible. Etant donné les quelques anomalies relatées pour la prosodie chez les personnes dyslexiques, les dispositifs d'aide à l'apprentissage dans ce domaine doivent aussi être conçus pour eux, tout en restant utiles aux apprenants sans pathologie.

La musique semble être un bon stimulateur pour l'aide au traitement du langage, notamment chez les personnes dyslexiques, et il est ici proposé de l'utiliser pour favoriser l'apprentissage d'une seconde langue. En effet, le cerveau se synchronisant à l'environnement, la perception d'une musique dotée d'un rythme régulier avec une pulsation à une fréquence susceptible d'accentuer un rythme cérébral globalement adapté à la durée des unités à extraire pour la tâche peut être utile.

L'hypothèse est qu'une séquence musicale de ce type peut constituer une amorce rythmique bénéfique, susceptible d'orienter et ajuster l'attention temporelle d'une manière adaptée à la taille des unités qui, dans le signal de parole, portent les indications pertinentes pour l'analyse d'indices prosodiques. Dans cette thèse, l'hypothèse est formulée en ces termes, qui relèvent de la psychologie cognitive, car les oscillations cérébrales ne sont pas analysées. Sur le plan de la psychologie cognitive, l'hypothèse au sujet d'un bénéfice apporté par l'amorce régulière à l'efficacité d'un entraînement basé sur la répétition des deux types de questions anglaises est fondée sur les notions d'attention dynamique (Jones, 1976), d'ajustement de l'attention aux régularités temporelles et d'anticipation temporelle sur la survenue d'événements sonores (Barnes & Jones, 2000). Comme dans les expériences de Jones *et al.* (2002), un préajustement rythmique de l'attention basé sur la perception du contexte (ici, l'amorce rythmique) pourrait peut-être permettre un traitement plus précis de



diverses caractéristiques de l'unité d'information devenue prioritaire, y compris ses caractéristiques fréquentielles.

Le choix d'un amorçage rythmique à 2 Hz s'appuie sur les résultats d'un ensemble d'expériences montrant l'influence positive de telles amorces, chez des personnes dyslexiques ou non, sur le traitement grammatical de phrases (Bedoin *et al.*, 2016b ; Canette *et al.*, 2020b, 2020c, Chern *et al.*, 2018 ; Fiveash *et al.*, 2020a ; Ladanyi *et al.*, 2020 ; Przybylski *et al.*, 2013). Or, l'analyse grammaticale s'effectue en grande partie sur la base de l'analyse de la parole en larges unités, les enveloppes d'amplitude qui rythment la prosodie. C'est pourquoi les amorces envisagées ici pour améliorer le traitement prosodique au cours des exercices d'un programme d'entraînement ont, comme dans les expériences qui viennent d'être citées, une pulsation lente et saillante, à 2 Hz.

Par ailleurs, le cadre théorique qui a inspiré cette hypothèse, comme les autres hypothèses dans l'ensemble de cette recherche, tire une partie de sa légitimité de sa compatibilité avec des règles du fonctionnement neurologique, sur le plan de la dynamique oscillatoire. C'est pourquoi, de façon plus spéculative et prudente, une réflexion s'est engagée sur le support neuro-fonctionnel des potentiels effets d'une amorce rythmique musicale à 2 Hz sur le traitement d'unités prosodiques. En tant que stimulation rythmique régulière et lente, on peut supposer qu'elle puisse déclencher ou renforcer des oscillations cérébrales lentes, directement utiles à l'extraction d'enveloppes de parole adaptées à l'analyse prosodique. Elle pourrait aussi favoriser la cascade oscillatoire qui sous-tend l'analyse du signal de parole pour assurer la détection de modulations brusques d'amplitude ou de fréquence fondamentale (F0) indiquant l'attaque de telles enveloppes. Cette hypothèse évoquant les oscillations cérébrales ne sera pas testée directement dans l'étude, qui n'analyse pas les oscillations. Il s'agit simplement d'une explication pour de potentiels effets des amorces sur l'évolution des performances des participants après un entraînement. Une telle interprétation est destinée à montrer que les potentiels effets sont compatibles avec le cadre théorique qui, en neuroscience cognitive, parle davantage de dynamique oscillatoire (i.e., *language oscillome*, Benitez-Burraco & Murphy, 2019) que de localisation cérébrale (i.e. « *spatial mapping is not explaining* », Poeppel *et al.*, 2008). L'hypothèse est aussi fondée sur une approche théorique plus large reconnaissant les liens étroits entre musique et langage, comme modélisés dans PRISM (Fiveash *et al.*, 2021).

Plus précisément pour le traitement prosodique des questions anglaises, le rythme optimal pour le traitement du changement d'intonation finale est lent, et un rythme à 2 Hz



paraissait approprié. En effet, la langue anglaise est accentuelle et fait alterner des syllabes faibles/fortes avec une syllabe accentuée environ toutes les 500 ms (Leong & Goswami, 2014a). Il se peut toutefois que cela soit un peu lent pour la durée de certains changements de hauteur de voix en fin de question, opérés sur une durée plus courte : celle d'une seule syllabe. Toutefois, une amorce rythmique à 2 Hz pourrait rester pertinente, car les oscillations delta exercent une modulation d'amplitude sur les oscillations thêta, compatibles avec la durée de la syllabe finale souvent porteuse de ce changement. Un ajustement des phases des ondes delta avec la musique permettrait, par une réaction en cascade, l'ajustement d'ondes plus rapides, rendant plus efficace le traitement de changements de hauteur de voix sur une à deux syllabes. Les ressources attentionnelles seraient ainsi guidées à un rythme adapté pour traiter de façon précise des modulations de fréquence importantes pour les questions. La concentration de l'attention sur le support de l'intonation finale des phrases pourrait conduire, à travers des exercices répétitifs, à distinguer les différentes sortes de questions anglaises, et à apprendre ces différences, probablement sur le mode d'un apprentissage statistique implicite. Ainsi, le choix du rythme de l'amorce musicale est basé sur les connaissances actuelles en neurosciences sur la dynamique de la cascade d'oscillations qui participe au traitement du signal de parole. Il est conforme aux modèles alliant l'attention temporelle à son support neuronal dynamique qui guident la thèse : AST (Poeppe, 2003) et TSF (Goswami, 2011).

Ce chapitre se donnait **trois principaux objectifs**. Le premier était d'évaluer l'effet d'une amorce rythmique à 2 Hz sur l'attention afin d'améliorer le traitement d'indices prosodiques lors d'un entraînement par répétition de questions avec deux types d'intonations, en imaginant un transfert possible au traitement ultérieur d'autres questions. Le deuxième objectif était de mesurer le possible effet renforçateur d'un retour visuel sur la qualité de la production de la prosodie anglaise après un tel entraînement basé sur l'imitation de phrases. Le choix d'utiliser un retour visuel était guidé par le bénéfice qu'un tel feedback apporte en rééducation des troubles de production de langage (Fabre *et al.*, 2016). Le contenu du retour visuel procuré est de nature acoustique), ce qui est assez innovant car la plupart des retours visuels utilisés en orthophonie renseignent sur des aspects articulatoires. Les résultats prometteurs de Guyot-Talbot *et al.* (2016) avec un retour visuel sur des aspects acoustiques ont fortement inspiré nos expériences. Il était espéré que la combinaison des deux types d'aide, associée à l'entraînement par répétitions, apporte un bénéfice supérieur à celui procuré en cas d'utilisation d'un seul des procédés. Le troisième et dernier objectif était de mesurer le bénéfice de cette association auprès d'adultes dyslexiques.

Afin de tester ces hypothèses, deux expériences ont été réalisées. L'**Expérience 13** comparait l'effet de trois programmes d'entraînement reposant sur l'apprentissage d'associations à un niveau implicite (l'un avec amorce rythmique seule, un autre avec retour visuel seul, et un autre combinant les deux aides) chez des adultes francophones normo-lecteurs. L'**Expérience 14** permettait d'observer l'effet de la combinaison des deux aides chez des adultes dyslexiques francophones pour vérifier la pertinence du dispositif le plus complet pour cette population atteinte d'une pathologie du langage.

## **7.6. Expérience 13 : comparaison de trois entraînements chez des adultes normo-lecteurs**

### **7.6.1. Principe de l'expérience et hypothèses**

L'Expérience 13 était guidée par les deux premiers objectifs du chapitre. Il s'agissait d'évaluer si une amorce rythmique à 2 Hz et un retour visuel accentuaient les bénéfices apportés par un entraînement fondé sur la répétition de phrases. Cet entraînement était destiné à faire apprendre, sans indication explicite, l'opposition entre les intonations montante et descendante des questions anglaises fermées/ouvertes. Pour cela, trois conditions d'entraînement par répétition ont été comparées : 1) l'une utilisait une amorce rythmique (32 secondes) à 2 Hz, 2) une autre utilisait un retour visuel direct sur la courbe d'intonation du locuteur superposée à celle du modèle et 3) la troisième combinait ces deux aides.

Les hypothèses suivantes ont été formulées :

- 1) Les trois aides devraient améliorer, après l'entraînement, les performances en production de questions anglaises sans modèle audio (changement final du F0 dans la bonne direction selon le type de question). Il devrait aussi améliorer la qualité de l'imitation après écoute d'un modèle audio de question. Dans ce dernier cas, la mesure évaluera entre les étapes pré- et post- entraînement, la diminution de la différence entre le modèle anglais et l'imitation produite par l'apprenant au niveau du changement de fréquence en fin de question (différence mesurée par un algorithme de *Dynamic Time Warping*, DTW, déformation temporelle dynamique).
- 2) L'aide apportée (observée à partir des tailles d'effet) par la combinaison de l'amorce rythmique et du retour visuel devrait être supérieure à celle apportée par l'amorce rythmique seule, elle-même supérieure au retour visuel seul.

## 7.6.2. Méthode

### 7.6.2.1. Participants

Quatre-vingt-dix normo-lecteurs (67 femmes, 22 hommes) âgés de 18 à 26 ans ( $M = 21.2$  ans,  $ET = 2$  mois) ont participé à l'étude. Ils étaient monolingues français, avec une vue normale ou corrigée et aucun déficit auditif connu. Ils n'avaient que peu ou pas pratiqué la musique et leur niveau d'anglais ne dépassait pas B2 (Cadre Européen Commun de Références pour les Langues, CECRL).

### 7.6.2.2. Matériel

#### Matériel de l'épreuves d'évaluation

**Epreuve de lecture à haute voix sans modèle audio.** Un ensemble de 20 phrases écrites courtes ( $< 1700$  ms) a été élaborer pour servir de support pour une lecture à haute voix. Elles étaient différentes de celles de l'épreuve de répétition et de celles de l'entraînement. Elles étaient de deux types : a) 10 questions ouvertes, b) 10 questions fermées. Elles ont été élaborées par paires avec le même mot à la fin des deux phrases d'une paire. Dans la paire, la question ouverte avait un contenu très proche de celui de la question fermée afin d'ajuster la comparaison des foyers intonatifs. Chaque phrase tenait sur une ligne avec une taille de caractères permettant une lecture confortable.

**Epreuve de répétition avec modèle audio.** La liste comportait 27 phrases anglaises, différentes de celles de l'épreuve de lecture et de l'entraînement, mais de structures comparables, et elles étaient de trois types : a) 9 phrases déclaratives, b) 9 questions fermées, c) 9 questions ouvertes. Elles étaient présentées à l'écrit avec les mêmes caractères que dans l'épreuve de lecture sans modèle audio.

Toutes les phrases ont été également enregistrées pour créer les fichiers-sons, en salle sourde anéchoïque à l'aide du logiciel RocMe! (Ferragne *et al.*, 2012). Les voix étaient celles de deux jeunes femmes anglophones natives parlant un anglais britannique standard. L'une des voix a été sélectionnée pour sa fréquence fondamentale ( $F_0$ ) basse, proche d'une voix masculine, en vue d'éviter un biais pour les participants masculins pouvant être plus à l'aise pour imiter une voix de hauteur proche de la leur.

## **Matériel pour l'entraînement**

**Amorces auditives.** Six amorces auditives de 33 secondes ont été sélectionnées. Trois ont été élaborées avec Ph. Lalitte (musicologue, Université de Bourgogne). Les trois autres étaient des séquences de bruits de l'environnement (bruit de la ville, bruit de marché et bruit de cour de récréation) non stressantes et dont il n'était pas possible d'extraire une régularité ni une parole intelligible. La durée de 33 secondes a été établie d'après la durée efficace pour produire un effet positif dans d'autres expériences sur l'aide au jugement grammatical (Fiveash *et al.*, 2020a).

**Matériel linguistique.** Vingt-sept phrases anglaises ont été élaborées, aucune n'étant présentée au participant dans les épreuves de répétition ou de lecture. Elles étaient de trois types : 9 phrases déclaratives, 9 questions fermées et 9 questions ouvertes. Enregistrées dans les mêmes conditions que celles de l'épreuve de répétition et avec les mêmes locutrices, elles respectaient les mêmes contraintes.

### **7.6.2.3. Procédure générale de l'expérience**

L'expérience se déroulait en trois étapes (Fig. 83) : les participants ont effectué une première fois les épreuves de répétition et de lecture (T1), puis réalisaient l'entraînement, et passaient enfin à nouveau les épreuves de répétition et de lecture (T2). Tout au long de l'expérience, les participants étaient assis confortablement en face de l'écran dans une salle calme. Ils portaient un casque audio aussi bien pour l'entraînement que pour les épreuves pré- et post-entraînement lorsque ces dernières présentaient des phrases de façon auditive. A cause des restrictions sanitaires liées à la crise Covid-19, pour les participants du groupe réalisant l'entraînement avec musique seule, l'expérience s'est déroulée à distance, au domicile des participants, en autonomie. Ils étaient suivis par téléphone depuis l'installation du logiciel Opensesame jusqu'à l'envoi des fichiers audio par mail et s'enregistraient à l'aide du dictaphone de leur téléphone. Les participants avaient pour consigne de démarrer l'enregistrement au début de chaque épreuve, créant ainsi 5 fichiers audio (2 pré-tests, 1 entraînement, 2 post-tests) qu'ils devaient envoyer par mail à la fin du protocole.

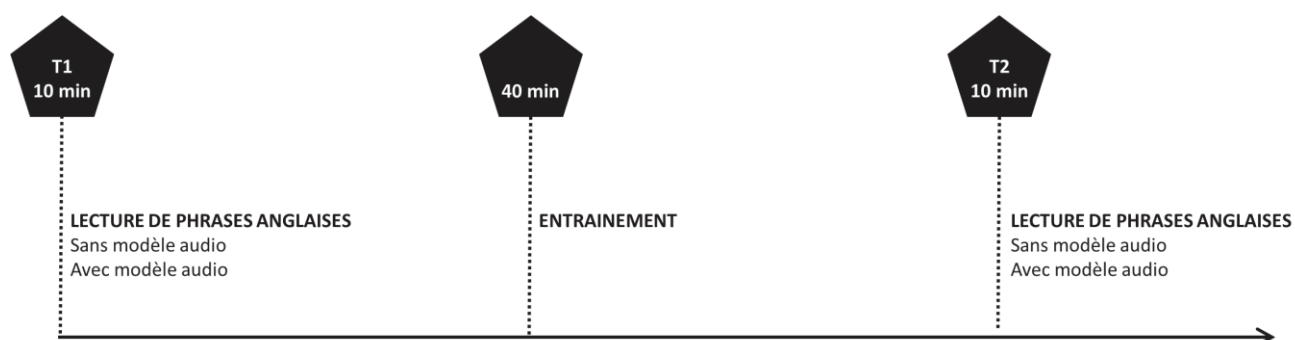


Figure 83. Procédure générale de l'expérience.

#### 7.6.2.4. Procédure des épreuves d'évaluation

La première épreuve était toujours la tâche de lecture de phrases anglaises sans modèle audio.

Dans l'épreuve de lecture à haute voix sans modèle audio, le participant devait lire aussi bien que possible chaque phrase anglaise présentée une à une à l'écran. L'ordre des phrases était pseudo-aléatoire pour éviter un regroupement de phrases d'une même catégorie, et il était identique pour tous. Sans contrainte de rapidité, le participant commençait l'enregistrement en appuyant sur le bouton d'enregistrement. La consigne était de prononcer la phrase avec le meilleur accent anglais possible. Il pouvait suivre l'avancement de l'épreuve grâce à un compteur indiquant à l'écran le nombre de phrases restantes. L'épreuve durait 5 minutes.

La seconde épreuve était la tâche de répétition avec modèle audio. Le participant écoutait attentivement dans un casque stéréo une phrase présentée simultanément à l'écrit sur l'écran. Sans contrainte de rapidité, le participant appuyait sur le bouton d'enregistrement et répétait la phrase entendue, ceci pour toutes les phrases de la liste, présentées dans un ordre aléatoirement contrôlé (comme pour l'épreuve précédente), identique pour tous. Aucun feedback n'était donné sur la qualité de la production et le participant ne pouvait pas se réécouter. Il pouvait suivre l'avancement de l'épreuve grâce à un compteur indiquant à l'écran le nombre de phrases restantes. La consigne demandait d'imiter au mieux ce qu'il entendait sans indication sur le paramètre le plus pertinent (qualité des phonèmes, prosodie, hauteur...). L'épreuve durait 5 minutes.

### 7.6.2.5. Procédure d'entraînement

Les participants ont été répartis aléatoirement en trois groupes indépendants de 30 participants chacun : 1) Groupe MR, amorce Musicale et Retour visuel (n = 30), 2) Groupe BR, amorce Bruit et Retour visuel (n = 30), 3) Groupe M, amorce Musicale *sans* retour visuel (n = 30). Il n'y avait pas de groupe sans amorce, l'amorce neutre était le bruit de l'environnement sans rythme dans le groupe BR.

Juste avant l'entraînement, un court texte devait être lu par les participants des groupes MR et BR (le groupe M ne réalisait pas cette étape) pour calibrer le micro et permettre au logiciel de repérer leur F0 : la réalisation des courbes du F0 représentant visuellement les modulations de hauteur de la voix de chacun était ainsi ensuite plus exacte, assurant un retour visuel précis et personnalisé.

Comme l'illustre la figure 84, l'entraînement commençait par l'écoute dans le casque d'une amorce (musique rythmée pour les groupes M et MR, bruits de l'environnement pour le groupe BR) de 33 secondes. La fin de l'amorce était immédiatement suivie d'une phrase anglaise présentée simultanément à l'écran et auditivement dans le casque. La consigne était de répéter la phrase entendue en l'imitant le mieux possible. Le participant choisissait le moment où il se sentait prêt et appuyait sur un bouton à l'aide de la souris pour déclencher l'enregistrement de sa voix imitant la phrase. Pour les groupes avec retour visuel (MR et BR), une représentation de la courbe du F0 de la phrase-modèle du locuteur anglophone (en rouge) et de celle de la répétition du participant (en bleu) apparaissait à l'écran dès la fin de l'imitation. Le participant était invité à les comparer visuellement en vue d'ajuster les courbes le mieux possible lors des prochaines répétitions. Le groupe M, qui n'avait pas de retour visuel, avait pour seule consigne d'enregistrer ses imitations orales sur le dictaphone.

Chaque phrase était écoutée et répétée cinq fois de suite, quel que soit le groupe. Après ces cinq écoutes-répétitions, une nouvelle amorce était entendue pour débiter un nouveau bloc d'entraînement sur une autre phrase. Dans tous les groupes, une amorce audio précédait chaque bloc (i.e. les 5 phrases-répétitions) : il s'agissait d'une musique rythmée pour les groupes M et MR, et d'une ambiance sonore non-rythmée dans le groupe BR. Le contenu précis des amorces variait entre deux blocs et chacune des trois amorces possibles était entendue six fois. Dans chaque condition, la présentation des trois types de phrases était aléatoire, mais plusieurs listes ont permis de faire en sorte que chacune des 27 phrases soit

précédée de chacune des 6 amorces possible (3 musiques et 3 bruits) à travers les participants. L'entraînement durait 40 min (Fig. 84).

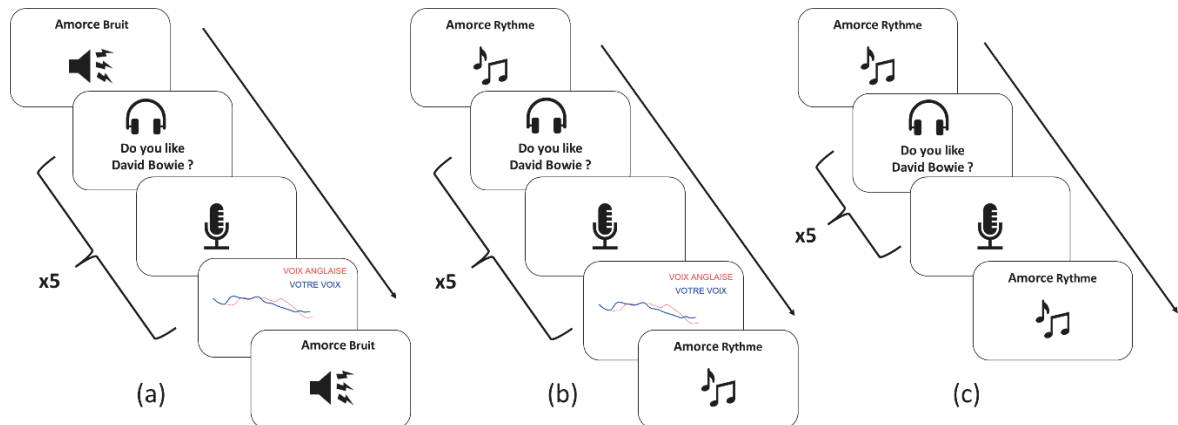


Figure 84. Déroulé de l'entraînement concernant la prosodie anglaise selon les différentes conditions d'amorçage et de retour visuel. (a) amorce bruit et retour visuel, groupe BR, (b) amorce musicale et retour visuel, groupe MR, et (c) amorce musicale sans retour visuel, groupe M.

### 7.6.3. Analyse des données

Pour l'épreuve de lecture de phrases sans modèle audio, la différence de F0 entre le début et la fin du segment de phrase portant le contour intonatif mesuré sur le spectrogramme avec Praat était calculée. Cela donnait une évaluation de la modulation de hauteur de voix, montante ou descendante, en fin de phrase produite. Cette évaluation a été réalisée par deux juges séparément, en se basant sur la courbe du F0 tracée par Praat et en vérifiant la conformité du début et de la fin de la courbe prise en compte (section finale de la phrase) avec les variations acoustiques de hauteur perçues dans un casque audio. En cas de différence d'appréciation entre les juges, le signal était réévalué et l'avis d'un troisième juge était demandé.

Pour l'épreuve de répétition de phrases, il était possible de faire une autre mesure, évaluant cette fois l'adéquation au modèle audio (phrase produite par un anglophone). Pour cela, la distance entre les F0 du modèle et celui du participant sur toute la phrase a été évaluée par un algorithme DTW (*Dynamic Time Wrapping*, déformation temporelle dynamique). Ces valeurs ont été moyennées pour chaque participant par type de phrase.

Afin d'évaluer l'effet du type d'entraînement, deux ANOVAs à mesures répétées ont été réalisées, l'une sur la différence de F0 en lecture, l'autre sur la distance DTW en répétition, avec le Type de phrase (*Lecture* : Montante, Descendante ; *Répétition* : Montante, Descendante, Plate) et la Session (T1, T2) comme facteurs intra-individuels, et le type d'Entraînement (MR, M, BR) comme facteur inter-individuel.

Les tailles d'effets ont été calculées à l'aide de l'êta-carré-partiel ( $\eta_p^2$ ) et interprétées selon Cohen (1988). Les comparaisons issues des interactions ont été calculées à l'aide du test  $t$  de Student bilatéral avec correction de Bonferroni et les tailles d'effet à l'aide du  $d$  de Cohen. Le seuil de significativité était placé à  $p \leq .05$ .

#### 7.6.4. Résultats

**Lecture à voix haute sans modèle audio.** L'ANOVA a mis en évidence un seul effet principal, celui du Type de phrase ( $F(1,87) = 143.42, p < .001, \eta_p^2 = 0.62$ ) car le F0 des questions descendantes baissait et celui des questions montantes augmentait à la fin des phrases. Elle a aussi révélé l'interaction Type de phrase  $\times$  Session ( $F(1,87) = 13.72, p < .001, \eta_p^2 = .14$ ) illustrée par la figure 85. Entre T1 et T2, la diminution du F0 en fin de phrase s'accroissait significativement pour les questions descendantes ( $t(89) = 2.97, p = .004, d = .35$ ). L'élévation du F0 en fin de phrase pour les questions montantes s'accroissait aussi mais façon non significative ( $t(89) = -1.02, p = .312, d = .10$ ). Cette configuration correspondait à ce qui se produisait sur l'ensemble des trois types d'entraînement. Il n'y avait pas d'autre interaction significative.

Le fait qu'un progrès entre T1 et T2 s'exprime forcément de façons opposées pour les questions montantes (élévation accrue du F0) et descendantes (baisse accrue du F0) a pu masquer certains effets, notamment l'effet principal de la Session, et peut-être certaines interactions impliquant le type d'entraînement. C'est pourquoi une ANOVA complémentaire sur les questions descendantes – seules questions dont la prosodie devait faire l'objet d'un apprentissage par les francophones – a été réalisée pour une étude plus détaillée des données.

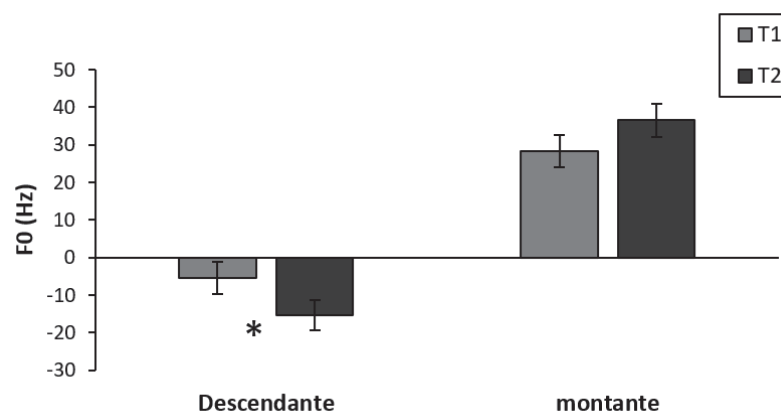


Figure 85. Différence de F0 (variation de hauteur perçue) entre début et fin de segment intonatif des questions chez les adultes contrôles, selon le type de question (descendante, montante), à T1 (gris) et T2 (noir). Les barres d'erreurs représentent l'erreur-type. L'étoile indique une différence significative.



Cette ANOVA confirmait l'effet de la Session sur les questions descendantes avec une baisse de F0 plus intense en fin de question à T2 qu'à T1 ( $F(1,87) = 8.65, p = .004, \eta_p^2 = 0.09$ ). L'interaction Session  $\times$  Entraînement n'était pas significative ( $F(2,87) = 2.25, p = .112, \eta_p^2 = .05$ ), mais comme une hypothèse portait sur les effets respectifs des trois entraînements, des tests  $t$  de Student ont été réalisés. Comme le montre la partie gauche de la figure 86, la baisse de F0 en fin de question descendante s'est significativement accentuée entre T1 et T2 pour l'entraînement M ( $t(29) = 2.04, p = .050, d = .34$ ) et l'entraînement MR ( $t(29) = 2.00, p = .055, d = .32$ ), mais pas pour l'entraînement BR avec retour visuel et amorce sonore sans rythme ( $t(29) = 1.51, p = .141, d = .36$ ). Ce dernier effet sur BR n'atteignait pas le seuil de significativité mais était de taille comparable, et même légèrement supérieure, aux effets de la session en conditions M et MR. L'absence de significativité de l'effet de session pour BR pourrait s'expliquer par la variabilité particulièrement forte des données à T1 dans cette condition.

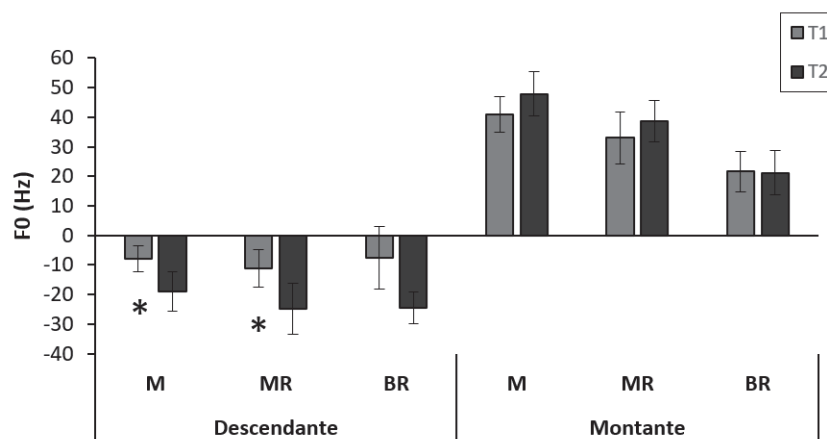


Figure 86. Différence de hauteur de F0 entre début et fin de segment intonatif des questions chez les normo-lecteurs, selon le type de question (descendante, montante), et le type d'entraînement (M, MR, BR), à T1 (gris) et à T2 (noir). Les barres d'erreurs représentent l'erreur-type. L'étoile indique une différence significative.

**Répétition d'un modèle audio.** L'ANOVA a montré l'effet principal du Type de phrase ( $F(2,174) = 108.82, p < .001, \eta_p^2 = .55$ ) avec une distance modèle-répétition (DTW) plus grande pour les phrases montantes que pour les phrases descendantes ( $t(89) = 12.12, p < .001, d = 1.26$ ) et les phrases plates ( $t(89) = 14.71, p < .001, d = 1.23$ ). L'effet principal du type d'Entraînement était significatif aussi ( $F(2,87) = 3.58, p = .032, \eta_p^2 = .076$ ) et il y avait une interaction Session  $\times$  Type de phrase ( $F(2,174) = 3.21, p = .043, \eta_p^2 = .036$ ). Celles-ci montraient que seule la DTW des phrases descendantes changeait en diminuant entre T1 et T2 ( $t(89) = 2.29, p = .024, d = .29$ ). L'interaction Session  $\times$  Type de phrase  $\times$  Entraînement n'était pas significative ( $F(4,174) = 1.07, p = .375$ ). L'observation de la figure 87, qui illustre

les résultats pour les questions descendantes, ainsi que l'importance de l'hypothèse à propos de ces phrases dont l'intonation était à apprendre, ont incité à réaliser une ANOVA complémentaire sur la DTW seulement pour ces questions descendantes, avec la Session comme facteur intra-individuel et le type d'Entraînement comme facteur inter-individuel.

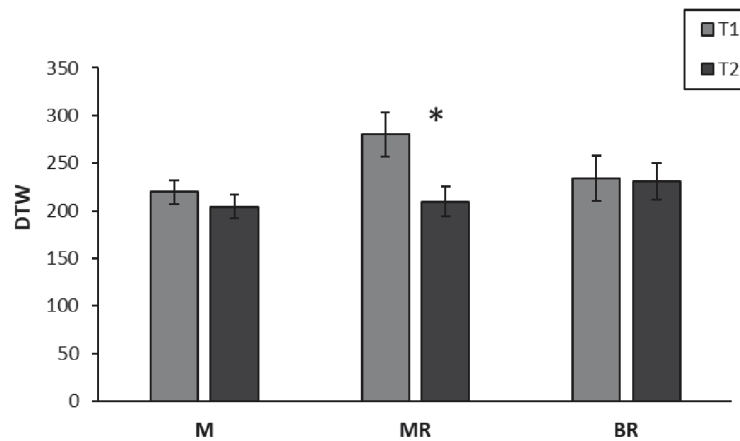


Figure 87. DTW moyenne des phrases descendantes selon le groupe défini par l'entraînement (M, MR, BR), à T1 (gris) et T2 (noir). Les barres d'erreurs représentent l'erreur-type.

Cette ANOVA sur les phrases descendantes a montré une interaction Session  $\times$  Entraînement s'approchant du seuil de significativité ( $F(2,87) = 2.61, p = .080, \eta_p^2 = .39$ ). Les test  $t$  de Student ont permis de préciser que cela s'expliquait par une diminution significative et de taille modérée de la DTW pour l'entraînement MR ( $t(29) = 2.70, p = .011, d = .66$ ), mais pas pour les entraînements M ( $t(29) = 1.49, p = .148, d = .24$ ) et BR ( $t(29) = 0.13, p = .895, d = .03$ ) (Fig. 87). Notons cependant que le groupe MR présentait une distance DTW plus élevée à T1 que le groupe M ( $t(43.86) = 2.33, p = .025, d = .60$ ) mais pas que le groupe BR ( $t(57.83) = 1.39, p = .171, d = .36$ ). Cela limitera l'interprétation de l'effet de la session pour ce groupe dont les faibles performances à T1 offraient une plus grande marge de progression possible.

Pour les phrases descendantes, des corrélations de Pearson entre la DTW à T1 et le progrès de la DTW (T1 moins T2) ont été calculées pour les entraînements M et MR, car tous deux montraient, pour l'entraînement, des tailles d'effet non négligeables (Fig. 88). Le progrès observé suite à l'entraînement MR était largement corrélé avec la DTW à T1 ( $r(28) = -0.80, p < .0001$ ) : la réduction de la distance DTW entre T1 et T2 était d'autant plus ample que la distance modèle-répétition était grande à T1. Le progrès observé suite à l'entraînement M était beaucoup plus modestement corrélé à la distance DTW à T1 ( $r(28) = -0.42, p = .022$ ). Le niveau de performance à T1 était différent entre ces groupes : la distance au modèle était plus grande pour le groupe MR que pour le groupe M et les performances présentaient une

plus grande hétérogénéité parmi les participants du groupe MR que du groupe M, comme l'illustre la figure 88. Une analyse complémentaire de la corrélation a été réalisée en sélectionnant dans le groupe MR les participants dont les performances présentaient la même étendue que dans le groupe M (avec une DTW entre 109 et 327). Dans ces conditions, la corrélation pour le groupe MR n'était plus significative mais restait de taille moyenne ( $r(28) = -0.30, p = .181$ ).

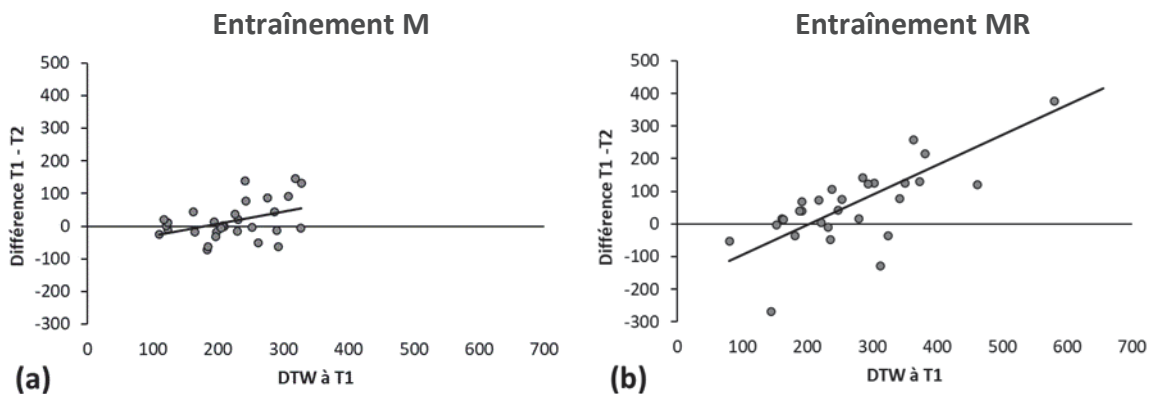


Figure 88. Corrélations entre la DTW à T1 (différence entre modèle et imitation avant l'entraînement) et l'amélioration de l'imitation à T2 (différence DTW T1 – DTW T2), pour les entraînements M (a) et MR (b).

### 7.6.5. Discussion intermédiaire

En résumé, dans l'Expérience 13, l'analyse des résultats apporte quelques éléments sur des progrès réalisés en production prosodique anglaise par les adultes normo-lecteurs après les entraînements très brefs proposés. L'intonation descendante des questions ouvertes était l'intonation la moins bien imitée avant l'entraînement. Quelle que soit la combinaison d'aides proposée, l'ensemble a montré après l'entraînement une diminution significative de la distance entre l'intonation produite et celle du modèle. De même, en l'absence de modèle, les questions ouvertes ont été produites en accentuant l'intonation descendante après l'entraînement. Ces changements se sont faits sélectivement pour les questions ouvertes, car pour les questions fermées il n'y a pas eu de changement d'intonation : celle-ci était et restait montante.

Un résultat surprenant était la distinction réalisée en production dans l'intonation entre les questions ouvertes et fermées dès T1, ceci même en l'absence de modèle audio. Les adultes normo-lecteurs ont peut-être mobilisé une distinction qui existe en français d'une façon très atténuée. Ils ont peut-être aussi déjà une expérience suffisante en anglais pour avoir développé une connaissance sur cette différence, ou leurs enseignants ont insisté sur ce point. Malgré cette compétence assez satisfaisante à l'initiale, l'intonation descendante des questions

ouvertes anglaises s'est accentuée après l'entraînement, et des progrès en prosodie sont donc possible chez des adultes apprenants tardifs d'une L2.

Si l'ensemble des entraînements a modulé la prosodie des productions, il y a en revanche peu d'arguments pour un effet plus fort de l'un des dispositifs d'aide utilisés. En lecture sans modèle audio, à l'issue d'une interaction non-significative, l'effet des entraînements M et MR qui utilisent l'amorçage musical sont les seuls à ressortir comme significatifs, mais l'étude de la taille des effets ne confirme pas qu'ils sont vraiment plus efficaces que l'entraînement utilisant le retour visuel seul (BR). En répétition d'un modèle audio, la combinaison de l'amorçage et du retour visuel (MR) a par ailleurs produit la seule amélioration significative sur les questions descendantes après l'entraînement. Cet effet de taille modéré est encourageant, mais reste à confirmer. Le groupe ayant reçu cet entraînement était en effet celui qui réussissait le moins bien l'imitation de ces questions avant le programme, offrant plus facilement une marge de progression possible. L'étude des corrélations a finalement apporté une information intéressante sur le plan des applications : les adultes normo-lecteurs les plus en difficulté à T1 peuvent tirer un bénéfice des apprentissages proposés, et en particulier à partir de celui associant l'amorçage rythmique et le retour visuel sur les informations acoustiques en prosodie.

## **7.7. Expérience 14 : Effet de l'entraînement avec amorçage rythmique et retour visuel chez des adultes dyslexiques**

### **7.7.1. Principe de l'expérience et hypothèses**

L'Expérience 14 permettait de répondre au troisième objectif du chapitre, qui était de mesurer le bénéfice de l'association amorce rythmique – retour visuel (condition MR de l'Expérience 13) auprès d'adultes dyslexiques. En effet, bien que quelques réserves soient à faire sur l'interprétation des résultats de l'Expérience 13 conduite chez des apprenants francophones sans pathologie, cette association des deux aides était potentiellement la plus robuste pour tenter d'améliorer l'apprentissage d'indices prosodiques en anglais. Il semblait important de proposer à chaque adulte dyslexique inclus dans l'étude le dispositif le plus prometteur. Le choix de ne pas tester d'emblée les deux autres conditions (avec amorce rythmique seule et avec amorce non-rythmique et retour visuel) auprès des adultes dyslexiques rend cette partie de l'étude encore incomplète et prospective.

Comme dans l'Expérience 13, une pulsation à 2 Hz dans l'amorce était destinée à encourager une synchronisation des ondes lentes – déficitaire en cas de dyslexie (Goswami, 2011) – avant l'écoute-répétition de phrases pendant l'entraînement. Il s'agit là du mécanisme neuronal sous-jacent supposé. Sur le plan cognitif, une distribution temporelle de l'attention plus adaptée à une séquenciation permettant d'extraire des unités prosodiques était alors espérée. Les fins de phrase portant l'intonation auraient ainsi des chances d'être mieux traitées pendant l'entraînement. Cela devait permettre de les traiter de façon plus profonde et précise pour leurs différentes dimensions, notamment le changement de hauteur. Le retour visuel, indice plus explicite mais imagé, était aussi destiné à attirer l'attention sur les indices prosodiques finaux permettant de différencier les types de phrases.

