

Université Lumière Lyon 2
Ecole doctorale Neurosciences et Cognition (NSCo)
Laboratoire Dynamique Du Langage, UMR 5596

THESE

Présentée en vue de l'obtention du grade de
DOCTEUR EN PSYCHOLOGIE DE L'UNIVERSITE DE LYON
Discipline : Sciences Cognitives mention Linguistique

Symphonie des oscillations cérébrales lors de la perception de parole: Études comportementale et en magnétoencéphalographie chez les enfants neurotypiques et dysphasiques.

Par Hélène Guiraud

Réalisée sous la direction des Docteurs Karim Jerbi et Véronique Boulenger

Soutenue publiquement le 15 décembre 2017

Devant le jury composé de :

Dr Mireille Besson (DR1)	LNC – CNRS, Aix-Marseille	Rapporteuse
Pr Séverine Samson (PR CE IUF)	PSITEC – Univ. Lille	Rapporteuse
Dr Olivier Bertrand (DR1)	DYCOG, CRNL – INSERM, Lyon	Examineur
Dr Benjamin Morillon (CR2)	INS – INSERM, Aix-Marseille	Examineur
Dr Karim Jerbi (CR1)	Dép. de Psychologie – Univ. Montréal	Directeur
Dr Véronique Boulenger (CR1)	DDL – CNRS, Lyon	Co-directrice

A mon papi,

Résumé

Les modèles actuels de perception de la parole suggèrent un couplage étroit entre les rythmes cérébraux, caractérisés par les oscillations neuronales, et le rythme de la parole, permettant de segmenter le flux verbal continu en unités linguistiques pertinentes pour la reconnaissance. En particulier, les modulations lentes d'amplitude de l'enveloppe temporelle de la parole, véhiculant l'information syllabique et prosodique, sont capables d'« entraîner » les oscillations corticales auditives dans la bande de fréquence thêta (4-7 Hz), échantillonnant le signal verbal en unités syllabiques. L'information temporelle qui caractérise la parole joue un rôle fondamental dans l'acquisition et le développement du langage ; un déficit de traitement des indices rythmiques de la parole a d'ailleurs été décrit dans les troubles développementaux du langage. L'objectif de ce travail de thèse était de mieux comprendre les processus neurocognitifs sous-tendant la perception du rythme de la parole naturelle chez l'enfant présentant un développement langagier typique ou atypique (dysphasie) dans trois études. Une première étude en magnétoencéphalographie (MEG) a permis de dévoiler la dynamique corticale oscillatoire chez des enfants francophones neurotypiques (8-13 ans) lors de l'écoute de phrases naturellement produites à un débit normal ou rapide. Nos résultats suggèrent l'existence de deux phénomènes d'« *entrainment* » des oscillations sur l'enveloppe temporelle de la parole à débit normal, l'un dans la bande thêta au sein des régions auditives droites, l'autre dans une bande centrée sur le débit syllabique moyen des stimuli dans les régions temporales antérieures gauches. Dans la condition de parole rapide, une synchronisation cortico-acoustique a été mise en évidence dans la bande thêta au sein des régions (pré)motrices gauches, reflétant le rôle de la voie dorsale d'intégration sensori-motrice dans les conditions d'écoute difficiles mais aussi dans le développement du langage oral. Les deux études suivantes ont été réalisées chez des enfants présentant une dysphasie expressive (8-13 ans) afin de tester l'hypothèse d'un trouble de traitement du rythme syllabique chez ces enfants, potentiellement sous-tendu par une dynamique corticale oscillatoire atypique. Dans une étude comportementale, nous avons évalué les capacités des enfants dysphasiques à décoder de la parole naturellement produite à débit normal ou rapide, ou accélérée artificiellement. Nous avons montré des performances réduites chez ces enfants, en regard d'enfants neurotypiques, pour traiter des phrases accélérées naturellement et artificiellement, suggérant un déficit d'extraction du rythme de la parole lorsque la fréquence des modulations de l'enveloppe temporelle augmente. Une étude en MEG, identique à celle

réalisée chez les enfants neurotypiques, nous a permis d'apporter de premiers éléments en faveur de cette interprétation en révélant un traitement cortical atypique de l'information syllabique dans la dysphasie, qui pourrait rendre compte des troubles phonologiques et morpho-syntaxiques souvent décrits dans ce trouble neuro-développemental. Une synchronisation réduite des oscillations thêta du cortex auditif a ainsi été mise en évidence chez les enfants dysphasiques par rapport à leurs pairs lors de la perception de parole à débit normal. L'absence d'alignement de l'activité oscillatoire des régions prémotrices sur l'enveloppe temporelle des phrases à débit rapide nous a en outre conduit à émettre l'hypothèse d'un dysfonctionnement de la voie dorsale chez ces enfants. Dans l'ensemble, ce travail de thèse fournit donc, pour la première fois à notre connaissance, des preuves expérimentales (i) de la synchronisation entre rythmes corticaux et rythme de la parole naturelle chez les enfants à développement langagier typique et (ii) d'une dynamique oscillatoire atypique lors de la perception de parole à débit normal et rapide chez les enfants dysphasiques.

Abstract

Current models of speech perception suggest a close correspondence between brain rhythms, characterized by neuronal oscillations, and speech rhythm, which would allow the brain to parse the incoming speech signal into relevant linguistic units for decoding. Slow amplitude modulations in speech temporal envelope, which convey syllabic and prosodic information, have been shown to entrain oscillatory activity of auditory cortex in the theta frequency band (4-7 Hz), sampling the acoustic signal into syllable-sized units. Temporal information in speech is a foundation for oral language acquisition and development; accordingly, deficits in processing speech rhythmic cues have been described in developmental language disorders. This thesis sought to throw light on the neurocognitive processes underlying the perception of natural speech in children with typical and atypical language development (Specific Language Impairment – SLI – or Developmental Language Disorder – DLD) in three experimental studies. In a first magnetoencephalography (MEG) study, we unraveled the oscillatory dynamics in a group of French-speaking typically-developing children aged 8 to 13 years old during listening to naturally-produced sentences either at a normal or fast rate. Our results suggested two types of entrainment of cortical oscillations on the temporal envelope of normal rate speech: the first one occurred in the theta band in right auditory cortex whereas the second one was found in a frequency band centered on the mean syllabic rate of our stimuli in left anterior temporal regions. As to the fast rate condition, we showed cortico-acoustic coupling in the theta band in left (pre)motor areas, reflecting the role of the sensorimotor dorsal pathway in challenging listening conditions as well as in language development. In two other studies, we tested the hypothesis of an impairment to process speech syllabic rhythm, potentially underpinned by atypical oscillatory cortical dynamics, in children with developmental language disorders mainly at the expressive level. In a behavioral study, we examined how French-speaking children with expressive DLD (8-13 years old) processed speech naturally produced at a normal or fast rate, or artificially accelerated. Our results showed poorer performance to decode fast sentences, either accelerated naturally or artificially, in these children as compared to their typically-developing peers, which suggests a deficit in extracting speech syllabic information with increased modulation frequency in the amplitude envelope. The last study, identical to the first one in MEG conducted in typically-developing children, provided the first piece of evidence in favor of this interpretation by showing atypical cortical processing of syllabic

information in children with DLD, which may account for the phonological and morpho-syntactic deficits frequently described in this developmental disorder. Reduced alignment of theta oscillatory activity in auditory cortex to normal rate speech has indeed been evidenced in children with DLD as compared to typically-developing children. Lack of synchronization of oscillations in left (pre)motor regions to amplitude envelope of fast rate sentences was also observed, which we interpreted as potential dysfunction of the dorsal stream in this population. To the best of our knowledge, the findings obtained in this thesis therefore provide first experimental evidence for (i) coupling between brain rhythms and rhythm of naturally produced speech in typically-developing children and (ii) atypical oscillatory cortical dynamics underlying normal and fast rate speech in children with developmental language disorders.

Financement et laboratoire de rattachement

Cette thèse a été financée par un contrat doctoral du LabEx ASLAN (ANR-10-LABEX-0081) de l'Université de Lyon dans le cadre du programme "Investissements d'Avenir" (ANR-11-IDEX-0007) du gouvernement français géré par l'Agence Nationale de la Recherche française. Elle a été réalisée au sein du laboratoire Dynamique Du Langage (DDL, CNRS/Université Lyon 2, UMR5596) à Lyon.

Les différentes études présentées ont été menées dans le cadre du projet ODYSSEE (*Oscillatory Dynamics of Speech Sensorimotor Ensembles*; PI: Véronique Boulenger; n° 11 JSH2 005 1) financé par l'Agence Nationale de la Recherche.

Au cours de ma thèse, une collaboration a par ailleurs été réalisée avec l'Université de Montréal et notamment le Centre de Recherche en Neuropsychologie et Cognition (CERNEC) et le laboratoire International de recherche sur le Cerveau, la Musique et le Son (BRAMS), afin de mener à bien les analyses en magnétoencéphalographie des études 1 et 3. Cette collaboration a donné lieu à un séjour doctoral In-doc de 6 mois (de décembre 2015 à mai-2016) financé par le programme *Erasmus Mundus, Student Exchange Network in Auditory Cognitive Neuroscience*.

Remerciements

Mes premiers remerciements vont à mes deux directeurs de thèse, Véronique et Karim.

Véronique, il y a 6 ans tu m'as fait découvrir les méandres du cerveau et du langage. Tu as su me guider tout au long de ces années, m'initier à la recherche en psychologie et me permettre de travailler avec les enfants ET la pathologie. Merci de m'avoir accompagnée dans tous les défis que l'on s'est lancé lors de l'élaboration de ce projet de thèse : la MEG, les enfants et les troubles du langage. Tu as toujours été extrêmement présente et bienveillante et je t'en remercie profondément.

Karim, tu as accepté de m'encadrer sur cette thèse et de me faire découvrir l'univers de la MEG et des oscillations corticales. Merci pour ta supervision, tes conseils et ta patience. Merci aussi de m'avoir accueillie au sein du « CocoLab » et permis de découvrir la douce vie Montréalaise, ses *party* et ses chercheurs. Ces 6 mois ont été pour moi très formateurs et une très belle expérience de vie. C'était vraiment pas si pire !

Merci à tous les deux d'avoir cru en moi !

Je remercie l'ensemble des membres de mon jury pour leur évaluation de ce travail de thèse. Merci au Docteur Mireille Besson et au Professeur Séverine Samson qui ont accepté de lire mon manuscrit, puis de venir des deux extrémités de la France pour ma soutenance. Je remercie également les Docteurs Olivier Bertrand et Benjamin Morillon d'avoir accepté d'examiner mon travail et de s'être également déplacé de plus ou moins loin.

Je tiens par ailleurs à remercier encore une fois le docteur Olivier Bertrand ainsi que le docteur Serge Pinto pour leurs conseils avisés et leur appréciation au cours de mes deux comités de suivi de thèse.

Un grand merci au laboratoire DDL, à sa directrice Sophie Kern et à tous ses membres pour m'avoir accueillie depuis ma première année de Master. Merci à l'équipe administrative, Linda et Rabia, pour votre accompagnement. Mes remerciements vont tout particulièrement à Nathalie et Manu pour leur aide précieuse au cours de l'élaboration du matériel expérimental et au cours de ces 4 années, ainsi qu'à l'équipe DENDY et tous ceux qui ont participé de près ou de loin à la réalisation de ce travail de thèse (Alice, Christian, Christophe C., Egidio,

Florence, François, Jen, Mélanie, Sébastien...). DDL est une belle expérience en tant que doctorant et j'ai été chanceuse d'en faire partie. Merci !

Enfin, mes années de thèse n'auraient pas été les mêmes sans mes acolytes de bureau, Agathe, Camille, Emilie, Leslie, Marie et Natacha. Merci pour tous ces moments partagés, les pauses thé, le chocolat, les verres après le boulot et encore le chocolat, votre soutien, votre aide et votre bonne humeur au cours de ces 4 années. C'était chat-rveilleux !

Je tiens également à remercier toutes les personnes extérieures à DDL qui ont contribué à ce travail : Ana-Sofia, merci pour ton aide précieuse avec Fieldtrip, ta disponibilité et ta patience lors de nos longues journées à essayer de comprendre les méandres de Matlab et Fieldtrip. Dimitri, merci pour ton apprentissage de l'analyse MEG et ta disponibilité lorsque j'étais bloquée dans mes analyses. Merci également à Hannu, pour la création d'une partie de mes scripts d'analyse, à Fanny pour ton aide si gentille avec la MEG et ses analyses, à Claude et Sébastien pour les acquisitions MEG au CERMEP, merci d'avoir été patients avec mes petits participants et merci pour vos conseils. Aussi, je tiens à remercier Damien Gouy d'avoir prêté sa voix pour nos stimuli, Lydie Merle pour son aide lors du recrutement de nos petits patients, Isabelle Peretz et le laboratoire Brams pour leur accueil à Montréal et enfin, les membres du CocoLab qui m'ont permis de passer un séjour tout aussi agréable que formateur à Montréal.