L'**hypothèse** était que, malgré leurs difficultés rythmiques, les adultes dyslexiques du petit groupe testé dans l'Expérience 14 allaient progresser après l'entraînement de type MR, choisi sur la base des résultats de l'Expérience 13 chez des adultes normo-lecteurs. Une amélioration des performances (F0 et DTW) à la suite de l'entraînement était attendue.

## **7.7.2. Méthode**

### **7.7.2.1. Participants**

Vingt dyslexiques (19 femmes, 3 hommes) de 18 ans à 28 ans ( $M = 21.4$  ans,  $ET = 2.3$  mois) ont participé à cette étude. Ils étaient monolingues français, avec une vue normale ou corrigée et aucun déficit auditif connu. Ils n'avaient que peu ou pas pratiqué de musique et leur niveau d'anglais ne dépassait pas B2 (Cadre Européen Commun de Références pour les Langues, CECRL). Ils avaient reçu un diagnostic durant l'enfance et ne présentaient pas d'autre trouble neuro-développemental (e.g. TDA-H). Un test de fluence en lecture (L'alouette, Lefavrais, 1967) vérifiait la persistance du déficit au moment de l'étude (selon les normes adultes établies par Cavalli *et al.*, 2018).

Ils ont été comparés aux 30 adultes normo-lecteurs de l'Expérience 13 (partie 7.6.2.1 de ce chapitre) ayant réalisé l'entraînement MR (Annexe 3).

### **7.7.2.2. Matériel**

Le matériel utilisé était le même que dans l'Expérience 13 décrit dans la partie 7.6.2.2 de ce chapitre pour la condition MR.

### 7.7.2.3. Procédure

La procédure était identique à celle de la condition MR de l'Expérience 13 décrite dans les parties 7.6.2.3, 7.6.2.4 et 7.6.2.5.

### 7.7.3. Analyses de données

Afin d'évaluer l'effet de l'entraînement dans le groupe dyslexique, deux ANOVAs à mesures répétées ont été réalisées, l'une sur la différence de F0 entre le début et la fin de la partie de phrase portant l'intonation (tâche de lecture sans modèle), l'autre et sur la distance (calculée par un algorithme DTW) entre l'intonation du modèle et de répétition, avec le Type de phrase (*Lecture* : Montante, Descendante ; *Répétition* : Montante, Descendante, Plate) et la Session (T1, T2) comme facteurs intra-individuels, et le Groupe (Normo-lecteur, Dyslexique) comme facteur inter-individuel. Comme dans l'Expérience 13, deux juges ont réalisé séparément ces mesures. En cas de différence d'appréciation, le signal était réévalué et l'avis d'un troisième juge était sollicité.

Les tailles d'effets ont été calculées à l'aide de l'éta-carré-partiel ( $\eta_p^2$ ) et interprétées selon Cohen (1988). Les comparaisons issues des interactions ont été calculées à l'aide du test  $t$  de Student bilatéral avec correction de Bonferroni et les tailles d'effet à l'aide du  $d$  de Cohen. Le seuil de significativité était placé à  $p \leq .05$ .

### 7.7.4. Résultats : Effet de l'entraînement MR chez les adultes dyslexiques – comparaison T1 – T2

**Lecture à haute voix sans modèle audio.** L'ANOVA a révélé un effet du Type de phrase ( $F(1,48) = 84.90, p < .001, \eta_p^2 = 0.64$ ), montrant que le F0 des questions descendantes baissait et celui des montantes augmentait en fin de phrase. L'interaction Type de phrase  $\times$  Session était significative ( $F(1,48) = 6.39, p = .015, \eta_p^2 = 0.12$ ). Elle s'expliquait surtout par l'accentuation significative de la baisse du F0 entre T1 et T2 pour les questions descendantes ( $t(49) = 2.22, p = .031, d = .30$ ). L'accentuation de l'augmentation finale du F0 observée (Fig. 89) n'était pas significative pour les questions montantes entre T1 et T2 ( $t(49) = -1.65, p = .105, d = .20$ ). Il n'y avait pas d'interaction impliquant le facteur Groupe, et la figure 89 représente les résultats pour les deux groupes (dyslexiques et normo-lecteurs) confondus.

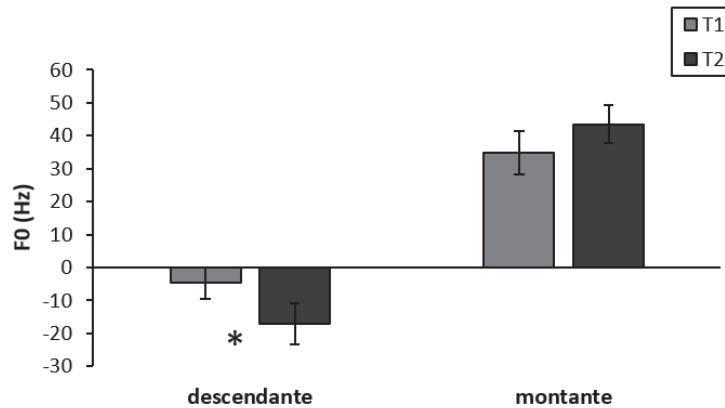


Figure 89. Différence de F0 entre début et fin de segment intonatif, selon le type de question (descendante, montante), à T1 (gris) et à T2 (noir), pour l'entraînement MR (normo-lecteurs et dyslexiques confondus). Les barres d'erreurs représentent l'erreur-type. L'étoile indique une différence significative.

**Répétition d'un modèle audio.** L'ANOVA a révélé un effet principal de la Session ( $F(1,48) = 8.18, p = .006, \eta_p^2 = .15$ ) montrant une diminution significative de la distance DTW passant de 312 ( $ET = 173.25$ ) à 267 ( $ET = 119.82$ ). De plus, l'interaction Session  $\times$  Type de phrase  $\times$  Groupe ( $F(2,96) = 3.08, p = .051, \eta_p^2 = .06$ ) était significative. Comme l'illustre la figure 90, pour le groupe dyslexique, la distance DTW diminuait significativement entre T1 et T2 pour les questions montantes ( $t(19) = 2.53, p = .021, d = .35$ ) et les questions descendantes ( $t(19) = 3.02, p = .007, d = .33$ ). En revanche, pour le groupe de normo-lecteurs, cette distance entre le modèle et la production (DTW) diminuait uniquement pour les phrases descendantes ( $t(29) = 2.70, p = .011, d = .66$ ). On notait que le groupe dyslexique présentait, à T1, une distance DTW marginalement plus grande que celle du groupe contrôle pour les phrases montantes ( $t(22,35) = 1.77, p = .090, d = .60$ ).

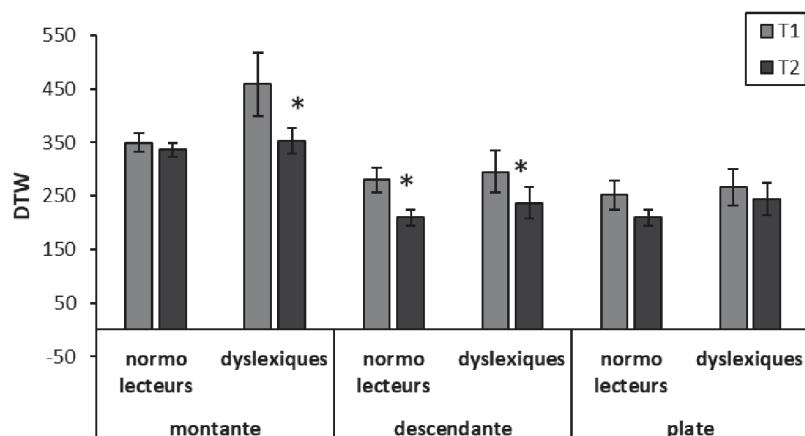


Figure 90. DTW moyenne selon le groupe, à T1 (gris) et à T2 (noir) selon le type de phrase. Les barres d'erreurs représentent l'erreur-type. L'étoile indique une différence significative.

La DTW ne permettant pas d'observer le sens (montant ou descendant) dans lequel les participants produisaient la phrase, une analyse complémentaire a été réalisée sur les différences de F0 entre le début et la fin de la syllabe portant l'intonation pour le groupe Dyslexique à l'aide du *t* de Student. Cette analyse plus ciblée sur les résultats présentés par la figure 91 (graphique à droite) montrait un changement dans l'orientation de la courbe d'intonation finale entre T1 et T2 uniquement pour les questions descendantes. En effet, à T1, les modulations du F0 en fin de question étaient proches de zéro et légèrement négatives. Entre T1 et T2, la descente classique du F0 pour les phrases descendantes apparaissait sélectivement pour les questions ouvertes (descendantes), et ce changement était un effet de grande taille ( $t(19) = 5.57, p < .001, d = 1.97$ ). Pour les questions fermées (montantes), l'accroissement final classique du F0 est aussi apparu, et cet effet était aussi de grande taille ( $t(19) = -9.81, p < .001, d = 3.25$ ).

**Comparaison visuelle des changements d'intonation entre T1 et T2 selon la tâche.** Cette analyse complémentaire portait sur la différence de F0 entre le début et la fin de la partie intonative des questions, comme celle réalisée pour les résultats de l'épreuve de lecture avec modèle. Une comparaison des figures présentant la configuration des résultats pour ces deux tâches chez les adultes dyslexiques est possible (Fig. 91). Ces configurations sont très différentes selon l'épreuve chez les adultes dyslexiques, ce qui n'était pas le cas chez les adultes normo-lecteurs (Expérience 13). En lecture sans modèle, les adultes dyslexiques ont produit à T1 une intonation montante quel que soit le type de question, et à T2 la différence s'est creusée entre les deux types de questions grâce à des questions montantes dont le F0 croissait davantage, et des questions descendantes dont le F0 décroissait très légèrement, opérant un léger changement d'orientation. En répétition avec modèle audio, la configuration était atypique à T1, avec une production descendante pour les deux types de questions. A T2, une distinction franche entre l'apparition d'une intonation montante pour les questions montantes et une intonation nettement descendante pour les questions descendantes est apparue.



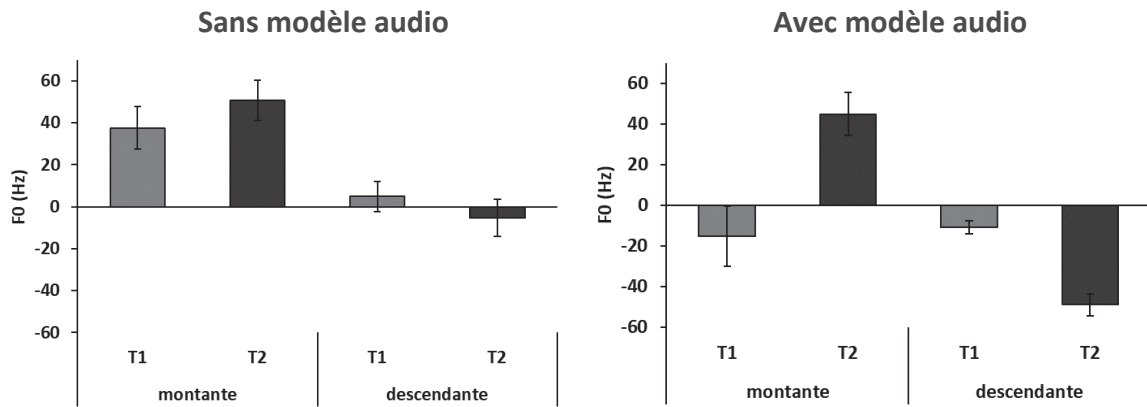


Figure 91. Différence de F0 entre le début et la fin de segment intonatif des questions montantes et descendantes, à T1 (gris) et à T2 (noir), en tâche de lecture sans modèle audio et avec modèle audio pour le groupe dyslexique. Les barres d'erreurs représentent l'erreur-type.

### 7.7.5. Discussion intermédiaire

En résumé, chez les adultes dyslexiques, les résultats montrent que des progrès sont possibles en prosodie anglaise grâce à un entraînement ciblé sur l'intonation des questions, malgré leur pathologie du langage. Dans l'épreuve de lecture sans modèle, ils ne se distinguent pas statistiquement des adultes normo-lecteurs : après l'entraînement, l'intonation descendante s'est accentuée sélectivement pour les questions ouvertes. Le phénomène est d'ampleur modeste, d'après la partie gauche de la figure 91, mais il s'agit d'un changement d'orientation de l'intonation. En répétition de modèle audio, les adultes dyslexiques ont amélioré de manière générale la qualité de leur imitation d'intonation, car la distance entre les modèles et les phrases produites a significativement diminué. Avec le même dispositif associant amorce rythmique et retour visuel, les normo-lecteurs réduisaient cette distance à T2 seulement pour les questions ouvertes, alors que les dyslexiques l'ont réduite aussi pour les questions fermées. Pour ces dernières, de moins bonnes performances chez les adultes dyslexiques que chez les normo-lecteurs avant l'entraînement pourraient en partie expliquer le progrès significatif à T2 seulement chez les dyslexiques, mais cette différence initiale n'était pas significative.

Un résultat étonnant a été relevé chez les adultes dyslexiques. L'intonation de leurs productions sans modèle audio à T1 était globalement correcte pour les questions anglaises, témoignant comme chez les adultes normo-lecteurs de connaissances sur la différence entre les deux types de questions. Par contre, leur intonation était indifférenciée et incorrecte pour ces mêmes questions en présence d'un modèle audio. Tout se passe comme si montrer ces informations auditives introduisait pour eux de la confusion, ce qui n'était pas le cas des adultes normo-lecteurs. La configuration correcte à T2, avec intonations respectivement

descendante et montante pour les questions ouvertes et fermées témoigne ainsi chez eux d'un véritable progrès après l'entraînement associant l'amorçage rythmique et le retour visuel.

### **7.8. Discussion sur l'aide à l'apprentissage de la prosodie anglaise (Expériences 13 et 14)**

Les Expériences 13 et 14 visaient à évaluer la possibilité de remédier, grâce à une séance unique d'entraînement intensif, à une difficulté constante rencontrée par les francophones en apprentissage de l'anglais L2 : interpréter et produire correctement des intonations différentes pour les questions fermées (montantes) et ouvertes (descendantes). Cette distinction acoustique nécessite une interprétation sémantique spécifique à l'anglais, différente de celle réalisée dans la L1 de nos participants. En effet, en français, l'intonation descendante existe pour certaines questions, mais plus rarement, et le ton descendant est davantage associé aux injonctions. Adopter une règle d'interprétation adaptée à l'anglais pour l'intonation descendante en fin de phrase est important pour participer à une conversation d'une manière adaptée du point de vue pragmatique. La difficulté des francophones à apprendre et appliquer cette règle prosodique pour les questions anglaises trouve sans doute en partie son origine dans le fait que les deux intonations finales (montante et descendante) existent dans leur L1, mais est associée à des modes de communication différents. Dans un tel cas, l'assimilation de la L2 aux règles de la L1 est favorisée, mais source de piège (Flege & Bohn, 2021). C'est pourquoi une aide ciblée sur ce type d'ambiguïté est souhaitable.

Les difficultés d'apprentissage d'une L2 sont accrues si l'apprenant est adulte, encore plus s'il est dyslexique. En effet, en cas de système linguistique mal établi dans la L1, la tendance habituelle à assimiler l'analyse de la parole en L2 à celle de la L1 se fait sur une base instable. C'est pourquoi l'Expérience 14 s'est aussi adressée à ce public, mais en lui fournissant uniquement le dispositif d'aide à l'apprentissage qui s'était avéré le plus efficace parmi les trois proposés à des adultes non-dyslexiques dans l'Expérience 13.

Le point commun aux trois dispositifs d'entraînement était la tâche dans les exercices. L'apprenant devait répéter des phrases anglaises avec la consigne de les imiter, sans indication explicite sur ce qu'il devait prendre en compte (la qualité des phonèmes, la hauteur de la voix, son timbre, la prosodie...), et les règles d'intonation ne lui étaient pas dites. Chaque phrase était entendue et imitée cinq fois pour encourager la précision, ce qui était une nouvelle subtilité par rapport au dispositif avec retour visuel qui a inspiré l'étude (Guyot *et al.*, 2016). L'hypothèse était qu'une inférence des règles intonatives dépendantes du type de

question était possible à partir de cet entraînement très court (45 mn). Cette brièveté était un défi pour les expériences. Il y avait deux critères de jugement pour l'efficacité de l'entraînement. Le premier d'attente d'intonations plus appropriées après l'entraînement qu'avant celui-ci, avec une généralisation vers la répétition de phrases nouvelles. Le deuxième était un transfert de compétences vers une tâche plus complexe de lecture à haute voix sans modèle audio. La répartition des adultes normo-lecteurs en trois groupes permettait d'évaluer trois types d'aide associés à l'exercice de répétition : une amorce rythmique Musicale avant les cinq répétitions de chaque phrase (M), ou une amorce sonore non-rythmique (Bruit) associée à un Retour visuel sur les courbe d'intonation de la production et du modèle (BR), ou la combinaison de ces deux aides (MR).

Les principaux résultats de l'Expérience 13 ont apporté quelques arguments en faveur de progrès après les entraînements, essentiellement à travers une accentuation de la descente du F0 en fin de questions ouvertes anglaises dans les tâches de lecture avec ou sans modèle audio. Ce résultat pour la tâche portant sur des phrases n'ayant pas fait l'objet d'un entraînement et produites sans support audio témoigne de transferts de compétence. La tâche de répétition d'un modèle audio était plus proche des exercices d'entraînement, mais les phrases étaient différentes et l'effet particulièrement bénéfique de l'entraînement MR est alors apparu. D'une manière générale, les deux conditions d'entraînement utilisant une amorce rythmique semblaient les plus efficaces, car elles produisaient des améliorations significatives. Le retour visuel seul produisait cependant aussi un effet, certes non-significatif, mais de taille équivalente, ce qui limite l'interprétation des effets des types d'entraînement.

L'objectif étant de proposer aux adultes dyslexiques (Expérience 14) le dispositif potentiellement le plus efficace, ils ont réalisé l'entraînement MR sur la base des indices suggérant son efficacité dans l'Expérience 13. Les résultats apportent des éléments encourageants sur l'efficacité de cette aide pour ce public. Ils n'en profitent pas moins que les normo-lecteurs dans la tâche de production sans modèle audio. Dans la tâche de répétition avec modèle audio, ce dispositif améliore significativement la qualité d'ajustement de leurs productions vocales aux modulations du F0 des questions-modèles ouvertes et fermées, alors que les progrès ne sont pas significatifs pour les questions fermées chez les normo-lecteurs.

Après une présentation et une courte discussion sur le niveau de base des apprenants normo-lecteurs pour la prosodie des questions anglaises, les effets des entraînements seront discutés en fonction des aides associées (amorçage et/ou retour visuel), pour aborder ensuite les difficultés particulières des apprenants dyslexiques et leur sensibilité à la combinaison des

deux aides. L'interprétation portera sur les mécanismes cognitifs sous-jacents aux bénéfices observés et reviendra sur l'ancrage de cette étude dans le cadre théorique initial.

### **7.8.1. Une certaine compétence initiale pour la prosodie des questions anglaises**

Il est étonnant de noter que, avant l'entraînement, les adultes des trois groupes sans pathologie produisaient déjà différemment l'intonation des questions anglaises ouvertes et fermées, y compris en l'absence de modèle audio. Leur faible niveau en anglais (ne dépassant pas le B2) ne les privait de l'acquisition de certaines connaissances en prosodie dans cette langue. Toutefois, la distinction entre l'intonation des deux types de questions était moins intense avant qu'après l'entraînement, comme en témoigne l'interaction entre le type de question et la session. L'origine de cette compétence avant l'entraînement interroge. Selon Flege et Bohn (2021), l'apprenant tardif d'une L2 a tendance à assimiler automatiquement les caractéristiques de la L2 à celles de la L1. Sans que les francophones en soient toujours conscients, certaines questions françaises ont une intonation descendante, et il n'est pas exclu que l'assimilation ait participé à établir une distinction entre deux intonations finales pour les deux types de questions anglaises. Cependant, en français, l'association des intonations descendantes/montantes aux questions ouvertes/fermées est moins claire et systématique qu'en anglais, et l'assimilation peut difficilement tout expliquer. En français, l'intonation descendante est en effet plus représentative des ordres, ou parfois de questions pour lesquelles le locuteur sous-entend qu'il connaît déjà la réponse. Comme l'assimilation peut difficilement expliquer la compétence initiale constatée, une autre explication est à rechercher du côté des apprentissages scolaires et des contextes de loisirs. Les jeunes adultes ont déjà eu des cours d'anglais, et certains ont pu retenir des explications explicites dispensées lors d'un enseignement conventionnel sur la différence prosodique pour les deux types de questions anglaises. Par ailleurs, même si leur niveau est faible (B2 maximum), ils ont depuis des années perçu la langue anglaise, en cours d'anglais, souvent lors de petits séjours linguistiques dans des Îles Britanniques, et en écoutant des séries ou vidéos sur internet en anglais.

### **7.8.2. Apprentissage implicite d'une particularité en prosodie anglaise**

Les résultats Expériences 13 et 14 ont montré qu'une accentuation (voire pour certains, un apprentissage) de la distinction prosodique entre les deux types de questions anglaises est possible à l'âge adulte, conformément à ce que prédit le modèle SLM-r (Flege & Bohn, 2021). Les données permettent ici d'ajouter que c'est possible grâce à une intervention ciblée, même très courte (45 mn). Ce résultat concorde avec un nombre croissant d'expériences

montrant des modifications significatives des compétences en langue seconde à partir d'entraînements brefs (Alves & Luchini, 2017 ; Callan *et al.*, 2003 ; Cebrian & Carlet, 2014 ; Guyot-Talbot *et al.*, 2016 ; Hardison, 2005 ; Hazan *et al.*, 2005 ; Lambacher *et al.*, 2002 ; Shinohara & Iverson, 2018 ; Thomson, 2011). Les résultats de l'Expérience 13 montrent que, quels que soient les dispositifs d'aide ajoutés à l'entraînement, les exercices demandant cinq imitations pour chaque phrase sont suivis d'une amélioration en prosodie. Plus précisément, la différence entre les intonations finales des deux types de questions s'est accrue. Ceci s'est essentiellement produit à travers l'accentuation de la baisse de F0 (i.e. diminution de la hauteur de la voix) à la fin des questions ouvertes. C'est un effet intéressant car c'est justement l'intonation descendante de certaines questions anglaises qui pourrait conduire les apprenants francophones à des erreurs d'interprétation, sur le plan pragmatique, par une confusion possible avec les ordres en français.

Etant donné qu'aucune règle n'était énoncée pendant l'entraînement, l'apprentissage s'est fait ici à travers la perception d'exemples et leur répétition, sur un mode pouvant être qualifié d'implicite. Cela ne signifie pas que les apprenants ne prenaient pas conscience d'une règle, ou ne cherchaient pas à l'inférer délibérément (ils le font sans doute, car habitués à ce type d'exercice typiquement académique). Cela veut dire qu'ils ne sont pas obligés d'avoir un enseignement explicite de la règle pour produire la différence de manière encore plus appropriée. Cette connaissance n'est pas forcément formalisée ni mémorisée comme une règle sur le plan cognitif, mais en tout cas ils appliquent plus nettement la différence entre les intonations des deux types de questions après l'entraînement. Ceci est cohérent avec Derwing *et al.* (2004) selon qui les apprentissages dans le domaine de la prosodie sont particulièrement efficaces s'ils sont basés sur un enseignement implicite plutôt qu'explicite.

Malgré les difficultés posées par l'apprentissage tardif d'une langue étrangère, certaines recherches montrent que des changements majeurs peuvent encore s'établir fondamentalement dans le système de connaissances et de compétences permettant le traitement de la parole dans une nouvelle langue chez des apprenants adultes (Alves & Luchini, 2017 ; Tamminen *et al.*, 2015 ; Wang *et al.*, 2003). Ici, l'entraînement était très ponctuel, et nous n'avons pas d'information sur la pérennité de l'effet. Par la suite, il conviendrait de la mesurer. Il était difficile de le faire dans cette étude, réalisée en partie à distance. De plus, la variabilité inter- mais aussi intra-individuelle concernant la présence d'informations anglophones dans le cadre de vie en pleine évolution de jeunes participants introduit certainement beaucoup de biais dans de telles évaluations à long terme. Evaluer si

l'effet se maintient deviendrait possible si l'entraînement était testé, dans une expérience ultérieure, au sein d'un groupe d'étudiants suivant au moins le même enseignement en anglais. Il deviendrait même alors possible de concevoir, avec les enseignants, les exercices susceptibles de renforcer ensuite certains acquis.

### **7.8.3. Effet d'amorçage musical rythmique sur la prosodie : liens avec les effets sur l'analyse grammaticale**

Plusieurs éléments dans les résultats de l'Expérience 13 ont montré que, parmi les trois conditions proposées, celles utilisant l'amorçage rythmique ou combinant le retour visuel et l'amorçage rythmique étaient les plus propices à l'amélioration de la distinction entre l'intonation des questions ouvertes vs. fermées, ceci en production de phrases. L'amorçage procuré par un rythme régulier issu du domaine musical sur une activité relevant du domaine linguistique est un argument de plus pour le cadre théorique général soutenant l'idée de ressources et de processus communs entre les traitements cognitifs de la musique et du langage (Fiveash *et al.*, 2021 ; Morillon *et al.*, 2015 ; Patel, 2011, 2012 ; Patel & Iversen, 2014). L'utilisation d'un rythme musical était aussi encouragée par les incitations associées au cadre théorique TSF (Goswami, 2011).

Les effets d'amorçage rythmique observés ici sur la prosodie se rapprochent particulièrement de ceux mis en évidence à partir d'un rythme musical identique (2 Hz) sur des tâches impliquant la sensibilité aux erreurs de grammaire dans plusieurs langues, chez des enfants (Bedoin *et al.*, 2016a,b, 2018 ; Canette *et al.*, 2020b ; Chern *et al.*, 2018 ; Fiveash *et al.*, 2020a ; Ladányi *et al.*, 2020 ; Przybylski *et al.*, 2013) comme des adultes (Canette *et al.*, 2019, 2020a) avec ou sans pathologies du langage. En effet, pour la prosodie comme pour la grammaire, sur le plan de l'analyse temporelle, il s'agit d'opérer une séquenciation à travers une large fenêtre temporelle (compatible avec les oscillations cérébrales delta ou thêta). De plus, les unités prosodiques telles que les enveloppes d'amplitude sont souvent décrites comme des indices fondamentaux pour l'analyse grammaticale des phrases (Cumming *et al.*, 2015 ; Gordon *et al.*, 2015). Les adultes exploitent par exemple les indices prosodiques (i.e. rythmiques) pour mieux saisir les organisations syntaxiques ambiguës (Dede, 2010). Ce rapprochement entre l'amorçage à partir d'une pulsation lente sur le traitement grammatical et le traitement prosodique était à l'origine de l'hypothèse d'un effet positif venant d'amorces rythmiques avec une pulsation à 2 Hz sur l'attention portée à des caractéristiques prosodiques des phrases. Après de telles amorces, à travers la répétition dans un exercice, l'extraction au moins implicite de régularités permettant d'associer deux types de questions à deux

intonations distinctes en anglais L2 était attendue. En d'autres termes, d'une manière générale, les amorces à 2 Hz étaient destinées à promouvoir la séquenciation et l'allocation de ressources pour l'analyse en grandes unités et pour la dimension prosodique. Cela semblait envisageable, car l'utilisation du rythme est parfois aussi bénéfique à l'apprentissage de compétences langagières que certaines prises en charge plus focalisées sur la compétence à apprendre (Bigand & Tillmann, 2021, Tillmann & Bedoin (soumis)).

#### **7.8.4. Un amorçage rythmique basé sur la hiérarchie des oscillations**

Cette étude recueillait uniquement des données comportementales. Des enregistrements EEG avec des analyses temps-fréquence seraient nécessaires pour en savoir plus sur les phénomènes liés aux oscillations cérébrales ayant peut-être joué un rôle dans l'amorçage rythmique utilisé. Le cadre théorique qui a abouti à la réalisation de cette recherche et qui a guidé les hypothèses ainsi que le choix du matériel était avant tout celui de l'attention temporelle, conçue comme déployée de façon cyclique dans le modèle de l'attention dynamique de Jones (1976, 2008 ; Jones & Boltz, 1989 ; Jones *et al.*, 2002). Ce cadre théorique intégrait aussi des notions de neuroscience cognitive, présentées dans le Chapitre 1, et évoquant une cascade d'oscillations cérébrales, des phénomènes de synchronisation entre ces oscillations et avec des rythmes externes pour expliquer les traitements temporels du signal musical ou de parole. Nous proposons ici de s'intéresser à cette perspective pour envisager comment l'amorçage rythmique proposé a pu prendre part aux progrès réalisés en prosodie par les participants. Les interprétations seront très spéculatives et assumées comme telles. Prendre en compte le support neuronal fonctionnel des mécanismes cognitifs impliqués dans l'analyse temporelle en prosodie est évidemment trop ambitieux pour une recherche basée sur les données comportementales. Cela paraissait toutefois cohérent avec le cadre théorique évoqué tout au long de cette thèse.

L'amorçage proposé ici avait la particularité de ne pas associer terme à terme les caractéristiques rythmiques précises d'une amorce et d'une cible, mais de créer un amorçage plus global. Celui-ci était destiné à déclencher ou renforcer des oscillations cérébrales lentes, de manière à prédisposer l'attention temporelle de façon favorable à l'analyse de grandes unités pour la prosodie dans le signal de parole. Il était aussi censé renforcer la cohérence entre ces oscillations lentes et plus rapides, pour guider l'extraction précise d'indices plus brefs, comme le début de changements dans les enveloppes de sons. Cet amorçage global était supposé déclencher ces états attentionnels favorables, cela même quelques instants après la fin de l'amorce, pour un retentissement sur le traitement de plusieurs phrases. Ce principe



d'amorçage est celui qui a été utilisé dans d'autres expériences pour tenter d'améliorer le traitement grammatical d'une série de plusieurs phrases (Bedoin *et al.*, 2016 ; Cannette *et al.*, 2019, 2020a, b ; Chern *et al.*, 2018 ; Fiveash *et al.*, 2020a, b ; Ladanyi *et al.*, 2021 pour une synthèse, voir Fiveash *et al.*, 2021 ou Tillmann & Bedoin, soumis).

Le choix d'un rythme à 2 Hz dans les amorces musicales était destiné à prédisposer la distribution cyclique de l'attention pour un découpage du signal selon des rythmes assez lents pour la prosodie. Le support musical avait l'avantage de permettre de diffuser un tel rythme de façon saillante, extrêmement régulière et plus facile à extraire que dans la parole. Les fenêtres d'analyse temporelle longues pré-ajustées grâce à ces amorces devaient élever une sensibilité à la récurrence des unités prosodiques. Elles seraient guidées par les oscillations delta, qui permettent des groupements temporels (*chunking*) en de larges unités compatibles avec les analyses prosodiques et syntaxiques (Boucher *et al.*, 2019). Ces rythmes cérébraux chercheraient activement à se synchroniser au rythme prosodique (Ghitza, 2017). Dans les Expériences 13 et 14, l'attention devait être ainsi globalement attirée par les amorces vers la présence d'une organisation au niveau d'analyse ciblée dans l'étude : la prosodie.

Le renforcement des oscillations delta favoriserait aussi de façon générale les prédictions temporelles (Zoefel, 2018), ce qui est globalement favorable à toute analyse du signal de parole. De manière implicite, cela a pu jouer un rôle : les questions ouvertes en anglais commencent par des mots comme *what, where, when, how*, et si les anticipations sont renforcées de manière générale, ils peuvent déclencher la prédiction de la survenue d'un événement final : une pente descendante de la hauteur de voix.

Lorsqu'une information est temporellement prédite, elle serait alors aussi encodée de manière plus précise pour toutes ses dimensions, notamment la hauteur (Jones *et al.*, 2002). Herbst et Obeser (2017) ont par exemple montré que la hauteur (aspect spectral) d'une note de musique était plus vite traitée si cette note pouvait être bien prédite sur le plan temporel à partir de la régularité des intervalles entre les notes précédentes. En améliorant cette prédictibilité par la stimulation du rythme delta, l'amorce rythmique dans les Expériences 13 et 14 a pu permettre un traitement plus précis des modulations du F0 dans une unité en fin de phrase.

Ce rythme delta (0,5-4 Hz) est aussi décrit comme présentant des liens étroits avec le rythme thêta (4-8 Hz), ce qui a pu participer à l'effet positif des amorces à 2 Hz dans cette étude. Dans les questions utilisées pour les Expériences 13 et 14, le changement de hauteur de



la voix en fin de phrase correspondait essentiellement à une modulation du F0 sur une syllabe. La détection d'un tel évènement est plus adaptée au rythme thêta, connu pour sa compatibilité avec les syllabes (Giraud *et al.*, 2007 ; Morillon *et al.*, 2010 ; Peele & Davis, 2012), qu'au rythme delta. Etant donné que la pulsation des amorces correspondait à la bande delta, cela pose question. Toutefois, deux éléments pourraient expliquer l'efficacité de ce rythme pour le traitement de modulations survenues sur une durée syllabique dans le signal.

Tout d'abord, les oscillations delta et thêta partageraient certaines propriétés. Lorsque surviennent des évènements auditifs saillants, comme les temps forts du rythme des amorces musicales, l'ensemble des oscillations lentes (delta et thêta) augmenterait leur tendance à s'adapter au rythme de l'information entrante (Gross *et al.*, 2013). La netteté des fluctuations temporelles dans les amorces inciterait ces oscillations à s'engager de façon générale dans une recherche de synchronisation (*entrainment*) au rythme des syllabes.

Ensuite, Gross *et al.* (2013) expliquent que les évènements abrupts (e.g. un changement brusque de hauteur en fin de question) augmentent aussi le couplage entre des rythmes cérébraux différents, favorisant la séquenciation du signal à des échelles multiples et cohérentes. C'est pourquoi, même si les oscillations delta pourraient être les plus adaptées à la pulsation des amorces, leur renforcement retentirait sur les oscillations thêta, adaptées aux syllabes porteuses de l'intonation finale dans les questions. Cela serait possible à travers les modulations phase-amplitude exercées par les oscillations delta sur les oscillations thêta. Delta et thêta rechercheraient dans le signal des éléments distincts et n'entreprendraient pas de relation d'emboîtement de phases (Cogan et Poeppel, 2011), mais la modulation d'amplitude exercée par delta sur thêta dans les aires temporelles permet aux premières et transférer leur renforcement aux secondes. Pour ces différentes raisons, stimuler les cycles attentionnels à 2 Hz pourrait non seulement accroître la sensibilité à l'organisation prosodique, mais aussi favoriser la séquenciation à des rythmes un peu plus rapides, compatibles avec la durée de changements intonatifs sur des syllabes finales, grâce aux relations en cascades des rythmes cérébraux. De façon très spéculative, il peut être supposé qu'en facilitant l'extraction des syllabes, de façon indirecte, l'amorce rythmique à 2 Hz pourrait rendre les apprenants plus attentifs à la dernière syllabe et à son contenu. Bien extraite dans ses limites temporelles, celle-ci pourrait ressortir comme clairement différente des précédentes, sur le plan de la fréquence fondamentale qui monte ou baisse dans les questions.

Il est d'ailleurs possible qu'un traitement raffiné de l'information au sein de la pente formée par le changement de hauteur soit accentué par le rythme cérébral bêta, car il est à la

fois l'objet de modulations d'amplitude par le rythme delta (Arnal *et al.*, 2015), mais aussi de modulation de phase de la part du rythme thêta, et il permettrait une analyse plus fine interne à la syllabe (Ghitza, 2011). Le renforcement de bêta, envisageable dans la dynamique oscillatoire à partir des amorces à 2 Hz, est aussi décrit comme présentant l'avantage de concourir au maintien des relations hiérarchiques, ce qui est fondamental pour la durée de l'effet d'amorçage temporel sur la parole qui suit la musique. L'effet d'amorçage rythmique musical observé dans les expériences trouve ainsi sa place dans les travaux des neurosciences cognitives sur la cascade oscillatoire et dans les modèles d'analyse multi-échelles (Benitez-Burraco & Murphy, 2019 ; Giraud & Poeppel, 2012 ; Hickok & Poeppel, 2007 ; Poeppel, 2003).

#### **7.8.5. Un dispositif adapté aux apprenants les plus en difficulté**

L'étude des corrélations a fait ressortir une particularité de l'apprentissage réalisé dans la condition associant amorçage rythmique et retour visuels par rapport à celle proposant seulement l'amorçage rythmique. Selon les corrélations, dans la condition combinant les deux aides, l'amélioration de l'imitation de l'intonation après l'entraînement était particulièrement efficace pour les apprenants qui étaient en grande difficulté à T1 avec l'intonation des questions anglaises. Cela constitue un argument pour appuyer des recommandations d'utilisation auprès de publics d'apprenants adultes de niveaux variés. La combinaison des deux aides est particulièrement préconisée pour les plus en difficulté sur le plan prosodique.

Si l'amorçage rythmique seul suffisait à améliorer significativement certaines performances, ce n'était pas le cas du retour visuel seul. Cette différence est à interpréter avec prudence, car elle n'était pas confirmée par l'étude de la taille des effets. Le retour visuel est un type d'outil ayant déjà fait ses preuves en orthophonie, y compris sous la forme de programmes courts, mais tout de même sur plusieurs séances (Fabre *et al.*, 2016 ; Girod-Roux *et al.*, 2019). Il s'agit le plus couramment d'un retour d'information sur l'aspect articulatoire de la production. Dans la présente étude, un point original était de donner une information sur le changement du F0, matérialisé par une courbe, avec peu d'explication sur ce qu'elle représentait pour ne pas biaiser l'attention trop explicitement vers l'intonation. Ce principe de retour visuel sur l'information acoustique est inspiré de rares travaux ayant mis au point une technique du même type. Guyot-Talbot *et al.* (2016) ont en effet entraînés des adultes à répéter des questions ouvertes et fermées anglaises en donnant eux aussi systématiquement les deux courbes d'intonation : celle du modèle et celle de l'imitation. Un effet bénéfique a été obtenu dans leur étude, mais il était, comme ici, de petite taille. Pour renforcer l'effet de ce

dispositif, notre condition avec retour visuel demandait cinq répétitions, mais cela n'a pas suffi à produire un progrès important chez les adultes de notre étude. La combinaison de cette aide à celle apportée par l'amorce rythmique était en revanche fructueuse.

Notons que les participants aux expériences de cette thèse avaient un niveau d'anglais assez faible, alors que ceux de Guyot-Talbot *et al.* (2016) étudiaient à l'Université pour devenir enseignants en anglais. S'ils présentaient de réelles lacunes en prosodie, ils étaient tout de même bien différents de nos participants et avaient sans doute été davantage exposés à cette langue. Or, nous avons vu que le niveau de performance de base en répétition de l'intonation était lié au bénéfice tiré de l'entraînement combinant retour visuel et amorçage rythmique. Certes, dans cette dernière condition, l'aide n'était pas seulement procurée par le retour visuel, mais elle y prenait part. La question de l'adéquation des types d'aide proposés en fonction du niveau initial est plus que jamais d'actualité.

#### **7.8.6. Association efficace de l'amorçage rythmique et du retour visuel dans le cadre théorique PRISM**

L'accroissement de la différence entre les deux types de questions produites après l'entraînement était surtout dû à l'accentuation de la baisse du F0 en fin de question ouverte, ce qui correspond à l'acquisition de ce qui est le moins familier pour un francophone. Ce changement, tout à fait positif, s'est produit de manière significative dans les groupes ayant bénéficié d'une amorce musicale rythmique.