Des remerciements tout particuliers aux enfants, et à leurs parents, qui ont accepté de jouer mes petits cobayes et sans qui ce travail n'aurait jamais pu voir le jour. Je remercie aussi vivement les orthophonistes et les services hospitaliers de Lyon qui nous ont permis de recruter les enfants, et notamment le service de Neurologie Pédiatrique, le service Épilepsie, Sommeil et Explorations Fonctionnelles Neuropédiatriques et le Centre de Référence des Troubles des Apprentissages de l'Hôpital Femme Mère Enfant de Bron ainsi que le service d'audiologie et d'explorations otoneurologiques de l'Hôpital Edouard Herriot de Lyon. Merci aux Docteurs Sonia Krifi-Papoz, Vania Herbillon, Aurélia Caillot-Bascoul et Sibylle Gonzalez-Monge pour votre engagement dans cette étude et votre amabilité. Un grand merci aux différentes associations qui ont cru en notre projet et diffusé notre requête de patients (Avenir Dysphasie, Dyspraxie France Dys et la Fédération Française des Dys), ainsi qu'aux écoles de Lyon qui ont accepté de me recevoir dans leurs classes pour « tester » leurs élèves.

Enfin, merci mille fois à tous ceux qui ont « subi » ces 4 années de thèse à mes côtés. Un grand merci à mes parents, qui m'ont soutenue et encouragée au cours de mes longues années de fac et ont participé à la relecture de ce manuscrit, à mes sœurs qui ont joué les cobayes plus d'une fois et à l'ensemble de ma famille et belle famille qui se demandaient si un jour j'entrerais dans le monde des adultes... Merci à tous les amis qui ont accepté de prendre part à mes « expériences » et rendent ma vie si belle, et tout particulièrement mes copines du Sud et de Lyon, Agathe, Leslie, Liu-Marie, Marie, Marine, Marion et Julie, qui ont toujours été là pour moi. Enfin un merci tout spécial à celui qui partage ma vie, Simon, et qui m'a supportée, dans ce travail et dans mes sautes d'humeur, tout au long de cette thèse. Merci d'être toi et de combler ma vie de bonheur.

Merci à tous ceux qui m'ont accompagnée dans cette folle aventure qu'est la thèse.

« Il faut un petit grain de folie pour donner des couleurs à la vie »

Justin Hurwitz

Table des matières

Préambule	18
INTRODUCTION GENERALE	22
I. La perception de la parole	23
1. Les unités perceptives de la parole	24
2. Le signal de parole : un signal quasi-rythmique	28
3. L'enveloppe temporelle du signal de parole	34
4. Variations du signal de parole	38
4.1. La parole rapide	40
4.2. Adaptation à la parole rapide	42
II. La perception de la parole dans le cerveau.....	46
1. Théories auditives, motrices et sensori-motrices de la parole	46
2. Bases cérébrales de la perception de la parole.....	50
3. Implication du système moteur dans la perception de la parole	54
4. Modèles neurobiologiques de la perception de la parole	63
III. Les rythmes cérébraux.....	69
1. Les oscillations cérébrales	69
2. La synchronisation oscillatoire	71
3. Oscillations cérébrales et attention dynamique	76
4. Rythmes cérébraux et prédiction de la parole	80
4.1. Modèles oscillatoires de la parole	80
4.2. Synchronisation oscillatoire sur le signal de parole.....	85
IV. Les rythmes chez l'enfant	99
1. Rythme de la parole et développement du langage	99
2. Dynamique cérébrale oscillatoire chez l'enfant.....	102
2.1. Dynamique cérébrale oscillatoire chez l'enfant au repos	102
2.2. Dynamique cérébrale oscillatoire chez l'enfant lors d'une stimulation auditive	105
V. La dysphasie.....	111
1. Définition	111
2. Classification des dysphasies.....	114
3. Etiologie des dysphasies	118
3.1. Les facteurs neurobiologiques	118
3.2. Les facteurs génétiques.....	123

3.3.	Les autres facteurs/causes.....	124
4.	Rythmes et dysphasie.....	126
4.1.	Troubles de discrimination des indices rythmiques de la parole	126
4.2.	Oscillations corticales et dysphasie	134
VI.	Objectifs et problématiques de la thèse	140
CONTRIBUTIONS EXPERIMENTALES		146
VII.	Dynamique oscillatoire lors de la perception de la parole chez les enfants neurotypiques.....	147
1.	Présentation de l'étude 1	148
2.	Synthèse des résultats de l'étude 1.....	180
VIII.	Perception de la parole rapide chez des enfants présentant une dysphasie expressive : études comportementale et en MEG.....	183
1.	Présentation de l'étude 2 : Étude comportementale	184
2.	Synthèse des résultats de l'étude 2.....	217
3.	Présentation de l'étude 3 : Étude en MEG	218
4.	Synthèse des résultats de l'étude 3.....	260
DISCUSSION GÉNÉRALE		262
IX.	Rythmes corticaux et perception de la parole chez les enfants neurotypiques.....	264
1.	Dynamique oscillatoire de la perception de la parole à débit normal.....	264
2.	Dynamique oscillatoire lors de la perception de la parole à débit rapide	272
3.	Perspectives chez les enfants neurotypiques	277
X.	Rythme et perception de la parole dans la dysphasie	283
1.	Capacités de perception de la parole rapide des enfants dysphasiques : étude comportementale.....	283
2.	Dynamique oscillatoire sous-tendant le traitement de la parole chez l'enfant dysphasique	288
2.1.	Le cas de la parole à débit normal	289
2.2.	Le cas de la parole à débit rapide	292
3.	Perspectives dans la dysphasie	295
3.1.	Analyses et études complémentaires	295
3.2.	Rôle du rythme dans la musique et le mouvement.....	298
XI.	Conclusion	304
BIBLIOGRAPHIE		306

ANNEXE 1. Oscillations et perception de la parole chez les adultes : étude en MEG	338
ANNEXE 2. Modulations de puissance dans les bandes beta et gamma basse chez les enfants neurotypiques	339
ANNEXE 3. Analyses de la fréquence fondamentale chez les enfants neurotypiques	340
ANNEXE 4. Modulations de puissance dans les bandes beta et gamma basse chez les enfants dysphasiques	342

Préambule

Une des grandes habiletés de l'Homme est d'interagir avec ses pairs grâce un outil unique et fascinant, le langage. Le langage peut se présenter sous différentes formes - gestuelle, graphique, etc - mais celle qui nous intéresse tout particulièrement ici est sa forme verbale, la parole. La parole est un signal extrêmement riche et complexe produit par l'appareil bucco-phonatoire d'un individu dans le but de communiquer sa pensée. La complexité de la parole tient notamment au fait qu'il s'agit d'un flux acoustique continu dont les frontières entre les mots et les phrases ne sont pas marquées physiquement (contrairement à l'écrit où les frontières sont facilement repérables). **Une des grandes questions qui anime la communauté (neuro)linguistique est alors de comprendre comment l'auditeur est capable de décoder le continuum de parole.** Un grand nombre de travaux de recherche ont ainsi vu le jour au début du XX^{ème} siècle dans le but d'identifier une unité perceptive de la parole, qui constituerait un élément de base pour la segmentation et le décodage de l'information verbale. Dans les années 70, la syllabe fut proposée comme candidat idéal pour assurer ce rôle (Massaro, 1974) ; depuis, les études menées chez l'enfant et l'adulte ont largement confirmé son caractère fondamental dans l'organisation temporelle de la parole. La syllabe véhicule notamment l'information rythmique de la parole : elle reflète la vitesse à laquelle un locuteur s'exprime (i.e. son débit de parole) et porte les différentes marques rythmiques telles que l'accentuation, les allongements de durée et l'intonation (Di Cristo, 2013) qui constituent, dès le plus jeune âge, des indices précieux pour l'auditeur. L'ensemble de ces fonctions confèrent à la syllabe un rôle particulier dans le décodage de la parole et notamment dans l'étape primaire de traitement, la segmentation du signal verbal continu en segments pouvant être facilement traités par le système perceptif (Cutler, 1992; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Segui, Dupoux, & Mehler, 1991).

Une nouvelle question se pose dès lors : **comment le système perceptif segmente-t-il le signal de parole ?** Si l'unité de base pour la segmentation est la syllabe, comment le cerveau repère-t-il chaque syllabe au sein de la chaîne parlée ? Le signal de parole est un signal hautement rythmique. Les segments de base de la parole apparaissent à intervalles quasi-réguliers au cours du temps, et l'auditeur va exploiter cette (quasi-)régularité temporelle pour découper le signal en unités linguistiques pertinentes pour la reconnaissance. Les syllabes s'enchaînent par exemple à une fréquence située entre 4 et 8 Hz en fonction des langues, c'est-à-dire qu'un locuteur va produire 4 à 8 syllabes par seconde dans ses énoncés

(Dellwo & Wagner, 2003; Pellegrino, Coupé, & Marisco, 2011). On sait depuis plus d'un siècle maintenant que notre système cérébral est lui aussi animé d'une certaine rythmicité. Les neurones génèrent en effet des potentiels électriques de façon périodique au repos (on parle d'activité oscillatoire intrinsèque ou endogène) et cette activité est amplifiée lorsqu'ils sont engagés dans une tâche cognitive particulière. Différents rythmes cérébraux ont été définis en fonction de la fréquence des oscillations corticales (delta, thêta, alpha, beta, gamma) chacun étant impliqué dans des fonctions cognitives particulières en fonction de la localisation anatomique (Buzsáki, 2006). La parole et le cerveau sont donc tous deux gouvernés par une organisation rythmique, doit-on alors y voir uniquement le fruit du hasard ou peut-on y trouver un lien particulier ? Les modèles neurolinguistiques récents ayant inspiré notre travail de recherche suggèrent l'existence d'une correspondance étroite entre les échelles temporelles de la parole et les rythmes corticaux, en particulier dans les régions auditives (Ghitza, 2011; Giraud & Poeppel, 2012; Poeppel, 2003), ce qui permettrait au cerveau de « résonner au rythme de la parole ». Le cerveau pourrait en effet synchroniser plus facilement certains de ses propres rythmes sur les rythmes phrastique, syllabique et phonémique au sein du signal verbal, facilitant par conséquent la segmentation du flux continu et donc le traitement du message linguistique. Comme nous le verrons, cette hypothèse fait l'objet d'un grand nombre d'études depuis quelques années, principalement chez l'adulte et en langue anglaise, et les données disponibles tendent à corroborer les prédictions des modèles. **Mais qu'en est-il de l'alignement entre les rythmes du cerveau et celui de la parole chez l'enfant ?**

Percevoir et traiter la parole, qui sont des processus qui nous semblent par ailleurs tout à fait triviaux, nécessitent ainsi une cascade de traitements très sophistiqués dans le cerveau. C'est en outre sans compter sur le fait que dans la vie quotidienne, le signal de parole n'est que très rarement perçu dans des conditions optimales et dans sa forme canonique. Nous sommes en permanence exposés à des bruits ambiants ou à des interlocuteurs différant de par leur âge, leur origine, leur manière de s'exprimer, leur débit etc., conduisant à de fortes variations notamment du rythme et de l'organisation du signal de parole. Toutefois, notre cerveau traite les différentes sources de variabilité avec une certaine aisance, sans que nous en ayons conscience. **Quels sont alors les secrets de notre cerveau pour faire face à la variabilité du signal de parole ?**

La variabilité existe également dans la manière dont le système langagier se développe chez l'individu. Malgré les qualités remarquables de notre système perceptif, nous ne sommes pas tous égaux face au signal complexe qu'est la parole. Certains enfants peuvent

ainsi présenter un développement langagier atypique et éprouver des difficultés particulières pour produire et/ou comprendre la parole. Ces difficultés sont durables et ont des répercussions fortes sur le développement de la personnalité de l'enfant, sa réussite scolaire, son intégration sociale mais aussi sur sa future insertion professionnelle. On estime que 4 à 5 % des enfants d'une tranche d'âge sont concernés par des troubles neurodéveloppementaux du langage. Dans ce travail, nous nous sommes intéressés à l'un de ces troubles, la dysphasie, qui est définie comme un trouble du langage oral affectant le versant expressif et/ou réceptif (Gérard, 1993; Christelle Maillart & Schelstraete, 2012). Deux grandes questions ont motivé notre recherche : **Comment le cerveau d'un enfant présentant un développement atypique du langage, et notamment une dysphasie, traite-t-il la parole et sa variabilité ? Comment suit-il le rythme de la parole ?**

S'il est illusoire de prétendre pouvoir répondre à l'ensemble de ces questions, l'objectif de ce travail de thèse est d'y apporter des éléments de réponse expérimentaux en examinant la réponse oscillatoire cérébrale, chez l'enfant présentant un développement typique et l'enfant présentant une dysphasie, à des variations rythmiques du signal de parole, plus précisément lors de la perception de parole produite à des débits différents.

La première partie de cette thèse, organisée en 6 chapitres, consiste en une Introduction théorique basée sur une revue de la littérature. Dans un premier chapitre sur la perception de la parole, nous présentons les caractéristiques du signal de parole, en nous attardant plus particulièrement sur les informations rythmiques et les variations de débit. Le deuxième chapitre porte sur les bases cérébrales de la perception de la parole et les théories et modèles neurobiologiques actuels. Nous évoquons ensuite la dynamique des rythmes cérébraux et son rôle dans le traitement de la parole chez l'adulte (chapitre III) et l'enfant présentant un développement typique (chapitre IV). Le chapitre V est dévolu à une présentation de la dysphasie et aux travaux portant sur le traitement du rythme dans ce trouble neurodéveloppemental. Enfin le chapitre VI expose la problématique de notre travail ainsi que nos principales hypothèses.

La deuxième partie de cette thèse, organisée en deux chapitres, présente nos Contributions expérimentales sous formes d'articles scientifiques. Le chapitre VII est axé sur la dynamique oscillatoire mesurée en magnétoencéphalographie (MEG) lors de la perception de la parole produite à des débits différents chez l'enfant à développement langagier typique. Cette étude fait l'objet d'un article soumis dans la revue *NeuroImage*. Dans le chapitre VIII,

nous présentons les résultats de deux études visant à examiner d'une part, les capacités de perception de la parole à débit normal et rapide d'enfants dysphasiques (étude comportementale faisant l'objet d'un article actuellement en révision dans la revue *PloS One*) et d'autre part, la dynamique oscillatoire sous-jacente (étude en MEG faisant l'objet d'un article en préparation).

Enfin, cette thèse se termine par une troisième partie consistant en une Discussion générale des différents résultats obtenus (chapitre IX pour les enfants neurotypiques et chapitre X pour les enfants dysphasiques), interprétés à la lumière de la littérature existante dans les domaines de la perception de la parole, du traitement de son rythme et des oscillations corticales. Des perspectives sont également évoquées, notamment concernant des analyses complémentaires nécessaires pour étayer nos interprétations, mais aussi concernant les liens entre rythmes verbaux et musicaux qui pourraient être exploités dans le cas de la dysphasie en particulier.

/

INTRODUCTION GENERALE

I. La perception de la parole

Le signal de parole est l'un des signaux les plus complexes que notre cerveau ait à traiter, correspondant à une onde sonore continue dont les fréquences varient entre 1 et 20 000 Hz. « C'est un signal acoustique continu, composé de sons successifs s'enchaînant les uns aux autres sans qu'aucune frontière physique explicite n'indique le début ou la fin des phonèmes, des syllabes et des mots composant la suite sonore. » (Billières, 2015).

Le codage de la parole est réalisé selon deux dimensions, la dimension spectrale ou fréquentielle et la dimension temporelle. L'information spectrale correspond à la distribution de l'énergie du signal sonore dans l'échelle de fréquence (spectre sonore), et permet de définir les différents formants¹ des unités de la parole (notamment les voyelles que nous définirons dans le chapitre I.1.) et leurs transitions (appelées transitions formantiques caractérisant l'articulation entre deux unités de parole ; Stevens & Klatt, 1974). L'information temporelle correspond quant à elle à l'organisation précise des différents éléments de la parole au cours du temps. Elle se rapporte aux modulations de l'amplitude ou de la fréquence du signal auditif sur une échelle de temps (Dehaene & Petit, 2009). Ces deux types d'information jouent un rôle essentiel dans le traitement de la parole, que ce soit au niveau du traitement auditif périphérique que cérébral. Dans ce travail de thèse, nous nous sommes particulièrement intéressés à l'information temporelle et à son rôle dans la perception de la parole. Notre objectif est de mieux comprendre comment le cerveau de l'auditeur va exploiter les différents indices rythmiques de la parole, c'est-à-dire les informations temporelles et prosodiques afin de découper le signal continu de parole en unités linguistiques et ainsi reconnaître la chaîne parlée.

Dans le premier chapitre de l'Introduction de cette thèse, nous définirons tout d'abord les principales unités permettant de traiter le signal de parole, les phonèmes et les syllabes, puis nous introduirons la notion de rythme en nous focalisant sur le rythme de la parole.

¹ Un formant est un maximum d'énergie du spectre sonore et correspond à la fréquence de résonance déterminée par la forme du tractus vocal. Il s'agit d'une fréquence renforcée dans le signal. Les formants donnent le timbre des voyelles et participent à celui des consonnes.

1. Les unités perceptives de la parole

Le signal de parole est produit par les différents muscles articulatoires du système bucco-phonatoire d'un locuteur (cordes vocales, langue, palais, lèvres...). C'est le recrutement préférentiel de tel ou tel muscle qui va permettre de former les différentes unités de base de la parole, les phonèmes et les syllabes.

Les phonèmes correspondent aux plus petites unités (segments) sonores de la chaîne parlée ; leur durée est comprise entre 60 et 150 ms, la majorité se situant autour de 50-80 ms (Ghitza & Greenberg, 2009; Greenberg, Hollenback, & Ellis, 1996; Rosen, 1992). En français, les phonèmes se composent de voyelles et de consonnes. Les voyelles et les consonnes se distinguent par le passage d'air ou non au sein de la cavité bucco-nasale. La production des voyelles résulte d'une libre circulation de l'air au niveau du tractus vocal, alors que la production des consonnes suppose un rétrécissement du tractus vocal en un point entraînant une interruption ou une diminution du passage de l'air.

Les syllabes sont les unités phoniques immédiatement supérieures aux phonèmes ; leur durée moyenne est comprise entre 150 et 300 ms, avec une moyenne autour de 200 ms (Greenberg & Arai, 2004; Greenberg, 1999; Greenberg et al., 1996; Rosen, 1992). Elles présentent un modèle structural hiérarchisé, organisé autour d'un noyau vocalique obligatoire auquel peut s'associer (de manière facultative) une attaque (une ou plusieurs consonnes initiales) et/ou une coda (une ou plusieurs consonnes suivant la voyelle). Le noyau vocalique et la coda forment la rime. Les voyelles occupent donc une place privilégiée dans la syllabe : le noyau vocalique constitue en effet l'élément central de la syllabe et est le seul élément indispensable pour sa formation.

La syllabe est avant tout une unité motrice correspondant à un pic d'expiration d'air phonatoire au niveau du conduit vocal, suite à la contraction des muscles respiratoires (Stetson, 1951)². Des propriétés dynamiques supplémentaires semblent également entrer en jeu dans la production des syllabes et notamment l'appareil articulatoire qui va permettre de former les différentes structures syllabiques possibles (consonne (C pour toute consonne) suivie d'une voyelle (V pour toute voyelle), CVC, VCC, etc., Meynadier, 2001; Tuller & Kelso, 1990). Au niveau phonétique, la syllabe est un groupement phonique saillant autour d'un sommet de sonorité (i.e. liée à l'impression auditive de prééminence d'un son par rapport

² Depuis cette définition de Stetson en 1951, des travaux ont montré que plusieurs syllabes pouvaient être produites en une seule impulsion expiratoire.

à d'autres ; Ladefoged, 1971) constitué par le noyau vocalique. Autrement dit, la syllabe est caractérisée par un pic de sonorité (la voyelle) autour duquel s'organisent les segments moins sonores. Toute rupture de ce patron de sonorité implique théoriquement une frontière syllabique (e.g., « tactique » [tak.tik]). La syllabe constitue une unité essentielle dans l'organisation de la parole et de la langue, aussi bien au regard de l'ensemble des langues, de l'acquisition, des pathologies, des erreurs de langage que du traitement psycholinguistique (Dolata, Davis, & MacNeilage, 2008; Meynadier, 2001). La syllabe, et notamment sa structure simple CV, est universelle et dominante dans 80% des langues (Dolata, Davis, & MacNeilage, 2008). Elle est de plus la première structure produite par l'enfant et la dernière qui résiste dans le cas d'une aphasie (Jakobson, 1969). On la retrouve également dans les erreurs de langage qui sont produites sur la base d'une structure syllabique. Meynadier (2001) cite ainsi diverses études en psycholinguistique révélant que les substitutions de phonèmes entre mots sont toujours réalisées sur des phonèmes à la même position syllabique (par exemple, dans *obera pastille*, la substitution se produit au niveau d'un phonème de première syllabe de chaque mot) et que les insertions de phonèmes génèrent toujours une syllabe admise dans la langue (par exemple, *électroc*nique, insertion de la syllabe *troc*). La syllabe joue également un rôle essentiel dans le traitement perceptif et phonologique de la parole, facilitant le décodage du signal (Massaro, 1974; Segui, Dupoux, & Mehler, 1991). Nous allons maintenant exposer les arguments en faveur d'un tel rôle de la syllabe, notamment dans la segmentation du signal continu de parole.

Les phonèmes composent donc les syllabes qui vont former des mots qui à leur tour formeront des phrases. L'ensemble de ces unités sont produites de façon continue par le locuteur et afin d'accéder à la signification de l'énoncé, l'auditeur va devoir segmenter le continuum de parole pour en isoler les différentes unités et ainsi établir une correspondance avec les représentations lexicales en mémoire. Plusieurs théories ont cherché à définir une unité perceptive de base permettant de décomposer et de traiter le signal de parole (Grossberg, 2003; Lahiri & Reetz, 2002; Marslen-Wilson & Warren, 1994; Massaro, 1974). Un grand nombre d'études considèrent la syllabe comme un bon candidat. Massaro (1974) est le premier à reconnaître la syllabe comme unité fondamentale de la reconnaissance de la parole. Le découpage syllabique permettrait au système perceptif de traiter des unités de parole à une vitesse compatible avec les capacités de traitement de l'information. Ce serait moins le cas pour les phonèmes qui, produits à un rythme plus rapide d'environ 10 à 20 par seconde,

imposent un défi plus difficile au système perceptif. Mehler et collaborateurs (1981) ont montré, dans une étude de perception de mots bi-syllabiques en français chez des adultes, que les séquences de phonèmes étaient plus faciles à reconnaître lorsqu'elles correspondaient à la première syllabe du mot que lorsqu'elles n'y correspondaient pas. Par exemple, la séquence « ba » était détectée plus rapidement dans le mot « balance » que dans « balcon » alors que la séquence « bal » était reconnue plus rapidement dans « balcon » que dans « balance » (voir aussi Segui et al., 1981). Ainsi, si la séquence phonémique cible ne correspond pas à la structure syllabique du stimulus entendu, les sujets doivent re-segmenter la séquence syllabique, allongeant alors le temps de traitement. Par ailleurs, les travaux de Segui et collaborateurs (1991) chez des adultes révèlent une primauté du traitement des syllabes sur celui des phonèmes. En effet, dans une tâche de détection de phonèmes, les auteurs montrent que les sujets francophones détectent plus rapidement les phonèmes présents au début d'une syllabe courte (CV) qu'au début d'une syllabe longue (CVC ou CCV).

L'ensemble de ces travaux confirme l'hypothèse de l'existence d'une unité naturelle de segmentation de la parole définie par la syllabe (Ridouane, Meynadier, & Fougeron, 2009) et impliquée dans le traitement phonologique (Ferrand, Segui, & Grainger, 1996). Tout locuteur posséderait la capacité de dénombrer intuitivement les syllabes d'un mot : ainsi, les enfants et les illettrés manipulent parfaitement les syllabes dans des séquences verbales, alors qu'il leur est plus difficile de manipuler les phonèmes (Bertoncini & Mehler, 1981; Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988; Bijeljac-Babic, Bertoncini, & Mehler, 1993; Eimas, 1999; Moraïs, Bertelson, Cary, & Alegria, 1986). Bijeljac-Babic et collaborateurs (1993) ont montré, grâce à la mesure de succion non nutritive (ou succion de haute amplitude), que des nouveau-nés français de 4 jours étaient capables de différencier des énoncés plurisyllabiques sur la base de leur nombre de syllabes, mais pas sur leur nombre de phonèmes. Bertoncini et Mehler (Bertoncini & Mehler, 1981) ont également mis en évidence que les très jeunes enfants âgés de moins de 2 mois discriminaient mieux des stimuli respectant la structure syllabique CVC du français (*e.g.*, « tap ») que des stimuli non syllabiques de type CCC enchaînant trois consonnes (*e.g.*, « pst »). Le fait d'associer les consonnes initiale et finale de la séquence « illégale » CCC à une voyelle, aboutissant à une structure de type VCCCV (*e.g.*, « upstu »), conduisait en revanche à de meilleurs taux de discrimination. Ainsi, l'enfant dès son plus jeune âge segmenterait la chaîne de parole en unités syllabiques ; ces segments constitueraient la base du développement de son système phonologique.

Cependant, ces résultats ne semblent pas pouvoir s'appliquer dans toutes les langues, suggérant une spécificité linguistique de l'effet syllabique (Cutler, Butterfield, & Williams, 1987; Cutler, Norris, & Segui, 1986). En effet, Cutler et collaborateurs (1986) ont réalisé une étude similaire à celle menée par Mehler et al. (1981) chez des locuteurs adultes anglophones, et n'ont pas observé l'effet de syllabification observé en français, suggérant une stratégie de traitement de la parole différente dans les deux langues.

Enfin, il faut noter que certains auteurs suggèrent que l'unité perceptive de base est le trait et non la syllabe (Lahiri & Reetz, 2002; Marslen-Wilson & Warren, 1994; Stevens, 2002; Streeter & Nigro, 1979). D'après ces auteurs, l'auditeur associerait la représentation des traits distinctifs du signal de parole, permettant de distinguer deux phonèmes (le voisement, la nasalité, l'antériorité, l'arrondissement et les différents lieux d'articulation), au lexique afin de décoder le message linguistique. Le traitement du signal de parole serait ainsi direct, sans passer par les représentations phonémiques et syllabiques.

A l'heure actuelle, aucun consensus ne semble encore se dégager quant à la nature de l'unité fondamentale de reconnaissance de la parole. Certains auteurs suggèrent ainsi l'existence de différentes unités perceptives dans le traitement de la parole dont l'implication dépend des facteurs phonologiques, lexicaux, grammaticaux et du contexte conversationnel dans lequel se trouve l'auditeur (Nguyen, 2005).