Ce résultat est particulièrement compatible avec l'explication proposée par le modèle PRISM pour rendre compte des effets bénéfiques du rythme musical sur le traitement du langage et comprendre les véritables mécanismes qui sous-tendent ces phénomènes (Fiveash *et al.*, 2021). Selon ce cadre théorique, les points communs entre le traitement du rythme de la musique et celui du langage sont 1) de reposer sur des traitements auditifs précis, 2) d'impliquer une synchronisation des oscillations cérébrales au rythme externe, 3) d'engager un couplage sensori-moteur.

Pour le point 1, dans les Expériences 13 et 14 le rythme des amorces musicales était extrêmement régulier, ce qui devait faciliter la précision du traitement temporel, et la pulsation à 2 Hz des amorces était d'une fréquence adaptée à la durée d'enveloppes dans la parole, ce qui devait permettre de les isoler ensuite avec précision pour l'analyse prosodique.

Pour le point 2, la synchronisation supposée des oscillations cérébrales delta avec le rythme à 2 Hz du signal musical devait s'établir facilement à partir de la tendance du rythme endogène delta à ajuster sa phase au rythme externe (Giraud & Poeppel, 2012 ; Gross *et al.*, 2013). A travers la cascade des oscillations cérébrales synthétisée par la figure 3, et particulièrement la modulation phase-amplitude exercée par le rythme delta sur les rythmes thêta (Gross *et al.*, 2013) et bêta (Arnal *et al.*, 2015), des fenêtres d'analyse temporelle auraient pu se pré-ajuster pour bien traiter, dans le signal de parole présenté ensuite, les portions de signal telles que les syllabes. Dans les Expériences 13 et 14, celles-ci portaient le changement de hauteur de voix à traiter avec précision pour sa direction, caractéristique d'une question ouverte ou fermée.

Pour le point 3, le couplage sensori-moteur évoqué par le cadre théorique correspond à l'activation d'aires cérébrales motrices par le signal de parole, régions d'où proviennent des rythmes endogènes lents capables de renforcer de façon top-down la dynamique oscillatoire par ailleurs évoquée par le signal. Ces influences descendantes seraient favorables à l'analyse temporelle précise et à échelles multiples de la parole (d'Ausilio *et al.*, 2009 ; Morillon *et al.*, 2010). Dans l'Expérience 13, l'analyse des données a montré que l'ajout du retour visuel à l'entraînement avec amorçage rythmique a augmenté l'effet positif du dispositif sur l'épreuve de répétition avec modèle audio. Autour du point 3 de PRISM, il est possible d'imaginer que des couplages sensori-moteurs aient été renforcés par l'ajout d'une information dans une modalité sensorielle supplémentaire – la vision – véhiculant une information sur ses aspects acoustiques, eux-mêmes liés aux composantes articulatoires (i.e. motrices) car issues de celles-ci. L'information visuelle fournie par le dispositif est en effet en accord avec l'aspect temporel de l'information acoustique, puisque les courbes représentent aussi le temps. Même s'il n'y a pour l'instant aucune preuve directe pour cette interprétation, le caractère audiovisuel du dispositif MR est sans doute le plus apte à accroître l'implication vertueuse des couplages sensorimoteurs, eux-mêmes déjà stimulés, dans cette condition, par les amorces musicales rythmiques.

### **7.8.7. Un dispositif d'amorçage et de retour visuel adapté aux apprenants dyslexiques**

Comme chez les normo-lecteurs, l'entraînement a permis des améliorations de l'intonation des phrases anglaises chez les dyslexiques. A l'âge adulte, il apparaît que les dyslexiques, malgré leurs difficultés de traitements des indices prosodiques longs (Goswami, 2011), sont capables de progresser dans un domaine qui nécessite d'extraire des unités assez durables

dans le signal de parole. En effet, dans l'épreuve de répétition avec modèle, l'intonation de leurs phrases devient plus proche de celles des modèles après l'entraînement. Ce progrès est d'autant plus intéressant qu'il concerne la capacité à produire des changements de F0, puisque l'amélioration s'applique seulement aux questions et non aux phrases dont l'intonation est plate. De plus, la réduction de la distance entre l'intonation entendue et l'intonation reproduite s'améliore chez eux aussi bien pour les questions ouvertes que pour les questions fermées, alors que le progrès est restreint aux questions ouvertes pour les adultes sans pathologie. La pente de la partie intonative finale des questions produites par les adultes dyslexiques, quasiment inexistante avant l'entraînement, présente après celui-ci la configuration classique des anglophones, avec une intonation montante pour les questions fermées et véritablement descendante pour les questions ouvertes.

Le dispositif associant amorçage rythmique et retour visuel qui leur a été proposé pourrait avoir attiré leur attention sur ce qui distingue ces types de questions, probablement en aidant à segmenter d'une manière appropriée à la prosodie, et sans doute en déclenchant la chaîne des oscillations/rythmes attentionnels emboîtés permettant de traiter de façon précise le contenu des syllabes. Le traitement de la parole s'est en tout cas, suffisamment bien réalisé pendant l'entraînement, et le retour visuel a sans doute aussi pallié des incertitudes, pour qu'un vrai changement en prosodie prenne place. Une restriction est à faire concernant l'interprétation pour l'Expérience 14. Le choix de proposer à tous les dyslexiques de l'échantillon le dispositif le plus prometteur pour les aider en prosodie prive pour l'instant de la possibilité de pousser plus loin la discussion sur ce qui leur est bénéfique dans l'entraînement.

Pour le public dyslexique, le progrès mesuré est prometteur pour ce qui est de la combinaison des deux aides (amorçage rythmique et retour visuel). C'est encourageant, car les personnes atteintes de cette pathologie déclarent souffrir de plus grandes difficultés que les autres en apprentissage tardif de l'anglais (Mazure-Palandre *et al.*, 2016). Ce résultat chez les adultes dyslexiques est en accord avec la corrélation observée dans l'Expériences 13, en dehors de la pathologie, entre l'intensité des progrès en intonation des questions anglaises et le faible niveau initial en prosodie anglaise.

Cependant, ce progrès traduit un apprentissage qui est encore fragile car, lorsqu'il s'agit de mobiliser une règle pour produire les questions ouvertes et fermées sans modèle audio (tâche de lecture), les dyslexiques ont amélioré leur production uniquement pour les phrases avec une intonation descendante. Cette limite peut être due à la brièveté de l'entraînement,

correcte pour les apprenants sans pathologie mais trop court en cas de dyslexie pour un transfert aussi éloigné. En effet, cette pathologie est caractérisée par une anomalie des oscillations lentes (Di Liberto *et al.*, 2018 ; Hämäläinen *et al.*, 2012 ; Lehongre *et al.* 2013), ce qui pourrait avoir créé une résistance à l'efficacité de la stimulation par amorçage rythmique. Selon Goswami (2012), des anomalies de la synchronisation de l'attention – et des oscillations cérébrales sous-jacentes – à un rythme lent (2 Hz) chez les dyslexiques rendraient difficile le traitement de l'enveloppe d'amplitude, socle de la prosodie (Goswami, 2011). L'entraînement étant efficace chez des normo-lecteurs, il se pourrait que les adultes dyslexiques aient simplement besoin d'un programme plus long, distribué sur plusieurs jours. Des progrès mieux transférés à diverses situations, dont celle de production sans modèle audio, seraient en accord avec les études montrant les bienfaits linguistiques d'un entraînement musical (i.e. prise en charge plus longue) chez des personnes atteintes de troubles du langage (François *et al.*, 2013) particulièrement lorsqu'il est centré sur le rythme (Flaugnacco *et al.*, 2015 ; Overy, 2008).

#### **7.8.8. Limites et perspectives pour des dispositifs d'aide à l'apprentissage brefs**

Cette étude a montré qu'un programme d'apprentissage ciblé sur un aspect prosodique de l'anglais L2 permettait des progrès chez de jeunes adultes normo-lecteurs ou dyslexiques. La simple répétition de phrases ne suffit sans doute pas mais, accompagnée de dispositifs pré-ajustant l'attention et les rythmes cérébraux de façon favorable à l'analyse temporelle nécessaire, elle peut générer des progrès. Le programme étant court (40 min), il pourrait trouver facilement sa place dans l'enseignement de l'anglais L2 pour des aspects résistant à l'enseignement explicite.

Une limite essentielle à cette étude exploratoire est l'incertitude quant à la durée des effets. Ce point est particulièrement sensible pour un entraînement aussi court. Dans le cas où l'on douterait qu'une régularité perçue, traitée et produite pendant 45 minutes, dans des conditions faisant tout pour favoriser l'apprentissage, puisse vraiment donner lieu à une connaissance stable, il conviendrait de proposer des moyens de pérenniser les progrès. Cela peut inclure une réflexion conduite avec les enseignants, pour bénéficier de leur expertise. Des explications explicites verbales et appuyée par des illustrations visuelles, comme celles des exercices ou plus proches de la réalité acoustique du signal (e.g. spectrogrammes) pourrait susciter l'intérêt d'apprenants adultes et les motiver. Cela leur permettrait de développer au sujet du point traité une connaissance explicite consciente. Une autre piste est d'imaginer de proposer ce type d'exercice assez régulièrement et de mesurer le renforcement de

l'apprentissage au fil du temps, ce qui serait possible dans le cadre d'une étude conduite avec un même groupe pendant un semestre. Le petit progrès observé pourrait aussi être présenté aux apprenants et utilisé pour stimuler leur motivation et les rassurer sur le fait que les adultes peuvent progresser en anglais L2, même s'ils sont atteints de dyslexie.

## Chapitre 8. Discussion générale

---

### 8.1. Rappel des objectifs, des choix théoriques et justification des choix méthodologiques

L'objectif principal de la thèse était la création et l'évaluation d'outils numériques pédagogiques pour aider à l'apprentissage en anglais L2 chez de jeunes adultes normo-lecteurs et dyslexiques.

Pour cela, nous avons élaboré et testé les effets de programmes d'entraînements perceptifs pour la phonologie et un entraînement par répétition pour la prosodie en anglais. Pour contribuer à ce domaine de recherche appliqué, nous avons choisi de concevoir des programmes d'exercices brefs, et d'ajouter des dispositifs de stimulation auditifs et visuels susceptibles d'optimiser leurs effets. L'influence de ces ajouts a été évaluée.

Plusieurs défis devaient être relevés. Le premier était de tester l'effet des entraînements auprès d'adultes normo-lecteurs, mais aussi d'adultes dyslexiques. Les compétences ciblées relevaient de domaines souvent déficitaires chez les personnes dyslexiques. Il s'agissait en effet d'apprendre de nouveaux phonèmes alors que leur système phonologique est souvent déficitaire (Dufor *et al.*, 2007 ; Fosker & Thierry, 2000 ; Hanley, 1997 ; Noordenbos *et al.*, 2013 ; Noordenbos & Serniclaes, 2015 ; Svensson & Jacobson, 2006 ; Van Beinum *et al.*, 2005 ; Vandermosten *et al.*, 2010 ; Witton *et al.*, 1998 ; Zabell & Everatt, 2002). De plus, les nouveaux phonèmes à apprendre se distinguaient par leurs caractéristiques de durée, et les personnes dyslexiques présentent des déficits des traitements temporels (Goswami *et al.*, 2002 ; Habib, 2021 ; Habib *et al.*, 2002 ; Leong & Goswami, 2014b ; Soltész *et al.*, 2013 ; Tallal *et al.*, 1996 ; Van Hirtum *et al.*, 2019b). Il s'agissait aussi d'apprendre des particularités de la prosodie anglaise, alors que les personnes dyslexiques présentent des difficultés d'ajustement de l'attention aux enveloppes de parole qui permettent de traiter la prosodie (Falk *et al.*, 2015 ; Falter & Noreika, 2014 ; Goswami, 2011 ; Peter & Stoel-Gammon, 2008 ; Van Hirtum *et al.*, 2019). Pour augmenter les chances de les aider, une réflexion était nécessaire sur ce qui pouvait rendre plus efficaces des entraînements perceptifs auditifs tels que ceux déjà proposés par des chercheurs, en français (Collet *et al.*, 2017) et en anglais (Cerbrian & Carlet, 2014 ; Guyot-Talbot *et al.*, 2016 ; Hu *et al.*, 2016 ; Krzonowski *et al.*, 2016 ; Lengeris, 2018 ; Shum *et al.*, 2021 ; Tamminen *et al.*, 2015). L'envie de proposer quelque chose d'original ne devait pas dispenser d'être attentive à la validité théorique des entraînements et des épreuves créées pour les évaluer. C'est pourquoi le choix des dispositifs



était guidé par les propositions de plusieurs modèles en sciences cognitives, à l'interface de la psychologie cognitive autour de l'attention temporelle, des neurosciences avec les oscillations cérébrales, de la neuropsychologie pour la dyslexie, et bien sûr de la phonétique et de la phonologie. S'adresser à des adultes constituait un autre défi, car les modifications corticales fonctionnelles seraient plus difficiles à provoquer que chez l'enfant, et les habitudes des adultes ont eu plus de temps pour créer et renforcer des automatismes inadaptés. Sur le plan pratique, les tests et les entraînements devaient enfin pouvoir être utilisés à domicile et répondre à des exigences de brièveté pour de futures utilisations dans des séances de cours en présentiel ou en distanciel.

Le premier dispositif imaginé était inspiré de recherches sur l'écoute dichotique. Il ne s'agissait pas d'écoute dichotique classique avec deux mots présentés simultanément. Un mot était présenté dans une oreille en concurrence avec un bruit blanc dans l'autre oreille, pour augmenter le poids de l'association privilégiée entre une oreille et l'hémisphère controlatéral, comme le font aussi les expériences classiques en écoute dichotique. Ce type de stimulation dichotique a été utilisé et comparé à une stimulation binaurale (le mot était alors diffusé dans les deux oreilles) lors d'exercices visant à faire apprendre de nouveaux contrastes phonémiques basés sur la durée. Dans les **Expériences 2, 3, 6, 7, 8 et 9**, il s'agissait de la durée du VOT pour le voisement des consonnes occlusives sourdes vs. sonores françaises et anglaises. L'assimilation de l'anglais à la règle française pour les durées de VOT conduit à des erreurs dont il s'agissait de détourner les apprenants. Les **Expériences 11 et 12** visaient à améliorer le traitement des voyelles anglaises /ɪ/ et /i:/ opposées par la durée, opposition phonologique nouvelle pour des francophones. Le principe était d'adresser plus directement les phonèmes porteurs d'un indice phonétique court à l'HG et les phonèmes porteurs d'un indice phonétique long à l'HD. Cette dissociation était réalisée en vue de tirer parti de la spécialisation du cortex auditif gauche pour la séquenciation du signal à une échelle temporelle très fine, et de celle du cortex auditif droit pour la séquenciation à une échelle plus large, hypothèse centrale du modèle AST, Poeppel, 2003 ; Poeppel *et al.*, 2008).

L'idée de ce dispositif s'appuyait sur un cadre théorique développé en psychologie cognitive autour de l'attention conçue comme diffusée de façon rythmique (théorie DAT, Jones, 1976 ; Barnes & Jones, 2000 ; Jones *et al.*, 2017) et les modèles de traitement temporel à échelles multiples du signal de parole tels que AST (Poeppel, 2003 ; Poeppel *et al.*, 2008). D'après le modèle neurolinguistique AST, les multiples fenêtres d'attention temporelle permettant d'analyser le signal de parole à plusieurs échelles reposeraient sur les oscillations

cérébrales étudiées en neurosciences cognitives. Une cascade d'oscillations cérébrales à des fréquences diverses, capables de s'inter-synchroniser et de se synchroniser au rythme du signal externe, déterminerait la taille des fenêtres d'analyse temporelle dans les hémisphères cérébraux complémentaires au niveau des cortex auditifs (Giraud & Poeppel, 2012 ; Hickok & Poeppel, 2000, 2007 ; Poeppel *et al.*, 2008). Dans cette recherche, il s'agissait de faire en sorte que les propositions d'aide soient cohérentes avec ce contexte théorique. Cette thèse ne prétend pas relever des neurosciences, et les résultats ne sont pas des analyses d'oscillations cérébrales. Les données sont comportementales et en potentiels évoqués (ERPs). Ce contexte théorique a été néanmoins exploré et décrit pour s'assurer, plus modestement, que les choix faits pour les entraînements étaient cohérents avec lui. Cela devait permettre d'inscrire cette recherche dans les idées développées en sciences cognitives autour d'un fonctionnement cérébral expliqué en termes de rythmes, d'oscillations et de synchronisations cérébrales (Benitez-Burraco & Murphy, 2019). Le cadre théorique TSF (Goswami, 2011) qui suppose à l'origine de la dyslexie des troubles de l'attention temporelle et des difficultés de synchronisation des oscillations cérébrales incitait aussi à procéder ainsi. C'est pourquoi un cadre théorique homogène a été détaillé dans le **Chapitre 1**, offrant un support aux expériences réalisées auprès des adultes normo-lecteurs et dyslexiques.

Le deuxième dispositif imaginé a été testé dans les **Expérience 13 et 14**. Il s'appuyait sur le même cadre théorique : celui du rôle de la synchronisation de l'attention temporelle aux rythmes externes pour traiter la parole avec précision. Cette fois, il s'agissait de stimuler directement les compétences rythmiques afin de favoriser une synchronisation attentionnelle favorable au traitement d'enveloppes de parole déterminant la prosodie. Pour cela, des amorces rythmiques musicales ont été utilisées en vue de renforcer l'efficacité d'un entraînement à la prosodie anglaise. Leur impact a été comparé, et aussi combiné, à celui d'une aide par retour visuel sur des données acoustiques (Guyot-Talbot *et al.*, 2016). Cette étude évoque les liens entre langage et rythme (Habib, 2021 ; Habib *et al.*, 2016 ; Tillmann & Bedoin, soumis).

Un dernier choix a été fait, sur le plan méthodologique. La plupart des épreuves évaluant les effets d'entraînement mesuraient des performances. Elles ont été élaborées spécifiquement pour cette recherche. Des expériences étaient dédiées à l'évaluation de certaines de leurs qualités psychométriques, en français (**Expérience 1**) et en anglais (**Expériences 4, 5, 10 et 13**), car il semblait important de dissocier les effets d'entraînement de ceux dus à la répétition des testings. Ils présentent des imperfections, mais que nous

préférences cerner et souligner. Les plus pertinentes pourront ainsi faire l'objet d'élaboration de normes pour produire des tests répondant aux besoins dans les domaines de l'éducation, de la formation et de l'orthophonie.

La discussion générale synthétisera en premier lieu les principaux résultats du travail de thèse en lien avec les hypothèses générales. Des rapprochements entre les différentes études permettront ensuite de préciser les apports au domaine de recherche, et de les discuter dans le cadre des théories qui les ont inspirées. Nous finirons par discuter des limites et des perspectives envisagées pour ce travail, tant sur le plan de la poursuite de travaux en recherche fondamentale que sur celui des applications.

## **8.2. Synthèse de la confrontation des hypothèses générales aux principaux résultats**

Les résultats de cette thèse ont apporté des arguments en faveur de l'efficacité d'entraînements brefs pour le renforcement du système phonologique de la L1 et l'apprentissage d'une L2. Certains effets sont imputables à l'entraînement perceptif (en phonologie) ou en production (prosodie), et sont à distinguer de ceux apportés par les dispositifs complémentaires de stimulation dichotique, d'amorçage rythmique et de retour visuel.

Les entraînements perceptifs pour l'apprentissage des nouveaux contrastes en L2 chez les adultes ont été suivis de progrès dépassant les effets test-retest, qu'il s'agisse de consonnes (Expériences 6, 7, 8, 9) ou de voyelles (Expériences 11 et 12), ce qui valide l'**Hypothèse générale 1**. C'est aussi le cas chez les adultes atteints de dyslexie, que ce soit pour le renforcement de leur système phonologique en L1 (Expériences 2 et 3) ou pour l'apprentissage d'une L2 (Expériences 8 et 12), conformément à l'**Hypothèse générale 4**. En ce qui concerne l'apprentissage du contraste de voisement pour des consonnes anglaises, la stimulation dichotique a optimisé les effets du programme d'entraînement par rapport à la stimulation binaurale chez des adultes normo-lecteurs et dyslexiques (Expériences 6, 7, 8, 9). Ce mode de présentation dichotique, dissociant la stimulation prioritaire des cortex auditifs gauche et droit en fonction de leurs compétences supposées pour traiter des indices acoustiques, respectivement courts et longs a aussi été utilisé avec succès pour la remédiation du traitement catégorielle de consonnes sourdes et sonores françaises (en L1) en cas de dyslexie (Expériences 2 et 3). Ce principe était basé sur l'idée centrale du modèle AST (Poeppel, 2003) : des fenêtres d'attention temporelle optimales pour le traitement de

phonèmes courts et longs seraient respectivement plus disponibles dans les cortex auditifs gauche et droit. Les résultats apportent ainsi des arguments en faveur de ce modèle, et donc de l'**Hypothèse générale 3**. Cependant, la stimulation dichotique a montré des limites. Peu d'éléments dans les résultats chez les adultes normo-lecteurs permettent de soutenir qu'elle augmentait l'impact de l'entraînement perceptif sur l'apprentissage de nouvelles voyelles anglaises opposées par la durée (Expérience 11). En revanche, les adultes dyslexiques ont, quant à eux, davantage progressé après l'entraînement si celui-ci avait été dispensé de manière dichotique (Expérience 12). L'**Hypothèse générale 2** est ainsi partiellement validée.

L'amorçage rythmique musical est un autre dispositif utilisé dans cette recherche pour accroître les effets d'un entraînement, cette fois basé sur des exercices de répétition et non exclusivement perceptif (Expériences 13 et 14). L'amorce rythmique était destinée à prédisposer l'attention à se synchroniser temporellement à de larges unités du signal de parole présenté ensuite. Cela devait permettre que celles-ci soient traitées avec une précision suffisante pour l'analyse de modulations de hauteur sous-jacentes à l'intonation finale de questions. Dans les amorces, la musique avait une pulsation à 2 Hz, rythme compatible avec les oscillations cérébrales delta. L'intonation finale des questions portait sur une syllabe, dont la durée devrait être mieux saisie par des oscillations thêta. Toutefois, un phénomène de cascade oscillatoire détaillé dans le Chapitre 1 est censé permettre aux oscillations delta de moduler l'amplitude des oscillations thêta, ce qui permettait de faire l'hypothèse d'une amélioration de l'extraction des syllabes. Les oscillations thêta modulant à leur tour l'amplitude des oscillations gamma, la cascade oscillatoire renforcée par les amorces à 2 Hz pourrait aussi favoriser un traitement précis des indices acoustiques intra-syllabiques par les oscillations gamma. L'hypothèse était que l'intonation montante des questions fermées et descendante pour les questions ouvertes pourrait être ainsi davantage prise en compte au cours des exercices de répétition. Les résultats ont montré que ces amorces rythmiques renforcent davantage les bénéfices apportés par l'entraînement portant sur l'intonation des deux types de questions en anglais chez des adultes normo-lecteurs (Expérience 13) ou dyslexiques (Expérience 14). Ces effets sont par ailleurs renforcés si un retour visuel permet en plus aux participants de comparer la courbe d'intonation du modèle à celle de sa propre répétition, et ce, même s'ils étaient dyslexiques. Ces résultats valident les **Hypothèses générales 5 et 6**.

### **8.3. Apprentissage de phonèmes anglais par des adultes francophones normo-lecteurs : origine et dépassement des difficultés**

#### **8.3.1. Origine des difficultés**

L'apprentissage tardif de l'anglais est particulièrement difficile pour les adultes francophones. Les similitudes apparentes de ces deux langues induisent l'utilisation massive de l'assimilation de phonèmes à ceux de la L1, selon les modèles d'apprentissage PAM (Best, 1995 ; Best *et al.*, 2001) et SLM-r (Flege, 1995a ; Flege & Bohn, 2021) (Expériences 5). Assimiler le traitement de l'anglais au français peut conduire à diverses erreurs. Deux étaient étudiées dans cette recherche.

La règle de voisement opposant les consonnes occlusives sourdes aux occlusives sonores françaises conduit à associer le trait sourd à un court VOT et le trait sonore à un long VOT, alors que l'inverse doit être fait pour les occlusives équivalentes en anglais. L'assimilation à la règle française pour le voisement conduit en particulier à percevoir les consonnes occlusives sonores anglaises comme sourdes (e.g. le /b/ anglais contient un VOT court et ressemble au /p/ français). Pour dépasser cette difficulté, de nouvelles routines doivent être adoptées. Elles se mettraient en place moins spontanément que pour des aspects complètement nouveaux dans une L2. Ici, la proximité apparente entre le voisement en français et en anglais, renforcée par l'usage de mêmes graphèmes pour des consonnes équivalentes dans les deux langues, peut retarder l'établissement de nouvelles routines. De nouveaux traitements en phonologie resteraient cependant possible toute la vie selon le modèle SLM-r (Flege & Bohn, 2021) et celui de la Perception Sélective Automatique (Strange, 2011). Le cas du voisement anglais pour des francophones est toutefois tellement source d'ambiguïté qu'il paraissait souhaitable d'imaginer un programme d'aide pour le dépasser. Une réflexion sur la routine de sélection perceptive en français s'imposait, en prenant en compte ce qu'elle implique sur le plan cérébral. La théorie AST (Boemio *et al.*, 2005 ; Poeppel, 2003 ; Poeppel *et al.*, 2008) suppose une dominance de l'HG pour le traitement des indices acoustiques courts, et une dominance hémisphérique droite pour celui des indices longs. Le traitement du voisement en français s'accompagnerait de l'établissement d'une routine neurocognitive associant le traitement du trait sonore à l'HD et celui du trait sourd à l'HG. Un traitement optimal des traits sonore et sourd en anglais demande des associations opposées. Il est compréhensible qu'un tel renversement soit difficile à opérer, et le dispositif de stimulation dichotique que nous avons mis en œuvre dans cette thèse visait à favoriser l'établissement d'une nouvelle routine.

La nécessité d'interpréter différemment les caractéristiques temporelles pour l'apprentissage de l'anglais L2 s'impose aussi pour les voyelles : il faut prendre en compte leur durée en anglais car certaines s'opposent par ce trait, ce qui n'est pas le cas en français. L'apprenant francophone doit ici fournir un grand effort pour cesser d'inhiber une information que l'expérience dans sa langue maternelle lui a appris à supprimer. Il s'agit d'un deuxième exemple de difficulté objectivement élevée, que ne rencontrent pas les apprenants dont la L1 est quantitative (i.e. où la durée des voyelles a une valeur phonologique). Là encore, le projet de la thèse était d'aider à développer une connaissance phonologique de la différence entre deux voyelles anglaises, l'une courte, l'autre longue, au moyen d'un entraînement conduisant à traiter de manière optimale chaque voyelle en l'adressant à l'hémisphère le plus compétent pour sa durée. L'hypothèse n'était pas seulement que la durée de ces voyelles serait mieux traitée, mais l'ensemble de leurs traits pourrait bénéficier d'un traitement plus précis dans de telles conditions, et l'apprentissage serait plus profond.

### **8.3.2. Dépassement des difficultés**

Malgré sa courte durée, l'entraînement perceptif proposé s'est avéré efficace en tant qu'aide à l'apprentissage. Il a permis d'améliorer la perception du voisement des consonnes anglaises (Expériences 7 et 8), ainsi que la perception et la production des voyelles courte/longue /ɪ/-i:/ (Expérience 11) en anglais. Il est donc possible d'aider des adultes dans ce difficile apprentissage. Ces résultats s'ajoutent aux données d'autres recherches, souvent dans d'autres langues, montrant les effets bénéfiques d'entraînements courts ciblant un contraste phonologique en L2 (Cebrian & Carlet, 2014 ; Collet *et al.*, 2012 ; Habib *et al.*, 2016 ; Kondaurova & Francis, 2010 ; Menning *et al.*, 2002).

Un résultat plus inédit est l'effet bénéfique supplémentaire apporté par le dispositif de stimulation dichotique pendant l'entraînement sur le voisement des consonnes (Expérience 6). La comparaison avec un groupe de participants réalisant les mêmes exercices avec une stimulation binaurale a fait ressortir l'intérêt de la présentation sélective des consonnes contenant un VOT court plus directement à l'HG et des consonnes contenant un VOT long plus directement à l'HD. Si l'entraînement avait cette composante dichotique, il était suivi d'une amélioration de plus grande taille en discrimination et d'un transfert de ce progrès à un couple de consonnes non entendu pendant l'entraînement, ainsi que d'une identification plus exacte et moins variable de la consonne sonore anglaise /g/ si son VOT était proche de la frontière. Il produisait alors aussi un transfert de compétences pour le seuil d'identification d'un couple de consonnes non entendu pendant l'entraînement (d/-t/), ainsi qu'une

identification plus exacte de l'occlusive sonore /d/. Ces améliorations étaient rendues plus convaincantes par l'insensibilité des épreuves de perception catégorielle à l'effet de répétition (Expérience 4). Cela montrait aussi un apprentissage profond.

Un autre intérêt de l'aide apportée par la stimulation dichotique est que son intensité ne dépendait pas de l'intensité du REA initial des participants (Expérience 6). Un test d'écoute dichotique (Bedoin *et al.*, 2010) évaluait si le participant présentait avant l'entraînement une dominance hémisphérique gauche pour le traitement d'une différence acoustique très brève entre deux consonnes (différence de lieu d'articulation). Seuls les participants qui ne présentaient pas de dominance hémisphérique droite pour ces indices brefs étaient retenus pour l'entraînement. L'étude des corrélations a montré que le fait de ne pas présenter initialement une forte propension de l'HG à s'investir fortement dans le traitement des indices brefs ne rend pas les individus moins réceptifs à l'aide apportée par l'entraînement perceptif dichotique. L'usage de ce dernier peut donc être envisagé assez largement.

Ces progrès sur le plan comportemental se doublaient d'une augmentation de la réaction électrophysiologique pré-attentionnelle à la survenue d'un /g/ parmi des /k/ en anglais, d'après l'augmentation d'amplitude de l'onde MMN (Expérience 7). L'apparition de cet effet pour une seule des deux consonnes peut paraître étonnante, mais d'autres expériences d'apprentissage ont montré des cas d'amélioration asymétrique pour un nouveau contraste (Ylinen *et.*, 2010).

Un autre résultat suggérait une modification du traitement des consonnes sur le plan de l'asymétrie hémisphérique engagée pour le traitement du voisement anglais, ce qui était vraiment attendu dans les hypothèses. En effet, l'onde MMN pour la consonne /k/ anglaise était dominante au niveau des électrodes posées à gauche avant l'entraînement, comme s'il s'agissait d'un /k/ français avec un VOT court. Cette asymétrie a disparu après l'entraînement, à cause d'une diminution de la MMN à gauche (Expérience 7). Cela pourrait être le signe d'un désengagement de l'HG dans le traitement de cette consonne. Après l'entraînement, cette consonne anglaise semble être traitée de manière plus adaptée à la durée de son VOT, plutôt que par simple assimilation à la consonne /k/ française. L'ensemble des données comportementales et électrophysiologiques plaide en faveur de l'hypothèse de l'efficacité de la présentation unilatérale des nouvelles consonnes en fonction de leur durée pour permettre d'apprendre à les traiter comme de nouveaux phonèmes.



Comme pour les consonnes, les épreuves utilisées aux étapes pré- et post-entraînement pour les voyelles étaient insensibles à l'effet de répétition (Expérience 10). Pour les voyelles aussi, l'entraînement perceptif a été couronné de quelque succès, avec une amélioration de grande taille en exactitude perceptive, ainsi qu'un transfert de compétences vers des épreuves de production où le /ɪ/ devenait significativement plus court et le /i:/ significativement plus long (Expérience 11). Par contre, très peu d'éléments dans les résultats montrent une optimisation de l'effet d'entraînement perceptif par le dispositif dichotique. La mesure d'un tel potentiel effet était rendue plus complexe par la différence avant l'entraînement entre les performances des groupes exercés de façon dichotique et binaurale. Ce résultat décevant pour la stimulation dichotique destinée à aider à mieux traiter la durée des voyelles conduit à conclure de façon modeste quant à l'efficacité de ce dispositif pour optimiser les traitements en fonctions des caractéristiques temporelles des phonèmes. Il peut toutefois en partie s'expliquer par des biais dans le matériel. Au-delà d'une différence de durée moyenne substantielle entre les deux voyelles, certains exemplaires de /ɪ/ dans les exercices étaient aussi longs que la durée moyenne des /i:/ et ils ne correspondaient pas à ce qui pouvait être particulièrement bien traité par une fenêtre attentionnelle plus étroite dans l'HG. Ainsi, pour certains mots contenant un /ɪ/, les oscillations gamma dominant dans le cortex auditif gauche auquel ils étaient adressés ne pouvaient pas vraiment permettre un excellent traitement. Il est possible que cela ait contribué à empêcher un bénéfice supplémentaire en cas de stimulation dichotique.

### **8.3.3. Contribution aux modèles d'apprentissage d'une L2 et aux modèles de traitement temporel de la parole**

Il existe des justifications aux difficultés des francophones à apprendre l'anglais en tant que seconde langue. Cependant, les résultats de nos études montrent qu'un apprentissage tardif est possible, conformément aux prédictions des modèle SLM-r (Flege, 1995a ; Flege & Bohn, 2021), PAM (Best, 1995 ; Best *et al.*, 2001) et ASP (Strange, 2011). Lors de l'apprentissage d'une L2, le modèle PAM suppose que le traitement des phonèmes d'une L2 se fait par assimilation aux catégories de la L1, sans possibilité de création de nouvelles catégories spécifiques à la L2. Le modèle détermine trois types de catégories : les sons de la L2 seront assimilés à une catégorie de la L1 ; les sons de la L2 ne sont pas catégorisables ; les sons de la L2 ne sont pas catégorisés comme relevant du langage. Le modèle SLM-r reconnaît lui aussi les mécanismes d'assimilation perceptive, mais admet la possibilité de la création de nouvelles



catégories phonémiques spécifiques à la L2. Les données des expériences réalisées dans cette recherche semblent aller dans le sens des propositions de ce modèle.

Le modèle SLM-r prédit aussi qu'un son de la L2 peut être dans un premier temps assimilé à une catégorie de la L1, tout en étant reconnu comme un exemplaire atypique, pour ensuite former une nouvelle catégorie phonémique distincte de celles de la L1 durant l'apprentissage. Certaines données de nos expériences appuient particulièrement ce dernier point. L'émergence d'une réponse électrophysiologique pré-attentionnelle pour les occlusives sonores anglaises à la suite de l'entraînement perceptif par stimulation dichotique sur le voisement en est une (Expérience 7). Le déplacement de la frontière anglaise dans les épreuves de perception catégorielle pour les consonnes en est une autre (Expérience 6). Les entraînements perceptifs semblent donc avoir permis aux apprenants de créer de nouvelles routines de sélection perceptive leur permettant de mieux repérer et traiter les indices acoustiques pertinents pour la langue anglaise, en s'écartant des habitudes liées à leur L1. Cela est conforme aux prédictions du modèle ASP (Strange, 2011), selon lequel que la capacité à mettre en place des routines de sélection perceptive reste disponibles pour l'acquisition d'une L2.

Le principe d'écoute dichotique pour des entraînements perceptifs s'appuyait sur un cadre théorique combinant psychologie cognitive et neurosciences cognitives. Le modèle d'attention temporelle AST (Poeppel, 2003) et les théories du traitement de la parole en lien avec les oscillations cérébrales (Giraud & Poeppel, 2012 ; Gross *et al.*, 2013 ; Lakatos *et al.*, 2005 ; Morillon *et al.*, 2010 ; Schroeder *et al.*, 2008 ; Spironelli & Angrilli, 2010) ont permis des hypothèses précises. Dans une démarche hypothético-déductive, ces prédictions ont été mises à l'épreuve des performances (et du recueil de données électrophysiologiques) dans les expériences réalisées. Les données qui confortent en partie ces hypothèses sont autant de nouveaux arguments favorables au contexte théorique qui l'a inspiré.

Dans les expériences classiques d'entraînement perceptif pour la phonologie, les exercices sont réalisés en écoute binaural (Callan *et al.* 2003 ; Collet *et al.*, 2017 ; Gaab *et al.*, 2007 ; Kangatharan *et al.*, 2021 ; Kondaurova & Francis, 2010 ; Krzonowski *et al.* 2016 ; Olson, 2021 ; Russo *et al.* 2005 ; Shinohara & Iverson 2018 ; Strange & Dittman, 1984 ; Tamminen *et al.*, 2015 ; Temple *et al.*, 2003). Des expériences complémentaires sont nécessaires pour confirmer et mieux cerner les apports potentiels du dispositif dichotique utilisé dans cette thèse. L'aide apportée à l'apprentissage d'une nouvelle association entre les durées de VOT et le voisement des consonnes contraste avec la faible efficacité pour

l'apprentissage des voyelles opposées par leur durée. Nous avons vu que cette absence d'effet sur les voyelles dans nos expériences pouvait être interprétée en s'inspirant de la durée des fenêtres d'analyse suggérées par les modèles cités. Les durées des voyelles courtes dans le matériel d'entraînement ne leur étaient pas suffisamment conformes (Expérience 11). Le rôle de ce biais reste à confirmer avec des expériences utilisant des voyelles d'une durée plus adaptées aux prédictions de ces modèles. Si ce rôle se confirmait, les résultats actuels sur les voyelles apporteraient un argument indirect pour le modèle AST (Poeppel, 2003). En effet, celui-ci ne décrit pas la différence entre les fenêtres d'analyse temporelle étroites et larges comme une différence relative, mais bien comme imposée par la durée des phases d'oscillations cérébrales distinctes dans les cortex auditifs gauche et droit. Les résultats présentés au sujet du dispositif dichotique sont à considérer comme une simple incitation à explorer une piste nouvelle pour l'aide à l'apprentissage en phonologie en L1 ou L2, inspirée par des modèles qui se développent en dehors du domaine des apprentissages.

Une hypothèse d'interprétation alternative pourrait être émise sur le mécanisme cognitif sous-jacent à l'aide apportée par le dispositif dichotique (Expérience 6). Le mode de présentation consistait à adresser un mot à une oreille et un bruit blanc à l'autre oreille afin de créer une concurrence entre les hémisphères pour le traitement du signal de parole. Or, le bruit a peut-être augmenté la difficulté des exercices, obligeant les participants à mobiliser une attention plus soutenue pour percevoir les mots. L'augmentation de la difficulté induite par le bruit pourrait ainsi élever le niveau général de l'attention et renforcer la profondeur du traitement réalisé sur les cibles. Il y a cependant, dans les résultats, un élément qui n'est pas compatible avec cette hypothèse, et dont seule des prédictions issues du modèle AST pourraient rendre compte. Ce résultat a été observé dans les mesures électrophysiologiques. Avant l'entraînement dichotique, le traitement du /k/ anglais (VOT long) provoquait une MMN plus ample au niveau de l'HG – spécialisé dans les indices courts –, suggérant l'assimilation au phonème /k/ français, qui s'écrit avec les mêmes lettres. Après l'entraînement dichotique, la MMN était bilatérale, de même amplitude sur les deux hémisphères, essentiellement grâce à sa diminution à gauche (Expérience 7). Cette modification de la dominance hémisphérique pour le traitement de la consonne sourde anglaise /k/ est cohérente avec un changement qualitatif dans la façon de la percevoir. Les résultats sont compatibles avec l'hypothèse inspirée par le modèle AST, selon laquelle le dispositif d'écoute dichotique est susceptible d'amener les apprenants à traiter les nouveaux phonèmes de façon optimale pour leur durée, et non par simple assimilation aux traitements

réalisés en L1 sur des phonèmes orthographiés de la même façon. Mieux traiter les consonnes anglaises en fonction de leur durée semble passer par un traitement du /k/ anglais (VOT long) impliquant moins l'HG, dans lequel les étroites fenêtres d'analyse seraient moins utiles que pour le /k/ français (VOT court). Cela va dans le sens des travaux expérimentaux sur la présence d'oscillations rapides dans le cortex auditif gauche (Belin *et al.*, 1998 ; Fiez *et al.*, 1995 ; Johnsrude *et al.* 1997 ; Zaehle *et al.*, 2004) et d'oscillations lentes à droite (Giraud *et al.*, 2007 ; Luo & Poeppel, 2012 ; Morillon *et al.*, 2012 ; Shtyrov *et al.*, 2000).