Le signal de parole correspond à une suite sonore continue que l'auditeur doit donc segmenter afin d'en extraire l'information linguistique. Pour cela, il va s'appuyer sur différentes unités perceptives, les phonèmes et les syllabes, permettant de découper le signal en segments facilement associables aux représentations lexicales stockées en mémoire. Ce niveau correspond au niveau segmental dans lequel la syllabe semble jouer un rôle primordial. Toutefois, afin de segmenter et de traiter le signal de parole, l'auditeur va aussi se baser sur un autre type d'informations présentes au niveau dit suprasegmental de la langue, les informations prosodiques. Ces informations correspondent aux « variations mélodiques et rythmiques indépendantes des propriétés phonologiques des segments » (Meynadier, 2001). La parole est en effet un signal qui présente une certaine rythmicité, puisqu'elle est constituée d'indices présentés à des intervalles quasi-réguliers. Cette quasi-rythmicité est notamment portée par la syllabe, qui va présenter les différentes marques prosodiques du signal de parole (allongement, accentuation, etc.). Ainsi, la syllabe se retrouve à tous les niveaux de la langue,

jouant un rôle d'unité perceptive/segmentale et prosodique. Dans la suite, nous nous intéresserons au niveau suprasegmental de la langue, et notamment à sa quasi-rythmicité.

2. Le signal de parole : un signal quasi-rythmique

Le rythme est omniprésent dans le comportement humain, non seulement au niveau linguistique mais aussi biologique, moteur et artistique. Selon Handel (1989), « l'organisation rythmique est inhérente à toute activité humaine. Le rythme crée, structure et exprime la qualité temporelle ». Il est ainsi défini par Malmberg (1967) comme « la perception d'évènements qui se répètent périodiquement selon un schéma ou un modèle ». Au niveau du signal de parole, le rythme est caractérisé comme « l'organisation temporelle des prééminences qui participent à la structuration prosodique des énoncés » (Astesano, 2001). Il faciliterait le traitement de la parole et notamment la segmentation lexicale et la compréhension (Cutler, 1992). La composante acoustique du rythme dans la parole correspond, d'après Valérie Padeloup (Padeloup, 2004), non seulement à l'organisation temporelle du signal de parole mais aussi aux éléments acoustiques (hauteur mélodique, intensité, timbre) pouvant participer seuls ou avec d'autres éléments, à la structuration rythmique de ce signal. Le rythme constitue ainsi un fondement de l'organisation prosodique de la parole.

Définissons tout d'abord la prosodie. **La prosodie** est couramment définie comme « le champ d'étude d'un ensemble de phénomènes, tels que l'accent, le rythme, les tons, l'intonation, la quantité, les pauses et le tempo, qui constituent ce qu'il est convenu d'appeler les éléments prosodiques ou les éléments suprasegmentaux du langage » (Di Cristo, 2013). Selon D'Imperio, Dittinger et Besson (2016), elle peut être considérée comme « un système phonologique complexe reposant sur trois sous-systèmes : l'intonation (modulations mélodiques de l'énoncé à caractère pragmatique), l'accentuation (prééminence au niveau du mot ou de la phrase) et le phrasé (structuration en constituants de nature prosodique dont les frontières sont signalées par une rupture rythmique et/ou intonative) ». Autrement dit, la prosodie désigne les modulations mélodiques et rythmiques produites dans un énoncé qui servent à lui donner un sens pragmatique et à marquer les frontières entre les constituants syntaxiques. Par exemple, elle permettra de déterminer si la phrase « La petite brise la glace » correspond à « La petite | brise la glace » ou à « La petite brise | la glace » (où | désigne la

rupture prosodique). Les informations prosodiques véhiculent des informations à la fois linguistiques (*e.g.*, syntaxiques), extralinguistiques (*i.e.* individu : fonction identificatrice) et paralinguistiques (*i.e.* attitudes et sentiments du locuteur) qui sont déterminantes pour l'interprétation des énoncés et de la conversation. La prosodie est en outre l'un des premiers aspects de la parole acquis par l'enfant, lui permettant d'accéder aux représentations lexicales (Jusczyk, 1995; Jusczyk, Cutler, & Redanz, 1993; Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998; Nazzi & Ramus, 2003; Port, 2003) et aux structures syntaxiques (Goswami, 2011; Nazzi & Ramus, 2003; Richards & Goswami, 2015). Nous reviendrons sur ce point dans le chapitre IV de cette Introduction.

La prosodie se caractérise principalement par trois paramètres acoustiques : l'intensité (percept de sonie), la durée des unités (percept de rythme) et la fréquence fondamentale (F0) (définie ci-après), dont les changements sont portés par le noyau vocalique de la syllabe. La F0 est le corrélat acoustique de la hauteur ou *pitch* d'un son, caractérisé par la fréquence de vibration de base des cordes vocales. Elle varie selon le sexe du locuteur, en raison des différences anatomiques et articulatoires chez les hommes et les femmes. Ainsi, chez les locuteurs masculins, la F0 moyenne est de 120 Hz en moyenne, alors qu'elle est d'environ 220 Hz chez les locuteurs féminins. La F0 est par ailleurs plus élevée les enfants (~320 Hz). Les cordes vocales vibrent de façon périodique, créant plusieurs sons qui se superposent pour former le signal sortant. La composante de base correspond à la F0 à laquelle se superpose un certain nombre d'ondes dont les fréquences sont des multiples de la F0. Le spectre de ces harmoniques forme ainsi les différents phonèmes de la parole.

Il n'existe pas de correspondance biunivoque entre un phénomène prosodique, l'accentuation par exemple, et un paramètre acoustique. Au contraire, un phénomène prosodique implique le plus souvent des changements sur plus d'un des trois paramètres, et la combinaison de ces indices peut varier en fonction du contexte mais aussi de la langue (voir le récent chapitre de D'Imperio et al., 2016).

L'**accentuation** correspond à l'augmentation de la durée, de l'intensité vocale (amplitude) et de la hauteur mélodique (F0) de certaines syllabes d'un énoncé oral ; c'est la mise en relief d'une syllabe par rapport à d'autres. Elle permet d'amplifier une unité de la chaîne linguistique et ainsi de découper la parole en groupes rythmiques, qui constituent des unités intonatives structurant l'énoncé et aidant ainsi l'auditeur à lui donner du sens (*e.g.*, « Le petit garçon » dans « Le petit garçon mange des céréales le matin »). On distingue des langues

à accent fixe (*e.g.*, français, finnois) où l'accent est toujours porté par la même syllabe au sein d'un mot (la dernière syllabe pour le français, la première pour le finnois) et les langues à accent libre (*e.g.*, anglais, espagnol), où la place de l'accent est déterminée par des facteurs morphologiques et où l'accent possède une valeur distinctive au sens où il permet d'accéder à la signification (*e.g.*, « *'import* » (importation) vs « *im'port* » (importer) en anglais). La structure syllabique joue un rôle majeur dans l'assignation de la place de l'accent dans les langues à accent libre. Ce sont en effet les syllabes « lourdes » (composées au moins d'une voyelle longue, d'une diphtongue³ ou d'une voyelle brève suivie d'une ou plusieurs consonnes tautosyllabiques⁴; *e.g.*, de type VCC ou CVC) qui attirent l'accent en comparaison des syllabes « légères » (*e.g.*, de type CCV, Meynadier, 2001).

L'**intonation** correspond quant à elle à la mélodie de la chaîne parlée, elle renvoie aux variations de la hauteur de la voix (i.e. modulations de F0) dans des unités supra-lexicales (phrases, énoncés) et leur apporte une valeur pragmatique (voir D'Imperio et al., 2016). L'intonation permet d'illustrer différents contenus sémantiques de la parole produite (Di Cristo, 2013) et de contraster, par exemple, une question (*e.g.*, « Tu as déjà mangé ? ») et une assertion (*e.g.*, « Tu as déjà mangé »).

Comme mentionné précédemment, la structuration prosodique de la parole repose également sur l'organisation temporelle, rythmique du signal. Là encore, la syllabe constitue l'unité de référence de cette organisation. Le rythme possède deux caractéristiques fondamentales qui sont la structure et la périodicité. La structure correspond à l'organisation interne d'une série successive d'éléments ou d'évènements (Fraisse, 1956). La périodicité quant à elle, dans sa définition la plus large, désigne un état, une qualité ou un fait récurrent, régulier, c'est-à-dire un pattern ou une structure répétée dans le temps et l'espace. « *What goes up must come down. [...] Without periodicity, there is no time; without time, there is no past, present, or future.* » (Buzsáki, 2006).

En musique, le rythme peut se définir comme une « alternance plus ou moins régulière de temps forts et de temps faibles, les temps forts étant assimilables à des syllabes accentuées, et les temps faibles à des syllabes inaccentuées » (Di Cristo, 2013). De cette alternance entre

³ Voyelle complexe qui change de timbre en cours d'émission (*e.g.*, *eight* en anglais contient la diphtongue [ei]).

⁴ Appartenant à la même syllabe.

temps forts et faibles au cours du temps émergent des motifs rythmiques qui vont créer une structure quasi-régulière.

Il est aujourd'hui bien connu que la distribution des syllabes au sein des énoncés diffère entre les langues (Abercrombie, 1967; Pike, 1945; Ramus, Nespor, & Mehler, 1999). On distingue classiquement trois types d'organisation du rythme dans les langues. Dans les langues dites à isochronie syllabique (*syllable-timed language*) telles que le français et l'italien, le rythme est régi par les syllabes dont la durée présente une grande régularité. Comme nous l'avons mentionné précédemment, la segmentation se fait alors en fonction du positionnement des syllabes. Les langues syllabiques se distinguent des langues dites à isochronie accentuelle (*stress-timed language*) comme l'anglais et l'allemand, où c'est l'accent tonique qui porte le rythme : il n'existe pas de régularité métronomique mais la durée des intervalles compris entre les syllabes accentuées est sensiblement la même (autrement dit, les syllabes accentuées sont produites à intervalles réguliers). Dans ces langues, la segmentation est réalisée sur la base de l'accent lexical. Un troisième type de classe rythmique, les langues moraiques ou moriques telles que le japonais et le hongrois, présente un rythme comparable à celui des langues syllabiques mais l'unité rythmique de base n'est pas la syllabe mais le more, qui est une unité de taille inférieure à celle de la syllabe (Ladefoged, 1975). La différence entre les langues accentuelles et syllabiques s'observe également dans la composition des syllabes (Dauer, 1983). Les langues accentuelles ont en général une plus grande variété de syllabes que les langues syllabiques et donc un inventaire de syllabes plus complexes. De plus, l'accent porte sur certaines syllabes tandis que les autres syllabes présentent un inventaire vocalique réduit, voire absent. Dans les langues syllabiques, l'accent et la structure syllabique sont généralement indépendants. En français, comme mentionné précédemment, s'il y a une accentuation, elle portera systématiquement sur la dernière syllabe du groupe rythmique qui sera alors plus longue (Meynadier, 2001). Cependant, l'accentuation ne joue pas de rôle dans l'accès au sens de l'énoncé. L'existence de trois classes rythmiques de langues est par exemple illustrée par le fait que nous présentons généralement des difficultés à segmenter les langues étrangères ayant une structure rythmique différente de notre langue maternelle (Otake & Cutler, 1996) ou encore par le fait que les nouveau-nés sont capables de différencier deux langues lorsqu'elles sont différentes en termes rythmiques mais pas lorsqu'elles appartiennent à la même classe (Nazzi et al., 1998).

Parmi les variables temporelles caractérisant le rythme de la parole, on compte la durée des unités linguistiques, les pauses et le débit.

Les **variations de durée** concernent généralement les phonèmes et les syllabes, qui peuvent alors être brefs ou longs. Tous les éléments constitutifs d'une syllabe peuvent ne pas être affectés par le même effet d'allongement : celui-ci peut être observé sur l'attaque ou la rime, ou encore, au sein de la rime, sur le noyau ou la coda. Dans certaines langues (*e.g.*, anglais, allemand, italien, finnois etc.), l'allongement des phonèmes peut signaler un changement au niveau lexical (i.e. il est pertinent en termes phonologiques). C'est le cas par exemple dans les mots italiens *fato* [fato] (« fait », consonne [t] brève) et *fatto* [fatto] (« destin », consonne [t] longue). En français, l'allongement d'un phonème n'a pas (ou plus) de telle valeur distinctive (autrefois, l'opposition entre « mètre », voyelle courte, et « maître », voyelle longue). Au niveau supralexical, l'allongement porte principalement sur la syllabe. L'un des effets les plus largement décrits est l'effet d'allongement final qui porte sur la syllabe finale d'une unité prosodique, contribuant à délimiter ces unités entre elles. C'est le cas du français où tout groupe rythmique se termine par une syllabe dont la durée est en moyenne deux fois plus longue que les syllabes internes du groupe qui sont de durées sensiblement égales (*e.g.*, « Le petit garçon »).

Les **pauses** participent également à la structuration des énoncés en ce qu'elles lient les groupes rythmiques successifs (Frasse, 1956). La pause est définie comme la manifestation physique (objective) ou perceptive (subjective) d'une interruption ponctuelle du flux de parole (Di Cristo, 2013). On distingue les pauses silencieuses (absence d'énergie acoustique) des pauses remplies, caractérisées par l'allongement exceptionnel d'une syllabe ou par la production d'interjections (*e.g.*, « euh », « hum » ; Campione & Véronis, 2004; Lacheret-Dujour & Beaugendre, 1999). Les pauses silencieuses précèdent et suivent toujours les groupes rythmiques qui coïncident avec l'effort respiratoire (les pauses correspondent alors à des prises de souffle). Certaines pauses silencieuses ne sont pas forcément associées à une prise de souffle, dans ce cas, elles visent à structurer le discours en signalant les frontières syntaxiques et les groupes sémantiques. Les pauses remplies peuvent refléter des hésitations mais aussi une activité cognitive telle que la recherche lexicale.

Enfin, le rythme de la parole est caractérisé par la vitesse à laquelle la parole est produite.

On définit le **débit de parole** comme « la manière plus ou moins rapide de s'exprimer oralement » (Trésor de la langue française). Le débit est défini à partir de la vitesse de parole ou vitesse d'élocution : celle-ci est mesurée en divisant le nombre de segments produits par le temps total de parole réalisé (incluant la phonation et les pauses), et est exprimée en nombre d'unités par secondes (Lacheret-Dujour & Beaugendre, 1999). Les unités mesurées correspondent généralement aux syllabes (la vitesse est alors exprimée en syllabes par seconde ou syll/s), mais on peut également se baser sur le nombre de phonèmes ou de mots produits par unité de temps.

$$\text{Débit de parole (vitesse d'élocution)} = \frac{\text{Nombre de syllabes}}{\text{Temps total de parole}}$$

On peut également mesurer la **vitesse d'articulation** (ou débit articulatoire), qui correspond à la vitesse à laquelle un locuteur articule un énoncé, les temps de pause étant décomptés (Grosjean & Deschamps, 1975). Elle dépend donc uniquement du temps de phonation (i.e. où le locuteur produit les sons de la parole). En général, tout silence dépassant 250 ms est considéré comme une pause (Grosjean & Deschamps, 1975). La vitesse d'articulation s'exprime également en syll/s et est calculée comme suit :

$$\text{Vitesse d'articulation} = \frac{\text{Nombre de syllabes}}{\text{Temps d'articulation}}$$

Où temps d'articulation = temps total de parole - temps total de pauses

Le débit de parole (syllabique) est fortement lié à la prosodie et dépend de facteurs individuels ou encore du contexte. En conversation spontanée, il atteindrait 5.29 syll/s en français selon les résultats de Grosjean et Deschamps (1975) et 5.73 syll/s selon Malécot et al. (1972). Vaxelaire (1993) décrit quant à elle une vitesse d'élocution de 4.2 à 4.6 syll/s à débit normal, et de 5.8 à 6.3 syll/s à débit rapide. En situation de lecture, Fougeron et Jun (1998) rapportent un débit moyen de 5.2 à 6 syll/s ; Padeloup (Padeloup, 2004, 2005) relève des valeurs de vitesse d'articulation de 4.2 syll/s en débit lent, 5.2 syll/s en débit normal et 6.8 syll/s en débit rapide. Le débit varie également selon les langues, en général de 4 à 8 syllabes/secondes, et semble dépendre de la complexité des structures syllabiques. Dellwo et Wagner (2003) ont mesuré, dans une tâche de lecture à débit variable (très lent, lent, normal, rapide et très rapide), une vitesse sensiblement plus rapide en français (5.86 syll/s) qu'en anglais (5.78 syll/s) dans une situation de débit normal. A partir d'un corpus de parole

multilingue composé de textes narratifs courts, Pellegrino et collègues (Pellegrino, Coupé, & Marisco, 2011) ont quant à eux rapporté un débit moyen de 7.18 syll/s en français et de 7.84 syll/s en japonais par exemple, alors qu'il était bien plus lent en vietnamien (5.22 syll/s).

Le rythme fait donc partie inhérente du signal de parole. Il constitue l'un des fondements de l'organisation prosodique de la parole, caractérisée par la quasi-périodicité des indices rythmiques principalement portés par la syllabe. La syllabe est en effet marquée par différentes variations prosodiques (accentuation, durée, etc.) qui vont permettre de structurer le discours et ainsi aider l'auditeur à accéder au sens de l'énoncé. La parole est donc un signal hautement temporel organisé en différentes échelles hiérarchisées autour des principales unités linguistiques (phonèmes, syllabes et mots) pertinentes pour la segmentation et l'identification du signal. Ces fenêtres d'intégration temporelle sont reflétées au niveau de la structure temporelle fine du signal de parole (Rosen, 1992), observable à partir de la représentation spectrale et qui englobe les informations temporelles rapides correspondant aux phonèmes (> 40 Hz), et au niveau de l'enveloppe d'amplitude ou enveloppe temporelle du signal, qui correspond aux fluctuations lentes marquées par les syllabes et les mots (< 20 Hz) et visible à partir de la représentation de l'onde sonore. Ces deux structures temporelles sont extraites en parallèle par le système auditif lors de la perception de la parole et jouent un rôle crucial dans son intelligibilité (Arnal & Giraud, 2016; Drullman, 1995). Dans ce travail de recherche, nous nous intéressons tout particulièrement à l'enveloppe temporelle du signal de parole, véhiculant l'information syllabique et donc l'information prosodique, rythmique.

3. L'enveloppe temporelle du signal de parole

L'enveloppe temporelle ou enveloppe d'amplitude de la parole correspond, au sens strict, à l'ensemble des modulations d'amplitude d'un signal de parole à une fréquence comprise entre 2 et 50 Hz (Figure 1; Rosen, 1992; voir aussi Drullman, Festen, & Plomp, 1994a, 1994b; Greenberg & Ainsworth, 2006). Les caractéristiques acoustiques de l'enveloppe telles que l'intensité, le temps de montée, le temps de descente, la durée et la fréquence de fluctuation déterminent, respectivement, les sensations de force, d'attaque, de chute, de longueur et de tempo (Rosen, 1992). Les fluctuations lentes d'amplitude de l'enveloppe véhiculent des informations prosodiques (suprasegmentales) telles que celles liées à l'accentuation et au rythme de la parole, mais permettent aussi d'identifier les unités

linguistiques contenues dans le signal (niveau segmental). L'enveloppe transmet en effet des informations comme le voisement (i.e. vibration périodique des cordes vocales), le mode d'articulation (e.g., changements de durée et de temps de montée entre consonnes fricative /ʃ/ et affriquée /tʃ/ dans « *chop* » (« couper ») et « *ship* » (« navire ») en anglais) ou encore des informations sur l'identité des voyelles (en particulier lorsque leur durée possède une valeur distinctive, e.g., « *ship* » vs « *sheep* » en anglais ; Rosen, 1992).

Il est largement reconnu que la composante dominante dans l'enveloppe temporelle se situe dans les modulations comprises entre 4 et 8 Hz. La puissance acoustique dans cette bande de fréquence émerge de la dynamique articuloire et plus précisément des cycles d'ouverture et de fermeture de la mâchoire nécessaires à la production, respectivement, des voyelles et consonnes au sein des syllabes (voir MacNeilage, 1998; MacNeilage & Davis, 2001). Ainsi, la fréquence d'oscillation de la mâchoire dominante dans l'enveloppe d'amplitude correspond au rythme syllabique (Houtgast & Steeneken, 1985); il est alors considéré que l'information syllabique domine dans l'enveloppe d'amplitude. Comme le montre la Figure 1, la structure syllabique est particulièrement reconnaissable dans l'enveloppe temporelle, grâce à son noyau vocalique qui forme une bouffée d'énergie au niveau du spectre du signal (pics signalés par une étoile sur la Figure 1). L'attaque des modulations successives dans l'enveloppe, précédant les pics d'amplitude, et son taux de variation correspondent au temps de montée de l'enveloppe (ou *rise time* en anglais) ; il jouerait un rôle primordial dans le traitement de la parole ainsi que dans son acquisition (Goswami, 2012; nous le verrons dans le chapitre IV). Le *rise time* signale en effet l'apparition d'une nouvelle syllabe dans la chaîne parlée et facilite donc la segmentation du signal de parole en syllabes et en mots. De plus, dans les langues accentuelles telles que l'anglais, le *rise time* est plus long pour une syllabe accentuée que pour une syllabe non accentuée, favorisant le traitement rythmique.

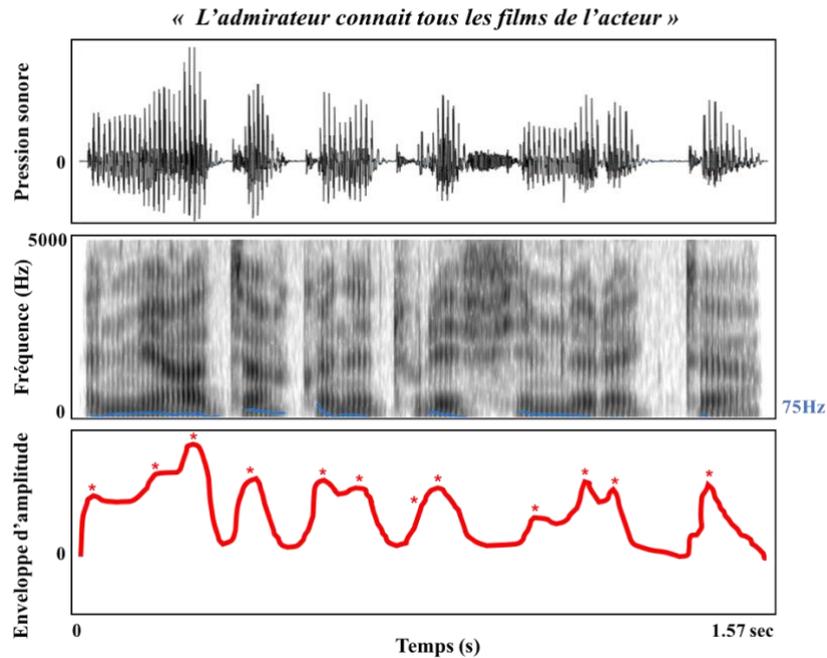


Figure 1 : Le signal de parole. Le panneau du haut montre les changements de pression sonore au cours du temps lors de la production de la phrase « L'admirateur connaît tous les films de l'acteur » produite par un locuteur masculin français. Au centre, le spectrogramme de ce signal est représenté, montrant la puissance acoustique dans les différentes bandes de fréquence au cours de la production de la phrase ; la F0 est notée en bleu, elle est de 75 Hz pour cette phrase. Le panneau du bas montre l'enveloppe temporelle de cette phrase ; elle est extraite à partir d'une transformée de Hilbert sur le signal de parole ré-échantillonné à 50 Hz. Le rythme syllabique est particulièrement visible, chaque syllabe étant marquée ici d'une étoile.

Un certain nombre d'études révèlent que l'intégrité de l'enveloppe temporelle du signal de parole est nécessaire à son intelligibilité. Dans un environnement silencieux, la compréhension d'un énoncé est en effet possible dans la mesure où l'information temporelle est préservée, même lorsque certaines bandes de fréquence sont supprimées (Drullman et al., 1994a, 1994b; Liégeois-Chauvel, Lorenzi, Trébuchon, Régis, & Chauvel, 2004; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). En utilisant la technique du *noise-vocoding* (i.e. dégradation du signal de parole consistant à retirer une part importante des informations spectrales ; voir Peelle & Davis 2012), Shannon et collaborateurs ont divisé des signaux de parole (consonnes a/C/a, voyelles h/V/d, et phrases) en 1 à 4 bandes de fréquence et extrait leur enveloppe temporelle. Pour cela, un filtre passe-bas (i.e. filtre laissant passer les basses fréquences et atténuant les hautes fréquences supérieures à la fréquence de coupure) a été appliqué, avec une fréquence de coupure de 16 Hz (i.e. les fréquences jusqu'à 16 Hz sont

maintenues alors que les fréquences supérieures à 16 Hz sont atténuées), 50 Hz, 160 Hz ou 500 Hz. Les enveloppes ont ensuite été utilisées pour moduler des bruits de bande passante comparable. Leurs résultats ont montré une diminution significative des performances de reconnaissance des consonnes et des phrases lorsque les fréquences comprises entre 16 Hz et 50 Hz dans l'enveloppe étaient supprimées, soit dans la condition du filtre passe-bas à 16 Hz. Aucune différence n'a été observée entre les conditions avec filtre à 50 Hz, 160 Hz et 500 Hz. Les auteurs ont conclu que la préservation de l'enveloppe d'amplitude du signal de parole est un prérequis à la bonne compréhension du message linguistique. Drullman et collaborateurs (Drullman et al., 1994a, 1994b) ont réalisé deux études similaires sur des phrases, masquées par du bruit, que les participants devaient restituer le plus fidèlement possible. Ils ont mis en évidence que l'intelligibilité était réduite (i.e. seuils de perception de la parole dans le bruit plus hauts) lorsque les modulations d'enveloppe supérieures à 4 Hz et 8 Hz étaient supprimées (avec des performances plus faibles dans la première condition), alors que la contribution de l'information au-dessus de 16 Hz était marginale (i.e. seuils constants pour les conditions de filtres passe-bas à 16, 32 et 64 Hz). Autrement dit, l'intelligibilité augmentait progressivement à mesure que les fluctuations comprises entre 4 et 8 Hz, et entre 8 et 16 Hz étaient conservées, alors que le fait de maintenir les modulations entre 16 et 32 Hz ne présentait aucune valeur ajoutée par rapport au fait d'inclure les modulations entre 32 et 64 Hz.

Ainsi, ces résultats indiquent que les informations contenues dans l'enveloppe temporelle importantes pour le traitement de la parole se situent entre 1 et 16 Hz, avec une composante dominante autour de 4 Hz qui correspond sensiblement au rythme syllabique (voir aussi Christiansen, Dau, & Greenberg, 2007; Greenberg & Arai, 2004; Greenberg, Arai, & Grant, 2006). Supprimer les modulations dans ces basses fréquences brouillerait les frontières entre les syllabes successives, impactant alors la segmentation et la compréhension. Plus récemment, Ghitza (2012) a d'ailleurs démontré que la suppression des modulations lentes de l'enveloppe temporelle entre 2 et 9 Hz diminuait significativement l'intelligibilité du signal de parole. Dans cette étude, les sujets avaient pour tâche d'écouter une suite de 10 chiffres et ensuite de composer sur un clavier les 5 derniers. Différentes variations du signal de parole étaient réalisées et, lorsque l'information rythmique correspondant aux bandes de fréquence supprimées était réintroduite, les performances des sujets étaient améliorées.