L'hypothèse alternative reste intéressante, mais ne peut pas expliquer ce phénomène. A propos du rôle du bruit, il est aussi possible de s'intéresser à l'hypothèse du poids attentionnel (Rees & Lavie, 1994) qui suppose que si la tâche centrale est complexe, les distracteurs sont bien inhibés. Cette théorie prédirait que le bruit présenté dans l'autre oreille que celle du mot cible serait un distracteur facilement supprimé et peu perturbant. Les exercices en anglais placent en effet les apprenants dans une situation de traitement difficile et impliquant fortement l'attention.

Sur le plan pragmatique, le dispositif de stimulation dichotique présente aussi un avantage : il est simple à mettre en place. Il ne requiert pas de matériel spécifique ou coûteux comme, par exemple, la stimulation transcrânienne magnétique (TMS) qui vise elle aussi à modifier l'implication de certaines aires cérébrales dans des traitements cognitifs.

Pour davantage inscrire ce travail dans le cadre théorique des oscillations cérébrales, des expériences avec des analyses en temps/fréquence de l'activité électrophysiologiques seraient nécessaires. Cela dépasse l'objectif de cette thèse. Une piste et une hypothèse peuvent être proposées. Une augmentation de la synchronisation des oscillations rapides dans le cortex auditif gauche à l'écoute de consonnes occlusives sonores anglaises (VOT court) et à l'écoute d'occlusives sourdes françaises (VOT court) pourrait être prédite chez les apprenants de l'anglais L2 améliorant leurs performances après un entraînement dichotique. L'hypothèse est que ce phénomène ne devrait pas se produire pour des consonnes sourdes anglaises ni pour des consonnes sonores françaises, toutes deux associées à un VOT long. La vérification de ce type d'hypothèse constituerait un argument supplémentaire pour la théorie AST.

## **8.4. Apprentissage de phonèmes par des adultes francophones dyslexiques : origine et dépassement des difficultés**

### **8.4.1. Origine des difficultés**

Les difficultés exprimées par les étudiants dyslexiques concernant l'apprentissage de l'anglais (Mazur-Palandre *et al.*, 2016) méritent d'être objectivées. Quelques données obtenues avant les entraînements concourent à confirmer qu'il s'agit de difficultés accrues par rapport à celles des étudiants normo-lecteurs (Expérience 5). Pour ce qui est de l'apprentissage de nouveaux phonèmes, ces difficultés sont cohérentes avec la persistance des troubles phonologiques documentée chez les adultes dyslexiques (Abadie *et al.*, 2016 ; Noordenbos *et al.*, 2013 ; Pennington *et al.*, 1990), mais aussi avec l'hypothèse d'un trouble des traitements temporels et de l'ajustement rythmique à l'origine des difficultés phonologiques et prosodiques des personnes dyslexiques (Goswami, 2011 ; Goswami & Leong, 2013 ; Lizarazu *et al.*, 2021 ; Muneaux *et al.*, 2004 ; Soltész *et al.*, 2013 ; Thomson *et al.*, 2006 ; Thomson & Goswami, 2008). Dans les expériences de cette recherche, les adultes dyslexiques ont présenté des difficultés plus grandes que les normo-lecteurs avant l'entraînement pour le traitement d'un indice de durée : le VOT en anglais, tout comme en français, avec un seuil d'identification inapproprié pour des consonnes opposées en voisement (Expérience 5). Ils percevaient aussi les voyelles courtes et longues avec moins d'exactitude et moins d'assurance d'après leurs réponses significativement plus lentes. Ils produisaient également des /i:/ anormalement courts. Ils semblaient ainsi moins bien saisir que les autres étudiants l'importance de marquer la différence entre les voyelles courtes et longues, qu'ils disposent seulement d'indices orthographiques ou d'un modèle audio (Expérience 12). Ces remarques contribuent à préciser la nature d'une partie des difficultés des étudiants dyslexiques en phonologie anglaise. Cela permet d'envisager de façon plus éclairée d'éventuels moyens de les aider. Ces données objectives montrent aussi que leurs plaintes au sujet de difficultés accrues en apprentissage de l'anglais sont légitimes.

Incidentement, il est apparu que l'épreuve dans laquelle ils ont le plus progressé après l'entraînement est celle qui les incitait à s'appuyer sur une analogie entre la longueur de la voyelle et la longueur du graphème qui le code orthographiquement (Expérience 13). L'entraînement comportait des éléments orthographiques, et il semble que les étudiants dyslexiques soient plus efficaces en production s'ils ont un support de cette nature. Celui-ci pourrait être fourni plus systématiquement dans les exercices pour ce type de public, et une

incitation à lire à haute voix des mots ou phrases contenant des /i/ et des /i:/ pourrait leur être particulièrement bénéfique.

Au-delà de difficultés avec les nouveaux phonèmes, les adultes dyslexiques ont aussi manifesté des anomalies pour le traitement d'aspects sonores relevant de la prosodie. L'Expérience 14 a montré qu'ils connaissent moins que les normo-lecteurs la différence entre les intonations différentes en anglais pour les questions ouvertes et fermées. Avant l'entraînement, ils produisaient les deux types de questions avec une intonation montante lorsqu'ils n'avaient pas de modèle audio, sans doute par assimilation au français. Par contre, en présence d'un modèle audio, l'intonation finale de leurs questions était toujours descendante. Cette production étrange de la prosodie pour tout type de questions, avant l'entraînement, suggère un fonctionnement véritablement atypique, souvent évoqué au sujet des troubles neuro-développementaux. Les adultes normo-lecteurs ne procédaient pas de la sorte. La survenue de cette anomalie de production en fonction de la présence d'un modèle audio illustre les difficultés persistantes des adultes dyslexiques pour les aspects liés aux sons de la langue. Au lieu de les aider, l'information procurée par un modèle audio les perturbait. Ils ont sans doute perçu que l'intonation descendait dans certaines questions, et ont généralisé abusivement à toutes les questions. Cela pourrait s'expliquer par une perception trop imprécise de la direction des modulations de hauteur de la voix. Sur le plan pédagogique, l'effet perturbant du modèle audio dans l'exercice de répétition avant l'entraînement montre que fournir de tels modèles audio, sans explicitation des différences, n'est pas la solution la plus appropriée pour guider les apprenants dyslexiques vers l'inférence d'une règle prosodique nouvelle.

#### **8.4.2. Dépassement des difficultés**

Malgré ces difficultés, il est possible pour des adultes dyslexiques francophones non seulement d'améliorer la perception catégorielle et donc le système phonologique de leur L1 (Expériences 2 et 3), mais aussi d'apprendre une L2 (Expériences 8, 12 et 14). Dans les deux cas, ils semblent réceptifs à la stimulation dichotique. Cette réceptivité s'est avérée accrue par rapport à celle des adultes normo-lecteurs pour les voyelles anglaises /i/ et /i:/ (Expérience 12).

Les données confirment tout d'abord qu'une remédiation du système phonologique de la langue maternelle est possible en cas de dyslexie grâce à un entraînement perceptif comme d'autres recherches l'ont montré (Collet *et al.*, 2012 ; Zoubrinetzky *et al.*, 2019), ceci cette

fois chez des adultes dyslexiques. L'efficacité de l'entraînement dichotique en français sur le traitement des indices courts, particulièrement avec une perception moins allophonique du /k/, est aussi cohérente avec un autre phénomène observé après l'entraînement des adultes dyslexiques. Ce phénomène est l'augmentation de la dominance de l'HG pour traiter le lieu d'articulation après l'entraînement, d'après l'augmentation du REA (*Right Ear Advantage*, i.e. avantage de l'HG) dans le test d'écoute dichotique pour les couples de mots qui diffèrent par le lieu d'articulation (Expérience 2). Les différences de lieu sont portées par des indices acoustiques très brefs (transitions de formants) et ce traitement a été davantage pris en charge par l'HG après l'entraînement dichotique en français. Cela constitue un effet de transfert de compétences pour un traitement temporel, ce qui ouvre des perspectives en termes d'aide à la remédiation en orthophonie.

Sur le plan cérébral, à la suite de cet entraînement, la sensibilité pré-attentionnelle des adultes dyslexiques s'est élevée pour le /g/ français, et une sensibilité plus tardive est apparue pour le /k/ français d'après l'étude des potentiels évoqués (Expérience 3). L'amplification de la MMN exclusivement pour la consonne dotée d'un VOT long (/g/) et non pour /k/ (VOT court) est cohérente avec le constat des difficultés que le traitement d'indices courts pose aux personnes dyslexiques (Tallal, 1980 ; Habib *et al.*, 2002). Cela va aussi dans le sens des données montrant une faible latéralisation des oscillations rapides dans l'HG en cas de dyslexie (Dushanova *et al.*, 2020 ; Lehongre *et al.*, 2013). Cette faible latéralisation gauche pourrait avoir fait perdre l'efficacité de l'adressage à l'HG des consonnes avec un VOT court pendant l'entraînement dichotique.

Pour l'apprentissage de nouveaux phonèmes en anglais L2, l'effet de l'entraînement dichotique ciblant les consonnes occlusives a eu un retentissement équivalent par rapport à celui produit chez les apprenants normo-lecteurs. Les progrès des adultes dyslexiques se sont aussi transférés à des mots non étudiés pendant l'entraînement (Expérience 8). La réceptivité des adultes dyslexiques à l'entraînement dichotique en anglais s'est en revanche montrée plus élevée que chez les adultes contrôles pour les voyelles (Expérience 12). Ils se sont aussi montrés plus réceptifs que les adultes normo-lecteurs à l'aide apportée pour l'apprentissage de la différence d'intonation des deux types de questions en anglais. Cela encourage à développer pour les personnes atteintes de dyslexie des aides particulières, car elles sont en mesure de dépasser certaines difficultés rencontrées dans cet apprentissage.

### **8.4.3. Contribution aux modèles d'apprentissage d'une L2 et de déficit des traitements temporels dans la dyslexie**

Selon le modèle SLM-r (Flege & Bohn 2021), l'apprentissage d'une L2 sera mieux réalisé si le système phonologique de la L1 est stable. En cas de dyslexie, et même chez l'adulte, ce système de la L1 ne l'est pas. Nos données ont confirmé les lacunes de ce système chez ces adultes pour le voisement des consonnes françaises (Expériences 1, 2 et 3). Comme ces anomalies étaient supposées avant la conception des programmes, nous souhaitions contribuer à stabiliser les connaissances phonologiques sur le voisement des consonnes françaises des participants dyslexiques à l'aide d'une série d'exercices (durée totale 2h30, répartie sur 5 jours), en amont d'un apprentissage sur le voisement en anglais (Expérience 2 et 3). L'objectif était surtout de renforcer leurs compétences en traitement des indices de durée du voisement propres au français, afin qu'elles ne risquent pas d'être contaminées par la stimulation dichotique qui allait, peu après, les inciter à engager différemment leurs hémisphères cérébraux pour le voisement de consonnes anglaises. Ce renforcement préalable a peut-être contribué aux progrès constatés en anglais.

Le fait d'avoir proposé un entraînement en français, avec le dispositif dichotique, seulement aux adultes dyslexiques a pu introduire un biais, qui doit être pris en compte pour des comparaisons entre les effets de l'entraînement en anglais chez les adultes normo-lecteurs et dyslexiques. Même si c'était dans des langues différentes, le groupe dyslexique a en effet bénéficié d'un total de deux semaines consécutives d'entraînement par stimulation dichotique basé sur la durée (5 heures au total) (Expériences 2, 3 et 8). Stimuler ainsi longuement les cortex auditifs des deux hémisphères de façon spécifiquement adaptée à la durée de leurs fenêtres d'analyse temporelle a pu avoir un effet particulièrement bénéfique sur les capacités de synchronisation au signal de parole. Cela a pu introduire un biais en faveur d'un effet plus fort de l'entraînement dichotique en anglais dans le groupe dyslexique. La succession des deux entraînements pour le voisement, d'abord en français puis en anglais, est probablement une combinaison intéressante à conseiller pour des aides proposées à des personnes en difficulté avec le langage.



## **8.5. Pré-ajustement de l'attention temporelle pour le traitement d'information prosodiques**

### **8.5.1. Intérêt d'un amorçage rythmique global pour le traitement du langage**

L'utilisation de l'amorçage rythmique musical global pour l'apprentissage d'une L2 est une application nouvelle pour ce principe. La plupart des expériences d'amorçage proposent des amorces dont la structure temporelle est spécifiquement adaptée à la structure de la séquence de parole (mot ou phrase) donnée à traiter immédiatement après (Cason & Schön, 2012 ; Cason *et al.*, 2015). Plusieurs expériences ont montré que l'écoute d'une séquence musicale au rythme très régulier, avec une pulsation à 2 Hz, peut influencer positivement le traitement grammatical ultérieur d'une série de phrases (Bedoin *et al.*, 2016 ; Bedoin *et al.*, 2018 ; Canette *et al.*, 2020 ; Chern *et al.*, 2018 ; Fiveash *et al.*, 2020a, 2020b ; Ladanyi *et al.*, 2020 ; Przybylski *et al.*, 2013). Il s'agit là d'un amorçage plus global. Une des interprétations possibles pour les effets de ce type d'amorçage rythmique sur les tâches de jugement grammatical est que le rythme lent de la musique induirait le préajustement de fenêtres d'analyse suffisamment larges pour être sensibles aux enveloppes de parole, unités prosodiques qui guident l'analyse grammaticale. Sur le plan cérébral, cela correspondrait au renforcement de la synchronisation d'oscillations cérébrales lentes (pour des synthèses, voir Fiveash *et al.*, 2021 ; Tillmann & Bedoin, soumis). Les Expériences 13 et 14 tentaient d'étendre l'utilisation de ce dispositif pour aider à mieux traiter des unités prosodiques en L2. Il ne s'agissait pas d'améliorer le traitement grammatical, mais de favoriser la perception d'unités prosodiques, qui participent par ailleurs aux traitements syntaxique et morphosyntaxique. Les effets observés en particulier chez les adultes dyslexiques sont encourageants pour l'aide potentielle dans le domaine de la prosodie.

Lorsqu'il s'agit de personnes dyslexiques, certaines précautions sont à prendre dans le choix de la musique utilisée en amorce. En effet, des données mesurant les oscillations cérébrales montrent que, à l'écoute d'une musique dont le rythme est très régulier, les oscillations se synchronisent à la pulsation de la musique aussi bien chez les adultes sans difficulté avec le langage que chez les adultes dyslexiques. Par contre, des tentatives de synchronisation à l'écoute d'une musique dont le rythme n'est pas régulier ne s'opèrent que chez les adultes normo-lecteurs (Fiveash *et al.*, 2020b). Ils parviennent ainsi à extraire les indices très subtils d'une certaine régularité. Ce n'est pas le cas des dyslexiques, ce qui montre une nouvelle facette de leurs difficultés rythmiques. Cela suggère que, pour un public dyslexique, une amorce rythmique peut être musicale, mais son efficacité pour synchroniser



les rythmes attentionnels et cérébraux dépend de la rigueur de sa régularité. Dans le contexte de notre étude, il était donc important de fournir des rythmes extrêmement réguliers et saillants afin d'espérer un effet d'amorçage global sur le traitement des indices prosodiques de questions anglaises. D'une manière plus générale, le succès de cette démarche d'utilisation de la musique et du rythme auprès de populations atteintes de troubles du langage pour aider au traitement du signal de parole apporte de nouveaux arguments pour les liens entre la musique et le langage. Ceux-ci sont parfois exploités différemment, à travers des programmes d'entraînements musicaux pour tenter de remédier plus durablement à des déficits langagiers (Flaugnacco *et al.*, 2015 ; Frey *et al.*, 2019 ; Habib, 2016 ; Thomson *et al.*, 2013)).

### **8.5.2. Rôle d'une cascade oscillatoire dans l'amorçage rythmique sur la prosodie**

L'effet d'amorçage rythmique musical sur l'apprentissage de la différence entre l'intonation des deux types de questions anglaises dans les Expériences 13 et 14 est cohérents avec la théorie des ressources partagées par le traitement de la musique et du langage (modèle PRISM, Fiveash *et al.*, 2021). Des similitudes entre ces deux domaines semblent conduire à mobiliser quelques mécanismes cognitifs identiques pour le traitement du signal musical et langagier. Les points communs qui peuvent expliquer ce partage sont notamment l'organisation hiérarchique du rythme musical, analogue à la prosodie de la parole (Hausen *et al.*, 2013 ; Lerdahl & Jackendoff, 1983).

Dans les Expériences 13 et 14, une réalisation correcte de la tâche de répétition pendant l'entraînement imposait d'être attentif aux unités prosodiques. La difficulté résidait en effet dans la capacité à percevoir la direction dans laquelle se produisait la modulation du F0 à la fin des phrases anglaises. Celle-ci montait ou descendait, alors qu'il s'agissait toujours d'une question, alors qu'en français l'intonation finale des questions est le plus souvent montante. Il s'agissait de sensibiliser les apprenants à cette différence, voire de les conduire à extraire une règle en prosodie au sujet des questions.

L'amorce musicale avec une pulsation à 2 Hz était destinée favoriser cette sensibilisation. Sa présentation visait à pré-ajuster une fenêtre attentionnelle de durée assez longue pour être adaptée à une séquenciation du signal de parole en unités larges, conformes à ce qui permet l'analyse prosodique. C'est en ce sens qu'il s'agissait d'un amorçage rythmique global. L'augmentation des performances après les entraînements (Expériences 13 et 14) qui proposaient ce type d'amorce suggère que l'amorçage rythmique global pourrait permettre un ajustement de l'attention temporelle pendant suffisamment longtemps pour réaliser

correctement plusieurs répétitions d'une même phrase. Une telle aide apportée à la réalisation des exercices pourrait expliquer que les participants ont mieux réussi la production de nouvelles questions anglaises après l'entraînement.

Ce progrès suggère aussi que les amorces rythmiques n'ont pas seulement aidé les adultes à mieux réaliser les exercices, mais qu'elles ont favorisé l'apprentissage d'une règle nouvelle en prosodie. Cette règle associe les directions de l'intonation finale à deux types de contenus de question : l'attente d'une réponse par oui/non, ou l'attente d'une réponse plus développée. Une hypothèse d'interprétation est que l'amorçage rythmique global induit une réflexion de la part de l'apprenant, sous la forme d'une recherche puis d'une découverte de l'origine de cette différence qui est de nature sémantique/pragmatique. Une hypothèse alternative est que l'amorce rythmique a favorisé des prédictions temporelles consistant à anticiper la survenue d'une enveloppe de parole descendante de quelques unités prosodiques après un mot interrogatif commençant par « W » (What/When/Where/Who), et la survenue d'une enveloppe ascendante si la phrase commençait par « Do ». La deuxième hypothèse décrit une règle un peu plus superficielle, qui pourrait émerger sur la base d'un apprentissage statistique réalisé lors des exercices qui proposaient de telles questions de façon répétitive. Sur le plan attentionnel, les amorces rythmiques auraient alors favorisé les anticipations et la séquenciation à un rythme lent permettant la perception précise du changement d'intonation attendu.

En sciences cognitives, une importance particulière est accordée à la validité neurologique des mécanismes cognitifs décrits pour expliquer des données comportementales. Les mécanismes attentionnels qui viennent d'être proposés concordent avec les recherches montrant que les oscillations cérébrales tentent de s'ajuster au rythme du langage (Giraud & Poeppel, 2012 ; Kösem & Wassenhove, 2017) comme à celui de la musique (Fujioka *et al.*, 2012 ; Jones, 1976 ; Nozaradan *et al.*, 2011, 2012, 2015). Cette synchronisation guiderait dans les deux cas les anticipations temporelles et la séquenciation (Arnal & Giraud, 2012 ; Calderone *et al.*, 2014 ; Poeppel & Assaneo, 2020 ; Tillmann, 2012). En affinant la précision temporelle des traitements, cette synchronisation améliorerait aussi globalement la perception auditive (Morillon & Baillet, 2017 ; Peelle & Davis, 2012 ; Van Wijk *et al.*, 2012 ; Yang *et al.*, 2018). Dans le cas précis du rythme à 2 Hz des amorces utilisées dans les Expériences 13 et 14, la perception de la pulsation musicale devrait renforcer la synchronisation des oscillations delta (1-3 Hz). Celles-ci sont décrites comme adaptées à la perception de la

prosodie (Calderone *et al.*, 2014 ; Keitel *et al.*, 2017), niveau d'organisation de la parole ciblé par les expériences.

Dans les questions présentées pendant l'entraînement, l'intonation finale portait sur la dernière syllabe, c'est-à-dire une unité trop courte par rapport à celles auxquelles les oscillations delta se synchronisent en priorité. La notion de cascade oscillatoire (Benitez-Burraco & Murphy, 2019) pourrait expliquer pourquoi ces amorces ont tout de même été efficaces. Un couplage est décrit entre les oscillations delta et thêta (4-8 Hz) (Gross *et al.*, 2013), et ces dernières ont des phrases d'une durée adaptée à l'extraction de syllabes (Luo & Poeppel, 2007 ; Morillon *et al.*, 2010 ; Poeppel *et al.*, 2008). C'est pourquoi le renforcement des oscillations delta par les amorces pourrait accentuer la synchronisation des oscillations thêta favorables à l'extraction et à l'analyse des syllabes. L'analyse détaillée du contenu des syllabes serait lui-même rendu possible par la relation de couplage thêta-gamma. Une fois renforcées, les phases des oscillations thêta moduleraient en effet les variations d'amplitude des oscillations gamma (Giraud & Poeppel, 2012), très rapides et adaptées à l'analyse fine du signal. Les variations du F0 dans la dernière syllabe seraient ainsi particulièrement bien traitées. Cette cascade oscillatoire, décrite en neurosciences, pourrait rendre compte de la façon dont des amorces dotées d'un rythme à 2 Hz réguleraient en amont les cycles d'attention temporelle nécessaires au traitement du signal de parole réalisé simultanément à plusieurs échelles (Hicoek *et al.*, 2015 ; Poeppel *et al.*, 2008).

Or, en cas de dyslexie, l'ajustement des oscillations delta à un rythme externe serait déficitaire (Soltész *et al.*, 2013). Pour le langage, cela entraverait chez les personnes dyslexiques l'ajustement des rythmes cérébraux aux enveloppes de parole (Falk *et al.*, 2015 ; Falter & Noreika, 2014 ; Goswami *et al.*, 2002 ; Peter & Stoel-Gammon, 2008), très utile pour l'analyse prosodique (Ghitza, 2011 ; Leong *et al.*, 2014). En proposant des amorces extrêmement régulières avant l'écoute de phrases qu'il est bon de traiter au niveau prosodique pour réaliser la tâche, l'Expérience 14 tentaient de déclencher tout de même chez les participants dyslexiques la synchronisation des oscillations delta et, par suite, la cascade oscillatoire permettant l'analyse des syllabes porteuses de l'intonation à traiter. Les résultats des Expériences 14 encouragent à explorer davantage cette piste pour inspirer des méthodes de remédiation dans le domaine de la prosodie chez des personnes présentant des pathologies du langage.

Dans une approche qui allie psychologie cognitive et neurosciences cognitives, ce travail est ainsi conforme à l'idée de Poeppel (2012) qui considère que *mapping is not*

*explaining* et propose de tenter de décrire les fondements des mécanismes cognitifs en évoquant davantage une dynamique oscillatoire (Benitez-Burraco & Murphy, 2019).

### **Mécanismes neurocognitifs à la base d'un meilleur traitement des fréquences après un amorçage rythmique.**

Dans le cadre général d'une conception dynamique de la distribution de l'attention dans le temps (Jones, 1976 ; Jones *et al.*, 2002), un rôle bénéfique des amorces rythmiques pour le traitement temporel est compréhensible. Cependant, une fois extraites, les unités du signal sonore semblent aussi être mieux traitées pour leur contenu. Dans les Expériences 13 et 14, il s'agit des variations de hauteur de la voix. En utilisant un matériel musical, Herbst et Obeser (2017) ont réalisé une expérience qui montre que les anticipations temporelles correctes permettant de traiter de manière qualitativement plus précise le contenu de l'information survenant au moment attendu. Une analogie peut être faite avec l'amélioration du traitement de la modulation du F0 dans la dernière syllabe des questions après un amorçage rythmique. L'expérience de Herbst et Obeser a montré que la hauteur (aspect spectral) d'une note de musique était plus vite traitée si cette note pouvait être bien prédite sur le plan temporel à partir de la régularité des intervalles entre les notes précédentes dans une séquence musicale. En améliorant les capacités d'anticipation par le pré-ajustement de fenêtres attentionnelles larges, peut-être par le biais de la stimulation d'un rythme cérébral lent, l'amorce rythmique aurait pu permettre un traitement plus précis des modulations du F0 dans une unité en fin de phrase.

Une autre interprétation est guidée par les travaux de Coffey *et al.* (2021). Ils suggèrent qu'un traitement particulièrement réussi du contenu fréquentiel d'un son est lié à un autre mécanisme de synchronisation des oscillations cérébrales. Un son dont la hauteur est par exemple de 300 Hz renforcerait dans le cerveau une bande de fréquence de cette même valeur, et ce renforcement se maintiendrait environ 100 ms après la fin du son. Concrètement, lors de l'écoute d'une phrase, si le F0 de la voix reste assez constant, il susciterait chez l'auditeur ce renforcement d'une bande de fréquence, qui servirait de base de comparaison (*template*) pour détecter très vite une déviance. Une telle déviance se produisait dans les questions présentées dans les Expériences 13 et 14, avec une élévation ou baisse marquée du F0 dans une syllabe après un début de phrase dont la hauteur était stable. La durée d'une syllabe est compatible avec les 100 ms pendant lesquelles le *template* resterait disponible. Le mécanisme et le rôle de ce phénomène électrophysiologique (FFR) sont encore mal compris,

mais il est cohérent avec les effets obtenus ici. Ils pourraient contribuer à l'expliquer, et les recherches à ce propos gagneraient à être suivies pour élargir le cadre théorique.

### **8.5.3. Apport du retour visuel pour un apprentissage prosodique en anglais**

Contrairement à l'amorçage rythmique, le retour visuel sur l'intonation pendant l'entraînement en prosodie dans les Expériences 13 et 14 n'a pas amélioré les performances des adultes normo-lecteurs. Un retour visuel du même type proposé à des adultes hispanophones apprenant l'anglais s'est pourtant montré efficace dans une autre expérience (Ramirez Verdugo, 2006). Il était proposé plus longtemps (10 semaines) que dans nos expériences, et la brièveté de l'entraînement n'est peut-être pas très adaptée à ce type de dispositif. Un retour visuel proposé pendant seulement une heure d'entraînement sur des voyelles a permis, dans une autre expérience, des progrès dans le traitement de ces voyelles (Kartushina *et al.*, 2015). Toutefois, le contenu du retour visuel était articulatoire, alors qu'il était acoustique dans nos expériences. Cette différence de contenu pourrait en partie expliquer qu'un entraînement bref utilisant un retour visuel ne soit efficace que s'il renvoie à une information articulatoire.

L'efficacité rapide du retour visuel articulatoire pourrait s'expliquer par l'activation supplémentaire qu'elle pourrait produire sur une boucle perceptivo-motrice déjà engagée dans la réalisation de la répétition. Le retour visuel sur des indices acoustiques est sans doute moins directement susceptible de renforcer l'activation d'une telle boucle. De plus, dans l'expérience de Kartushina *et al.* (2015), les apprenants s'exerçaient sur une seule voyelle pendant une heure, tandis que les participants des Expériences 13 et 14 se sont exercés sur deux types de questions en seulement 45 min.

Peu d'études s'intéressent aux effets du retour visuel sur des aspects acoustiques, privilégiant plutôt les retours visuels articulatoires (Cleland, *et al.*, 2015 ; Fabre *et al.*, 2016 ; Gallagher, 2013). L'entraînement utilisant le retour visuel sur les modulations de hauteur de la voix a tout de même exercé un certain effet, positif, dans les Expériences 13 et 14. L'apport de ce dispositif à l'entraînement en prosodie est apparu à travers une accentuation des progrès suivant l'entraînement lorsque le retour visuel s'ajoutait à l'amorce rythmique. Ce bénéfice apporté par le retour visuel combiné aux amorces rythmiques concernait les adultes normo-lecteurs et les adultes dyslexiques. Il est apparu comme d'autant plus grand que les apprenants avaient un faible niveau pour le traitement des questions anglaises avant l'entraînement. Il s'agit ainsi d'un dispositif qui pourrait être particulièrement adapté à des apprenants en grande

difficulté avec l'apprentissage de l'anglais L2. Guyot-Talbot *et al.* (2016) ont montré l'efficacité d'un entraînement avec ce même retour visuel, mais sans amorces rythmiques. Leurs participants étaient cependant très différents de ceux des Expériences 13 et 14. Il s'agissait d'étudiants se destinant à devenir enseignants en anglais. Ils n'étaient pas bilingues, mais leur niveau était meilleur que celui des participants aux expériences de cette thèse.

Ainsi, contrairement à des adultes ayant déjà atteint un bon niveau d'anglais, un public d'apprenants avec un niveau faible semble peu réceptif au dispositif de retour visuel seul. Par contre, la combinaison de ce retour visuel sur des données acoustiques combiné à l'utilisation d'amorces rythmiques est apparu comme efficace pour aider des adultes ayant un faible niveau en anglais à apprendre une nouvelle règle en prosodie dans cette langue, qu'ils soient normo-lecteurs ou dyslexiques.

## **8.6. Limites**

Le cadre théorique général à l'origine des hypothèses relevait de la psychologie cognitive avec les concepts d'attention temporelle et de fenêtres d'analyse, mais ces concepts ont aussi une validité sur le plan neuronal. Les recherches explorant les phénomènes de synchronisation des oscillations cérébrales ont pour cela été exposées dans le cadre théorique des travaux réalisés dans la thèse. La complexité de ce domaine nécessitait une description détaillée et étayée d'arguments expérimentaux pour un exposé cohérent et aussi didactique que possible. La thèse n'a cependant pas la prétention de relever des neurosciences. La synthèse proposée dans la première partie de la thèse est destinée à l'usage de ceux qui, non formés à ce sujet, souhaiteraient puiser des idées pour imaginer des dispositifs d'aide aux apprentissages dans le domaine du langage. Une limite de la thèse est l'absence d'expérience analysant les oscillations cérébrales elles-mêmes. Une piste d'hypothèse pour une telle expérience a simplement été proposée. L'objectif de la thèse était plus modeste. Le choix méthodologique de compléter les données comportementales en se tournant en priorité vers l'analyse des potentiels évoqués était motivé par le souhait de tester si la sensibilité pré-attentionnelle aux nouveaux phonèmes se modifiait après l'entraînement.

D'autres limites sont à pointer au niveau du matériel linguistique utilisé dans les entraînements. Le programme d'exercices portant sur le voisement des consonnes françaises et anglaises gagnerait à proposer des mots produits par une plus grande variété de locuteurs (Expériences 2, 3, 6, 7, 8 et 9). Même si la variabilité des stimuli était élevée et offrait pour

les phonèmes ciblés des contextes vocaliques ou consonantiques diversifiés, le nombre de locuteurs pour les voix entendues a été augmenté pour l'entraînement sur les voyelles (trois locuteurs) mais reste insuffisant. Des expériences testant des protocoles d'entraînement perceptifs HVPT ont montré l'intérêt d'élever cette variabilité pour l'apprentissage de nouveaux contrastes de consonnes en L2 (Lively *et al.*, 1993), pour augmenter la généralisation des compétences (Sadakata & McQueen, 2013), les transférer vers la production (Strange & Dittman, 1984) et assurer la persistance des progrès (Cebrian & Carlet, 2014). Dans les exercices proposés, il serait bon d'introduire plus de voix pour tenter de renforcer les effets observés.

La brièveté des entraînements proposés était délibérée. Elle a pu toutefois ne pas convenir à certains apprenants et générer une hétérogénéité des effets des programmes. La question des rythmes et des styles d'apprentissage est au cœur des débats sur l'école et sur son caractère inclusif (Wanlin *et al.*, 2019), mais elle reste pertinente pour des apprenants adultes (Berdal-Masuy *et al.*, 2004). Elle a cependant peu de place dans les universités. La mise à disposition d'outils d'entraînement auprès des étudiants pourrait être un moyen d'offrir une aide pour compenser des compétences hétérogènes initiales en L2, ou hétérogènes quant à la vitesse de généralisation des acquis. Il s'agirait de permettre à ceux qui en ont le plus besoin de réaliser des exercices en autonomie, comme cela se fait déjà dans certaines universités (e.g. Paris-Cité) pour la formation des futurs enseignants en anglais. Pour que les étudiants les utilisent, il faut que leur contenu soit adapté à un public adulte, comme c'est le cas dans les programmes proposés ici. La brièveté des exercices et un autre aspect attractif. Si les résultats pouvaient être répliqués, et les programmes d'entraînement améliorés, il serait possible de proposer aux enseignants qu'ils les utilisent auprès des étudiants, cette fois de façon régulière. Des données montrent en effet que des entraînements suivis un peu plus longtemps débouchent sur davantage de progrès (Olson, 2019, 2021).

Une autre limite concerne le rôle potentiel des informations orthographiques données aux apprenants pendant l'apprentissage sur la durée des voyelles anglaises (Expériences 10, 11 et 12). Elles permettaient à l'apprenant de faire un lien entre la durée des voyelles et la longueur des graphèmes et de se baser ensuite sur l'analogie entre les deux pour produire des /ɪ/ d'une durée différente des /i:/. Une partie des effets observés à la suite des entraînements perceptifs, dichotique ou binaural, peut être attribuée à cette composante orthographique. Toutefois, les données ont montré que les participants extraient quand même des informations sur la durée des voyelles dans les stimuli auditifs. Ils les utilisaient pour



améliorer leurs productions, puisque leurs /i:/ étaient plus longs lorsqu'ils disposaient d'un modèle audio en plus du support orthographique dans les épreuves de répétition.

D'autres limites concernent les adultes dyslexiques (Expériences 2, 3 et 8). Seul l'entraînement avec le dispositif dichotique leur a été proposé. La comparaison avec le dispositif binaural n'a porté que sur les adultes normo-lecteurs. Cela limite les interprétations. Il s'agissait d'un choix guidé par le souhait de leur proposer les entraînements potentiellement les plus robustes. Au-delà de cette considération essentiellement d'ordre éthique, et compte tenu de l'effet positif des programmes qui leur ont été proposés, il serait souhaitable de tester des groupes d'adultes dyslexiques contrôles pour qui l'entraînement aux nouveaux phonèmes serait binaural. Pour la prosodie (Expérience 14), des groupes dyslexiques avec un entraînement renforcé seulement par des amorces rythmiques ou seulement par le retour visuel devraient aussi être testés. Une autre limite à propos des participants dyslexiques est l'absence de données sur la nature de leurs principaux déficits. Les déficits phonologiques seraient les plus persistants chez les adultes dyslexiques (Abadie *et al.*, 2016 ; Noordenbos *et al.*, 2013 ; Pennington *et al.*, 1990). Réaliser un bilan auprès d'adultes dyslexiques pour cerner les déficits cognitifs sous-jacents à leurs difficultés en lecture est difficile, notamment parce qu'ils mettent en place des compensations qui les masquent. Des outils existent cependant, en orthophonie (e.g. ECLAT 16+, Gola-Asmussen *et al.*, 2010), et un bilan orthophonique et neuropsychologique réalisé auprès de 30 étudiants dyslexiques et 30 étudiants normo-lecteurs a confirmé, au niveau du groupe, le maintien de performances pathologiques en métaphonologie, mais aussi un empan visuo-orthographique plus petit, une mémoire à court terme faible, une impulsivité dans les situations d'attention soutenue, et la moitié des étudiants dyslexiques ne traitaient pas spontanément l'information visuelle globale (Abadie & Bedoin, 2016). Ce type d'exploration des compétences des participants dyslexiques procurerait des informations permettant de découvrir si certaines caractéristiques cognitives sont prédictives d'une bonne réceptivité aux aides proposées. Les compétences rythmiques et le niveau de pratique musicale gagneraient aussi à être évalués, pour les mêmes raisons.

Des limites existent aussi concernant les épreuves utilisées pour évaluer les apprenants aux étapes pré- et post-entraînement. Les épreuves de traitement catégoriel des consonnes sont insensibles à l'effet test-retest, mais leur fidélité est faible (Expérience 1). Pour les épreuves de perception et de production des voyelles, certaines mesures sont insensibles à l'effet test-retest, mais pas toutes, ce qui limite l'interprétation des effets, malgré une bonne fidélité (Expérience 11). Des différences entre les performances des groupes entraînés avec le



dispositif dichotique ou binaural, dès T1, ont aussi créé de telles limites (Expérience 11). Quant aux normes, elles sont encore à l'état d'ébauche car la taille des échantillons est faible (Expériences 1, 4 et 10). Il serait aussi intéressant d'inclure des adultes non-étudiants (normo-lecteurs et dyslexiques) et même de créer des normes séparées pour les étudiants et les non-étudiants. Cela permettrait de disposer de normes plus représentatives. Il manque aussi des normes réalisées auprès de locuteurs anglophones pour les épreuves de perception catégorielle afin de permettre des comparaisons à ce qui est attendu chez des locuteurs natifs.

Tester des participants dyslexiques poursuivant des études universitaires crée une limite pour la généralisation des effets au sein de cette pathologie. Les données confirment certaines lacunes par rapport au groupe des normo-lecteurs, par exemple pour l'intonation des questions anglaises (Expérience 14), mais le niveau était tout de même globalement bon avant l'entraînement, avec par exemple une différence déjà produite entre les voyelles courtes et longues (Expérience 12). Encore peu de jeunes dyslexiques accèdent aux études supérieures universitaires (Colé & Sprenger-Charolles, 2021). Ceux qui entreprennent ces études pourraient être ceux dont la rééducation orthophonique a particulièrement bien fonctionné, ou dont la dyslexie est moins sévère. Cela pose une limite pour la généralisation des résultats à la population dyslexique. Prolonger la recherche avec de jeunes adultes ayant fait moins d'études permettrait sans doute de s'adresser à une population qui a encore plus besoin d'aide, et pour laquelle la marge de progression pourrait aussi être plus large. Nous souhaitons, avant la pandémie, proposer ces entraînements à des participants plus jeunes, dans un lycée agricole où beaucoup d'élèves sont dyslexiques (15% dans le lycée partenaire). Cela nous aurait permis de limiter le biais de recrutement. La crise sanitaire liée au COVID-19 a empêché la réalisation de cette partie de l'étude.

Enfin, à cause de la pandémie de COVID-19, certains protocoles ont été réalisés à distance (Expériences 11 et 13). Cela rend compte au moins en partie de différences observées entre certains groupes avant l'entraînement. Toutefois, ces contraintes ont aussi permis de réfléchir aux alternatives et de mettre en place des protocoles entièrement à distance. En connaissant mieux les limites de ce mode d'évaluation et de conduite d'entraînement, il devient plus facilement envisageable de créer des dispositifs d'aide pour différents publics, non-mobiles, géographiquement éloignés ou dispensés d'assiduité à l'université.

## 8.7. Perspectives pour de futures recherches et applications

Cette recherche ouvre des perspectives et encourage à déployer des efforts pour se donner les moyens de contribuer à une meilleure réussite pour un apprentissage réputé, à juste raison, particulièrement difficile : celui de l'anglais L2 pour des apprenants francophones tardifs, surtout s'ils sont dyslexiques.

Une des pistes est d'imaginer quelles populations pourraient bénéficier de programmes d'entraînement brefs, mais ciblés, tels que ceux évalués ici.