Comme nous venons de le voir, les indices rythmiques du signal de parole, et notamment l'intégrité des modulations d'amplitude de basse fréquence présentes dans l'enveloppe temporelle, revêtent une importance capitale pour son intelligibilité. Les indices rythmiques qui semblent les plus importants à l'intelligibilité d'un énoncé correspondent aux informations syllabique et prosodique. Dans la vie quotidienne, la parole est rarement perçue dans sa forme canonique et ces indices rythmiques contribuent largement à la variabilité du signal de parole (spécificité du locuteur, variantes régionales, sociales, psychologiques). S'il est presque utopique de dresser un inventaire exhaustif des sources de variation de la parole et de leurs conséquences sur la perception et la compréhension, dans la prochaine partie, nous en présenterons un aperçu en insistant sur les variations portées par les indices rythmiques qui nous intéressent dans cette thèse.

4. Variations du signal de parole

En tant qu'auditeurs, nous faisons face à une grande variabilité ou pourrait-on dire à de grandes *variabilités* des sons de parole (voir Meunier, 2001, 2005). Labov (1986) a identifié différentes sources de variation inhérentes à la parole telles que les variations dues à l'accent, au rythme ou aux interactions sociales. Rossi (1989) a distingué ces phénomènes inscrits dans le fonctionnement de la parole (spécificité du locuteur, variantes régionales, sociales...) d'autres sources additionnelles telles que les bruits extérieurs. Ces différentes variations sont en premier lieu observées au sein d'un même locuteur : la production d'un même mot deux fois dans les mêmes conditions résultera en deux réalisations physiquement différentes en termes de durée et d'amplitude par exemple. Ceci est dû au fait que la production de la parole implique une activité musculaire complexe basée sur des gestes articulatoires difficiles à reproduire à l'identique. Toutefois, l'auditeur sera capable de faire abstraction de ces variations pour reconnaître les deux productions comme désignant le même mot. Parmi les sources de variation inhérentes de la parole, on compte également la coarticulation, qui correspond au chevauchement et à l'interaction des différents articulateurs lors de la production de segments phonétiques successifs. Dans une syllabe de type CV, les gestes articulatoires associés à la consonne initiale et ceux associés à la voyelle sont partiellement superposés. Ainsi, la coarticulation correspond à l'influence d'un son de parole sur un autre. L'assimilation est un autre phénomène de variation : elle correspond au transfert d'une caractéristique (trait phonétique) d'un son sur un son immédiatement voisin. Par exemple, le

/b/ final du mot « robe » sera prononcé /b/ (consonne voisée) dans « robe grise » mais /p/ (consonne non voisée) dans « robe sale ».

La variabilité de la parole provient également des différences entre locuteurs. D'un locuteur à l'autre, le flux de parole fluctue, modifiant la structure du signal et ses différents indices rythmiques (débit, accentuation et intonation). Ces variations sont principalement liées aux différences anatomiques au niveau du conduit vocal et aux stratégies articulatoires sous-jacentes chez les différents locuteurs. En effet, l'anatomie de l'organe bucco-phonatoire varie selon l'âge et le sexe et donne lieu à une grande variabilité d'un locuteur à l'autre. L'un des paramètres les plus affectés par l'âge et le sexe est la fréquence fondamentale qui correspond à la hauteur de la voix (cf. partie I.2 de cette Introduction). Les hommes et les femmes, dont la taille des résonateurs varie, produisent également des voyelles différentes (i.e. valeurs de formant plus élevées chez les femmes). De plus, le signal de parole peut varier au cours de l'énoncé d'un même locuteur, selon la situation dans laquelle il se trouve (discussion entre amis, à caractère professionnel, conférence, lecture à voix haute...) ou selon son état émotionnel (anxiété, stress, fatigue, joie...). En fonction de la situation de communication, l'articulation peut en effet être soutenue ou relâchée, le débit de parole plus ou moins rapide, les syllabes plus ou moins accentuées, l'intonation plus ou moins forte selon les hypothèses que le locuteur formule sur la probabilité d'être compris par ses interlocuteurs (cf. théorie de l'hyper- / hypoarticulation, Lindblom, 1990). Par exemple, lors d'une conférence, le locuteur va s'appliquer, cherchant à optimiser la compréhension pour les auditeurs. Il va parler plutôt lentement, avec une dynamique et une force articulatoire élevée, réalisant des pauses régulières et longues afin de faciliter l'encodage. Lors d'un discours informel entre amis, le débit de parole sera en revanche plus rapide, les pauses moins fréquentes et moins longues, les gestes articulatoires seront plus économiques, certains mots seront même restructurés et de nouveaux mots pourront apparaître (Browman & Goldstein, 1992; Grosjean & Deschamps, 1975; Kohler, 1992; Lucci, 1983).

Enfin, les sons de parole sont soumis à des variations régionales et sociales. Au niveau régional, prenons l'exemple du français standard qui distingue les voyelles /e/ et /ɛ/ en position finale de mots (e.g., « épée » /epe/ et « épais » /epɛ/). En français méridional, cette distinction contrastive n'est pas établie, seule la voyelle /e/ est réalisée ; ceci se retrouve au niveau perceptif puisque les locuteurs du français méridional considèrent la paire /epe/-/epɛ/ comme étant composée de mots identiques (voir par exemple Dufour, Nguyen, & Frauenfelder, 2007). Le contexte sociologique constitue également une source importante de

variation. Le phonème /r/ diffère par exemple en fonction de la classe sociale des locuteurs : il est plus marqué chez les New-Yorkais aisés, alors reflet d'un certain prestige, que dans la classe plus populaire (Labov, 1986).

Outre les variations inhérentes aux sons de parole, d'autres sources extérieures peuvent être relevées, en particulier la présence de bruit. Il est en effet rare que la parole soit perçue dans le silence absolu ; les échanges verbaux se font le plus souvent dans un environnement bruyant, que le bruit provienne d'autres conversations ou de la rue (voitures, engins de chantier, musique etc.). Pourtant, nous sommes capables de suivre la conversation et d'isoler la voix de notre interlocuteur parmi d'autres pour identifier le contenu de l'information qui nous est destinée.

En dépit de toutes ces variations, l'auditeur est capable de s'adapter : il appréhende les différentes sources de variabilité intra- et interindividuelle, ainsi que le bruit, pour extraire l'information linguistique pertinente et la décoder. Face à un locuteur s'exprimant avec un accent étranger marqué, ou avec un débit de parole élevé, l'auditeur aura tout d'abord des difficultés à comprendre le message linguistique transmis, mais au bout de quelques instants, la compréhension sera facilitée. Ce processus d'adaptation progressif aux variations de débit ou de prononciation de la parole illustre parfaitement la remarquable flexibilité du système perceptif de l'homme à normaliser un large éventail de paramètres du signal acoustique afin de rendre le discours accessible et compréhensible. Pour cela, les auditeurs ajusteraient automatiquement leurs frontières phonémiques à celles de leurs interlocuteurs afin d'élaborer une représentation phonologique qui puisse être traitée par le système perceptif (Francis, Baldwin, & Nusbaum, 2000; Golomb, Peelle, & Wingfield, 2007; Guiraud, Ferragne, Bedoin, & Boulenger, 2013).

Nous allons maintenant nous intéresser plus spécifiquement aux variations du rythme de la parole et à la manière dont les auditeurs y font face.

4.1. La parole rapide

Lorsqu'on s'intéresse aux variations de débit de parole en situation expérimentale, deux types de stimuli peuvent être étudiés : la parole produite naturellement à un débit rapide et la parole compressée temporellement.

Lorsqu'un locuteur augmente son débit de parole naturellement, des modifications au niveau temporel et spectral sont observées. La longueur des indices acoustiques est réduite

et les gestes articulatoires sont effectués plus rapidement, ce qui les rend moins précis. Il se produit alors des phénomènes de réduction (Kohler, 1992) : la structure acoustique des segments est altérée, les différences entre les phonèmes sont réduites, les phénomènes de coarticulation et d'assimilation sont accentués. En outre, en français et en anglais, les voyelles et les syllabes non accentuées (pour l'anglais) présentent une réduction plus importante que les consonnes et les syllabes accentuées (Duez, 1995, 1998, 1999; Max & Caruso, 1997; Van Son & Pols, 1996, 1999), rendant le patron prosodique plus proéminent (Janse, Nootboom, & Quené, 2003).

La parole compressée temporellement correspond quant à elle à une réduction artificielle, linéaire ou non, de la durée des indices acoustiques d'un signal de parole préalablement enregistré. Il s'agit d'une modification de la structure temporelle du signal sans que la composition spectrale soit affectée. Cette accélération du signal de parole est réalisée à l'aide d'un algorithme. Elle peut être linéaire, c'est-à-dire que tous les segments de la parole sont réduits de manière identique, ou alors non linéaire, dans ce cas, la réduction varie selon le type d'unités linguistiques, afin de refléter les réductions temporelles observées lors de l'accélération naturelle.

La parole naturellement produite à un débit rapide implique donc des modifications au niveau temporel (réduction de certains indices acoustiques) mais aussi spectral (coarticulation et assimilation), alors que la parole compressée ne présente que des modifications temporelles, que la compression soit linéaire ou non. Cette spécificité de variation sur deux niveaux de la parole naturelle rapide peut constituer une difficulté supplémentaire pour l'auditeur, complexifiant le traitement de la parole. Janse (2004) a comparé le traitement de phrases naturellement produites à un débit rapide (8.5 syllabes/s), compressées de manière linéaire (au même débit) et compressées de manière sélective (suivant le même pattern de réduction temporelle que la parole naturelle rapide). Elle a observé, dans une tâche de détection de phonèmes dans des phrases, que les adultes néerlandais avaient plus de difficultés à traiter la parole naturelle rapide que la parole compressée linéairement ; les performances des participants ne différaient en revanche pas significativement entre la parole naturelle rapide et la parole compressée sélectivement. La parole naturelle rapide et la parole compressée « mimant la structure temporelle » de la parole accélérée naturellement étaient en outre jugées moins agréables à écouter que la parole compressée linéairement. Ainsi, de par les changements temporels et le plus fort recouvrement articulatoire entre les segments, la

parole produite naturellement à un débit rapide constitue un signal particulièrement difficile à décoder pour l'auditeur. Les résultats de Janse (2004), combinés à ceux obtenus dans une étude précédente (Janse et al., 2003), montrent également que la présentation de parole compressée plus comparable à la parole naturelle rapide n'améliore pas l'intelligibilité, au contraire. Les auteurs ont interprété l'ensemble de leurs résultats comme reflétant le fait que les changements temporels associés à une accélération naturelle du débit découlent plus de contraintes articulatoires que d'une stratégie de communication en faveur de l'auditeur (comme suggéré par la théorie de l'hyper-/hypoarticulation de Lindblom, 1990). Autrement dit, la manière dont les locuteurs augmentent leur débit ne serait pas forcément la plus favorable à la compréhension ni la préférée des auditeurs.

4.2. Adaptation à la parole rapide

Malgré les variations rythmiques auxquels l'auditeur est continuellement confronté, il est tout à fait capable de s'adapter et de comprendre le message linguistique. Un certain nombre d'études ont ainsi examiné le phénomène d'adaptation à la parole rapide. Les auteurs ont rapporté que les jeunes adultes et les personnes âgées sont parfaitement capables de s'adapter à de la parole compressée temporellement de manière linéaire, et ce jusqu'à un taux de compression correspondant à 65% de la durée originale du signal naturel produit à un débit normal (Adank & Janse, 2009; Dupoux & Green, 1997; Golomb et al., 2007; Janse, 2004; Janse et al., 2003; Peelle & Wingfield, 2005). Dupoux et Green (1997) ont demandé à de jeunes adultes américains (anglophones) d'écouter 20 phrases, présentées par blocs de 5, compressées linéairement à 38% ou 45% de leur durée originale, et d'écrire les mots entendus dont ils pouvaient se rappeler. Les sujets étaient divisés en deux groupes : le 1^{er} groupe entendait des phrases produites par un locuteur masculin, le 2^{ème} des phrases produites par un locuteur féminin. Les résultats ont d'abord montré de meilleures performances dans la condition de compression à 45% que dans celle à 38% (i.e. compression plus forte rendant la compréhension plus difficile). Ils ont ensuite décrit une augmentation des scores de rappel entre les trois premiers blocs de phrases, soit après l'écoute de 5 à 15 phrases et ce, quel que soit le genre du locuteur. Le processus d'adaptation se met donc en place très rapidement et indépendamment des caractéristiques acoustico-phonétiques des locuteurs. Toutefois, ce processus était plus rapide pour les phrases compressées à 45% (augmentation des performances entre le premier et le deuxième bloc puis stabilisation) que pour celles compressées à 38% (augmentation des scores entre le premier et le deuxième bloc, puis entre

le deuxième et le troisième bloc). L'adaptation est donc directement liée au taux de compression et par conséquent au débit de parole. Dans deux autres expériences utilisant le même paradigme, les auteurs ont montré que (1) l'adaptation pouvait être transférée d'un locuteur à un autre (i.e. elle se produit même lorsqu'un changement de locuteur intervient pour le troisième bloc) et (2) un changement de taux de compression (phrases compressées à 38% dans les deux premiers blocs, puis à 50% dans le troisième bloc) n'entraînait pas de retour des performances à la ligne de base, même si un léger déclin des scores de rappel était observé pour les deux premières phrases suivant le changement ; l'adaptation était toutefois plus rapide que lors de la première exposition. Ces résultats suggèrent donc qu'une fois l'adaptation mise en place, elle se maintient à long terme, reflétant un processus d'apprentissage dans lequel, afin de s'ajuster au changement de débit, les paramètres phonétiques sont extraits puis stockés pour être utilisés ultérieurement.

La plupart des études évaluant l'adaptation aux variations de débit de parole ne se sont concentrées que sur la parole compressée temporellement (i.e. accélérée de manière artificielle). Si cette méthode a permis de révéler l'existence de processus cognitifs robustes impliqués dans l'adaptation à la parole rapide, il est crucial d'examiner une situation plus écologique : la parole naturelle produite à un débit rapide. Comme mentionné précédemment, la compression de la parole n'altère en effet que la structure temporelle du signal, alors que la parole accélérée naturellement entraîne des modifications temporelles mais aussi spectrales suite aux différentes stratégies articulatoires du locuteur pour accélérer son discours (i.e. phénomènes de réduction, coarticulation etc., Duez, 1999; Kohler, 1992). Adank et Janse (2009) ont comparé l'adaptation à ces deux types de parole chez des adultes néerlandais dans une tâche de vérification de phrases produites soit à débit naturel normal (4.7 syll/s), soit à débit naturel rapide (10.2 syll/s), soit de phrases compressées (pour atteindre le débit naturel rapide). Les trois conditions de débit étaient présentées dans trois blocs successifs. Les sujets étaient divisés en deux groupes afin d'évaluer un possible effet de transfert entre les conditions de parole naturelle rapide et compressée : un premier groupe était exposé à la parole naturelle normale, puis naturelle rapide, puis compressée ; le second groupe était exposé à la parole naturelle normale, puis compressée et enfin naturelle rapide. Afin d'évaluer l'adaptation au cours du temps, chaque condition de débit présentée aux sujets était divisée en 10 mini-blocs de 6 phrases chacun. L'analyse des scores de bonnes réponses et des temps de réponse a d'abord montré des performances plus faibles et des temps de réponse plus longs pour la parole naturelle rapide que pour la parole naturelle normale ; les temps de réponse

étaient également plus longs pour la parole naturelle rapide comparée à la parole compressée (en accord avec Janse, 2004, Figure 2). Les auteurs ont ensuite observé la mise en place d'un processus d'adaptation pour les deux types de parole accélérée (naturellement et artificiellement), qui était toutefois plus rapide lorsqu'il s'agissait de parole compressée (i.e. augmentation des performances entre les deux premiers mini-blocs (6 à 12 phrases), alors que l'adaptation requérait l'écoute de 12 à 18 phrases, soit trois mini-blocs, pour la parole naturelle rapide, Figure 2). Enfin, un transfert d'apprentissage entre les deux types de parole accélérée a été mis en évidence : les sujets étaient plus performants pour traiter la parole naturelle rapide s'ils avaient au préalable été exposés à la parole compressée, alors que l'inverse n'était pas vrai (Figure 2). L'adaptation préalable à de la parole compressée semble donc faciliter la compréhension ultérieure de parole naturelle rapide. Dans l'ensemble, cette étude révèle donc que les adultes s'adaptent à la parole accélérée à la fois naturellement et artificiellement par compression temporelle de manière assez rapide et efficace, mais de façon sensiblement différente : l'adaptation à la parole naturelle rapide semble être plus difficile puisque plus longue à mettre en place, résultant probablement des modifications supplémentaires au niveau de la composante spectrale du signal de parole (Janse, 2004). Dans une étude chez des enfants francophones âgés de 8-9 ans (CE2), nous avons observé des résultats sensiblement similaires : les enfants présentaient en effet une amélioration de leurs temps de réponse après l'écoute de 5 phrases accélérées naturellement et compressées temporellement (Guiraud, Ferragne, Bedoin, & Boulenger, 2013). Bien que l'adaptation se soit mise en place aussi rapidement pour les deux types de parole accélérée, une analyse des scores d'erreurs a néanmoins montré de moins bonnes performances pour la parole naturelle rapide. S'adapter à la parole naturelle rapide pourrait donc reposer sur des processus cognitifs spécifiques, recrutant potentiellement des aires cérébrales supplémentaires sous-tendant le traitement des aspects temporel et spectral du signal. Nous le verrons dans la suite de cette Introduction.

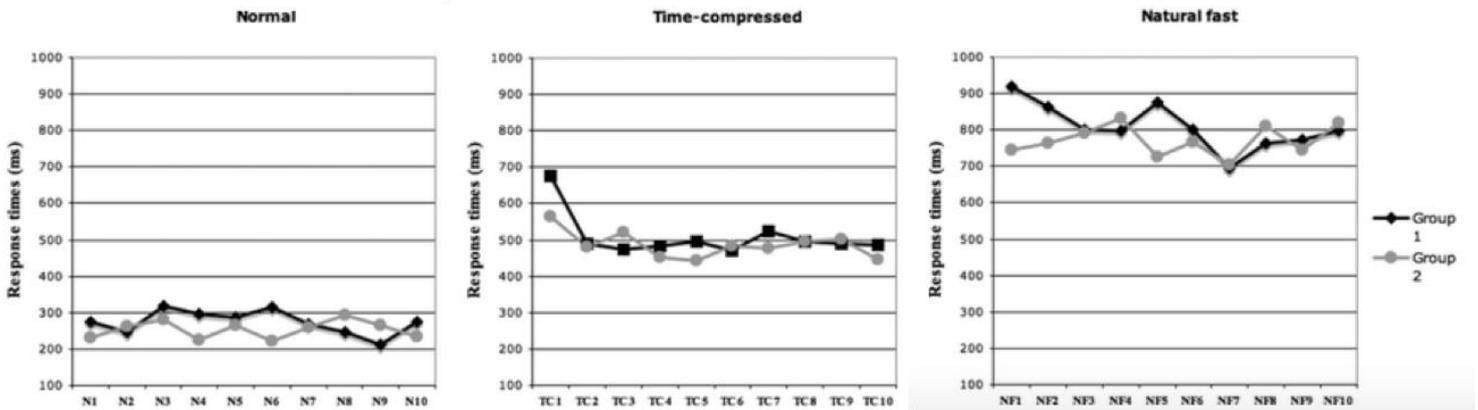


Figure 2 : Temps de réponse moyens en millisecondes pour chacun des 10 blocs de 6 phrases perçues dans chacune des conditions de débit de parole : à gauche, parole produite à un débit naturel normal (4.7 syll/s) ; au milieu, parole compressée linéairement ; à droite, parole produite naturellement à un débit rapide (10.6 syll/s). Pour chaque condition de débit de parole, les résultats sont présentés pour les deux groupes de sujets : le groupe 1 entendait d'abord les phrases naturelles normales, puis naturelles rapides et enfin compressées ; le groupe 2 entendait d'abord les phrases naturelles normales, puis compressées, puis naturelles rapides. Figure empruntée à Adank & Janse (2009).

Dans cette partie, nous avons vu que le signal de parole était un signal très variable, dépendant des spécificités du locuteur et de la situation dans laquelle il se trouve. L'une des sources de variations qui nous intéresse tout particulièrement dans ce travail de thèse est la variation de débit de parole et notamment le cas de la parole rapide. Une telle variation affecte tout particulièrement le rythme du signal et l'information syllabique et, comme révélé par plusieurs études, constitue une difficulté pour l'auditeur dont le système doit rapidement mettre en place des processus d'adaptation. Nous allons maintenant nous intéresser aux processus corticaux mis en jeu lors du traitement du signal de parole plus ou moins variable.

II. La perception de la parole dans le cerveau

Comment sommes-nous capables d'extraire les invariants phonologiques du signal de parole et d'en décoder le contenu linguistique, malgré la grande variabilité des productions ? Un certain nombre de théories ont vu le jour afin de mieux comprendre les processus cognitifs et neurobiologiques mis en jeu dans le décodage des informations phonétiques. Dans ce chapitre, nous présenterons d'abord les principales théories de la perception de la parole, à savoir celles qui défendent la nature auditive ou motrice du traitement. Après un bref aperçu des bases cérébrales du langage, nous nous intéresserons ensuite plus précisément au rôle du système moteur dans la perception de la parole. Enfin, nous exposerons les modèles neurobiologiques récents proposés pour rendre compte de la manière dont est implémentée la parole dans le cerveau.

1. Théories auditives, motrices et sensori-motrices de la parole

Deux grands types de théories s'affrontent pour rendre compte des processus mis en jeu dans le décodage du signal de parole. L'élément crucial sur lequel elles se distinguent tient aux mécanismes mis en jeu lorsque nous percevons autrui parler : alors que les théories dites « auditives » supposent l'utilisation de mécanismes perceptuels purement auditifs, les théories dites « motrices » suggèrent l'implication de processus de contrôle articulatoire et plus généralement du système moteur.

Les tenants **des théories auditives** postulent que la perception de la parole repose uniquement sur les processus auditifs. Elle dépendrait des mêmes mécanismes généraux (i.e. non spécifiques à la parole) que ceux sous-tendant l'audition et l'apprentissage perceptif, et résulterait d'une analyse spectro-temporelle du signal acoustique (Diehl, Lotto, & Holt, 2004; Massaro, 1974; Ohala, 1996). Autrement dit, les invariants phonémiques seraient récupérés et accessibles directement dans le signal acoustique, qui serait alors traité uniquement par le système auditif. Il existerait une organisation hiérarchique très structurée des indices acoustiques (voisement, transitions formantiques etc.) du signal de parole dans laquelle les caractéristiques acoustiques invariantes seraient reliées aux traits phonétiques, quels que soient le locuteur ou le contexte phonétique (Stevens & Blumstein, 1978). Une multitude

d'indices acoustiques serait ainsi appariée aux catégories perceptives telles que les phonèmes et les syllabes. Par exemple, le simple contraste phonétique /aba/ - /apa/ serait distingué par pas moins de 16 indices acoustiques différents. Dans cette perspective, la perception de la parole est donc considérée comme la reconnaissance de patterns acoustiques complexes mettant principalement en jeu le système auditif (Diehl et al., 2004 pour une revue). Ces approches auditives se basent principalement sur le fait que le phénomène de perception catégorielle⁵ (i.e. capacité à mieux discriminer, à distance acoustique constante, deux phonèmes appartenant à deux catégories phonologiques différentes que deux phonèmes appartenant à la même catégorie ; Liberman et al., 1957) ne serait ni spécifique à la parole (nous percevons également les sons non verbaux de manière catégorielle ; Pisoni, 1977; Stevens & Klatt, 1974) ni à l'homme (la perception catégorielle a été décrite chez les animaux pour des sons verbaux et non verbaux; Kuhl & Miller, 1975, 1978). Ainsi, la perception de la parole reposerait alors sur des mécanismes généraux communs à ceux impliqués dans le traitement auditif.

La théorie motrice de la perception de la parole se base quant à elle sur l'existence d'une relation non linéaire entre les propriétés acoustiques du signal de parole et les représentations phonétiques stockées en mémoire, qui serait notamment liée aux phénomènes de coarticulation (Liberman & Mattingly, 1985; Liberman & Whalen, 2000; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Selon cette théorie, les caractéristiques invariantes de la perception de la parole seraient les gestes articulatoires du locuteur. Plus précisément, les invariants phonémiques seraient extraits, non pas dans le signal acoustique

⁵Lors de la perception de la parole, l'auditeur va devoir s'appuyer sur des indices perceptifs particuliers afin de découper le signal en différentes unités/segments capables de s'apparier avec les représentations lexicales déjà mémorisées. Le traitement perceptif serait régi par le phénomène de perception catégorielle. « La perception catégorielle est le phénomène par lequel les différences acoustiques entre sons de la parole ne sont pas perceptibles sauf s'ils appartiennent à des catégories phonétiques différentes » (Liberman, Harris, Hoffman, & Griffit, 1957). Elle est classiquement observée lors d'une tâche de discrimination ou d'identification auditive de deux syllabes appartenant à un continuum dont les extrémités correspondent à deux syllabes qui ne diffèrent que par un trait phonétique distinctif donné (e.g., /ba/ et /pa/ qui diffèrent par leur délai d'établissement du voisement (*voice onset time* ou VOT) correspondant au temps entre la barre d'explosion de la consonne et le voisement de la voyelle). Entre ces extrémités, plusieurs syllabes intermédiaires sont créées en faisant varier le VOT. Ces tâches permettent de localiser avec précision la frontière catégorielle, c'est-à-dire le point sur le continuum où les auditeurs considèrent deux syllabes comme faisant partie de la même catégorie ou non (La Guitton et al., 1997; Liberman et al., 1957). Ainsi, dans les sons de la parole, il existerait des frontières phonétiques nettes, perçues inconsciemment et permettant de distinguer les différents phonèmes de la parole. Ce processus est par ailleurs présent chez le nourrisson, lui permettant de discriminer les contrastes phonétiques de n'importe quelle langue de façon catégorielle (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Puis vers l'âge de 10-12 mois, l'enfant se spécialise pour les phonèmes de sa langue maternelle (Trehub, 1976; voir Christophe, Pallier, Bertoncini, & Mehler, 1991 pour une revue).

comme proposé par les tenants des approches auditives, mais grâce à un appariement entre les gestes articulatoires du locuteur et les représentations motrices activées par le signal de parole perçu chez l'auditeur. Autrement dit, lors de l'écoute d'un interlocuteur, nous récupérons non pas le signal acoustique de la parole, mais les commandes motrices transmises aux articulateurs (i.e. langue, lèvres, cordes vocales) et nécessaires à la production des sons. Liberman et ses collègues suggèrent ainsi que la communication est possible grâce à la parité entre les représentations phonétiques et motrices invariantes entre les communicants (locuteur et auditeur) : « *Communication can succeed only if the two parties have a common understanding about what counts – what counts for the speaker must count for the listener.* » (Liberman & Whalen, 2000). En outre, les processus de perception et de production doivent être liés ; leurs représentations doivent être, en un certain sens, identiques.