Le dispositif de stimulation dissociée des hémisphères cérébraux selon la durée des informations à extraire afin qu'elles soient traitées avec précision pourrait être envisagé pour la remédiation de troubles du langage en dehors de la dyslexie. Les adultes dyslexiques ont réalisé des progrès rapidement grâce à un entraînement qui utilisait cette technique de stimulation. Des patients dont le système phonologique est imparfait pour d'autres raisons pourraient se le voir proposer. Les enfants atteints de TDL, par exemple, ou malentendants. Cela demanderait des adaptations, s'il s'agit de participants malentendants, car les déficits auditifs sont parfois asymétriques. Un ajustement du volume sonore adapté à l'une et l'autre oreille serait nécessaire pour une aide adaptée. La presbycusie, qui concerne cette fois davantage les seniors, est un autre handicap qui entraîne des erreurs d'identification de phonèmes et limite fortement la communication en isolant le patient. Les exercices imaginés seraient peut-être mieux réussis par les personnes qui en souffrent, si des phonèmes fréquemment confondus étaient adressés prioritairement à l'oreille et à l'hémisphère les plus indiqués pour un traitement temporel adapté. Cela pourrait favoriser, le temps de l'exercice, un traitement plus précis de ces phonèmes pour leurs différentes dimensions, temporelles mais aussi spectrales. Ces moments pourraient être l'occasion de mieux sélectionner les indices acoustiques à prendre en compte, et sur lesquels continuer, ensuite, à davantage porter attention. Cette sélection permettrait de créer de nouvelles routines d'analyse du signal de parole (Strange, 2011) pour une meilleure adaptation aux limites imposées par la presbycusie.

Une autre population qui pourrait bénéficier de ces apprentissages serait des francophones plus jeunes (en formation initiale). Le projet de cette thèse était de proposer les entraînements sur les consonnes et les voyelles à des lycéens. Cependant ce projet a été retravaillé suite à la pandémie de COVID-19 et à la fermeture des écoles. Cette partie du travail a été remplacée par l'étude sur l'aide à l'apprentissage de la prosodie. Travailler chez l'enfant permettrait de tester les outils d'aide à l'apprentissage auprès d'une population chez

qui des habitudes de perception et de production inadaptées de l'anglais sont moins installées que chez l'adulte.

Un autre prolongement envisagé est d'évaluer un potentiel transfert de compétences depuis la perception vers la production de mots à la suite des entraînements sur les consonnes anglaises. Un test a déjà été élaboré pour cela. Il permet d'enregistrer les productions de mots anglais après ces entraînements. Cela permettrait d'observer si l'apprentissage réalisé est suffisamment profond pour permettre ce type de transfert, qui s'est par exemple produit à la suite de l'entraînement portant sur les voyelles anglaises. Les épreuves élaborées pour tester la production sont des tâches de lecture avec et sans modèle audio. Deux hypothèses peuvent être émises sur ce transfert. La première est qu'avant l'entraînement les participants pourraient produire les consonnes sonores anglaises avec un long VOT négatif dans les deux tâches, mais qu'ils produiraient déjà une aspiration (caractérisant les sourdes anglaises) dans la tâche avec modèle audio. De dernier trait est une particularité saillante de l'anglais pour les francophones, et l'adopter l'implique pas d'inhiber un trait, comme c'est le cas pour le long VOT négatif des sonores en français. L'autre hypothèse est qu'après l'entraînement les consonnes sourdes pourraient être produites plus correctement qu'avant celui-ci, dans les deux tâches. Pour les consonnes sonores anglaises, le VOT négatif pourrait être réduit pour sa durée, ou pour son intensité. Les résultats d'une étude pilote auprès d'adultes normolecteurs montrent qu'après l'entraînement avec dispositif dichotique, les participants produisent en anglais une aspiration pour les occlusives sourdes, réduisent l'intensité du VOT négatif pour les occlusives sonores tout en augmentant pour celles-ci l'intensité de la voyelle suivante, ce qui crée l'illusion d'un VOT négatif absent.

Par ailleurs, une analyse complémentaire est en cours à propos de l'effet de l'entraînement sur la qualité des voyelles anglaises /i/ et /i:/ en production. Elle évalue le poids accordé aux indices de durée et aux indices spectraux (différence entre les F1 et F2) qui différencient ces voyelles, avant et après l'entraînement. L'hypothèse pourrait suivre la prédiction de Flege (1997) selon laquelle il est possible d'apprendre deux nouveaux phonèmes dans une L2 en considérant l'un comme un exemplaire d'un phonème de la L1 et l'autre comme relevant d'une autre catégorie. Dans le cas d'un entraînement perceptif ciblant ces deux voyelles et plaçant l'apprenant dans des conditions optimales pour bien les traiter perceptivement (dispositif dichotique), l'apprenant pourrait développer de telles nouvelles connaissances phonologiques. Le /i:/ pourrait être assimilé au /i/ français, considéré comme un exemplaire simplement atypique par sa durée. Si le dispositif permet un traitement précis

des phonèmes, il pourrait conduire l'apprenant à découvrir que le contenu formantique du /ɪ/ n'est pas celui d'un /i/ français, mais est proche de celui du /e/ français. Une assimilation du /ɪ/ au /e/ ne pourrait pas s'expliquer par la prise en compte des informations orthographiques mises à disposition pendant l'entraînement, mais par une analyse fine des composantes spectrales. Dans ce cas, les /ɪ/ et /i:/ anglais ne seraient plus représentés comme des exemplaires d'une même catégorie.

Pour évaluer cela, une nouvelle expérience de *oddball* passif en EEG a été réalisée et les données sont acquises auprès de 20 participants normo-lecteurs, avant et après un entraînement dichotique ciblant les voyelles /ɪ/ et /i:/. Les données sont en cours d'analyse. Plusieurs conditions permettent d'isoler la sensibilité à la durée des voyelles en maintenant constant le contenu formantique, et la sensibilité au contenu formantique en égalisant la durée des stimuli standards et déviants. L'hypothèse théorique suppose que l'entraînement dichotique ne conduit pas à accorder un poids démesuré aux indices de durée, mais assure un traitement plus précis des phonèmes, y compris pour leurs caractéristiques formantiques. L'hypothèse opérationnelle est que la MMN devrait augmenter à T2, particulièrement lorsque l'information formantique du /ɪ/ est présente dans le déviant.

Ces données pourront être mises en relation avec celles acquises en production, avant et après l'entraînement. La distance entre les F1 et F2 du /ɪ/ produit pourrait se rapprocher de celle du /e/, qui est aussi typique du /ɪ/ anglais. Cette modification n'est pas attendue pour le /i:/. Si l'entraînement dichotique permet d'accéder à une analyse plus fine du contenu des voyelles, le contenu formantique du /i:/ pourrait aussi changer pour se rapprocher de la prononciation avec diphtongaison du /i:/ produit par des anglophones (Krzonowski *et al.*, 2018). Cette dernière hypothèse s'appuie aussi sur le fait que, pendant l'entraînement, les /i:/ sont présentés à l'oreille gauche-HD, et que l'HD est censé être particulièrement compétent pour le traitement des indices spectraux (Zatorre, 1997). Dans une expérience, Ylinen *et al.* (2017) ont eux aussi fait varier ces paramètres temporels et formantiques, mais ceci dans des entraînements perceptifs. Ils ont montré une amélioration de la perception des voyelles anglaises chez des apprenants finois (augmentation de la MMN) après l'entraînement, de manière asymétrique puisque la MMN augmentait davantage si le déviant était /i:/ parmi de /ɪ/ que dans le cas inverse. Dans l'expérience en cours que nous évoquons ici, l'entraînement n'est pas réalisé sur des stimuli modifiés, ce sont les épreuves pré- et post-entraînement qui proposent des stimuli artificiellement transformés pour leur durée ou leur contenu formantique.

Une extension envisagée de cette recherche est de proposer les entraînements dichotiques de manière adaptée à des anglophones apprenant le français. Un entraînement sur le voisement des consonnes françaises est déjà au point (Expériences 2 et 3). Sur le modèle des Expériences 13 et 14, il serait aussi possible d'utiliser l'amorçage rythmique musical et le retour visuel pour aider des personnes en difficulté avec la perception de la différence d'intonation entre les questions et les ordres en français. Un entraînement spécifique à cette différence est déjà testé au sein de notre équipe auprès d'enfants dont la pathologie touche le langage : des enfants atteints d'un TDL ou de troubles du spectre autistique. L'aide à l'apprentissage de la durée des voyelles pourrait aussi être appliquée à d'autres langues quantitatives, comme l'allemand, à destination d'apprenants dont la langue native n'est pas quantitative, comme le français.

La dernière perspective de ce travail est en cours de réalisation. Il s'agit de réaliser le développement économique des outils d'aide à l'apprentissage. Pour cela, nous nous faisons accompagner par une Société d'Accélération du Transfert de Technologies (SATT PULSALYS) de l'Université de Lyon. Cet accompagnement permettra le développement technologique et commercial des outils.

## **8.8. Conclusion**

L'ensemble des résultats de cette recherche apporte de nouveaux arguments au modèle SLM-r (Flege & Bohn, 2021) et à la possibilité pour des apprenants tardifs de progresser dans l'apprentissage de nouveaux phonèmes ou de caractéristiques prosodiques dans une L2. Ils montrent aussi que des adultes, normo-lecteurs ou dyslexiques, peuvent faire des progrès dans ces domaines après des entraînements brefs conçus à partir de prédictions inspirées du modèle AST (Poeppel, 2003 ; Poeppel *et al.*, 2008). Ce modèle suppose une implication de l'HG et de l'HD dans les traitements temporels du signal de parole, avec des rôles distincts et complémentaires sur le plan de la fréquence d'échantillonnage de l'information. L'origine possible de cette distinction pourrait venir des différences de fréquence des oscillations cérébrales, plus rapide dans le cortex auditif gauche que dans le cortex auditif droit, selon un modèle du fonctionnement neuro-cognitif centré sur les rythmes cérébraux (Giraud & Poeppel, 2012 ; Hickok & Poeppel, 2000, 2007 ; Poeppel *et al.*, 2008). Le dispositif dichotique testé dans la thèse visait à exploiter cette asymétrie hémisphérique pour faire percevoir aux apprenants de manière optimale le voisement de consonnes anglaises ou la durée de voyelles anglaises lors d'exercices. Après un entraînement réalisé avec ce dispositif,

les premiers résultats présentés sont encourageants, car les performances s'améliorent un peu plus qu'avec un dispositif binaural, un transfert de compétences s'est aussi produit, la sensibilité pré-attentionnelle aux phonèmes appris s'est élevée et l'engagement respectif des hémisphères gauche et droit pour le traitement des nouvelles consonnes s'est modifié. L'ensemble suggère que le dispositif consistant à adresser les phonèmes prioritairement à un hémisphère en fonction de ses compétences pour le traitement d'indices acoustiques d'une durée conforme à celle des traits phonétiques du phonème est une piste à explorer pour de futurs dispositifs d'aide à l'apprentissage. L'hypothèse d'un bénéfice apporté par une amorce rythmique au traitement précis de l'intonation des questions anglaises et à l'apprentissage d'une particularité en prosodie anglaise s'appuyait sur les mêmes cadres théoriques et les données montrent l'effet positif de telles amorces, que les adultes soient normo-lecteurs ou dyslexiques. Cela procure, indirectement et de façon encore modeste, quelques éléments de plus en faveur de l'intérêt d'une prise en compte des apports des neurosciences au sujet des oscillations cérébrales, y compris lorsqu'il s'agit d'imaginer des moyens d'améliorer les apprentissages. L'approche du fonctionnement cérébral centrée sur une dynamique oscillatoire organisée en cascade et sur son rôle dans la synchronisation de l'attention au signal perçu est source d'idées nouvelles pour des applications dans le domaine de l'aide au traitement du langage (Fiveash *et al.*, 2021 ; Habib, 2021 ; Bigand & Tillmann, 2021; Tillmann & Bedoin, soumis) ; elle l'est aussi dans le domaine de l'aide aux apprentissages et à la remédiation. Cependant, les hypothèses initiales de la thèse ne sont que partiellement vérifiées, notamment à cause de l'absence d'effet du dispositif d'écoute dichotique sur les bénéfices apportés par un entraînement perceptif portant sur les voyelles. D'autres expériences d'entraînement perceptif avec dispositif dichotique sont nécessaires pour tenter de répliquer les effets, les tester sur d'autres aspects temporels en phonologie, les adapter et les tester auprès d'apprenants anglophones pour le français L2, tester à nouveau les effets potentiels sur l'apprentissage des durées des voyelles avec des stimuli mieux choisis pour la durée des phonèmes. La généralisation des résultats serait prématurée, et il faut être prudent sur leur portée actuelle. Nous avons vu que des hypothèses alternatives pour l'effet du dispositif dichotique existent, et d'autres peuvent survenir. Des expériences destinées à répondre plus directement à des questions sur les oscillations cérébrales sont aussi à prévoir en mesurant celles-ci directement. Les premiers résultats sont toutefois encourageants et s'inscrivent dans une dynamique de recherche appliquée fondée sur la recherche fondamentale en vue d'aider les apprenants adultes, atteints de troubles du langage ou non, dans l'apprentissage d'une nouvelle langue.

## Références

---

- Abadie, R., & Bedoin, N. (2016). Les étudiants dyslexiques à l'Université. *Neurologies*, 19(192), 298–303.
- Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *The Journal of Neuroscience*, 28(15), 3958-3965. <https://doi.org/10.1523/JNEUROSCI.0187-08.2008>
- Abrams, D.A., Nicol, T., Zecker, S., & Kraus, N. (2009) Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience*, 29, 7686–7693. <https://doi.org/10.1523/JNEUROSCI.5242-08.2009>
- Albouy, P., Benjamin, L., Morillon, B., & Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Sciences*, 367, 1043-1047
- Allen, E. J., Burton, P. C., Olman, C. A., & Oxenham, A. J. (2017). Representations of pitch and timbre variation in human auditory cortex. *Journal of Neuroscience*, 37(5), 1284-1293. <https://doi.org/10.1523/JNEUROSCI.2336-16.2016>
- Alonso-Búa, B., Díaz, F., & Ferraces, M. J. (2006). The contribution of AERPs (MMN and LDN) to studying temporal vs. linguistic processing deficits in children with reading difficulties. *International Journal of Psychophysiology*, 59(2), 159–167. <https://doi.org/10.1016/j.ijpsycho.2005.03.020>
- Alves, U., & Luchini, P. (2017). Effects of perceptual training on the identification and production of word-initial voiceless stops by Argentinian learners of English. *A Journal of English Language*, 70(3), 15-32. <https://doi.org/10.5007/2175-8026.2017v70n3p15>
- Anderson, J. M., Gilmore, R., Roper, S., Crosson, B., Bauer, R. M., Nadeau, S., Beversdorf, D. Q., Cibula, J., Rogish, M., Kortencamp, S., Hughes, J. D., Gonzalez Rothi, L. J., & Heilman, K. M. (1999). Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model. *Brain and Language*, 70(1), 1-12. <https://doi.org/10.1006/brln.1999.2135>
- Araujo, S., Faisca, L., Bramao, I., Reis, A., & Petersson, K. M. (2015). Lexical and sublexical orthographic processing: An ERP study with skilled and dyslexic adult readers. *Brain and Language*, 141, 16-27. <https://doi.org/10.1016/j.bandl.2014.11.007>
- Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085. <https://doi.org/10.1093/cercor/bhu103>
- Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398. <https://doi.org/10.1016/j.tics.2012.05.003>



- Artieda, J., Pastor, M. A., Lacruz, F., & Obeso, J. A. (1992). Temporal discrimination is abnormal in Parkinson's disease. *Brain*, *115*(1), 199–210. <https://doi.org/10.1093/brain/115.1.199>
- Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, *52*(4), 1135–1145. <https://doi.org/10.1111/j.1467-8624.1981.tb03159.x>
- Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, *4*(2), 1–10. <https://doi.org/10.1126/sciadv.aao3842>
- Assaneo, M.F., Rimmele, J.M., Orpella, J., Ripollés, P., de Diego-Balaguer, R., & Poeppel, D. (2019). The lateralization of speech-brain coupling is differently modulated by intrinsic auditory and top-down mechanisms. *Frontiers in Integrative Neuroscience*, *13*(28), 1-11. <https://doi.org/10.3389/fnint.2019.00028>
- Atienza, M., Cantero, J. L., & Dominguez-Marin, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning and Memory*, *9*(3), 138–150. <https://doi.org/10.1101/lm.46502>
- Badin, P., Tarabalka, Y., Elisei, F., & Bailly, G. (2010). Can you “read” tongue movements? Evaluation of the contribution of tongue display to speech understanding. *Speech Communication*, *52*(6), 493–503. <https://doi.org/10.1016/j.specom.2010.03.002>
- Bahl, M., Plante, E., & Gerken, L. A. (2009). Processing prosodic structure by adults with language-based learning disability. *Journal of Communication Disorders*, *42*(5), 313–323. <https://doi.org/10.1016/j.jcomdis.2009.02.001>
- Baker, W., & Trofimovich, P. (2005). Interaction of native- and second- language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*(1), 1-27. <https://doi.org/10.1177/00238309050480010101>
- Baker, W., Trofimovich, P., Flege, J. E., Mack, M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech*, *51*(4), 317-342. <https://doi.org/10.1177/0023830908099068>
- Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, *42*(3), 254-311. <https://doi.org/10.1006/cogp.2000.0738>
- Bedoin, N. (2003). Sensitivity to voicing similarity in printed stimuli: Effect of a training programme in dyslexic children. *Journal of Phonetics*, *31*(3-4), 541-546. [https://doi.org/10.1016/S0095-4470\(03\)00044-5](https://doi.org/10.1016/S0095-4470(03)00044-5)
- Bedoin, N. (2016). Apports de la neuropsychologie dans les troubles spécifiques du langage écrit. In M. Kremer, E. Lederlé, & Ch. Maeder (Eds.), *Le Guide de l'Orthophoniste* (71 pages), Volume 3, Chapitre 1. Lavoisier.



- Bedoin, N. (2017). Rééquilibrer les analyses visuo-attentionnelles globales et locales pour améliorer la lecture chez des enfants dyslexiques de surface. *ANAE, 19, Tome III*(148), 276-294.
- Bedoin, N., Abadie, R., Krzonowski, J., Ferragne, E., & Marcastel, A. (2019). A combined forced-attention dichotic listening - Go/Nogo task to assess response inhibition and interference suppression: An auditory Event-Related Potential investigation. *Neuropsychology, 33*(8), 1136-1150. <https://doi.org/10.1037/neu0000586>
- Bedoin, N., Besombes, A. M., Escande, E., Dumont, A., Lalitte, P., & Tillmann, B. (2018). Boosting syntax training with temporally regular musical primes in children with cochlear implants. *Annals of Physical and Rehabilitation Medicine, 61*(6), 365–371. <https://doi.org/10.1016/j.rehab.2017.03.004>
- Bedoin, N., Besombes, A.-M., Escande, E., Dumont, A., & Tillmann, B. (2016a). Déficiences morphosyntaxiques chez des enfants sourds implantés : Hypothèses et pistes de remédiation par le rythme musical. *Rééducation Orthophonique, 268*, 261-288.
- Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016b). Temporally regular musical primes facilitate subsequent syntax processing in children with Specific Language Impairment. *Frontiers in Neuroscience, 10*(245), 1-11, <https://doi.org/10.3389/fnins.2016.00245>
- Bedoin, N., Ferragne, E., Lopez, C., Herbillon, V., De Bellescize, J., & des Portes, V. (2011). Atypical hemispheric asymmetries for the processing of phonological features in children with rolandic epilepsy. *Epilepsy and Behavior, 21*(1), 42-51. <https://doi.org/10.1016/j.yebeh.2011.02.026>
- Bedoin, N., Ferragne, E., & Marsico, E. (2010). Hemispheric asymmetries depend on the phonetic feature: A dichotic study of place of articulation and voicing in French stops. *Brain and Language, 115*(2), 133-140. <https://doi.org/10.1016/j.bandl.2010.06.001>
- Bedoin N., & Krifi, S. (2009). The complexity of phonetic features organization in reading. In F. Pellegrino, E. Marsico, I. Chitoran, C. Coupé, *Approaches to phonological complexity*. Berlin: Phonology & Phonetics.
- Bedoin, N., Krzonowski, J., & Ferragne, E. (2013). How voicing, place, and manner of articulation differently modulate event-related potentials associated with response inhibition. *Proceedings of INTERSPEECH*, (août 2013) (oral communication), Lyon, France.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309–312. <https://doi.org/10.1038/35002078>
- Belin, P., Zilbovicius, M., Crozier, S., Thivard, L., Fontaine, A., Masure, M. C., Samson, Y. (1998). Lateralization of speech and auditory temporal processing. *Journal of Cognitive Neuroscience, 10*, 536-540. <https://doi.org/10.1162/089892998562834>.
- Benikos, N., Johnstone, S. J., & Roodenrys, S. J. (2013). Varying task difficulty in the Go/Nogo task: The effects of inhibitory control, arousal, and perceived effort on ERP

- components. *International Journal of Psychophysiology*, 87(3), 262-272.  
<https://doi.org/10.1016/j.ijpsycho.2012.08.005>
- Benítez-Burraco, A., & Murphy, E. (2019). Why brain oscillations are improving our understanding of language. *Frontiers in Behavioral Neuroscience*, 13, 110.  
<https://doi.org/10.3389/fnbeh.2019.00190>
- Berdal-Masuy, F., Briet, G., & Pairon, J. (2004). Apprendre seul, à son rythme et encadré. *Etudes de Linguistique Appliquée*, 2(134), 173-190.
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, 2, 1-12.  
<https://doi.org/10.3389/fpsyg.2011.00094>
- Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, 25(3-4), 399-410.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (Issue January 1995, pp. 171–204).
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109(2), 775–794.  
<https://doi.org/10.1121/1.1332378>
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-O. Bohn and M. J. Munro (Eds.) *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Benjamins, Amsterdam.
- Bewick, V., Cheek, L., & Ball, J. (2004). Statistics review 13: Receiver Operating Characteristic curves. *Critical Care*, 8(6), 508-512. <https://doi.org/10.1186/cc3000>
- Bhide, A., Power, A., & Goswami, U. (2013). A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, 7(2), 113-123. <https://doi.org/10.1111/mbe.12016>
- Bigand, E., & Tillmann, B. (2021). Near and far transfer: Is music special? *Memory & Cognition*. Advance online publication. <https://doi.org/10.3758/s13421-021-01226-6>
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10(5), 512-28. <https://doi.org/10.1093/cercor/10.5.512>
- Bishop, D. V. M., Hardiman, M. J., & Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Developmental Science*, 14(2), 402-416.  
<https://doi.org/10.1111/j.1467-7687.2010.00990.x>

- Blumstein, S. E., Baker, E., & Goodglass, H. (1977a). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, *15*, 19–30.
- Blumstein, S. E., Cooper, W. E., Zurif, E. B., & Caramazza, A. (1977b). The perception and production of Voice-Onset Time in aphasia. *Neuropsychologia*, *15*(3).  
[https://doi.org/10.1016/0028-3932\(77\)90089-6](https://doi.org/10.1016/0028-3932(77)90089-6)
- Blumstein, S. E., Myers, E. B., & Rissman, J. (2005). The perception of voice onset time: An fMRI investigation of phonetic category structure. *Journal of Cognitive Neuroscience*, *17*(9), 1353-1366. <https://doi.org/10.1162/0898929054985473>
- Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, *8*(3), 389-395.  
<https://doi.org/10.1038/nn1409>
- Boets, B., de Beeck, H., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., Bulthé, J., Snaert, S., Wouters, J., & Ghesquière, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, *342*, 1251-1254.  
<https://doi.org/10.1126/science.1244333>
- Boets, B., Vandermosten, M., Poelmans, H., Luts, H., Wouters, J., & Ghesquière, P. (2011). Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Research in Developmental Disabilities*, *32*, 560-570.  
<https://doi.org/10.1016/j.ridd.2010.12.020>
- Bogacka, A. (2004). On the perception of English high vowels by Polish learners of English. In C. U. Press. (Ed.), *University of Cambridge second postgraduate conference in language research* (pp. 43-50). Cambridge, UK: Cambridge University Press.
- Bogliotti, C. (2005). *Perception catégorielle et perception allophonique: incidences de l'âge, du niveau de lecture et des couplages entre prédispositions phonétiques*. [Thèse de doctorat, Université Paris-Diderot - Paris VII].
- Bogliotti, C., Messaoud-Galusi, S., & Serniclaes, W. (2002). Relations entre la perception catégorielle de la parole et l'apprentissage de la lecture. Proceedings of the *XXIVèmes Journées d'Etudes sur la Parole*, Nancy (24-27 juin 2002), 197-200.
- Böhning, D., Böhning, W., & Holling, H. (2008). Revisiting Youden's index as a useful measure of the misclassification error in meta-analysis of diagnostic studies. *Statistical Methods in Medical Research*, *17*(6), 543-554.  
<https://doi.org/10.1177/0962280207081867>
- Bonacina, S., Cancer, A., Lanzi, P. L., Lorusso, M. L., & Antonietti, A. (2015). Improving reading skills in students with dyslexia: The efficacy of a sublexical training with rhythmic background. *Frontiers in Psychology*, *6*, 1-8.  
<https://doi.org/10.3389/fpsyg.2015.01510>
- Bosse, M. L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, *104*(2), 198-230.  
<https://doi.org/10.1016/j.cognition.2006.05.009>

- Bosse, M. L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: A cross-sectional study. *Journal of Research in Reading*, 32, 230-253. <https://doi.org/10.1111/j.1467-9817.2008.01387.x>
- Boucher, V. J., Gilbert, A. C., & Boutheina, J. (2019). The role of low-frequency neural oscillations in speech processing: Revisiting delta entrainment. *Journal of Cognitive Neuroscience*, 31(8), 1205-1215. <https://doi.org/10.1162/jocn.a.01410>
- Bouhon, M. (2018). *Evaluation d'exercices audio-visuels et dichotique pour l'apprentissage de consonnes et voyelles anglaises : Un pack en perception et production de parole*. Mémoire de Master 2 de Neuropsychologie et Neurosciences Cliniques, Université Lumière Lyon 2.
- Boutros, N. N., Korzyukov, O., Jansen, B., Feingold, A., & Bell, M. (2004). Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. *Psychiatry Research*, 126(3), 203–215. <https://doi.org/10.1016/j.psychres.2004.01.007>
- Bradlow, A. R. (2008). Training non-native language sound patterns: Lessons for training Japanese adults to the /r/-/l/ contrast. In J. G. Hansen Edwards and M. L., Zampini (Eds). *Phonology and second language acquisition*. Benjamins.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tokhura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101(4), 2299-2310. <https://doi.org/10.1121/1.418276>
- Broyd, S. J., Johnstone, S. J., Barry, R. J., Clarke, A. R., McCarthy, R., Selikowitz, M., & Lawrence, C. A. (2005). The effect of methylphenidate on response inhibition and the event-related potential of children with attention deficit/hyperactivity disorder. *International Journal of Psychophysiology*, 58(1), 47-58. <https://doi.org/10.1016/j.ijpsycho.2005.03.008>
- Bruni, O., Ferri, R., Novelli, L., Finotti, E., Terribili, M., Troianiello, M., Valente, D., Sabatello, U., & Curatolo, P. (2009). Slow EEG amplitude oscillations during NREM sleep and reading disabilities in children with dyslexia. *Developmental Neuropsychology*, 34(5), 539-551. <https://doi.org/10.1080/87565640903133418>
- Bryden, M. P., & Sprott, D. A. (1981). Statistical determination of degree of laterality. *Neuropsychologia*, 19, 571–581. [https://doi.org/10.1016/0028-3932\(81\)90023-3](https://doi.org/10.1016/0028-3932(81)90023-3)
- Buchman, A. S., Garron, D. C., Trost-Cardamone, J. E., Wichter, M. D., & Schwartz, M. (1986). Word deafness: One hundred years later. *Journal of Neurology, Neurosurgery and Psychiatry*, 49(5), 489-499. <https://doi.org/10.1136/jnnp.49.5.489>
- Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, 12, 679–690. <https://doi.org/10.1162/089892900562309>

- Buzsáki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679), 1926–1929. <https://doi.org/10.1126/science.1099745>
- Cabrera, L., & Gervain, J. (2020). Speech perception at birth: The brain encodes fast and slow temporal information. *Science Advances*, 6, 1-9. <https://doi.org/10.1126/sciadv.aba7830>
- Calderone, D. J., Lakatos, P., Butler, P. D., & Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends in Cognitive Sciences*, 18(6), 300–309. <https://doi.org/10.1016/j.tics.2014.02.005>
- Calet, N., Gutiérrez-Palma, N., Defior, S., & Jiménez-Fernández, G. (2019). Linguistic and non-linguistic prosodic skills in Spanish children with developmental dyslexia. *Research in Developmental Disabilities*, 90, 92-100. <https://doi.org/10.1016/j.ridd.2019.04.013>
- Callan, D. E., Tajima, K., Callan, A. M., Kubo, R., Masaki, S., & Akahane-Yamada, R. (2003). Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *NeuroImage*, 19(1), 113-124. [https://doi.org/10.1016/S1053-8119\(03\)00020-X](https://doi.org/10.1016/S1053-8119(03)00020-X)
- Cancer, A., Bonacina, S., Antonietti, A., Salandi, A., Molteni, M., Lorusso, M. L., & Bella, S. D. (2020). The effectiveness of interventions for developmental dyslexia: Rhythmic reading training compared with hemisphere-specific stimulation and action video games. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.01158>
- Canette, L.-H., Bedoin, N., Lalitte, P., Bigand, E., & Tillmann, B. (2020a). The regularity of rhythmic primes influences syntax processing in adults. *Auditory Perception & Cognition*, 2(3), 163-179. <https://doi.org/10.1080/25742442.2020.1752080>
- Canette, L.-H., Fiveash, A., Krzonowski, J., Corneyllie, A., Lalitte, P., Thompson, D., Trainor, L., Bedoin, N., & Tillmann, B. (2020b). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, 138. <https://doi.org/10.1016/j.neuropsychologia.2019.107324>.
- Canette, L.-H., Lalitte, P., Bedoin, N., Pineau, M., Bigand, E., & Tillmann, B. (2020c). Rhythmic and textural musical sequences differently influence syntax and semantic processing in children. *Journal of Experimental Child Psychology*, 191. <http://doi.org/10.1016/j.jecp.2019.104711>.
- Caplan, D., Gow, D., & Makris, N. (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*, 45(2), 293-298. <https://doi.org/10.1212/wnl.45.2.293>
- Capliez, M. (2011). Typologie des erreurs de production d'anglais des francophones : Segments vs. suprasegments. *Cahiers de l'Aplut*, XXX(3), 44-6.
- Carlet, A., & Cebrian, J. (2019). Assessing the effect of perceptual training on L2 vowel identification, generalization and long-term Effects. In A. M. Nyvad, M. Hejná, A. Højen, A. B. Jespersen, & M. H. Sørensen (Eds), *A sound approach to language matters: In honor of Ocke-Schwen Bohn* (91-119). <https://doi.org/10.7146/aul.322.218>



- Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica*, *155*, 43–50. <https://doi.org/10.1016/j.actpsy.2014.12.002>
- Cason, N., Marmursztejn, M., D’Imperio, M., & Schön, D. (2019). Rhythmic abilities correlate with L2 prosody imitation abilities in typologically different languages. *Language and Speech*, *63*(1), 149-165. <https://doi.org/10.1177/0023830919826334>
- Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652-2658. <https://doi.org/10.1016/j.neuropsychologia.2012.07.018>
- Cassim, R., Tallcott, J. B., & Moores, E. (2014). Adults with dyslexia demonstrate large effects of crowding and detrimental effects of distractors in a visual tilt discrimination task. *PLoS One*, *9*(9). <https://doi.org/10.1371/journal.pone.0106191>
- Catts, H. (1989). Speech production deficits and reading disabilities. *Journal of Speech and Hearing Research*, *54*, 422-428. <https://doi.org/10.1044/jshd.5403.422>
- Cavalli, E., Colé, P., Leloup, G., Poracchia-George, F., Sprenger-Charolles, L., & El Ahmadi, A. (2018). Screening for dyslexia in French-speaking university students: An evaluation of the detection accuracy of the Alouette test. *Journal of Learning Disabilities*, *51*(3), 268-282. <https://doi.org/10.1177/0022219417704637>
- Cebrian, J., & Carlet, A. (2014). Second-language learners’ identification of target-language phonemes: A short-term phonetic training study. *Canadian Modern Language Review*, *70*(4), 474-499. <https://doi.org/10.3138/cmlr.2318>
- Cheour, M., Alho, K., Sainio, K., Rinne, T., Reinikainen, K., Pohjavuori, M., Renlund, M., Aaltonen, O., Eerola, O., & Näätänen, R. (1996). The ontogenetically earliest discriminative response of the human brain. *Psychophysiology*, *33*, 478-481. <https://doi.org/10.1111/j.1469-8986.1996.tb01074.x>
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*(5), 351-353. <https://doi.org/10.1038/1561>
- Cheour, M., Korpilahti, P., Martynova, O., & Lang, A. H. (2001). Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. *Audiology and Neuro-Otology*, *6*(1), 2-11. <https://doi.org/10.1159/000046804>
- Chern, A., Tillmann, B., Vaughan, C., & Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of Experimental Child Psychology*, *173*, 371–379. <https://doi.org/10.1101/193961>
- Chládková, K., Escudero, P., & Lipski, S. C. (2013). Pre-attentive sensitivity to vowel duration reveals native phonology and predicts learning of second-language sounds. *Brain and Language*, *126*(3), 243-252. <https://doi.org/10.1016/j.bandl.2013.05.020>

- Cho, T. & Ladefoged, P. (1999). Variations and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27, 207-229. <https://doi.org/10.1006/jpho.1999.0094>
- Chobert, J., François, C., Velay, J. L., & Besson, M. (2014). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, 24(4), 956-967. <https://doi.org/10.1093/cercor/bhs377>
- Christmann, C. A., Berti, S., Steinbrink, C., & Lachmann, T. (2014). Differences in sensory processing of German vowels and physically matched non-speech sounds as revealed by the mismatch negativity (MMN) of the human event-related brain potential (ERP). *Brain and Language*, 136, 8-18. <https://doi.org/10.1016/j.bandl.2014.07.004>
- Cleland, J., Scobbie, J. M., & Wrench, A. A. (2015). Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics and Phonetics*, 29(8-10), 575-597. <https://doi.org/10.3109/02699206.2015.1016188>
- Cogan, G. B., & Poeppel, D. (2011). A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *Journal of Neurophysiology*, 106(2), 554-563. <https://doi.org/10.1152/jn.00075.2011>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Colé, P., Cavalli, E. & Duncan, L. (2020). Les compensations de l'adulte dyslexique de niveau universitaire. In Colé, P., Cavalli, E. & Duncan, L. (Eds). *La dyslexie à l'âge adulte : approche neuropsychologique* (pp. 287-324). De boeck.
- Colé, P., & Sprenger-Charolles, L. (2021). *La dyslexie : de l'enfant à l'adulte*. Dunod.
- Collet, G., Colin, C., & Serniclaes, W. (2017). Remédiation audiophonologique de la dyslexie: (RapDys©), un logiciel visant à réorganiser la perception allophonique des sons de la parole en perception phonémique. *ANAE - Approche Neuropsychologique Des Apprentissages Chez l'Enfant*, 29(148), 257-263.
- Collet, G., Colin, C., Serniclaes, W., Hoonhorst, I., Markessis, E., Deltenre, P., & Leybaert, J. (2012). Effect of phonological training in French children with SLI: Perspectives on voicing identification, discrimination and categorical perception. *Research in Developmental Disabilities*, 33(6), 1805-1818. <https://doi.org/10.1016/j.ridd.2012.05.003>
- Corriveau, K., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex*, 45(1), 119-130. <https://doi.org/10.1016/j.cortex.2007.09.008>
- Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic auditory processing skills and Specific Language Impairment: A new look at an old hypothesis. *Journal of Speech, Language, and Hearing Research*, 50, 647-666. [https://doi.org/10.1044/1092-4388\(2007/046\)](https://doi.org/10.1044/1092-4388(2007/046))

- Crombie, M. A. (1997). The effects of specific learning difficulties (dyslexia) on the learning of a foreign language in school. *Dyslexia*, 3, 27-47. [https://doi.org/10.1002/\(SICI\)1099-0909\(199703\)3:1<27::AID-DYS43>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1099-0909(199703)3:1<27::AID-DYS43>3.0.CO;2-R)
- Crothers, J. (1978). Typology and universals of vowel systems. In J. H. Greenberg, C. A. Ferguson, E. A. Moravcsik (Eds), *Universals of human language volume 2: Phonology*. Stanford, California: Stanford University Press, pp. 93-152.
- Cuetos, F., Martínez-García, C., & Suárez-Coalla, P. (2018). Prosodic Perception Problems in Spanish Dyslexia. *Scientific Studies of Reading*, 22(1), 41-54. <https://doi.org/10.1080/10888438.2017.1359273>
- Cumming, R., Wilson, A., & Goswami, U. (2015). Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: A new look at a perceptual hypothesis. *Frontiers in Psychology*, 6, 972. <https://doi.org/10.3389/fpsyg.2015.00972>
- Cutini, S., Szűcs, D., Mead, N., Huss, M., & Goswami, U. (2016). Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. *NeuroImage*, 143, 40-49. <https://doi.org/10.1016/j.neuroimage.2016.08.012>
- Cutler, A., & Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, 20(1), 1-10. <https://doi.org/10.1177/002383097702000101>
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, 19(5), 381-385. <https://doi.org/10.1016/j.cub.2009.01.01->
- Danon-Boileau, L., & Barbier, D. (2000). *Play on*. Audim!edia: CD-Rom PC.
- De Bode, S., Sininger, Y., Healy, E. W., Mathern, G. W., & Zaidel, E. (2007). Dichotic listening after cerebral hemispherectomy: Methodological and theoretical observations. *Neuropsychologia*, 45, 2461-2466. <https://doi.org/10.1016/j.neuropsychologia.2007.03.026>
- De Boysson-Bardies, B. (2007). Comment la parole vient à l'enfant. *Revue Française de Psychanalyse*, 71, 1473-1480.
- De Bree, E., Wijnen, F., & Zonneveld, W. (2006). Word stress production in three-year-old children at risk of dyslexia. *Journal of Research in Reading*, 29(3), 304-317. <https://doi.org/10.1111/j.1467-9817.2006.00310.x>
- De Martino, S., Espesser, R., Rey, V., & Habib, M. (2001). The "temporal processing deficit" hypothesis in dyslexia: New experimental evidence. *Brain and Cognition*, 46(1-2), 104-108. [https://doi.org/10.1016/s0278-2626\(01\)80044-0](https://doi.org/10.1016/s0278-2626(01)80044-0).
- Dede, G. (2010). Utilization of prosodic information in syntactic ambiguity resolution. *Journal of Psycholinguist Research*, 39(4), 345-374. <https://doi.org/0.1007/s10936-009-9139-x>



- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, 8(4), 919-924. <https://doi.org/10.1097/00001756-199703030-00021>
- Delahaie, M., Sprenger-charolles, L., Serniclaes, W., Billard, C., Tichet, J., Pointeau, S., & Vol, S. (2004). Perception catégorielle dans une tâche de discrimination de phonèmes et apprentissage de la lecture. *Revue Française de Pédagogie*, 147, 91-105.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655-679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- Desai, R., Liebenthal, E., Waldron, E., & Binder, J. R. (2008) Left posterior temporal regions are sensitive to auditory categorization. *Journal of Cognitive Neuroscience*, 20(7), 1174-1188. <https://doi.org/10.1162/jocn.2008.20081>
- Di Liberto, G. M., Peter, V., Kalashnikova, M., Goswami, U., Burnham, D., & Lalor, E. C. (2018). Atypical cortical entrainment to speech in the right hemisphere underpins phonemic deficits in dyslexia. *NeuroImage*, 175, 70-79. <https://doi.org/10.1016/j.neuroimage.2018.03.072>
- Diehl, S. F. (1999). Listen and learn? A software review of Earobics®. *Language, Speech, and Hearing Services in Schools*, 30(1), 108-116. <https://doi.org/10.1044/0161-1461.3001.108>
- Diehl, J. J., Bennetto, L., Watson, D., Gunlogson, C., & McDonough, J. (2008). Resolving ambiguity: A psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain and Language*, 106(2), 144-152. <https://doi.org/10.1016/j.bandl.2008.04.002>
- DiFino, S. M., & Lombardino, L. J. (2004). Language learning disabilities: The ultimate foreign language challenge. *Foreign Language Annals*, 37(3), 390-400. <https://doi.org/10.1111/j.1944-9720.2004.tb02697.x>
- Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*, 81, 181-187. <https://doi.org/10.1016/j.neubiorev.2017.02.011>
- Dinkage, K. T. (1971). The inability to learn a foreign language. In G. Blaine & C. C. McArthur (Eds), *Emotional problems of the student* (pp. 185-206). New-York: Appleton-Century-Crofts.
- DiStefano, M., Marano, E., & Viti, M. (2004). Stimulus-dominance effects and lateral asymmetries for language in normal subjects and in patients with a single functional hemisphere. *Brain and Cognition*, 56, 55-62. <https://doi.org/10.1016/j.bandc.2004.05.006>
- Donchin, E. (1981). Surprise! ... Surprise? *Psychophysiology*, 18(5), 493-513. <https://doi.org/10.1017/S0140525X0133011X>

- Dorsaint-Pierre, R., Penhune, V. B., Watkins, K. E., Neelin, P., Lerch, J. P., Bouffard, M., & Zatorre, R. J. (2006). Asymmetries of the planum temporal and Heschl's gyrus: Relationship to language lateralization. *Brain*, *129*(5), 1164-1176. <https://doi.org/10.1093/brain/awl055>
- Downey, D. M., Snyder, L. E., & Hill, B. (2000). College students with dyslexia: Persistent linguistic deficits and foreign language learning. *Dyslexia*, *6*, 101-111. [https://doi.org/10.1002/\(SICI\)1099-0909\(200004/06\)6:2<101::AID-DYS154>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-0909(200004/06)6:2<101::AID-DYS154>3.0.CO;2-8)
- Dufor, O., Serniclaes, W., Sprenger-Charolles, L., & Demonet, J. F. (2007). Top-down processes during auditory phoneme categorization in dyslexia: A PET study. *NeuroImage*, *34*, 1692-1707. <https://doi.org/10.1016/j.neuroimage.2006.10.034>
- Dushanova, J., Lalova, Y., Kalonkina, A., & Tsokov, S. (2020). Speech-brain frequency entrainment of dyslexia with and without phonological deficits. *Brain Sciences*, *10*(12), 1-23. <https://doi.org/10.3390/brainsci10120920>
- Elliot, L. L., Longinotti, C., Meyer, D., Raz, I., & Zucker, K. (1981). Developmental differences in identifying and discriminating CV syllables. *Journal of the Acoustical Society of America*, *70*(3), 669-677. <https://doi.org/10.1121/1.386929>
- Escudero, P., Sisinni, B., & Grimaldi, M. (2014). The effect of vowel inventory and acoustic properties in Salento Italian learners of Southern British English vowels. *Journal of Acoustical Society of America*, *135*, 1577-1584. <https://doi.org/10.1121/1.4864477>
- Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *The Journal of the Acoustical Society of America*, *121*(6), 3814. <https://doi.org/10.1121/1.2722209>
- Fabre, D., Hueber, T., Canault, M., Bedoin, N., Acher, A., Bach, C., Lamborion, L., & Badin, P. (2016). Apport de l'échographie linguale à la rééducation orthophonique, Actes des XVIèmes Rencontres Internationales d'Orthophonie : " Orthophonie et technologies innovantes" (pp. 199-225). OrthoEdition.
- Fabre, D., Hueber, T., Canault, M., Bedoin, N., Bach, C., Labourion, L., Badin, P., Fabre, D., Hueber, T., Canault, M., Bedoin, N., Acher, A., & Topouzkhaniyan, S. (2016). Apport de l'échographie linguale à la rééducation orthophonique. XVIèmes Rencontres Internationales d'Orthophonie : « Orthophonie et technologies innovantes », Déc 2016, (pp. 199-225), Paris, France.
- Falk, S., Lanzilotti, C., & Schön, D. (2017). Tuning neural phase entrainment to speech. *Journal of Cognitive Neuroscience*, *29*(8), 1378-1389. [https://doi.org/10.1162/jocn\\_a\\_01136](https://doi.org/10.1162/jocn_a_01136)
- Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology*, *6*, 1-12. <https://doi.org/10.3389/fpsyg.2015.00847>

- Falter, C., & Noreika, V. (2014). Time processing in developmental disorders: A comparative view. In D. Lloyd & V. Arstila (Eds), *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*. MIT Press.
- Fernandez, M. A., & Smith, M. L. (2000). Comparing the fused dichotic words test and the intracarotid amobarbital procedure in children with epilepsy. *Neuropsychologia*, *38*, 1216-1228.
- Ferragne, E., Flavier, S., & Fressard, C. (2013) ROCme! software for the recording and management of speech corpora. *14th Annual Conference of the International Speech Communication Association*, Aug 2013, Lyon, France.  
<http://www.ddl.cnrs.fr/Download/Index.asp?Langue=FR&Page=ROCme>
- Fiez, J. A., Raichle, M.E., Miezin, F.M., Petersen, S.E., Tallal, P., & Katz, W.F. (1995). Studies of auditory and phonological processing effects of stimulus characteristics and task demands. *Journal of Cognitive Neuroscience*, *7*, 357-375.  
<https://doi.org/10.1162/jocn.1995.7.3.357>
- Fiveash, A., Bedoin, N., Gordon, R. L., & Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*, *35*(8), 771-791. <https://doi.org/10.1037/neu0000766>
- Fiveash, A., Bedoin, N., Lalitte, Ph., & Tillmann, B. (2020a). Rhythmic priming of grammaticality judgments in children: Duration matters. *Journal of Experimental Child Psychology*. 104885. <https://doi.org/10.1016/j.jecp.2020.104885>
- Fiveash, A., Schön, D., Canette, L. H., Morillon, B., Bedoin, N., & Tillmann, B. (2020b). A stimulus-brain coupling analysis of regular and irregular rhythms in adults with dyslexia and controls. *Brain and Cognition*, *140*, 105531.  
<https://doi.org/10.1016/j.bandc.2020.105531>
- Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015) Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE* *10*(9).  
<https://doi.org/10.1371/journal.pone.0138715>
- Flege, J. E. (1987). The production of « new » and « similar » phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47-65.  
[https://doi.org/10.1016/S0095-4470\(19\)30537-6](https://doi.org/10.1016/S0095-4470(19)30537-6)
- Flege, J. E. (1995a). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press.
- Flege, J. E. (1995b). Two procedures of training a novel second language phonetic contrast. *Applied Psycholinguistics*, *16*, 425-442.
- Flege, J. E., & Bohn, O.-S. (2021). The Revised Speech Learning Model. In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress*, 84-118. Cambridge University Press. <https://doi.org/10.1017/9781108886901.003>

- Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437-470. <https://doi.org/10.1006/jpho.1997.0052>
- Fosker, T., & Thierry, G. (2004). P300 investigation of phoneme change detection in dyslexic adults. *Neuroscience Letters*, 357, 171-174. <https://doi.org/10.1016/j.neulet.2003.12.084>
- Fosker, T., & Thierry, G. (2005). Phonological oddballs in the focus of attention elicit a normal P3b in dyslexic adults. *Cognitive Brain Research*, 24(3), 467-475. <https://doi.org/10.1016/j.cogbrainres.2005.02.019>
- Franceschini, S., Gori, S., Ruffino, M., Viola, S., Molteni, M., & Facoetti, A. (2013). Action video games make dyslexic children read better. *Current Biology*, 23(6), 462-466. <https://doi.org/10.1016/j.cub.2013.01.044>
- François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, 23(9), 2038-2043. <https://doi.org/10.1093/cercor/bhs180>
- Fraser, J., Goswami, U., & Conti-Ramsden, G. (2010). Dyslexia and Specific Language Impairment: The role of phonology and auditory processing. *Scientific Studies of Reading*, 14(1), 8-29. <https://doi.org/10.1080/10888430903242068>
- Frey, A., François, C., Chobert, J., Velay, J. L., Habib, M., & Besson, M. (2019). Music training positively influences the preattentive perception of voice onset time in children with dyslexia: A longitudinal study. *Brain Sciences*, 9(4). <https://doi.org/10.3390/brainsci9040091>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815-836. <https://doi.org/10.1098/rstb.2005.1622>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127-138. <https://doi.org/10.1038/nrn2787>
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, 32(5), 1791-1802. <https://doi.org/10.1523/JNEUROSCI.4107-11.2012>
- Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology and Neuroscience*, 25, 295-310. <https://doi.org/10.17943007>
- Gabay, Y., & Holt, L. L. (2015). Incidental learning of sound categories is impaired in developmental dyslexia. *Cortex*, 73, 131-143. <https://doi.org/10.1016/j.cortex.2015.08.008>

- Galantucci, B., Fowler, C., & Golstein, L. (2009). Perceptuomotor compatibility effects. *Attention, Perception, & Psychophysics*, *71*, 1138-1149. <https://doi.org/10.3758/APP.71.5.1138>
- Gallagher, L. (2013). The effectiveness of ultrasound technology as a visual biofeedback tool on the productive speech intelligibility of adolescents and young adults with a hearing. Proceedings from *Evidence Based Practices of Clinicians*. London, ON: The University of Western Ontario.
- Gandour, J., Tong, Y., Talavage, T., Wong, D., Dziedzic, M., Xu, Y., Li, X., & Lowe, M. (2007). Neural basis of first and second language processing of sentence-level linguistic prosody. *Human Brain Mapping*, *28*(2), 94-108. <https://doi.org/10.1002/hbm.20255>
- Ganschow, L., & Myer, B. (1988). Profiles of frustration: Second language learners with specific learning disabilities. In J. Lalande (Ed.), *Shaping the future of foreign language education* (pp. 32-53). Lincolnwood, IL: National Text.
- Gauger, L. M., Lombardino, L. J., & Leonard, C. M. (1997). Brain morphology in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1272-1284. <https://doi.org/10.1044/jslhr.4006.1272>
- Geschwind, N., & Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region. *Science*, *161*(3837), 186-187. <https://doi.org/10.1126/science.161.3837.186>
- Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, *2*, 1-13. <https://doi.org/10.3389/fpsyg.2011.00130>
- Ghitza, O. (2017). Acoustic-driven delta rhythms as prosodic markers. *Language, Cognition and Neuroscience*, *32*(5), 545-561. <https://doi.org/10.1080/23273798.2016.1232419>
- Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, *66*(1-2), 113-126. <https://doi.org/10.1159/000208934>
- Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron*, *56*(6), 1127-1134. <https://doi.org/10.1016/j.neuron.2007.09.038>
- Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., & Kleinschmidt, A. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*, *84*(3), 1588-1598. <https://doi.org/10.1152/jn.2000.84.3.1588>
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, *15*(4), 511-517. <https://doi.org/10.1038/nn.3063>



- Girod-Roux, M., Hueber, T., Fabre, D., Gerber, S., Canault, M., Bedoin, N., Acher, A., Béziaud, N., Truy, E., & Badin, P. (2020). Rehabilitation of speech disorders following glossectomy, based on ultrasound visual illustration and feedback. *Clinical Linguistics and Phonetics*, 34(9), 826–843. <https://doi.org/10.1080/02699206.2019.1700310>
- Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology*, 32, 401-424. [https://doi.org/10.1016/0022-0965\(81\)90105-3](https://doi.org/10.1016/0022-0965(81)90105-3)
- Gola-Asmussen, C., Lequette, C., Pouget, G., Rouyet, C., & Zorman, M. (2010). *ECLAT 16+*. CeFoCOP/Université de Provence Aix-Marseille I - Cognisciences LSE Université Pierre Mendès France Grenoble  
[http://www.cognisciences.com/article.php3?id\\_article=86](http://www.cognisciences.com/article.php3?id_article=86)
- Golumbica, E. M. Z., Poeppel, D., & Schroeder, C. E. (2012). Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. *Brain and Language*, 122(3), 151-161. <https://doi.org/10.1016/j.bandl.2011.12.010>. Temporal
- Gordon, R. L., Jacobs, M. S., Schuele, C. M., & McAuley, J. D. (2015). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Annals of New-York Academy of Sciences*, 1337, 16-25.  
<https://doi.org/10.1111/nyas.12683>
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3-10. <https://doi.org/10.1016/j.tics.2010.10.001>
- Goswami, U. (2012). Entraining the brain: Applications to language research and links to musical entrainment. *Empirical Musicology Review*, 7(1-2), 57-63.  
<https://doi.org/10.18061/1811/52980>
- Goswami, U., Huss, M., Mead, N., Fosker, T., & Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: Significant longitudinal relations with word reading and reading comprehension. *Cortex*, 49(5), 1363-1376. <https://doi.org/10.1016/j.cortex.2012.05.005>
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, 4(1). <https://doi.org/10.1515/lp-2013-0004>
- Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(16), 10911-10916. <https://doi.org/10.1073/pnas.122368599>
- Gow, D. W., & Gordon, P. C. (1993). Coming to terms with stress: Effects of stress location in sentence processing. *Journal of Psycholinguistic Research*, 22(6), 545-578.  
<https://doi.org/10.1007/BF01072936>
- Grahn, J. A. (2009). The role of the basal ganglia in beat perception: Neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences*, 1169, 35-45. <https://doi.org/10.1111/j.1749-6632.2009.04553.x>

- Grenon, I., Kubota, M., & Shappard, C. (2019). The creation of a new vowel category by adult learners after adaptative phonetic training. *Journal of Phonetics*, 72, 17-34. <https://doi.org/10.1016/j.wocn.2018.10.005>
- Grimaldi, M., Sisinni, B., Gili Fivela, B., Invitto, S., Resta, D., Alku, P., & Brattico, E. (2014). Assimilation of L2 vowels to L1 phonemes governs L2 learning in adulthood: A behavioral and ERP study. *Frontiers in Human Neuroscience*, 8, 1-14. <https://doi.org/10.3389/fnhum.2014.00279>
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLoS Biology*, 11(12). <https://doi.org/10.1371/journal.pbio.1001752>
- Guiraud, H., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Caillot-Bascoul, A., Gonzalez-Monge, S., & Boulenger, V. (2018). Speak slower! How children with language disorder deal with speech rate. *PLoS One*, 13(1). <https://doi.org/10.1371/journal.pone.0191808>.
- Guyon, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, 107, 2711-2724. <https://doi.org/10.1121/1.428657>
- Guyot-Talbot, A., Heidlmayr, K., Ferragne, E., & Mann, T. (2016, Juillet). *Entraînements à la prosodie des questions ouvertes et fermées de l'anglais chez des apprenants francophones*. Journée d'Etude de la Parole. Paris, France.
- Habib, M. (2021). The neurological basis of developmental dyslexia and related disorders: A reappraisal of the temporal hypothesis, twenty years on. *Brain Sciences*, 11(6). <https://doi.org/10.3390/brainsci11060708>
- Habib, M., & Robichon, F. (1996). Les mécanismes cérébraux de la lecture : Un modèle en neurologie cognitive. *Médecine/Sciences*, 12(6-7), 707. <https://doi.org/10.4267/10608/811>
- Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2016). Music and dyslexia: A new musical training method to improve reading and related disorders. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00026>
- Habib, M., Rey, V., Daffaure, V., Camps, R., Espesser, R., Joly-Pottuz, B., & Démonet, J. F. (2002). Phonological training in children with dyslexia using temporally modified speech: A three-step pilot investigation. *International Journal of Language and Communication Disorders*, 37(3), 289-308. <https://doi.org/10.1080/13682820210136278>
- Hacking, J. F., Smith, B. L., & Johnson, E. M. (2017). Utilizing electropalatography to train palatalized versus unpalatalized consonant productions by native speakers of American English learning Russian. *Journal of Second Language Pronunciation*, 3(1), 9-33. <https://doi.org/10.1075/jslp.3.1.01hac>
- Halle, M., Hughes, G. W., & Radley, J. P. (1957). Acoustic properties of stop consonants. *Journal of the Acoustical Society of America*, 29(1), 107-116.

- Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain*, *116*(1), 243-266. <https://doi.org/10.1093/brain/116.1.243>
- Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., & Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: An MEG study. *NeuroImage*, *59*(3), 2952-2961. <https://doi.org/10.1016/j.neuroimage.2011.09.075>
- Hanley, J. R. (1997). Reading and spelling impairments in undergraduate students with developmental dyslexia. *Journal of Research in reading*, *20*(1), 22-30. <https://doi.org/10.1111/1467-9817.00017>
- Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., & Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage*, *185*, 96-101. <https://doi.org/10.1016/j.neuroimage.2018.10.037>
- Hardison, D. (2003). Acquisition of second-language speech: Effects of visual cues, context and talker variability. *Applied Psycholinguistics*, *24*, 495-522. <https://doi.org/10.1017/S0142716403000250>
- Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal*, *22*(2), 175-190. <https://doi.org/10.1558/cj.v22i2.175-190>
- Hary, J. M., & Massaro, D. W. (1982). Categorical results do not imply categorical perception. *Perception & Psychophysics*, *32*(5), 409-418. <https://doi.org/10.3758/BF03202770>
- Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology*, *4*, 1-16. <https://doi.org/10.3389/fpsyg.2013.00566>
- Hazan, V., & Barrett, S. (2000). The development of perceptual cue-weighting in children aged 6 to 12. *Journal of Phonetics*, *28*(4), 377-396. <https://doi.org/10.1006/jpho.2000.0121>
- Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360-378. <https://doi.org/10.1016/j.specom.2005.04.007>
- Heidlmayr, K., Ferragne, E., & Isel, F. (2021). Neuroplasticity in the phonological system: The PMN and the N400 as markers for the perception of non-native phonemic contrasts by late second language learners. *Neuropsychologia*, *156*, 107831. <https://doi.org/10.1016/j.neuropsychologia.2021.107831>
- Heiervang, E., Hugdahl, K., Steinmetz, H., Smievoll, A. I., Stevenson, J., Lund, A., Ersland, L., & Lundervold, A. (2000). Planum temporale, planum parietale and dichotic listening in dyslexia. *Neuropsychologia*, *38*(13), 1704-1713. [https://doi.org/10.1016/S0028-3932\(00\)00085-3](https://doi.org/10.1016/S0028-3932(00)00085-3)



- Heim, S., Eulitz, C., & Elbert, T. (2003). Altered hemispheric asymmetry of auditory P100m in dyslexia. *European Journal of Neuroscience*, *17*(8), 1715-1722. <https://doi.org/10.1046/j.1460-9568.2003.02596.x>
- Helland, T., Asbjornsen, A. E., Hushovd, A. E., & Hugdahl, K. (2008). Dichotic listening and school performance in dyslexia. *Dyslexia*, *14*, 42-53. <https://doi.org/10.1002/dys>
- Hepper, P.G. (1991) An examination of fetal learning before and after birth. *The Irish Journal of Psychology*, *12*(2), 95-107, <https://doi.org/10.1080/03033910.1991.10557830>
- Herbst, S. K., & Obleser, J. (2017). Implicit variations of temporal predictability: Shaping the neural oscillatory and behavioural response. *Neuropsychologia*, *101*, 141-152. <https://doi.org/10.1016/j.neuropsychologia.2017.05.019>
- Hesling, I., Dilharreguy, B., Peppé, S., Amirault, M., Bouvard, M., & Allard, M. (2010). The integration of prosodic speech in high functioning Autism: A preliminary fMRI study. *PLoS ONE*, *5*(7). <https://doi.org/10.1371/journal.pone.0011571>
- Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception: Acoustic rhythmic entrainment induces subsequent perceptual oscillation. *Psychological Science*, *26*(7), 1006-1013. <https://doi.org/10.1177/0956797615576533>
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*(4), 131-138. [https://doi.org/10.1016/S1364-6613\(00\)01463-7](https://doi.org/10.1016/S1364-6613(00)01463-7)
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech understanding. *Nature*, *8*, 393-402. <https://doi.org/10.1038/nrn2113>
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177-180. <https://doi.org/10.1126/science.182.4108.177>
- Hirst, D. J. (2006). Prosodic aspects of speech and language. In K. Brown (Ed), *Encyclopedia of Language and Linguistics*, 2nd Ed. (pp. 539-546), Oxford: Elsevier.
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P. & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, *104*, 353-366. <https://doi.org/10.1016/j.jecp.2009.07.005>
- Hornickel, J., & Kraus, N. (2013). Unstable representation of sound: A biological marker of dyslexia. *Journal of Neuroscience*, *33*(8), 3500-3504. <https://doi.org/10.1523/JNEUROSCI.4205-12.2013>
- Horváth, J., Winkler, I., & Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, *79*(2), 139-147. <https://doi.org/10.1016/j.biopsycho.2008.04.001>

- Houtgast, T., & Steeneken, H. J. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *Journal of the Acoustical Society of America*, 77(3), 1069-1077. <https://doi.org/10.1121/1.392224>
- Hu, W., Mi, L., Yang, Z., Tao, S., Li, M., Wang, W., Dong, Q., & Liu, C. (2016). Shifting perceptual weights in L2 vowel identification after training. *PLoS ONE*, 11(9), 1-14. <https://doi.org/10.1371/journal.pone.0162876>
- Huang, B. H., & Jun, S. A. (2011). The effect of age on the acquisition of second language prosody. *Language and Speech*, 54(3), 387-414. <https://doi.org/10.1177/0023830911402599>
- Hueber, T. (2013). Ultraspeech-player: Intuitive visualization of ultrasound articulatory data for speech therapy and pronunciation training. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, Août 2013 (pp. 75–753). Lyon, France.
- Hutsler, J., & Galuske, R. A. W. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends in Neurosciences*, 26(8), 429-435. [https://doi.org/10.1016/S0166-2236\(03\)00198-X](https://doi.org/10.1016/S0166-2236(03)00198-X)
- Hwang, H., & Lee, H. (2015). The effect of high variability phonetic training on the production of English vowels and consonants. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.
- Iliadou, V., Kaprinis, S., Kandylis, D., & Kaprinis, G. S. (2010). Hemispheric laterality assessment with dichotic digits testing in dyslexia and auditory processing disorder. *International Journal of Audiology*, 49(3), 247-252. <https://doi.org/10.3109/14992020903397820>
- Inceoglu, S. (2015). Effects of perceptual training on second language vowel perception and production. *Applied Psycholinguistics*, 37(5), 1175-1199. <https://doi.org/10.1017/S0142716415000533>
- Ioup, G. (2008). Exploring the role of age in the acquisition of a second language phonology. In J. G. Hansen Edwards and M. L., Zampini (Eds). *Phonology and second language acquisition*. Benjamins.
- Iverson, P., & Evans, B.G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *The Journal of the Acoustical Society of America*, 126(2), 866-877. <https://doi.org/10.1121/1.3148196>
- Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience*, 1(2), 136-152. <https://doi.org/10.1162/jocn.1989.1.2.136>
- Jackson, C. N., & O'Brien, M. G. (2011). The interaction between prosody and meaning in second language speech production1. *Die Unterrichtspraxis/Teaching German*, 44(1), 1-11. <https://doi.org/10.1111/j.1756-1221.2011.00087.x>

- Jacquemin, L., Mertens, G., Schlee, W., Van de Heyning, P., & Gilles, A. (2019). Literature overview on P3 measurement as an objective measure of auditory performance in post-lingually deaf adults with cochlear implant. *International Journal of Audiology*, *58*(12), 816-823. <http://doi.org/10.1080/14992027.2019.1654622>
- Jakoby, H., Goldstein, A., & Faust, M. (2011). Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology*, *48*(11), 1517-1531. <https://doi.org/10.1111/j.1469-8986.2011.01227.x>
- Jamieson, D. G., & Morosan, D. E. (1986). Training nonnative speech contrasts in adults: Acquisition of the English /delta-/ /theta/ contrast by francophones. *Perception and Psychophysics*, *40*, 205-215. <https://doi.org/10.3758/bf03211500>
- Jamison, H. L., Watkins, K. E., Bishop, D. V. M., & Matthews, P. M. (2006). Hemispheric specialization for processing auditory nonspeech stimuli. *Cerebral Cortex*, *16*(9), 1266-1275. <https://doi.org/10.1093/cercor/bhj068>
- Jiang, J., Alwan, A., Keating, P. A., Auer, Jr., E. T., & Bernstein, L. E. (2002). On the correlation between facial movements, tongue movements and speech acoustics. *Journal on Applied Signal Processing*, *11*, 1174-1188. <https://doi.org/10.1155/S1110865702206046>
- Joanisse, M. F., Zevin, J. D., & McCandliss, B. D. (2007). Brain mechanisms implicated in the preattentive categorization of speech sounds revealed using fMRI and a short-interval habituation trial paradigm. *Cerebral Cortex*, *17*(9), 2084-2093. <https://doi.org/10.1093/cercor/bhl124>
- Johnson, B. W., McArthur, G., Hautus, M., Reid, M., Brock, J., Castles, A., & Crain, S. (2013). Lateralized auditory brain function in children with normal reading ability and in children with dyslexia. *Neuropsychologia*, *51*(4), 633-641. <https://doi.org/10.1016/j.neuropsychologia.2012.12.015>
- Johnsrude, I. S., Zatorre, R. J., Milner, B. A., & Evans, A. C. (1997). Left-hemisphere specialization for the processing of acoustic transients. *Neuroreport*, *8*(7), 1761-1765. <https://doi.org/10.1097/00001756-199705060-00038>
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*(5), 323-355. <https://doi.org/10.1037/0033-295X.83.5.323>
- Jones, M. R. (2008). Musical time. In I. C. S. Hallam and M. Thaut (Eds). *The Oxford Handbook of Music Psychology*. Oxford Psychology Press, Oxford, 81-92.
- Jones, M. R. (2016). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds), *The Oxford Handbook of Music Psychology* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198722946.013.13>
- Jones, M. R. (2019). *Time will tell: A theory of dynamic attending*. Oxford University Press. <https://doi.org/10.1093/oso/9780190618216.001.0001>

- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459-491. <https://doi.org/10.1037/0033-295X.96.3.459>
- Jones, A., Hsu, Y. F., Granjon, L., Waszak, F. (2017). Temporal expectancies driven by self- and externally generated rhythms. *Neuroimage*, 156, 352-362. <https://doi.org/10.1016/j.neuroimage.2017.05.042>
- Jones, M., Moynihan, H., Mackenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, 13(4), 313-319. <https://doi.org/10.1111/1467-9280.00458>
- Jordan, G. (2004). *Theory construction in second language acquisition* (Vol. 8). John Benjamins Publishing.
- Jurek, L., Longuet, Y., Baltazar, M., Amestoy, A., Schmitt, V., Desmurget, M., Geoffroy, M.M. (2019). How did I get so late so soon? A review of time processing and management in autism. *Behavioral Brain Research*, 374. <https://doi.org/10.1016/j.bbr.2019.112121>
- Jusczyk, P. W., Kennedy, L. J., Rosner, B. S., & Reed, M. A. (1989). Could temporal order differences underlie 2-month-olds' discrimination of English voicing contrasts? *Journal of the Acoustical Society of America*, 85(4), 1741-1749. <https://doi.org/10.1121/1.397963>
- Kail, M., & Fayol, M. (2000). *L'Acquisition du langage*, vol. 1. Le Langage en émergence, 2.
- Kamiyama, T. (2004). Perception of foreign accentedness in L2 prosody and segments: L1 Japanese speakers learning L2 French. *Speech Prosody*, 721-724.
- Kangatharan, J., Giannakopoulou, A., & Uther, M. (2021). The effect of HVP training in vowel perception on bilingual speech production, *Research in English Language Teaching*, 2(1), 15-32. <https://doi.org/10.18488/journal.179.2021.21.15.32>
- Kartushina, N., Hervais-adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America*, 138, 817-832. <https://doi.org/10.1121/1.4926561>
- Kartushina, N., & Martin, C.D. (2019). Talker and acoustic variability in learning to produce nonnative sounds: Evidence from articulatory training. *Language Learning*, 69, 71-105. <https://doi.org/10.1111/lang.12315>
- Keitel, A., Ince, R. A. A., Gross, J., & Kayser, C. (2017). Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks. *NeuroImage*, 147, 32-42. <https://doi.org/10.1016/j.neuroimage.2016.11.062>
- Kern, P., Assaneo, M. F., Endres, D., Poeppel, D., & Rimmele, J. M. (2021). Preferred auditory temporal processing regimes and auditory-motor synchronization. *Psychonomic Bulletin and Review*, 1860-1873. <https://doi.org/10.3758/s13423-021-01933-w>

- Kershner, J. R. (2019). Neuroscience and education: Cerebral lateralization of networks and oscillations in dyslexia. *Laterality: Asymmetries in Body, Brain and Cognition*, 25(1), 109-125. <https://doi.org/10.1080/1357650X.2019.1606820>
- Kimura, D. (1967). Functional asymmetry of the brain in dichotic listening. *Cortex*, 3(2), 163-178. [https://doi.org/10.1016/s0010-9452\(67\)80010-8](https://doi.org/10.1016/s0010-9452(67)80010-8)
- Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *The Journal of the Acoustical Society of America*, 124(6), 3959-3971. <https://doi.org/10.1121/1.2999341>
- Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, 38(4), 569-587. <https://doi.org/10.1016/j.wocn.2010.08.003>.
- Kondaurova, B., Van Orden, G., Smith, S., Green, P., & Haith, M. (1990). Phonological processing skills in adult dyslexics. *Child Development*, 61, 1753-1778. <https://doi.org/10.2307/1130836>
- Korpilahti, P., Lang, H., & Aaltonen, O., 1995. Is there a late-latency mismatch negativity (MMN) component? *Electroencephalography and Clinical Neurophysiology*, 4(95). [https://doi.org/10.1016/0013-4694\(95\)90016-G](https://doi.org/10.1016/0013-4694(95)90016-G)
- Kösem, A., & Van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience*, 32(5), 536-544. <https://doi.org/10.1080/23273798.2016.1238495>
- Kotz, S. A., & Gunter, T. C. (2015). Can rhythmic auditory cueing remediate language related deficits in Parkinson's disease? *Annals of the New York Academy of Sciences*, 1337, 62-68. <https://doi.org/10.1111/nyas.12657>
- Kotz, S. A., Gunter, T. C., & Wonneberger, S. (2005). The basal ganglia are receptive to rhythmic compensation during auditory syntactic processing: ERP patient data. *Brain and Language*, 95, 70-71. <https://doi.org/10.1016/j.bandl.2005.07.039>
- Kotz, S. A., & Schmidt-Kassow, M. (2015). Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex*, 68, 48-60. <https://doi.org/10.1016/j.cortex.2015.02.021>
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599-605. <https://doi.org/10.1038/nrn2882>
- Krzonowski, J. (2020). *Perception et production de voyelles de l'anglais par des apprenants francophones : effet d'entraînements en perception et en production*. Thèse de Doctorat en psychologie. Université Lyon 2.



- Krzonowski, J., Ferragne, E., & Pellegrino, F. (2016). Perception et production de voyelles de l'anglais par des apprenants francophones : effet d'entraînements en perception et en production. *Journée d'étude de La Parole*, 491-499.
- Krzonowski, J., Pellegrino, F., Ferragne, E. (2018). Étude acoustique de la production de voyelles de l'anglais par des apprenants francophones. *Journée d'étude de La Parole*, 523-531. <https://doi.org/10.21437/jep.2018-60>
- Kujala, T., Myllyviita, K., Tervaniemi, M., Alho, K., Kallio, J., & Näätänen, R. (2000). Basic auditory dysfunction in dyslexia as demonstrated by brain activity measurements. *Psychophysiology*, 37(2), 262-266. <https://doi.org/10.1017/S0048577200990140>
- Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., & Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(5). <https://doi.org/10.1002/wcs.1528>
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, 94(3), 1904-1911. <https://doi.org/10.1152/jn.00263.2005>
- Lambacher, S. G., Martens W., L., Kakehi, K., Marasinghe, C. A., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, 26(2), 227-247. <https://doi.org/10.1017/S0142716405050150>
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119-159.
- Lazzaro, I., Gordon, E., Whitmont, S., Meares, R., & Clarke, S. (2001). The modulation of late component event related potentials by pre-stimulus EEG theta activity in ADHD. *International Journal of Neuroscience*, 107(3-4), 247-264. <https://doi.org/10.3109/00207450109150688>.
- Lehongre, K., Morillon, B., Giraud, A. L., & Ramus, F. (2013). Impaired auditory sampling in dyslexia: Further evidence from combined fMRI and EEG. *Frontiers in Human Neuroscience*, 7, 1-8. <https://doi.org/10.3389/fnhum.2013.00454>
- Lengeris, A. (2018). Computer-based auditory training improves second-language vowel production in spontaneous speech. *The Journal of the Acoustical Society of America*, 144(3), 165-171. <https://doi.org/10.1121/1.5052201>
- Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *The Journal of the Acoustical Society of America*, 128(6), 3757-3768. <https://doi.org/10.1121/1.3506351>
- Leonard, C. M., Voeller, K. K. S., Morris, M. K., Garofalakis, M., Lombardino, L. J., Honeyman, J. C., Mao, J., Agee, O. F., Staab, E. V., Alexander, A. W., Andersen, H. G., & Hynd, G. W. (1993). Anomalous cerebral structure in dyslexia revealed with magnetic

- resonance imaging. *Archives of Neurology*, 50(5), 461-469.  
<https://doi.org/10.1001/archneur.1993.00540050013008>
- Leong, V., & Goswami, U. (2014a). Assessment of rhythmic entrainment at multiple timescales in dyslexia: Evidence for disruption to syllable timing. *Hearing Research*, 308, 141-161. <https://doi.org/10.1016/j.heares.2013.07.015>
- Leong, V., & Goswami, U. (2014b). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience*, 8, 1-14. <https://doi.org/10.3389/fnhum.2014.00096>
- Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *The Journal of the Acoustical Society of America*, 136(1), 366-381. <https://doi.org/10.1121/1.4883366>
- Leppänen, P. H. T., Eklund, K. M., & Lyytinen, H. (1997). Event-related brain potentials to change in rapidly presented acoustic stimuli in newborns. *Developmental Neuropsychology*, 13(2), 175-204. <https://doi.org/10.1080/87565649709540677>
- Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure. *Music Perception: An Interdisciplinary Journal*, 1(2), 229-252.
- Levy, E. S. (2009). On the assimilation-discrimination relationship in American English adults' French vowel learning. *The Journal of the Acoustical Society of America*, 126(5), 2670-2682. <https://doi.org/10.1121/1.3224715>
- Lewis, J. P., & Frick, R. W. (1999). Row blindness in Gestalt grouping and developmental dyslexia. *Neuropsychologia*, 37(3), 385-393. [https://doi.org/10.1016/s0028-3932\(98\)00028-1](https://doi.org/10.1016/s0028-3932(98)00028-1)
- Li, A., & Post, B. (2014). L2 acquisition of prosodic properties of speech rhythm. *Studies in Second Language Acquisition*, 36(2), 223-255.  
<https://doi.org/10.1017/S0272263113000752>
- Li, K., Zhang, S., Li, M., Lo, W. K., & Meng, H. (2010). Detection of intonation in L2 English speech of native Mandarin learners. *7th International Symposium on Chinese Spoken Language Processing, ISCSLP*. Nov. 2010, (pp. 69-74). Tainan, Taiwan.  
<https://doi.org/10.1109/ISCSLP.2010.5684846>
- Liberman, A., Cooper, F., Shankweiler, D., Studdert-Kennedy, M., (1967). Perception of the speech code. *Psychological Review*, 74, 431-461. <https://doi.org/10.1037/h0020279>
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358-368. <https://doi.org/10.1037/h0044417>
- Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36. [https://doi.org/10.1016/0010-0277\(85\)90021-6](https://doi.org/10.1016/0010-0277(85)90021-6)

- Liberman, A.M., & Mattingly, I.G. (1989). A specialization for speech perception. *Science*, 243(4890), 489-494. <https://doi.org/10.1126/science.2643163>
- Liberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187-196. [https://doi.org/10.1016/S1364-6613\(00\)01471-6](https://doi.org/10.1016/S1364-6613(00)01471-6)
- Liebenthal, E., Binder, J. R., Spitzer, S. M., Possing, E. T., & Medler, D. A. (2005). Neural substrates of phonemic perception. *Cerebral Cortex*, 15(10), 1621-1631. <https://doi.org/10.1093/cercor/bhi040>
- Liégeois-Chauvel, C., De Graaf, J. B., Laguitton, V., & Chauvel, P. (1999). Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cerebral Cortex*, 9(5), 484-496. <https://doi.org/10.1093/cercor/9.5.484>
- Liégeois-Chauvel, C., Giraud, K., Badier, J. M., Marquis, P., & Chauvel, P. (2001). Intracerebral evoked potentials in pitch perception reveal a functional asymmetry of the human auditory cortex. *Annals of the New York Academy of Sciences*, 930, 117-132. <https://doi.org/10.1111/j.1749-6632.2001.tb05728.x>
- Liégeois-Chauvel, C., Lorenzi, C., Trébuchon, A., Régis, J., & Chauvel, P. (2004). Temporal envelope processing in the human left and right auditory cortices. *Cerebral Cortex*, 14(7), 731-740. <https://doi.org/10.1093/cercor/bhh033>
- Lijffijt, M., Lane, S. D., Meier, S. L., Boutros, N. N., Burroughs, S., Steinberg, J. L., Gerard Moeller, F., & Swann, A. C. (2009a). P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory. *Psychophysiology*, 46(5), 1059-1068. <https://doi.org/10.1111/j.1469-8986.2009.00845.x>
- Lijffijt, M., Moeller, F.G, Boutros, N. N., Steinberg, J. L., Meier, S. L., Lane, S. D., & Swann, A. C. (2009b). Diminished P50, N100 and P200 auditory sensory gating in bipolar I disorder. *Psychiatry Research*, 167(3), 191-201. <https://doi.org/10.1016/j.psychres.2008.04.001>
- Lisker, L. (1978). In qualified defense of VOT. *Language and Speech*, 21(4), 375-383. <https://doi.org/10.1177/002383097802100413>
- Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *WORD*, 20(3), 384-422. <https://doi.org/10.1080/00437956.1964.11659830>
- Lisker, L., & Abramson, A. S. (1967). Some effects of contexts of context on Voice Onset Time in English stops. *Language and Speech*, 10, 1-28. <https://doi.org/10.1177/002383096701000101>
- Lisker, L., Liberman, A. M., Erickson, D. M., Dechovitz, D., & Mandler, R. (1977). On pushing the voice-onset-time (VOT) boundary about. *Language and Speech*, 20, 209-216. <https://doi.org/10.1177/002383097702000303>
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/ II: The role of phonetic environment and talker variability in learning



- new perceptual categories. *Journal of the Acoustical Society of America*, 94(3), 1242-1255. <https://doi.org/10.1121/1.408177>
- Lizarazu, M., Lallier, M., Bourguignon, M., Carreiras, M., & Molinaro, N. (2021). Impaired neural response to speech edges in dyslexia. *Cortex*, 135, 207-218. <https://doi.org/10.1016/j.cortex.2020.09.033>
- Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., & Carreiras, M. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. *Human Brain Mapping*, 36(12), 4986-5002. <https://doi.org/10.1002/hbm.22986>
- Llinas, R. (1993). Is dyslexia a dyschronia? *Annals New York Academy of Sciences*, 682(1), 48-56. <https://doi.org/10.1111/j.1749-6632.1993.tb22958.x>
- London, J. (2012). Three things linguists need to know about rhythm and time in music. *Empirical Musicology Review*, 7(1-2), 5-11. <https://doi.org/10.18061/1811/52973>
- Lum, J. A., Ullman, M. T., & Conti-Ramsden, G. (2013). Procedural learning is impaired in dyslexia: Evidence from a meta-analysis of serial reaction time studies. *Research in Developmental Disabilities*, 34(10), 3460-3476. <https://doi.org/10.1016/j.ridd.2013.07.017>
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001-1010. <https://doi.org/10.1016/j.neuron.2007.06.004>
- Luo, H., & Poeppel, D. (2012). Cortical oscillations in auditory perception and speech: Evidence for two temporal windows in human auditory cortex. *Frontiers in Psychology*, 3, 1-10. <https://doi.org/10.3389/fpsyg.2012.00170>
- Luo, Y., Zhang, Y., Feng, X., & Zhou, X. (2010). Electroencephalogram oscillations differentiate semantic and prosodic processes during sentence reading. *Neuroscience*, 169(2), 654-664. <https://doi.org/10.1016/j.neuroscience.2010.05.032>
- Lyons, M., Schoen Simmons, E., & Paul, R. (2014). Prosodic development in middle childhood and adolescence in high-functioning autism. *Autism Research*, 7(2), 181-196. <https://doi.org/10.1002/aur.1355>
- Maasen, B., Groenen, P., Crul, T., Assman-Hulsmans, C., & Gabreels, F. (2001). Identification and discrimination of voicing and place-of-articulation in developmental dyslexia. *Clinical Linguistics and Phonetics*, 15(4), 319-339. <https://doi.org/10.1080/02699200010026102>
- Macmillan, N., & Creelman, C. (2005). *Detection Theory: A user's guide*. London, UK: Lawrence Erlbaum.
- Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

- Magnan A. Ecalte, J., Veuillet, E., & Collet, L. (2004). The effects of an audio-visual training program in dyslexic children. *Dyslexia*, 10(2), 131-140. <https://doi.org/10.1002/dys.270>
- Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199-211. <https://doi.org/10.1162/089892906775783660>
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, N.J: Routledge.
- Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science*, 271(5245), 77-81. <https://doi.org/10.1126/science.271.5245.77>
- Marshall, C. R., Harcourtbrown, S., Ramus, F., & Van Der Lely, H. K. J. (2009). The link between prosody and language skills in children with specific language impairment SLI and/or dyslexia. *International Journal of Language and Communication Disorders*, 44(4), 466-488. <https://doi.org/10.1080/13682820802591643>
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314-324. [doi:10.3758/s13428-011-0168-7](https://doi.org/10.3758/s13428-011-0168-7)
- Mazur-Palandre, A., Abadie, R., & Bedoin, N. (2016). Etudiants dyslexiques à l'Université : Spécificité des difficultés ressenties et évaluation des déficits. *Développements*, 18-19, 139-177.
- McAnally, K. I., & Stein, J.F. (1997) Scalp potentials evoked by amplitude modulated tones in dyslexia. *Journal of Speech, Language, and Hearing Research*, 40, 939-945. <https://doi.org/10.1044/jslhr.4004.939>.
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization on spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, 2, 89-108. <https://doi.org/10.3758/cabn.2.2.89>
- McCann, J., & Peppé, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language and Communication Disorders*, 38(4), 325-350. <https://doi.org/10.1080/1368282031000154204>
- McClelland, J. L., Fiez, J. A., & McCandliss, B. D. (2002). Teaching the /r-l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology and Behavior*, 77, 657-662. [https://doi.org/10.1016/s0031-9384\(02\)00916-2](https://doi.org/10.1016/s0031-9384(02)00916-2)
- McKibbin, K., Elias, I. J., Saucier, D. M., & Engebregston, D. (2003). Right hemispheric dominance for processing extended non-linguistic frequency transitions. *Brain and Cognition*, 53, 322-326. [https://doi.org/10.1016/s0278-2626\(03\)00135-0](https://doi.org/10.1016/s0278-2626(03)00135-0)

- McMurray, B., & Aslin, R. N. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, *95*, 15-26. <https://doi.org/10.1016/j.cognition.2004.07.005>
- Medina, V., Hoonhorst, I., Bogliotti, C., & Serniclaes, W. (2010). Development of voicing perception in French: Comparing adults, adolescents, and children. *Journal of Phonetics*, *38*(4), 493-503. <https://doi.org/10.1016/j.wocn.2010.06.002>
- Medina, V., & Serniclaes, W. (2005). Late development of the categorical perception of speech sounds in pre-adolescent children. *ZAS Papers in Linguistics*, *42*, 13-31. <https://doi.org/10.21248/zaspil.42.2005.271>
- Mendonça Alves, L., Reis, C., & Pinheiro, Â. (2014). Prosody and reading in dyslexic children. *Dyslexia*, *21*(1), 35-49. <https://doi.org/10.1002/dys.1485>
- Menning, H., Imaizumi, S., Zwitserlood, P., & Pantev, C. (2002). Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrasts of the Japanese language. *Learning and Memory*, *9*(5), 253-267. <https://doi.org/10.1101/lm.49402>
- Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, *343*, 1006-1010. <https://doi.org/10.1126/science.1245994>
- Milne, R. D., Nicholson, T., & Corballis, M. C. (2003). Lexical access and phonological decoding in adult dyslexic subtypes. *Neuropsychology*, *7*(3), 362-368. <https://doi.org/10.1037/0894-4105.17.3.362>
- Mokari, G.P., & Werner, S. (2017). Perceptual training of second-language vowels: Does musical ability play a role? *Journal of Psycholinguistic Research*, *47*(1), 95-112. <https://doi.org/10.1007/s10936-017-9517-8>
- Moll, K., & Jones, M. (2013). Naming fluency in dyslexic and nondyslexic readers: Differential effects of visual crowding in foveal, parafoveal, and peripheral vision. *The Quarterly Journal of Experimental Psychology*, *66*(11), 2085-2091. <https://doi.org/10.1080/17470218.2013.840852>
- Moore, E., Cassim, R., & Talcott, J. B. (2011). Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrimental effects of distractors in visual search tasks. *Neuropsychologia*, *49*, 3881-3890. <https://doi.org/10.1016/j.neuropsychologia.2011.10.005>
- Moore, E., Tsouknida, E., & Romani, C. (2015). Adults with dyslexia can use cues to orient and constrain attention but have a smaller and weaker attention spotlight. *Vision Research*, *111*, 55-65. <https://doi.org/10.1016/j.visres.2015.03.019>
- Moreno, S., Lee, Y., Janus, M., & Bialystok, E. (2015). Short-term second language and music training induces lasting functional brain changes in early childhood. *Child Development*, *86*(2), 394-406. <https://doi.org/10.1111/cdev.12297>

- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712-723. <https://doi.org/10.1093/cercor/bhn120>
- Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(42). <https://doi.org/10.1073/pnas.1705373114>
- Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology*, *31*, 230-238. <https://doi.org/10.1016/j.conb.2014.12.005>
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(43), 18688-18693. <https://doi.org/10.1073/pnas.1007189107>
- Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Bénar, C. G., & Giraud, A. L. (2012). Asymmetric function of theta and gamma activity in syllable processing: An intracortical study. *Frontiers in Psychology*, *3*, 1-9. <https://doi.org/10.3389/fpsyg.2012.00248>
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, *1337*(1), 26-31. <https://doi.org/10.1111/nyas.12629>
- Morrison, S. (1998). Computer applications: Earobics Pro. *Child Language Teaching and Therapy*, *14*, 279-284. <https://doi.org/10.1191/026565998669702253>
- Moulay, E. (2016). *Evaluation et remédiation de déficit phonologique chez des étudiants dyslexiques : Perception catégorielle et expérience en ERP*. Mémoire de Master 2 de Neuropsychologie, Université de Lille.
- Muneaux, M., Ziegler, J. C., Truc, C., Thomson, J., & Goswami, U. (2004). Deficits in beat perception and dyslexia: Evidence from French. *NeuroReport*, *15*(8), 1255-1259. <https://doi.org/10.1097/01.wnr.0000127459.31232.c4>
- Näätänen, R. (1992). *Attention and brain function*. Psychology Press.
- Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*, 313-329. [https://doi.org/10.1016/0001-6918\(78\)90006-9](https://doi.org/10.1016/0001-6918(78)90006-9)
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432-434. <https://doi.org/10.1038/385432a0>
- Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human

- brain? *Neuroscience Letters*, 98(2), 217-221. [https://doi.org/10.1016/0304-3940\(89\)90513-2](https://doi.org/10.1016/0304-3940(89)90513-2)
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing; A review. *Clinical Neurophysiology*, 118(12), 2544-2590. <https://doi.org/10.1016/j.clinph.2007.04.026>
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4), 375-425. <https://doi.org/10.1111/j.1469-8986.1987.tb00311.x>
- Narain, C., Scott, S.K., Wise, R.J.S., Rosen, S., Leff, A., Iversen, S.D., & Matthews, P.M. (2003). Defining a left-lateralized response specific to intelligible speech using fMRI. *Cerebral Cortex*, 13(12), 1362-1368. <https://doi.org/10.1093/cercor/bhg083>
- Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., & Mooney, R. (2013). A circuit for motor cortical modulation of auditory cortical activity. *Journal of Neuroscience*, 33(36), 14342-14353. <https://doi.org/10.1523/JNEUROSCI.2275-13.2013>
- Nenonen, S., Shestakova, A., Huotilainen, M., & Näätänen, R. (2005). Speech-sound duration processing in a second language is specific to phonetic categories. *Brain and Language*, 92(1), 26-32. <https://doi.org/10.1016/j.bandl.2004.05.005>
- Nguyen, N. (2005). La perception de la parole. In N. Nguyen, S. Wauquier and J. Durand (Eds). *Phonologie et phonétique* (pp. 425-447). Hermès
- Nickels, S., Opitz, B., & Steinhauer, K. (2013). ERPs show that classroom-instructed late second language learners rely on the same prosodic cues in syntactic parsing as native speakers. *Neuroscience Letters*, 557, 107-111. <https://doi.org/10.1016/j.neulet.2013.10.019>
- Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex*, 47(1), 117-127. <https://doi.org/10.1016/j.cortex.2009.08.016>
- Nicolson, R. I., Fawcett, A. J., Berry, E. L., Jenkins, I. H., Dean, P., & Brooks, D. J. (1999). Association of abnormal cerebellar activation with motor learning difficulties in dyslexic adults. *The Lancet*, 353(9165), 1662-1667. [https://doi.org/10.1016/S0140-6736\(98\)09165-X](https://doi.org/10.1016/S0140-6736(98)09165-X)
- Noordenbos, M. W., Segers, E., Serniclaes, W., & Verhoeven, L. (2013). Neural evidence of the allophonic mode of speech perception in adults with dyslexia. *Clinical Neuropsychology*, 124, 1151-1162. <https://doi.org/10.1016/j.clinph.2012.12.044>
- Noordenbos, M. W., Segers, E., Serniclaes, W., Mitterer, H., & Verhoeven, L. (2012). Allophonic mode of speech perception in Dutch children at risk for dyslexia: A longitudinal study. *Research in Developmental Disabilities*, 33(5), 1469-1483. <https://doi.org/10.1016/j.ridd.2012.03.021>



- Noordenbos, M., & Serniclaes, W. (2015). Categorical perception in dyslexia: A meta-analysis. *Scientific Studies of Reading, 19*, 340-359. <https://doi.org/10.1080/10888438.2015.1052455>
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience, 31*(28), 10234-10240. <https://doi.org/10.1523/JNEUROSCI.0411-11.2011>
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience, 32*(49), 17572-17581. <https://doi.org/10.1523/JNEUROSCI.3203-12.2012>
- Nozaradan, S., Zerouali, Y., Peretz, I., & Mouraux, A. (2015). Capturing with EEG the neural entrainment and coupling underlying sensorimotor synchronization to the beat. *Cerebral Cortex, 25*(3), 736-747. <https://doi.org/10.1093/cercor/bht261>
- Oades, R. D., Dittmann-Balcar, A., Schepker, R., Eggers, C., & Zerbin, D. (1996). Auditory event-related potentials (ERPs) and mismatch negativity (MMN) in healthy children and those with attention-deficit or Tourette/tic symptoms. *Biological Psychology, 43*, 163-185.
- Obleser, J., Zimmermann, J., Van Meter, J. & Rauschecker, J.P. (2007). Multiple stages of auditory speech perception reflected in event-related fMRI. *Cerebral Cortex, 17*(10), 2251-2257. <https://doi.org/10.1093/cercor/bhl133>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.
- Olson, D.J. (2019). Feature acquisition in second language phonetic development: Evidence from phonetic training. *Language Learning, 69*(2), 366-404. <https://doi.org/10.1111/lang.12336>
- Olson, D. J. (2021). Phonetic feature size in second language acquisition: Examining VOT in voiceless and voiced stops. *Second Language Research, 1*-28. <https://doi.org/10.1177/02676583211008951>
- Ortega-Llebaria, M., Faulkner, A., & Hazan, V. (2001). Auditory-visual L2 speech perception: Effects of visual cues and acoustic-phonetic context for Spanish learners of English. *International Conference on Auditory-Visual Speech Processing*, 149-154.
- Oshio, K. I. (2011). Possible functions of prefrontal cortical neurons in duration discrimination. *Frontiers in Integrative Neuroscience, 5*, 1-2. <https://doi.org/10.3389/fnint.2011.00025>
- Overath, T., Kumar, S., Von Kriegstein, K., & Griffiths, T. D. (2008). Encoding of spectral correlation over time in auditory cortex. *Journal of Neuroscience, 28*(49), 13268-13273. <https://doi.org/10.1523/JNEUROSCI.4596-08.2008>

- Overy, K. (2008). Classroom rhythm games for literacy support. In T. R. Miles, J. Westcombe, & D. Ditchfield (Eds), *Music and Dyslexia: a positive approach* (pp. 26-44). John Wiley. <https://doi.org/10.1002/9780470988183>
- Overy, K., Norton, A. C., Cronin, K. T., Gaab, N., Alsop, D. C., Winner, E., & Schlaug, G. (2004). Imaging melody and rhythm processing in young children. *NeuroReport*, *15*(11), 1723-1726. <https://doi.org/10.1097/01.wnr.0000136055.77095.f1>
- Pagliarini, E., Scocchia, L., Granocchio, E., Sarti, D., Stucchi, N., Guasti, M.T. (2020). Timing anticipation in adults and children with Developmental Dyslexia: evidence of an inefficient mechanism. *Scientific Report*, *10*(1). <https://doi.org/10.1038/s41598-020-73435-z>
- Patel, A. D. (2008). Science & music: Talk of the tone. *Nature*, *453*(7196), 726-727. <https://doi.org/10.1038/453726a>
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*, 142. <https://doi.org/10.3389/fpsyg.2011.00142>
- Patel, A. D. (2012). The OPERA hypothesis: Assumptions and clarifications. *Annals of the New York Academy of Sciences*, *1252*(1), 124-128. <https://doi.org/10.1111/j.1749-1566.2011.06426.x>
- Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience*, *8*. <https://doi.org/10.3389/fnsys.2014.00057>
- Paulesu, E., Démonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S. F., Cossu, G., Habib, M., Frith, C. D., Frith, U. (2001). Dyslexia: Cultural diversity and biological unity. *Science*, *291*(5511), 2165-2167. <https://doi.org/10.1126/science.1057179>
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 1-17. <https://doi.org/10.3389/fpsyg.2012.00320>
- Penhune, V. B., Zatorre, R. J., MacDonald, J. D., & Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex*, *6*(5), 661-672. <https://doi.org/10.1093/cercor/6.5.661>
- Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child development*, *61*, 1753-1778.
- Pernet, C., Poline, J., Demonet, J., & Rousselet, G. (2009). Brain classification reveals the right cerebellum as the best biomarker of dyslexia. *BMC neuroscience*, *10*(1), 67. <https://doi.org/10.1186/1471-2202-10-67>



- Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics*, 22(3), 171-198. <https://doi.org/10.1080/02699200701799825>
- Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 564-573. <https://doi.org/10.1037/0096-1523.16.3.564>
- Plaisant, O. (1989). Who I am? A dyslexic. *Annales Médico-Psychologiques*, 147, 205-209.
- Plante, E., Bahl, M., Vance, R., & Gerken, L. A. (2010). Children with specific language impairment show rapid, implicit learning of stress assignment rules. *Journal of Communication Disorders*, 43(5), 397-406. <https://doi.org/10.1016/j.jcomdis.2010.04.012>
- Poeppel, D. (2001). Pure word deafness and the bilateral processing of the speech code. *Cognitive Science*, 25(5), 679-693. [https://doi.org/10.1016/S0364-0213\(01\)00050-7](https://doi.org/10.1016/S0364-0213(01)00050-7)
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as "asymmetric sampling in time". *Speech Communication*, 41, 245-255. [https://doi.org/10.1016/S0167-6393\(02\)00107-3](https://doi.org/10.1016/S0167-6393(02)00107-3)
- Poeppel, D. (2012). The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology*, 29(1-2), 34-55. <https://doi.org/10.1080/02643294.2012.710600>
- Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, 28, 142-149. <https://doi.org/10.1016/j.conb.2014.07.005>
- Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 21(6), 322-334. <https://doi.org/10.1038/s41583-020-0304-4>
- Poeppel, D., Boemio, A., Simon, J., Sauvé, K., Depireux, D., Ribary, U., & Llinas, R. (2000). High-frequency response asymmetry to auditory stimuli of varying spectral complexity. *Society for Neuroscience, New Orleans*.
- Poeppel, D., Guillemin, A., Thompson, J., Fritz, J., Bavelier, D., & Braun, A. R. (2004). Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia*, 42(2), 183-200. <https://doi.org/10.1016/j.neuropsychologia.2003.07.010>
- Poeppel, D., Idsardi, W. J., & Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 1071-1086. <https://doi.org/10.1098/rstb.2007.2160>
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10), 2128-2148. <https://doi.org/10.1016/j.clinph.2007.04.019>

- Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., Kotz, S. A., & Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, *27*(1), 121-131. <https://doi.org/10.1037/a0031277>
- Pulvermüller, F., Huss, M., Kherif, F., Del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(20), 7865-7870. <https://doi.org/10.1073/pnas.0509989103>
- Ramirez Verdugo, D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness*, *15*(3), 141-159. <https://doi.org/10.2167/la404.0>
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841-865. <https://doi.org/10.1093/brain/awg076>
- Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781-791. [https://doi.org/10.1016/S0926-6410\(03\)00202-7](https://doi.org/10.1016/S0926-6410(03)00202-7)
- Richards, S., & Goswami, U. (2015). Auditory processing in Specific Language Impairment (SLI): Relations with the perception of lexical and phrasal stress. *Journal of Speech and Hearing Research*, *58*, 1292–1305. <https://doi.org/10.1044/2015>
- Rimol, L. M., Eichele, T., & Hugdahl, K. (2006). The effect of voice-onset-time on dichotic listening with consonant-vowel syllables. *Neuropsychologia*, *44*(2), 191-196. <https://doi.org/10.1016/j.neuropsychologia.2005.05.006>
- Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, *1077*(1), 135-143. <https://doi.org/10.1016/j.brainres.2006.01.043>
- Robb, S. L., Carpenter, J. S., & Burns, D. S. (2011). Reporting guidelines for music-based interventions. *Journal of Health Psychology*, *16*(2), 342-352. <https://doi.org/10.1177/1359105310374781>
- Roberts, C. A., Fairclough, S., Fisk, J. E., Tames, F. T., & Montgomery, C. (2013). Electrophysiological indices of response inhibition in human polydrug users. *Journal of Psychopharmacology*, *27*(9), 779-789. <https://doi.org/10.1177/0269881113492899>
- Robichon, F., & Habib, M. (1998). Abnormal callosal morphology in male adult dyslexics: Relationships to handedness and phonological abilities. *Brain and Language*, *62*(1), 127-146. <https://doi.org/10.1006/brln.1997.1891>
- Robin, D. A., Tranel, D., & Damasio, H. (1990). Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain and Language*, *39*(4), 539–555. [https://doi.org/10.1016/0093-934X\(90\)90161-9](https://doi.org/10.1016/0093-934X(90)90161-9)

- Rojas, D. E. (2016, April 25). *Les français et les langues étrangères*. Retrieved February 17, 2017, from La revue des Ressources, Web site : <http://www.ressources.org>
- Rolka, E. J., & Silverman, M. J. (2015). A systematic review of music and dyslexia. *Arts in Psychotherapy, 46*, 24-32. <https://doi.org/10.1016/j.aip.2015.09.002>
- Romani, C., Di Betta, A., Tsouknida, E., & Olson, A. (2008). Lexical and nonlexical processing in developmental dyslexia: A case for different resources and different impairments. *Cognitive Neuropsychology, 25*, 798-830. <https://doi.org/10.1080/02643290802347183>
- Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 336*(1278), 367-373. <https://doi.org/10.1098/rstb.1992.0070>
- Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., & Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research, 156*(1), 95-103. <https://doi.org/10.1016/j.bbr.2004.05.012>
- Sadakata, M., & McQueen, J. M. (2013). High stimulus variability in nonnative speech learning supports formation of abstract categories: Evidence from Japanese geminates. *Journal of the Acoustical Society of America, 134*(2), 1324-1335. <https://doi.org/10.1121/1.4812767>
- Saloranta, A., Alku, P., & Peltola, M. S. (2020). Listen-and-repeat training improves perception of second language vowel duration: Evidence from mismatch negativity (MMN) and N1 responses and behavioral discrimination. *International Journal of Psychophysiology, 147*, 72-82. <https://doi.org/10.1016/j.ijpsycho.2019.11.005>
- Saloranta, A., Heikkola, M. L., & Peltola, M. S. (2022). Listen-and-repeat training in the learning of non-native consonant duration contrasts: Influence of consonant type as reflected by MMN and behavioral methods. *Journal of Psycholinguistic Research*. <https://doi.org/10.1007/s10936-022-09868-6>
- Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology/ Evoked Potentials, 62*(6), 437-448. [https://doi.org/10.1016/0168-5597\(85\)90054-1](https://doi.org/10.1016/0168-5597(85)90054-1)
- Samson, S., Ehrlé, N., & Baulac, M. (2001). Cerebral substrates for musical temporal process. *Annals of the New York Academy of Sciences, 930*, 166-178. <https://doi.org/10.1111/j.1749-6632.2001.tb05732.x>
- Schmidt, A. M. (2012). Effects of EPG treatment for English consonant contrasts on L2 perception and production. *Clinical Linguistics and Phonetics, 26*(11-12), 909-925. <https://doi.org/10.3109/02699206.2012.718036>
- Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability. *NeuroReport, 20*(1), 31-36.

- Schonwiesner, M., RübSamen, R., & Von Cramon, D. Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *European Journal of Neuroscience*, 22(6), 1521-1528. <https://doi.org/10.1111/j.1460-9568.2005.04315.x>
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 1-16. <https://doi.org/10.1016/j.tins.2008.09.012>.
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, 12(3), 106-113. <https://doi.org/10.1016/j.tics.2008.01.002>
- Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, 20(2), 172-176. <https://doi.org/10.1016/j.conb.2010.02.010>
- Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1998). Auditory processing and dyslexia: Evidence for a specific speech processing deficit. *NeuroReport*, 9(2), 337-340. <https://doi.org/10.1097/00001756-199801260-00029>
- Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (1999). The role of phonological awareness, speech perception, and auditory temporal processing for dyslexia. *European Child and Adolescent Psychiatry*, 8, 28-34. <https://doi.org/10.1007/pl00010690>
- Schulte-Körne, G., Deimel, W., Bartling, J., & Remschmidt, H. (2001). Speech perception deficit in dyslexic adults as measured by mismatch negativity (MMN). *International Journal of Psychophysiology*, 40(1), 77-87. [https://doi.org/10.1016/S0167-8760\(00\)00152-5](https://doi.org/10.1016/S0167-8760(00)00152-5)
- Schwartz, J., & Tallal, P. (1980). Rate of acoustic change may underlie hemispheric specialization for speech perception. *Science*, 207(4437), 1380-1381. <https://doi.org/10.1126/science.7355297>
- Scott *et al.*, (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123, 2400-2406. <https://doi.org/10.1093/brain/123.12.2400>
- Serniclaes, W. (1975). Perceptual processing of acoustic correlates of the voicing feature. Proceedings of the *Speech Communication Seminar*, Stockholm. Almqvist & Wiksell: Upsalla.
- Serniclaes, W. (2018). Allophonic theory of dyslexia: A short overview. *JSM Communication Disorders*, 2, 1010. <https://doi.org/10.1002/brb3.2194>
- Serniclaes, W., López-Zamora, M., Bordoy, S., & L. Luque, J. (2021). Allophonic perception of VOT contrasts in Spanish children with dyslexia. *Brain and Behavior*, 11(6), 1-13. <https://doi.org/10.1002/brb3.2194>

- Serniclaes, W., & Seck, M. (2018). Enhanced sensitivity to subphonemic segments in dyslexia: A new instance of allophonic perception. *Brain Sciences*, 8(4). <https://doi.org/10.3390/brainsci8040054>
- Serniclaes, W., & Sprenger-Charolles, L. (2003). Categorical perception of speech sounds and dyslexia. *Current Psychology Letters*, 1(10). <https://doi.org/10.4000/cpl.379>
- Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Demonet, J. F. (2001). Perceptual discrimination of speech sounds in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, 44(2), 384-399. [https://doi.org/10.1044/1092-4388\(2001/032\)](https://doi.org/10.1044/1092-4388(2001/032))
- Serniclaes, W., Van Heghe, S., Mousty, P., Carré, R., & Sprenger-Charolles, L. (2004). Allophonic mode of speech perception in dyslexia. *Journal of Experimental Child Psychology*, 87(4), 336-361. <https://doi.org/10.1016/j.jecp.2004.02.001>
- Serniclaes, W., Ventura, P., Morais, J., & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition*, 98(2), 35-44. <https://doi.org/10.1016/j.cognition.2005.03.002>
- Sharma, A., & Dorman, M. F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America*, 106, 1078-1083. <https://doi.org/10.1121/1.428048>
- Shinohara, Y., & Iverson, P. (2018). High variability identification and discrimination training for Japanese speakers learning English /r/-/l/. *Journal of Phonetics*, 66, 242-251. <https://doi.org/10.1016/j.wocn.2017.11.002>
- Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R. J., & Näätänen, R. (2000). Discrimination of speech and of complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *NeuroImage*, 12(6), 657-663. <https://doi.org/10.1006/nimg.2000.0646>
- Shum, K.K.M., Au, T.K.F., Romo, L.F., & Jun, S.A. (2021). Learning challenging L2 sounds via computer training: High-variability perceptual training for children and adults. *Language Learning and Development*, 1-16. <https://doi.org/10.1080/15475441.2021.1876699>
- Simon, C., & Fourcin, A. J. (1978). Cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America*, 63(3), 925-935. <https://doi.org/10.1121/1.381772>
- Simos, P. G., Molfese, D. L., & Brenden, R. A. (1997). Behavioral and electrophysiological indices of voicing-cue discrimination: Laterality patterns and development. *Brain and Language*, 57, 122-150. <https://doi.org/10.1006/brln.1997.1836>
- Smith, J. L., Johnstone, S. J., & Barry, R. J. (2004). Inhibitory processing during the Go/NoGo task: An ERP analysis of children with attention-deficit/hyperactivity disorder. *Clinical Neurophysiology*, 115(6), 1320-1331. <https://doi.org/10.1016/j.clinph.2003.12.027>



- Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research*, 43(2), 219-234. <https://doi.org/10.1007/BF00309831>
- Snowling, M. (1987). *Dyslexia: A cognitive developmental perspective*. Basil Blackwell.
- Soltész, F., Szucs, D., Leong, V., White, S., & Goswami, U. (2013). Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. *PLoS ONE*, 8(10). <https://doi.org/10.1371/journal.pone.0076608>
- Soroli, E., Szenkovits, G., & Ramus, F. (2010). Exploring dyslexics' phonological deficit III: Foreign speech perception and production. *Dyslexia*, 16, 318-340. <https://doi.org/10.1002/dys.415>
- Sparks, R. L., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal*, 75, 3-16. <https://doi.org/10.2307/329830>
- Sparks, R. L., Ganschow, L., Kenneweg, S., & Miller, K. (1991). Use of an Orton-Gillingham approach to teach a foreign language to dyslexic/learning-disabled students: Explicit teaching of phonology in a second language. *Annals of Dyslexia*, 41, 96-118. <https://doi.org/10.1007/BF02648080>
- Spironelli, C., & Angrilli, A. (2010). Developmental aspects of language lateralization in delta, theta, alpha and beta EEG bands. *Biological Psychology*, 85(2), 258-267. <https://doi.org/10.1016/j.biopsycho.2010.07.011>
- Sprenger-Charolles, L., & Serniclaes, W. (2004). Nature et origine de la dyslexie développementale : l'hypothèse phonologique (pp 113-146). In S. Valdois, P. Colé, D. David, (Eds), *Apprentissage de la lecture et dyslexies développementales : De la théorie à la pratique orthophonique et pédagogique*. Solal.
- Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, 38(4), 387-401. [https://doi.org/10.1016/0013-4694\(75\)90263-1](https://doi.org/10.1016/0013-4694(75)90263-1)
- Stevens, C., Fanning, J., Coch, D., Sanders, L., & Neville, H. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. *Brain Research*, 1205, 55-69. <https://doi.org/10.1016/j.brainres.2007.10.108>
- Stevens, K. N., & Klatt, D. H. (1974). Role of formant transition in the voiced-voiceless distinction for stops. *Journal of the Acoustical Society of America*, 55, 653-659. <https://doi.org/10.1121/1.1914578>
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39(4), 456-466. <https://doi.org/10.1016/j.wocn.2010.09.001>

- Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics*, 36(2), 131-145. <https://doi.org/10.3758/BF03202673>
- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: the re-education of selective perception. In J. G. Hansen Edwards and M. L., Zampini (Eds). *Phonology and second language acquisition*. Benjamins.
- Strouse, A., Ashmead, D. H., Ohde, R. N., & Grantham, D. W. (1998). Temporal processing in the aging auditory system. *Journal of Acoustical Society of America*, 104(4), 2385-2399. <https://doi.org/10.3342/kjorl-hns.2011.54.9.585>
- Suemitsu, A., Dang, J., Ito, T., & Tiede, M. (2015). A real-time articulatory visual feedback approach with target presentation for second language pronunciation learning. *The Journal of the Acoustical Society of America*, 138(4), 382-387. <https://doi.org/10.1121/1.4931827>
- Svensson, I., & Jacobson, C. (2006). How persistent are phonological difficulties? A longitudinal study of reading retarded children. *Dyslexia*, 12, 3-20. <https://doi.org/10.1002/dys.296>
- Swan, D., & Goswami, U. (1997). Phonological awareness deficits in developmental dyslexia and the phonological representations hypothesis. *Journal of Experimental Child Psychology*, 66, 18-41. <https://doi.org/10.1006/jecp.1997.2375>
- Taimi, L., Jähi, K., Alku, P., Peltola, M.S. (2014). Children learning a non-native vowel – The effect of a two-day production training. *Journal of Language Teaching Research*, 5, 1229-1235. <https://doi.org/10.4304/jltr.5.6.1229-1235>
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, 9, 182-198.
- Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., Schreiner, C., Jenkins, W. M., & Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271, 81-84. <https://doi.org/10.1126/science.271.5245.81>
- Tamminen, H., Peltola, M. S., Kujala, T., & Näätänen, R. (2015). Phonetic training and non-native speech perception - New memory traces evolve in just three days as indexed by the mismatch negativity (MMN) and behavioural measures. *International Journal of Psychophysiology*, 97(1), 23-29. <https://doi.org/10.1016/j.ijpsycho.2015.04.020>
- Teki, S., Grube, M., & Griffiths, T. D. (2011). A unified model of time perception accounts for duration-based and beat-based timing. *Frontiers in Integrative Neuroscience*, 5, 1-7. <https://doi.org/10.3389/fnint.2011.00090>
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., Obrig, H., & Wartenburger, I. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *Journal of Neuroscience*, 29(47), 14726-14733. <https://doi.org/10.1523/JNEUROSCI.1246-09.2009>



- Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., & Gabrieli, J. D. E. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Sciences*, *100*(5), 2860-2865. <https://doi.org/10.1073/pnas.0030098100>
- Theunissen, F., & Miller, J. P. (1995). Temporal encoding in nervous systems: A rigorous definition. *Journal of Computational Neuroscience*, *2*(2), 149-162. <https://doi.org/10.1007/BF00961885>
- Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Computer Assisted Language Instruction Consortium Journal*, *28*(3), 744-765. <https://doi.org/10.11139/cj.28.3.744-765>
- Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading*, *29*(3), 334-348. <https://doi.org/10.1111/j.1467-9817.2006.00312.x>
- Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology Paris*, *102*(1-3), 120-129. <https://doi.org/10.1016/j.jphysparis.2008.03.007>
- Thomson, J. M., & Goswami, U. (2010). Learning novel phonological representations in developmental dyslexia: Associations with basic auditory processing of rise time and phonological awareness. *Reading and Writing*, *23*(5), 453-473. <https://doi.org/10.1007/s11145-009-9167-9>
- Thomson, J. M., Leong, V., & Goswami, U. (2013). Auditory processing interventions and developmental dyslexia: A comparison of phonemic and rhythmic approaches. *Reading and Writing: An Interdisciplinary Journal*, *26*(2), 139-161. <https://doi.org/10.1007/s11145-012-9359-6>
- Thompson, E. C., Woodruff Carr, K., White-Schwoch, T., Tierney, A., Nicol, T., & Kraus, N. (2016). Hemispheric asymmetry of endogenous neural oscillations in young children: Implications for hearing speech in noise. *Scientific Reports*, *6*, 1-6. <https://doi.org/10.1038/srep19737>
- Tillmann, B. (2012). Music and language perception: Exceptions, structural integration, and cognitive sequencing. *Topics in Cognitive Science*. 1-17. <https://doi.org/10.1111/j.1756-8765.2012.01209.x>
- Tillmann, B., & Bedoin, N. (soumis). *Pathologies et troubles du développement de l'enfant*. (Chapitre d'ouvrage). Dunod.
- Torppa, R., Faulkner, A., Laasonen, M., Lipsanen, J., & Sammler, D. (2020). Links of prosodic stress perception and musical activities to language skills of children with cochlear implants and normal hearing. *Ear and Hearing*, *41*(2), 395-410. <https://doi.org/10.1097/AUD.0000000000000763>

- Toyomura, A., Fujii, T., & Kuriki, S. (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *NeuroImage*, 57(4), 1507-1516. <https://doi.org/10.1016/j.neuroimage.2011.05.039>
- Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, 6(7), 669-673. <https://doi.org/10.1038/nn1084>
- Tremblay, A., Broersma, M., Coughlin, C. E., & Choi, J. (2016). Effects of the native language on the learning of fundamental frequency in second-language speech segmentation. *Frontiers in Psychology*, 7, 1-15. <https://doi.org/10.3389/fpsyg.2016.00985>
- Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research*, 45, 564-572. [https://doi.org/10.1007/978-1-4020-8265-8\\_100733](https://doi.org/10.1007/978-1-4020-8265-8_100733)
- Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of the Acoustical Society of America*, 102(6), 3762-3773. <https://doi.org/10.1121/1.420139>
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis, A. B. (2001). Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear and Hearing*, 22(2), 79-90. <https://doi.org/10.1097/00003446-200104000-00001>
- Turkeltaub, P. E., & Coslett, H. B. (2010). Localization of sublexical speech perception components. *Brain and Language*, 114(1), 1-15. <https://doi.org/10.1016/j.bandl.2010.03.008>
- Valdois, S., Bosse, M. - L., Ans, B., Carbonnel, S., Zorman, M., & David, D. (2003). Phonological and visual processing deficits can dissociate in developmental dyslexia: Evidence from two case studies. *Reading and Writing: An Interdisciplinary Journal*, 16, 541-572. <https://doi.org/10.1023/A:1025501406971>
- Van Beinum, F. J., Schwippert, C. E., Been, P. H., van Leeuwen, T. H., & Kuijpers, C. T. L. (2005). Development and application of a /bAk-/dAk/ continuum for testing auditory perception within the Dutch longitudinal dyslexia study. *Speech Communication*, 47(1-2), 124-142. <https://doi.org/10.1016/j.specom.2005.04.003>
- Van Hirtum, T., Ghesquière, P., Wouters, J. (2019a). Atypical neural processing of rise time by adults with dyslexia. *Cortex* 113, 128-140.
- Van Hirtum, T., Moncada-Torres, A., Ghesquière, P., & Wouters, J. (2019b). Speech envelope enhancement instantaneously effaces atypical speech perception in dyslexia. *Ear and Hearing*, 40(5), 1242-1252. <https://doi.org/10.1097/AUD.0000000000000706>
- Van Summers, W., & Leek, M.R. (1992). The role of spectral and temporal cues in vowel identification by listeners with impaired hearing. *Journal of Speech, Language, and Hearing Research*, 35(5), 1189-1199.

- Van Wijk, B. C. M., Beek, P. J., & Daffertshofer, A. (2012). Neural synchrony within the motor system: What have we learned so far? *Frontiers in Human Neuroscience*, *6*, 1-15. <https://doi.org/10.3389/fnhum.2012.00252>
- Vandermosten, M., Boets, B., Luts, H., Poelmans, H., Golestani, N., Wouters, J., & Ghesquière, P. (2010). Adults with dyslexia are impaired in categorizing speech and nonspeech sounds on the basis of temporal cues. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10389-10394. <https://doi.org/10.1073/pnas.0912858107>
- Vandermosten, M., Cuynen, L., Vanderauwera, J., Wouters, J., & Ghesquière, P. (2017). White matter pathways mediate parental effects on children's reading precursors. *Brain and Language*, *173*, 10-19. <https://doi.org/10.1016/j.bandl.2017.05.002>
- Vandermosten, M., Hoeft, F., & Norton, E. S. (2016). Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: A review and quantitative meta-analysis. *Current Opinion in Behavioral Sciences*, *10*, 155-161. <https://doi.org/10.1016/j.cobeha.2016.06.007>
- Vanrullen, R., & Dubois, J. (2011). The psychophysics of brain rhythms. *Frontiers in Psychology*, *2*, 203. <https://doi.org/10.3389/fpsyg.2011.00203>
- Venables, P. H. (1964). Input dysfunction in schizophrenia. *Progress in Experimental and Personality Research*, *72*, 1-47
- Viemeister, N. F., & Wakefield, G. H. (1991). Temporal integration and multiple looks. *Journal of the Acoustical Society of America*, *90*, 858-865. <https://doi.org/10.1121/1.401953>
- Virtala, P., Talola, S., Partanen, E. & Kujala, T. (2020). Poor neural and perceptual phoneme discrimination during acoustic variation in dyslexia. *Scientific Reports*, *10*, 8646. <https://doi.org/10.1038/s41598-020-65490-3>
- Voyer, D., & Techentin, C. (2009). Dichotic listening with consonant–vowel pairs: The role of place of articulation and stimulus dominance. *Journal of Phonetics*, *37*, 162-172. <https://doi.org/10.1016/j.wocn.2008.12.001>
- Wade-Woolley, L., & Wood, C. (2006). Prosodic sensitivity and reading development. *Journal of Research in Reading*, *29*(3), 253-257. <https://doi.org/10.1111/j.1467-9817.2006.00306.x>
- Wallet, L. (2015). La phonologie de laboratoire : à l'interface phonologie-phonétique. *Langages*, *2*(198), 133-150. <https://doi.org/10.3917/lang.198.0133>
- Wan, C. Y., Rüber, T., Hohmann, A., & Schlaug, G. (2010). The therapeutic effects of singing in neurological disorders. *Music Perception*, *27*(4), 287-295. <https://doi.org/10.1525/mp.2010.27.4.287>

- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113(2), 1033-1043. <https://doi.org/10.1121/1.1531176>
- Wanlin, P., Dessart, P. & Crahay, M. (2019). Hétérogénéité des niveaux et rythmes des élèves : dilemmes d’enseignants du primaire en formation et titulaires dans deux contextes. *Les Cahiers du CERFEE*, 54. <https://doi.org/10.4000/edso.7518>
- Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., & Kraus, N. (2009). Relating structure to function: Heschl’s gyrus and acoustic processing. *Journal of Neuroscience*, 29(1), 61-69. <https://doi.org/10.1523/JNEUROSCI.3489-08.2009>
- Werker, J. F., & Tees, R. C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 41(1), 48-61. <https://doi.org/10.1037/h0084150>
- Wexler, B. E., & Halwes, T. (1983). Increasing the power of dichotic methods: The fused rhymed words test. *Neuropsychologia*, 21(1), 59-66. [https://doi.org/10.1016/0028-3932\(83\)90100-8](https://doi.org/10.1016/0028-3932(83)90100-8)
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701-702. <https://doi.org/10.1038/nn1263>
- Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., Stein, J. F., & Green, G. G. R. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Current Biology*, 8(14), 791-797. [https://doi.org/10.1016/S0960-9822\(98\)70320-3](https://doi.org/10.1016/S0960-9822(98)70320-3)
- Wolff, P. H. (2002). Timing precision and rhythm in developmental dyslexia. *Reading and Writing*, 15(1-2), 179-206. <https://doi.org/10.1023/a:1013880723925>
- Wood, C. C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *Journal of the Acoustical Society of America*, 60, 1381-1389. <https://doi.org/10.1121/1.381231>
- Wronka, E., Kaiser, J., & Coenen, A. M. L. (2008). The auditory P3 from passive and active three-stimulus oddball paradigm. *Acta Neurobiologica*, 68(3), 362-372.
- Wu, Y., Gendrot, C., Hallé, P., & Adda-Decker, M. (2015). On improving the pronunciation of French /r/ in Chinese learners by using real-time ultrasound visualization. *ICPhS 2015, 18<sup>th</sup> International Congress of Phonetic Sciences* (Août 2015), Glasgow, UK.
- Yang, Y., Dewald, J. P. A., van der Helm, F. C. T., & Schouten, A. C. (2018). Unveiling neural coupling within the sensorimotor system: directionality and nonlinearity. *European Journal of Neuroscience*, 48(7), 2407-2415. <https://doi.org/10.1111/ejn.13692>
- Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2010). Training the brain to weight speech cues differently: A study of

- Finnish second-language users of English. *Journal of Cognitive Neuroscience*, 22(6), 1319-1332. <https://doi.org/10.1162/jocn.2009.21272>
- Zabell, C., & Everatt, J. (2002). Surface and phonological subtypes of adult developmental dyslexia. *Dyslexia*, 8, 160-177. <https://doi.org/10.1002/dys.223>
- Zaehle, T., Wüstenberg, T., Meyer, M., & Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: A sparse temporal sampling fMRI study. *European Journal of Neuroscience*, 20(9), 2447-2456. <https://doi.org/10.1111/j.1460-9568.2004.03687.x>
- Zatorre, R. J. (1989). Perceptual asymmetry on the dichotic fused word test and cerebral speech lateralization determined by carotid sodium amytal test. *Neuropsychologia*, 27, 1207-1219. [https://doi.org/10.1016/0028-3932\(89\)90033-x](https://doi.org/10.1016/0028-3932(89)90033-x)
- Zatorre, R.J., (1997). Cerebral correlates of human auditory processing: Perception of speech and musical sounds. In J.Syka (Ed.), *Acoustical Signal Processing in the Central Auditory System* (pp. 453– 468). Plenum Press, New York.
- Zatorre, R. J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences*, 930(1), 193-210.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946-953. <https://doi.org/10.1093/cercor/12.2.140>
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 6(1), 37-46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)
- Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, 14(4), 1908-1919. <https://doi.org/10.1523/jneurosci.14-04-01908.1994>
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058), 846-849.
- Zhang, Y., & Ding, H. (2020). The effect of ambiguity awareness on second language learners' prosodic disambiguation. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.573520>
- Zhang, S., Li, K., Lo, W. K., & Meng, H. (2010). Perception of English suprasegmental features by non-native Chinese learners. *Proceedings of the International Conference on Speech Prosody*, 1-4.
- Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences of the United States of America*, 113(19), 5212-5217. <https://doi.org/10.1073/pnas.1603984113>

Zoefel, B. (2018). Speech entrainment: Rhythmic predictions carried by neural oscillations. *Current Biology*, 28(18), 1102-1104. <https://doi.org/10.1016/j.cub.2018.07.048>

Zoubinetzky, R., Collet, G., Nguyen-Morel, M. A., Valdois, S., & Serniclaes, W. (2019). Remediation of allophonic perception and visual attention span in developmental dyslexia: A joint assay. *Frontiers in Psychology*, 10, 1-16. <https://doi.org/10.3389/fpsyg.2019.01502>

## **Annexes**

---



# Annexe 1. Fiche de passation adressée aux participants de l'Expérience 10

## Fiche d'instructions

Bonjour, tout d'abord merci d'avoir accepté de participer à cette étude, c'est très important pour nous !

Nous aurions préféré vous rencontrer pour vous faire passer cette expérience, mais Covid-19 oblige, nous avons dû adapter le protocole. Par ce fait il implique une plus grande vigilance de votre part quant aux exercices à réaliser.

### Étape n°1 : Vérification du matériel

Êtes-vous muni d'un **ordinateur** (hors Chromebook) et d'un **casque stéréo** ? Si oui, vous pouvez participer à cette étude.

Pour vérifier si votre casque fait stéréo :

- Mettez votre casque dans le bon sens (L = gauche ; R = droite)
- En suivant ce lien : <https://youtu.be/6TWJaFD6R2s> ; vérifiez que lorsque l'enceinte "Left" s'anime, vous n'entendez le son que dans votre oreille gauche ; et que lorsque l'enceinte "Right" s'anime, vous n'entendez le son que dans votre oreille droite.
  - ⇒ Si vous entendez le son dans vos deux oreilles : vous n'avez pas de casque stéréo et vous ne pouvez pas participer à cette étude.
  - ⇒ Si vous entendez une grande différence en termes de volume sonore entre vos deux oreilles : vous ne pouvez pas participer à notre étude.

Si tout est correct pour cette étape, passons à la suivante.

### Étape n°2 : Installation des programmes

Téléchargez le dossier "Expérimentation" et mettez-le sur votre bureau :

- Lien Windows :
  - Lancez l'exécuteur " opensesame\_3.3.5-py37-win64-1.exe »
  - Acceptez que le programme modifie votre ordinateur
  - Lancez l'installation, le logiciel se lancera ensuite automatiquement
- Lien Mac :
  - Lancez l'exécuteur " opensesame\_3.2.8-py2.7-macos-1.dmg"
  - Une fenêtre s'ouvre, déplacez l'icône OpenSesame sur l'icône des applications
  - L'installation commence alors et le logiciel se lancera automatiquement ensuite
- Lien Ubuntu / Linux :

- Allez sur <https://osdoc.cogsci.nl/3.3/download/#ubuntu> après avoir installé le logiciel (nous ne sommes pas des habituées de ce système)

Si vous avez réussi à ouvrir le logiciel, d'abord c'est top ! Sinon appelez-nous, nous vous aiderons avec plaisir !

### **Étape n°3 : Informations à savoir avant de commencer**

Compte-tenu de la crise sanitaire actuelle, cette expérience ne peut pas être réalisée en présentiel.

La durée totale de ce protocole est d'environ 30mn : 15mn x 2.

Le silence est de rigueur pour cette expérience, veuillez à vous installer dans une pièce calme.

Le logiciel OpenSesame utilisé pour effectuer les tâches est un peu capricieux, il peut mettre beaucoup de temps à s'ouvrir. Afin de ne pas le surcharger et de ne pas l'ouvrir plusieurs fois, double-cliquez une seule fois sur le fichier, puis patientez.

Pour nommer vos fichiers, il vous faudra votre CODESUJET qui va, ou vous sera, transmis par mail.

### **Étape n°4 : Exercices du jour (Première Session – Jour 1)**

- **Exercice de perception** : l'exercice se déroule en 2 temps, le fichier utilisé est le même, seule la consigne va différer :
  - Mettez votre casque stéréo dans le bon sens.
  - Sur le bureau de votre ordinateur, ouvrez le dossier "Expérimentation".
  - Double-cliquez sur "Voc\_voyelles", laissez le temps au programme de s'ouvrir, ne recliquez pas).
  - Une fois le programme ouvert, cliquez sur la grosse flèche verte en haut à gauche.
  - Une petite fenêtre s'ouvre : cliquez sur "OK".
  - Nommez votre fichier de la manière suivante : Voc\_CODESUJET-1.
  - Cliquez sur "Enregistrer" : l'exercice se lance !
  - ⓪ La 1<sup>ère</sup> consigne est la suivante : vous allez entendre un mot dans votre casque, et sur votre écran vous allez voir apparaître **2 mots écrits de part et d'autre d'une image** (un mot à gauche et un mot à droite). Appuyez sur la touche correspondant au mot écrit.
  - Ⓛ La 2<sup>nde</sup> consigne est la suivante : vous allez entendre deux mots dans votre casque, et sur votre écran va apparaître **une image**. Vous devez décider à l'aide du clavier quel mot (le 1<sup>er</sup> ou le 2<sup>nd</sup>) correspond à l'image.
  
- **Exercice de production** : nous allons vous demander de vous enregistrer, avec votre téléphone portable. Il vous est demandé de dire les mots anglais avec le meilleur accent possible (là encore ne vous mettez pas de pression, nous n'écoutons pas les enregistrements mais nous les analysons sous forme de spectre).
  - Mettez votre casque stéréo dans le bon sens.
  - Sur le bureau de votre ordinateur, ouvrez le dossier "Expérimentation".

- Double-cliquez sur “Production“ “, laissez le temps au programme de s’ouvrir, ne recliquez pas).
  - Une fois le programme ouvert, cliquez sur la grosse flèche verte en haut à gauche.
  - Une petite fenêtre s’ouvre : cliquez sur “OK“.
  - Nommez votre fichier de la manière suivante : Production\_CODESUJET-1.
  - Cliquez sur “Enregistrer“, l’exercice se lance !
  - La consigne est la suivante : Vous allez voir **un mot écrit**. Lisez-le une fois silencieusement puis à haute voix avec le *meilleur accent possible*.
  - Sur votre téléphone, ouvrez l’enregistreur vocal (dictaphone) et appuyez sur “lecture“.
  - Pensez à marquer une pause entre chaque mot (cela nous aidera pour notre analyse).
  - Une fois l’exercice terminé, mettez cet enregistrement sur votre ordinateur et renommez-le : HV\_CODESUJET-1.
- La première partie de l’exercice est terminée ! La deuxième partie va se lancer, vous pouvez d’ailleurs voir la consigne sur votre écran.
- Sur votre téléphone, ouvrez de nouveau l’enregistreur vocal (dictaphone) et appuyez sur “lecture“. Pensez à marquer une pause entre chaque mot (cela nous aidera pour notre analyse).
  - La consigne est la suivante : Vous allez entendre une voix **prononcer un mot** anglais puis vous le verrez **écrit**. Vous devrez le répéter avec le meilleur accent possible.
  - Une fois l’exercice terminé, mettez cet enregistrement sur votre ordinateur et renommez-le : Imi\_CODESUJET-1.

Vous avez fini la première session ! Vous pouvez envoyer les fichiers par mail à l’expérimentatrice.

Fichiers à envoyer :

- Voc\_CODESUJET-1
- HV\_CODESUJET-1
- Imi\_CODESUJET-1

### **Étape n°5 : Exercices du jour (Deuxième Session – Jour 5)**

Cette étape suit exactement la même procédure que l’**étape n°4**, vous devez effectuer exactement les mêmes exercices.

**Sauf**, nommez vos fichiers de la manière suivante :

Voc\_CODESUJET-2.

HV\_CODESUJET-2.

Imi\_CODESUJET-2.

Une fois cela terminé vous avez fini la deuxième session et donc l’expérimentation ! Vous pouvez envoyer les fichiers par mail à l’expérimentatrice qui vous expliquera l’étude que nous menons.

Fichiers à envoyer :

- Voc\_CODESUJET-2
- HV\_CODESUJET-2
- Imi\_CODESUJET-2

Nous vous remercions de nouveau pour votre participation à cette expérience, en espérant que ces instants n'ont pas été désagréables pour vous !

A bientôt peut-être.

## Annexe 2. Fiche de passation adressée aux participants des Expériences 11 et 12

### Fiche d'instructions

Bonjour, tout d'abord merci d'avoir accepté de participer à cette étude, c'est très important pour nous !

Nous aurions préféré vous rencontrer pour vous faire passer cette expérience, mais Covid-19 oblige, nous avons dû adapter le protocole. Par ce fait il implique une plus grande vigilance de votre part quant aux exercices à réaliser.

#### Étape n°1 : Vérification du matériel

Êtes-vous muni d'un **ordinateur** (hors Chromebook) et d'un **casque stéréo** ? Si oui, vous pouvez participer à cette étude.

Pour vérifier si votre casque fait stéréo :

- Mettez votre casque dans le bon sens (L = gauche ; R = droite)
- En suivant ce lien : <https://youtu.be/6TWJaFD6R2s> ; vérifiez que lorsque l'enceinte "Left" s'anime, vous n'entendez le son que dans votre oreille gauche ; et que lorsque l'enceinte "Right" s'anime, vous n'entendez le son que dans votre oreille droite.
  - ⇒ Si vous entendez le son dans vos deux oreilles : vous n'avez pas de casque stéréo et vous ne pouvez pas participer à cette étude.
  - ⇒ Si vous entendez une grande différence en termes de volume sonore entre vos deux oreilles : vous ne pouvez pas participer à notre étude.

Si tout est correct pour cette étape, passons à la suivante.

#### Étape n°2 : Installation des programmes

Téléchargez le dossier "Expérimentation" et mettez-le sur votre bureau :

- Lien Windows :
  - Lancez l'exécuteur " opensesame\_3.3.5-py37-win64-1.exe »
  - Acceptez que le programme modifie votre ordinateur
  - Lancez l'installation, le logiciel se lancera ensuite automatiquement
- Lien Mac :
  - Lancez l'exécuteur " opensesame\_3.2.8-py2.7-macos-1.dmg"
  - Une fenêtre s'ouvre, déplacez l'icône OpenSesame sur l'icône des applications
  - L'installation commence alors et le logiciel se lancera automatiquement ensuite
- Lien Ubuntu / Linux :
  - Allez sur <https://osdoc.cogsci.nl/3.3/download/#ubuntu> après avoir installé le logiciel (nous ne sommes pas des habituées de ce système)

Si vous avez réussi à ouvrir le logiciel, d'abord c'est top ! Sinon appelez-nous, nous vous aiderons avec plaisir !

### Étape n°3 : Informations à savoir avant de commencer

Compte-tenu de la crise sanitaire actuelle, cette expérience ne peut pas être réalisée en présentiel.

La durée totale de ce protocole est d'environ 1h15 : 15mn x 5.

Le silence est de rigueur pour cette expérience, veuillez à vous installer dans une pièce calme.

Le logiciel OpenSesame utilisé pour effectuer les tâches est un peu capricieux, il peut mettre beaucoup de temps à s'ouvrir. Afin de ne pas le surcharger et de ne pas l'ouvrir plusieurs fois, double-cliquez une seule fois sur le fichier, puis patientez.

Pour nommer vos fichiers, il vous faudra votre CODESUJET qui va, ou vous sera, transmis par mail.

### Étape n°4 : Exercices du jour (Première Session – Jour 1)

- **Exercice de perception** : l'exercice se déroule en 2 temps, le fichier utilisé est le même, seule la consigne va différer :
  - Mettez votre casque stéréo dans le bon sens.
  - Sur le bureau de votre ordinateur, ouvrez le dossier "Expérimentation".
  - Double-cliquez sur "Voc\_voyelles", laissez le temps au programme de s'ouvrir, ne recliquez pas).
  - Une fois le programme ouvert, cliquez sur la grosse flèche verte en haut à gauche.
  - Une petite fenêtre s'ouvre : cliquez sur "OK".
  - Nommez votre fichier de la manière suivant : Voc\_CODESUJET-1.
  - Cliquez sur "Enregistrer" : l'exercice se lance !
  - Ⓞ La 1<sup>ère</sup> consigne est la suivante : vous allez entendre un mot dans votre casque, et sur votre écran vous allez voir apparaître **2 mots écrits de part et d'autre d'une image** (un mot à gauche et un mot à droite). Appuyer sur la touche correspondant au mot écrit.
  - Ⓞ La 2<sup>nde</sup> consigne est la suivante : vous allez entendre deux mots dans votre casque, et sur votre écran va apparaître **une image**. Vous devez décider à l'aide du clavier quel mot (le 1<sup>er</sup> ou le 2<sup>nd</sup>) correspond à l'image.
  
- **Exercice de production** : nous allons vous demander de vous enregistrer, avec votre téléphone portable. Il vous est demandé de dire les mots anglais avec le meilleur accent possible (là encore ne vous mettez pas de pression, nous n'écoutons pas les enregistrements mais nous les analysons sous forme de spectre).
  - Mettez votre casque stéréo dans le bon sens.
  - Sur le bureau de votre ordinateur, ouvrez le dossier "Expérimentation".
  - Double-cliquez sur "Production", laissez le temps au programme de s'ouvrir, ne recliquez pas).
  - Une fois le programme ouvert, cliquez sur la grosse flèche verte en haut à gauche.
  - Une petite fenêtre s'ouvre : cliquez sur "OK".

- Nommez votre fichier de la manière suivante : Production\_CODESUJET-1.
- Cliquez sur “Enregistrer“, l’exercice se lance !
- La consigne est la suivante : Vous allez voir **un mot écrit**. Lisez-le une fois silencieusement puis à haute voix avec le *meilleur accent possible*.
- Sur votre téléphone, ouvrez l’enregistreur vocal (dictaphone) et appuyez sur “lecture“. Pensez à marquer une pause entre chaque mot (cela nous aidera pour notre analyse).
- Une fois l’exercice terminé, mettez cet enregistrement sur votre ordinateur et renommez-le : HV\_CODESUJET-1.

La première partie de l’exercice est terminée ! La deuxième partie va se lancer, vous pouvez d’ailleurs voir la consigne sur votre écran.

- Sur votre téléphone, ouvrez de nouveau l’enregistreur vocal (dictaphone) et appuyez sur “lecture“. Pensez à marquer une pause entre chaque mot (cela nous aidera pour notre analyse).
- La consigne est la suivante : Vous allez entendre une voix **prononcer un mot** anglais puis vous le verrez **écrit**. Vous devrez le répéter avec le meilleur accent possible.
- Une fois l’exercice terminé, mettez cet enregistrement sur votre ordinateur et renommez-le : IMI\_CODESUJET-1.

Vous avez fini la première session ! Vous pouvez envoyer les fichiers par mail à l’expérimentatrice.

Fichiers à envoyer :

- Voc\_CODESUJET-1
- HV\_CODESUJET-1
- IMI\_CODESUJET-1

### Étape n°5 : Entraînements (Jours 2, 3 et 4)

Les entraînements se déroulent en 2 parties : une partie apprentissage de vocabulaire (chaque mot vous sera présenté 2 fois) suivie par une partie exercice. Le tout dure environ 15 minutes.

Chaque jour, vous allez devoir faire ceci :

- Mettez votre casque stéréo dans le bon sens.
- Sur le bureau de votre ordinateur, ouvrez le dossier “Expérimentation“, puis “Entraînement“.
- Double-cliquez sur le fichier correspondant au jour de votre entraînement (laissez au programme le temps de s’ouvrir)
- Une fois le programme ouvert, cliquez sur la grosse flèche verte en haut à gauche.
- Une petite fenêtre s’ouvre, cliquez sur “OK“
- Nommez votre fichier de la manière suivante : Voy\_CODESUJET\_JourX (à la place du X indiquez le numéro du jour).
- Cliquez sur “Enregistrer“.

Une fois les trois jours d’entraînement effectués, envoyez les trois fichiers des trois jours par mail à l’expérimentatrice.



### **Étape n°6 : Exercices du jour (Deuxième Session – Jour 5)**

Cette étape suit exactement la même procédure que l'**étape n°4**, vous devez effectuer exactement les mêmes exercices.

**Sauf**, nommez vos fichiers de la manière suivante :

Voc\_CODESUJET-2.

HV\_CODESUJET-2.

IMI\_CODESUJET-2.

Une fois cela terminé vous avez fini la deuxième session et donc l'expérimentation ! Vous pouvez envoyer les fichiers par mail à l'expérimentatrice qui vous expliquera l'étude que nous menons.

Fichiers à envoyer :

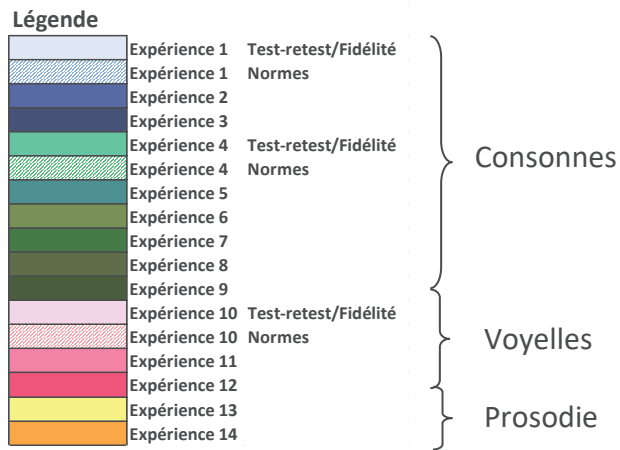
- Voc\_CODESUJET-2
- HV\_CODESUJET-2
- IMI\_CODESUJET-2

Nous vous remercions de nouveau pour votre participation à cette expérience, en  
espérant que ces instants n'ont pas été désagréables pour vous !

A bientôt peut-être.

### Annexe 3. Tableau de répartition des 291 participants à travers les expériences

Un total de 224 adultes normo-lecteurs et 67 adultes dyslexiques ont participé aux expériences de la thèse. Certains adultes ont participé à plusieurs expériences dont les numéros sont rappelés dans la légende. Le tableau permet, grâce aux couleurs, de comprendre comment les groupes de participants ont été constitués pour chacune des expériences en vue d'optimiser la disponibilité des 291 participants.



Groupes indépendants	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	Exp 9	Exp 10	Exp 11	Exp 12	Exp 13	Exp 14
12 normo-lecteurs														
12 dyslexiques														
9 dyslexiques														
6 dyslexiques														
16 normo-lecteurs														
4 normo-lecteurs														
20 normo-lecteurs														
14 normo-lecteurs														
20 normo-lecteurs														
20 normo-lecteurs														
20 normo-lecteurs														
20 dyslexiques														
60 normo-lecteurs														
30 normo-lecteurs														
20 dyslexiques														

Total normo-lecteurs n = 224 ; dyslexiques n = 67

## Annexe 4. Fiche d'information

Notice d'information participants majeurs  
Conception et validation d'aides par programmes informatisées pour l'anglais L2  
APIPAL2 (sans EEG)

Version 2.0 du 05/04/2021

**Promoteur :**

AM Business  
6 rue des bons enfants  
63120 COURPIERE

Représenté par :

Madame Marion RANVIER  
Web accessibility & CS foundation Director  
11 Place Antonin Poncet  
69002 Lyon

**Investigateur principal :**

**Margot Bouhon**

Doctorante CIFRE en Sciences Cognitives  
11 Place Antonin Poncet 69002 Lyon

Email : [margot.bouhon@contentsquare.com](mailto:margot.bouhon@contentsquare.com)

*Madame, Monsieur,*

*Votre médecin ou un investigateur vous a présenté l'étude APIPAL2 et sollicite votre accord pour que vous y participiez. Ce document a pour objectif de vous donner toutes les informations relatives à cette étude de façon à vous permettre d'exercer au mieux votre liberté de décision. Ce document est obligatoire et son contenu est défini par le **Code de la Santé Publique, article L 1122-1** régissant les recherches impliquant la personne humaine.*

*Il décrit précisément l'étude et mentionne toutes les autorisations réglementaires obtenues pour sa mise en œuvre.*

*Avant de prendre une décision, il est important que vous lisiez attentivement ces pages qui vous apporteront les informations nécessaires concernant les différents aspects de cette **étude**. Vous devez conserver ce document. N'hésitez pas à poser des questions si vous ne comprenez pas certains éléments.*

*La signature du formulaire de consentement devenue obligatoire par l'application du code de la Santé Publique (livre I, titres 2 et 3 du CSP), n'affecte aucunement vos droits légaux.*

*Votre participation est entièrement volontaire.*

A travers cette recherche nous souhaitons créer des alternatives aux méthodes classiques d'apprentissage de l'anglais tout en respectant les programmes scolaires et les possibilités de mise en place au sein des universités. Les méthodes actuelles sont en effet insuffisantes et non adaptées, en particulier en cas de dyslexie ou de difficultés d'apprentissage de cette langue. En parallèle, les nouveaux tests mis en place dans cette recherche pour mesurer l'efficacité des entraînements permettront par la suite aux professionnels de l'éducation et de la santé de disposer d'outils précis pour le dépistage d'élèves dont les difficultés en anglais s'expliquent par des traitements inappropriés des sons de la parole, que cela soit associé à une dyslexie ou non. L'objectif est alors de proposer un dispositif d'apprentissage adapté à chaque apprenant. C'est pourquoi il est important pour nous de faire participer à cette recherche aussi bien des adultes sans pathologie que des adultes dyslexiques.

L'objectif est d'améliorer la perception des sons anglais à travers des entraînements adaptés aux apprenants francophones, y compris en cas de dyslexie. Nous avons montré dans de précédentes études (Moulay & Bedoin, 2016 ; Bouhon, Charron, Foureix & Bedoin, 2017 ; Bouhon & Bedoin, 2018 ; Lecuelle, Ferragne & Bedoin, 2019) que ces entraînements permettaient non seulement une amélioration de la perception en anglais, mais aussi en français, ainsi que des progrès en production de phrases/mots anglais, y compris pour des phrases/mots n'ayant pas été traités lors de l'entraînement.

Nous souhaitons avoir au total, toutes populations confondues, 300 participants.

### **Le protocole**

Nous envisageons que cette recherche permette un meilleur traitement des sons anglais ainsi qu'une meilleure production. Vous n'encourez aucun risque en participant à ce protocole. Sauf en cas d'abandon de la procédure, nous vous demanderons de réaliser ces tests avec le plus de sérieux possible.

Si vous ne participez pas à cette étude, il vous est possible d'utiliser vos propres outils d'apprentissage ou ceux proposés par les sites internet et l'université.

Durant l'étude nous vous demandons de continuer toutes vos activités comme habituellement. Nous souhaitons que notre recherche reflète l'utilisation qui en sera faite par la suite.

### **Méthodologie et déroulement de l'expérimentation**

Cette étude a lieu à distance ou dans les locaux de l'équipe CAP du CRNL.

Ce lieu dispose des moyens humains, matériels et techniques adaptés à la recherche et compatibles avec les impératifs de sécurité des personnes qui s'y prêtent.

Lors de la séance expérimentale, des tâches à réaliser vous seront expliquées et illustrées avec des exemples. Ces tâches impliqueront d'écouter des sons dans un casque audio et éventuellement de voir des images sur un écran, ainsi que de produire des phrases/mots anglais. Vous répondrez à des questions en appuyant sur les touches d'un clavier ou d'un boîtier de réponses, en pointant, ou en produisant de la parole.

Cette étude aura une durée de .....

Vous recevrez à la fin de cette étude une indemnité de 10 euros en cadeaux

### **Contraintes et désagréments**

Cette expérience peut entraîner de la fatigue, dans une mesure comparable à celle de la réalisation de tout exercice classique.

### **Quels sont vos droits ?**

#### ***o Participation volontaire***

Vous êtes entièrement libre d'accepter ou de refuser de participer à cette recherche sans que cela modifie la qualité des soins auxquels vous avez droit, ou les relations existant avec votre médecin ou l'investigateur.

Si vous décidez de participer à cette recherche, mais que vous changez d'avis au cours de celle-ci, vous pouvez à tout moment demander d'interrompre votre participation à l'étude sans aucun préjudice, sans justification de votre part et sans que votre responsabilité soit engagée. Dans ce cas, vos données recueillies jusque-là seront utilisées dans les résultats de l'étude.

D'autre part, s'il le juge nécessaire pour votre bien, l'investigateur pourra modifier votre suivi et vous pourrez continuer à bénéficier pleinement de sa compétence.

Plus généralement, votre participation à cette étude ne décharge en aucune façon le promoteur et les investigateurs de leurs devoirs envers vous.

À tout moment, toutes les informations que vous souhaiteriez obtenir ultérieurement concernant cette recherche vous seront communiquées dans la mesure du possible par votre médecin et/ou par l'investigateur. Vous serez tenu informé de toute nouvelle donnée importante concernant l'étude à laquelle vous acceptez de participer.

L'investigateur, tout comme le promoteur, peut interrompre à tout moment votre participation à l'étude s'il juge que cela est dans votre intérêt, ou arrêter l'étude dans sa globalité pour des raisons médicales, administratives ou autres.

Pour pouvoir participer à cette étude, vous devez nécessairement être affilié à un régime d'assurance maladie telle que celui de la Sécurité Sociale.

#### ***o Confidentialité et protection des données***

Dans le cadre de la recherche non interventionnelle à laquelle nous vous proposons de participer, un traitement informatique de vos données personnelles va être mis en œuvre pour permettre d'analyser les résultats de la recherche au regard de son objectif précité. Le traitement est fondé sur l'exécution d'une mission d'intérêt public dont est investi le responsable de traitement.

Le responsable du traitement est Margot Bouhon dont les coordonnées figurent sur la première page de ce document.

Le délégué à la protection des données (DPO) est Clarisse Marthe-Knaëbel qui peut être contacté à l'adresse e-mail suivante : [privacy@contentsquare.com](mailto:privacy@contentsquare.com) ou par courrier postal à AM business, 11 place Antonin Poncet, 69002 Lyon.

Pour l'analyse, et dans la mesure où ces données sont nécessaires à la recherche, les catégories de données personnelles vous concernant qui sont collectées et traitées dans le cadre de la recherche sont les suivantes :

- les données relatives à vos habitudes de vie.

Ces données seront identifiées de façon confidentielle par un code alphanumérique.

Les données personnelles vous concernant et que nous aurons recueillies lors de votre participation seront uniquement transmises à la responsable scientifique, et à ses collaborateurs habilités au sein de la société. Vos données seront également transmises au promoteur. Vos données pourront également, dans des conditions assurant leur confidentialité, être transmises aux autorités de santé françaises ou étrangères et aux scientifiques habilités. Des garanties de transferts appropriées (via "Firefox Send") seront mises en œuvre pour assurer la sécurité des données transférées hors UE. Vos données personnelles seront conservées pour une durée de 5 ans et archivées pour une durée de 15 ans puis supprimées.

Vos données personnelles seront collectées et traitées conformément au Règlement Général sur la Protection des Données 2016/679 du 27 avril 2016 (« RGPD »), à la loi Informatique et Libertés du 6 janvier 1978, et à la méthodologie de référence MR003 de la Commission Nationale de l'Informatique et des Libertés (CNIL) pour laquelle AM Business a signé un engagement de conformité.

- **Exercer vos droits**

Vous pourrez également, à tout moment, exercer votre droit d'accès, de rectification, de suppression, de limitation et d'opposition au traitement vous concernant en faisant la demande auprès du DPO dont les coordonnées figurent ci-dessus.

Si vous avez des questions ou des réclamations au sujet du traitement de vos données au cours de cette étude, vous pouvez contacter Margot Bouhon par voie électronique : [margot.bouhon@contentsquare.com](mailto:margot.bouhon@contentsquare.com).

Si vous estimez, après avoir contacté Margot Bouhon, que vos droits sur vos données ne sont pas respectés, vous pouvez adresser une réclamation à la CNIL : <https://www.cnil.fr/fr/webform/adresser-une-plainte>.

La publication des résultats de la recherche ne comportera aucun résultat individuel identifiant. Les résultats pourront faire l'objet de présentations scientifiques ou de publications dans la presse médicale et scientifique, mais en aucun cas votre nom ou toute autre donnée identifiante n'apparaîtra. Vous avez la possibilité d'être informé des résultats globaux en fin d'étude, conformément au dernier alinéa de l'article L1122-1 du code de la santé publique. Votre participation à la présente recherche est totalement libre. Vous pouvez quitter l'étude à tout moment, sans encourir aucune responsabilité ni subir aucun préjudice. Pour cela, il vous suffira d'en informer Margot Bouhon. En cas d'interruption de l'étude, les informations vous concernant seront

conservées sauf opposition de votre part (et dans ce cas, elles ne pourront être supprimées qu'à condition que cela ne compromette pas gravement la réalisation des objectifs de la recherche).

Par ailleurs, sauf opposition expresse de votre part adressée au responsable scientifique, Margot Bouhon, dont les coordonnées figurent sur la première page de ce document, vos données recueillies dans le cadre de cette étude pourront être réutilisées lors de recherches ultérieures exclusivement à des fins scientifiques. Seules les données agrégées issues de la présente recherche pourront être réutilisées pour des projets ultérieurs. Aucune réutilisation de donnée à caractère personnel ne sera faite. Aussi, le RGPD et la Loi informatique et Libertés ne s'appliquent pas dans un tel contexte.

○ **Dispositions réglementaires**

Le Comité de Protection des Personnes SUD-EST V a émis un avis favorable à la réalisation de cette étude le 28/04/21. Enfin, cette recherche respecte le règlement général sur la protection des données.

L'investigateur doit vous fournir toutes les explications nécessaires concernant cette recherche. Vous avez le droit d'arrêter votre participation à quelque moment que ce soit, et quel que soit le motif ; vous continuerez à bénéficier du suivi médical et cela n'affectera en rien votre surveillance future.

**Qui pouvez-vous contacter pour toute question ?**

Si vous avez des questions concernant l'étude, n'hésitez pas à nous les poser. Nous pouvons vous donner les informations complémentaires que vous souhaitez. Les noms et numéros de téléphone des personnes à contacter sont les suivants :

Investigateur coordonnateur de l'étude

**Margot Bouhon**

Thèse CIFRE en Sciences Cognitives, ED Neurosciences et Cognition, Université Claude Bernard Lyon1,  
Centre de Recherche en Neurosciences de Lyon (CRNL), équipes CAP et TRAJECTOIRES

Neurocampus, Ch Le Vinatier, 95 Bd Pinel 69500 Bron

AM Business, 11 place Antonin Poncet, 69002, Lyon

Email : [margot.bouhon@contentsquare.com](mailto:margot.bouhon@contentsquare.com)

Nous vous remercions de l'attention que vous avez portée à la lecture de cette notice. Une copie de ce document vous sera remise pour que vous puissiez bénéficier de l'ensemble des informations concernant votre participation à l'étude.

Lorsque vous aurez lu cette note d'information, il vous sera proposé, si vous êtes d'accord, de donner votre consentement écrit en signant le formulaire préparé à cet effet.



## Annexe 5. Formulaire de consentement

### FORMULAIRE DE CONSENTEMENT DE PARTICIPATION A UNE RECHERCHE

AM Business (représenté par Marion RANVIER, 11 place Antonin Poncet, 69002 Lyon) vous propose de participer à une recherche impliquant la personne humaine intitulée **CONCEPTION ET VALIDATION D'AIDES PAR PROGRAMMES INFORMATISES POUR L'ANGLAIS L2 (APIPAL2)** organisée au Centre de Recherche en Neurosciences de Lyon (CRNL), sous la responsabilité scientifique de Margot Bouhon. Cette recherche a pour objectif scientifique général de créer des alternatives aux méthodes classiques d'apprentissage de l'anglais tout en respectant les programmes scolaires et les possibilités de mise en place au sein des universités. Les méthodes actuelles sont en effet insuffisantes et non adaptées en cas de dyslexie ou de difficultés d'apprentissage de cette langue. En parallèle, les nouveaux tests mis en place dans cette recherche pour mesurer l'efficacité des entraînements permettront par la suite aux professionnels de l'éducation et de la santé de disposer d'outils précis pour le dépistage d'élèves dont les difficultés en anglais s'expliquent par des traitements inappropriés des sons de la parole, que cela soit associé à une dyslexie ou non. L'objectif est alors de proposer une prise en charge adaptée à chaque apprenant. C'est pourquoi il est important pour nous de faire participer à cette recherche aussi bien des adultes sans pathologie que des adultes dyslexiques.

L'objectif est d'améliorer la perception des sons anglais à travers des entraînements adaptés aux apprenants francophones, y compris en cas de dyslexie. Nous avons montré dans de précédentes études (Moulay & Bedoin, 2016 ; Bouhon, Charron, Foureix & Bedoin, 2017 ; Bouhon & Bedoin, 2018 ; Lecuelle, Ferragne & Bedoin, 2019) que ces entraînements permettaient non seulement une amélioration de la perception en anglais, mais aussi en français, ainsi que des progrès en production de sons/phrases anglais, y compris pour des mots/phrases n'ayant pas été traités lors de l'entraînement.

Ce protocole a reçu un avis favorable du Comité de Protection des Personnes Sud-Est V, le 28/04/21.

Votre consentement ne décharge pas les organisateurs de la recherche de leur responsabilité et vous conserverez tous vos droits garantis par la Loi du 9 août 2004 et la Loi informatique et liberté du 6 Janvier 1978 modifiée.

A l'issue de la recherche, une information sur les résultats globaux de l'étude pourra vous être donnée, conformément à la loi du 4 mars 2002.

Veillez noter que la loi interdit votre participation à une recherche si vous n'êtes pas affiliés à la Sécurité Sociale.

### Consentement :

Je soussigné(e) \_\_\_\_\_ déclare accepter participer à la recherche impliquant la personne humaine intitulée : « APIPAL2 » organisée par le CRNL dans les conditions précisées ci-dessous et pour laquelle le CPP Sud-Est V a émis un avis favorable le 28/04/21.

Cette étude est régie par le Code de la Santé Publique (loi du 9 août 2004, livre Ier, titre II)

J'ai reçu et j'ai bien compris les informations ci-dessus, et il m'a été précisé que :

- je dois être affilié à un régime de Sécurité Sociale pour pouvoir participer
- je suis libre d'accepter ou de refuser ainsi que d'arrêter à tout moment ma participation ; j'en informerai alors Margot Bouhon
- j'accepte le traitement des données personnelles qui me concernent en conformité avec les dispositions de la loi n°78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés et du règlement européen n°2016/679 sur la protection des données.

Le responsable de traitement est Margot Bouhon, doctorante en Sciences Cognitives. Je pourrai à tout moment demander des informations à Margot Bouhon (joignable [margot.bouhon@contentsquare.com](mailto:margot.bouhon@contentsquare.com)) et exercer auprès d'elle mon droit d'accès, de rectification, d'opposition et d'effacement ainsi que mon droit de demander la limitation de l'usage de mes données, la portabilité et de définir le sort de mes données en cas de décès. Je pourrai retirer mon consentement à tout moment auprès d'elle. En cas de retrait de mon consentement et/ou d'exercice du droit à l'effacement de mes données, j'ai été informé que ce retrait n'a pas d'incidence sur les activités menées et sur l'utilisation des données obtenues sur la base du consentement éclairé exprimé avant que celui-ci n'ait été retiré. Je consens à ce que les données antérieurement collectées pourront être conservées et utilisées dans le cadre de la présente recherche.

La publication des résultats de la recherche ne comportera aucun résultat individuel identifiant.

Si je le désire, je peux être tenu au courant des résultats globaux de la recherche en m'adressant à Margot Bouhon.

Mon consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités. Je conserve tous mes droits garantis par la loi.

Fait à .....,

le .....

**NOM ET SIGNATURE DE L'INVESTIGATEUR**

Margot Bouhon

A handwritten signature in black ink, appearing to be 'Margot Bouhon', enclosed within a hand-drawn oval.

**SIGNATURE DU SUJET PRECEDEE DE LA MENTION LU ET APPROUVE**