L'un des arguments majeurs en faveur de la théorie motrice tient au phénomène de coarticulation : Liberman et collègues (1967) prennent ainsi l'exemple du phonème /d/ qui présente des valeurs de deuxième formant différentes lorsqu'il se trouve dans les contextes /di/ « *dee* » et /du/ « *doo* ». Dans les deux cas, la commande neuromotrice est identique, invariante : pour les auteurs, c'est cette invariance qui serait exploitée par l'auditeur pour comprendre le phonème. Liberman et Mattingly (1985) ont par ailleurs défendu l'idée que la perception de la parole repose sur des processus radicalement différents de ceux sous-tendant le traitement d'autres types de sons. Dans cette optique, il existerait donc un module spécifique dédié à la perception de la parole, présent uniquement chez l'homme et inné. Ce postulat a été vivement débattu et critiqué et est aujourd'hui largement abandonné (voir par exemple Diehl et al., 2004 et Galantucci, Fowler, & Turvey, 2006).

La théorie motrice implique donc des interactions fortes entre perception et production de la parole, et par conséquent l'existence d'une convergence dynamique implicite entre les participants d'une conversation, se traduisant par une parole fluide, une coordination respiratoire, un débit de parole semblable et un alignement des représentations linguistiques, ce qui garantirait un échange de bonne qualité (Garrod & Pickering, 2004; Pickering & Garrod, 2006). La découverte des neurones miroirs dans les années 90 a permis de soutenir l'idée d'une telle implication du système moteur, articulatoire, dans la perception de la parole. Les neurones miroirs, décrits pour la première fois dans l'aire F5 du cortex prémoteur ventral du macaque (Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Vittorio Gallese, Fadiga, Fogassi, & Rizzolatti, 1996) puis dans le cortex pariétal postérieur (Fogassi et al., 2005; Gallese, Fadiga, Fogassi, & Rizzolatti, 2002), s'activent durant l'exécution d'une action

mais aussi lors de l'observation de cette action réalisée par un autre individu, ou encore de l'écoute du son qui lui est associé (*e.g.*, casser une cacahuète, Kohler et al., 2002; voir Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010 pour des revues). Il est d'ailleurs intéressant de noter que l'aire prémotrice F5 chez le singe possède des similitudes cytoarchitectoniques et fonctionnelles avec l'aire de Broca chez l'homme (voir Rizzolatti & Arbib, 1998). Un système miroir a également été décrit chez l'homme, puisque les études d'imagerie cérébrale ont révélé des activations communes au sein des cortex prémoteur et pariétal lors de l'exécution et de l'observation d'actions (*e.g.*, Buccino et al., 2001, 2004). Le système miroir, en appariant production et perception, formerait ainsi un système impliqué dans la compréhension du sens des actions d'autrui et jouerait un rôle dans la communication intentionnelle (Rizzolatti & Craighero, 2004). En 1998, Rizzolatti et Arbib ont en outre suggéré une implication du système miroir dans l'émergence du langage chez l'homme et donc dans la perception de la parole (Rizzolatti & Arbib, 1998; voir aussi Arbib, 2005; Gentilucci & Corballis, 2006). Le système miroir permettrait en effet d'établir le lien de parité entre l'auteur et le destinataire d'un message, pré-requis à l'émergence de toute communication volontaire, qui constitue le fondement de la théorie motrice de la perception de la parole (Lieberman & Whalen, 2000).

L'hypothèse d'un rôle primordial du système moteur dans la perception de la parole fait toutefois débat et certains auteurs suggèrent que les gestes articulatoires de la parole ne peuvent permettre à eux seuls la perception du signal acoustique de parole (Lotto, Hickok, & Holt, 2009; Rauschecker & Scott, 2009). C'est pourquoi les études actuelles penchent plutôt en faveur de l'existence d'une interaction étroite entre les systèmes moteur et auditif dans la perception de la parole, définissant les unités verbales non pas comme purement acoustiques ou purement motrices, mais comme sensori-motrices (Hickok & Poeppel, 2007; Schwartz, Basirat, Ménard, & Sato, 2012).

Le modèle de la perception pour le contrôle de l'action (*PACT: Perception for Action Control Theory*; Schwartz et al., 2012) est l'un des modèles défendant le caractère sensori-moteur des unités de la parole. Il est principalement fondé sur l'existence de connaissances procédurales motrices et auditives mises en place lors de l'acquisition de la parole et partagées par le locuteur et l'auditeur. Deux principes guident ce modèle : la perception implique une connaissance de l'action, et l'action vise à réaliser des gestes structurés par la perception. La perception de la parole est donc considérée ici comme un

ensemble de mécanismes permettant de comprendre mais aussi de contrôler la parole en tant que système de communication. Ce modèle s'inspire de différents postulats sur les processus mis en jeu lors du développement du système langagier. Premièrement, l'acquisition du langage est permise par la co-structuration des représentations sensorielles et motrices (apprentissage sensori-moteur par imitation) permettant le développement de cartes sensori-motrices de la parole. Deuxièmement, la production est structurée par la perception, c'est-à-dire que les gestes articulatoires sont façonnés par rapport à leurs propriétés perceptives auditives et visuelles. Enfin, le système perceptif se construit à partir du système bucco-phonatoire et selon ses contraintes anatomiques. Ainsi, selon le modèle PACT, lors de la perception d'un signal de parole, les représentations articulatoires de l'auditeur contraignent le traitement et l'interprétation des entrées sensorielles à l'aide de mécanismes de simulation motrice interne.

Après avoir passé en revue les principales théories de la perception de la parole, nous allons présenter les substrats cérébraux sous-tendant les processus de perception de la parole.

2. Bases cérébrales de la perception de la parole

La recherche sur les fonctions motrices, sensorielles et cognitives du cerveau a débuté au XIX^{ème} siècle avec la méthode dite anatomo-clinique, qui consistait à réaliser des descriptions *post-mortem* (autopsies) de cerveaux de patients atteints de troubles cognitifs. Ces études, à la base de la neuropsychologie, ont permis de fournir les premières descriptions fonctionnelles du cerveau. Puis l'avènement des techniques d'imagerie cérébrale, permettant de « voir le cerveau fonctionner » chez le sujet sain et donc de ne plus se baser uniquement sur le cerveau lésé, a permis de dresser des cartes fonctionnelles plus précises.

A partir de l'étude de l'organisation cellulaire et des différentes fonctions cérébrales, Brodmann (1909) a proposé une carte cyto-architectonique du cortex cérébral, le découpant en 52 régions appelées aires de Brodmann (Figure 3). Chaque région serait par ailleurs impliquée dans une fonction particulière. Par exemple, l'aire 4 dans le gyrus précentral correspond au cortex moteur primaire et joue un rôle dans le contrôle et l'exécution des mouvements volontaires. L'aire 6, située en avant de l'aire 4, correspond au cortex prémoteur et participe à la planification des mouvements.

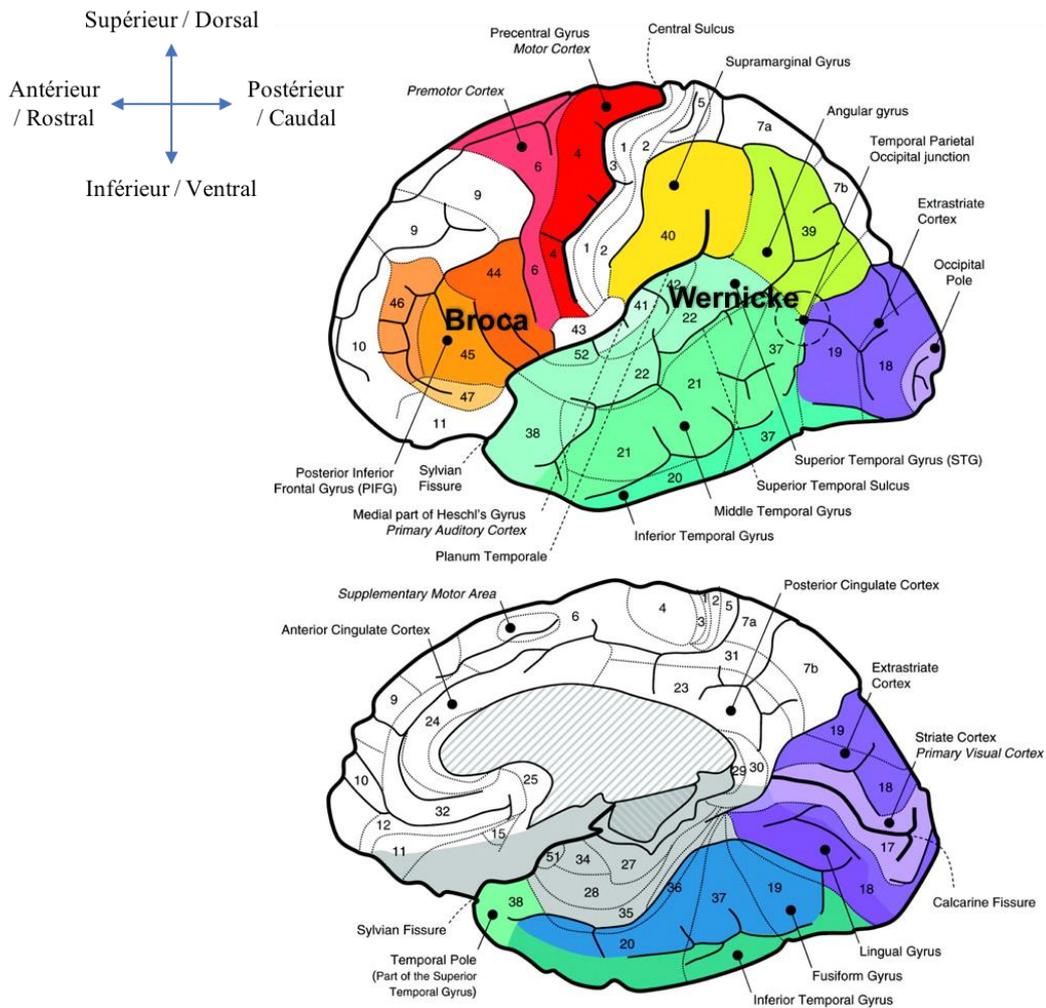


Figure 3 : Cartographie de Brodmann et principales régions cérébrales impliquées dans le traitement du langage. Vue latérale de l'hémisphère gauche (en haut) et vue médiale de l'hémisphère droit (en bas). La numérotation correspond à la cartographie cyto-architectonique de Brodmann. Les régions colorées correspondent aux régions impliquées dans le système langagier. Figure adaptée de Démonet, Thierry, & Cardebat (2005).

Concernant le langage, les premières études lésionnelles ont montré que des régions spécifiques sont impliquées dans la perception et la production de la parole. Paul Broca, un neurochirurgien français, découvrit en 1861 l'une des zones principales responsable du traitement du langage chez l'homme, suite à l'étude du cerveau d'un de ses patients, Monsieur Leborgne, qui ne pouvait plus que prononcer la syllabe « tan ». Ce patient comprenait parfaitement ce qu'on lui disait, mais était incapable de produire une phrase complète ou d'exprimer ses idées par écrit. Il ne présentait pourtant aucune atteinte motrice oro-faciale. L'autopsie de ce patient, suivie de celle de patients similaires, révéla une lésion au sein du gyrus frontal inférieur gauche (au niveau de la troisième circonvolution). Cette zone fut alors

associée à la production du langage et nommée « aire de Broca » (Figure 3), le trouble du langage résultant de sa lésion étant désigné sous le nom d'aphasie de Broca (non fluente). Une dizaine d'années plus tard, en 1874, Carl Wernicke, un neurologue allemand, révéla l'existence d'une autre zone impliquée dans le traitement du langage, mais cette fois dans sa compréhension. Il décrivit des patients, présentant une lésion au sein de la partie postérieure du cortex temporal supérieur gauche, qui s'exprimaient avec un discours incohérent et dénué de sens. Cette aire fut nommée « aire de Wernicke » et associée à la compréhension de la parole (Figure 3) ; une atteinte de cette aire conduit à une aphasie de Wernicke (fluente).

Depuis, l'organisation anatomique du langage a été précisée, permettant de réfuter une dichotomie fonctionnelle stricte des aires de Broca et Wernicke lors de la production et de la perception de la parole ; d'autres régions ont par ailleurs été associées aux différentes fonctions du langage oral (Figure 3). L'aire de Broca serait divisée en deux unités fonctionnelles : la partie postérieure du gyrus frontal inférieur (IFG postérieur, aire 44 de Brodmann) serait impliquée dans les processus phonologiques et la production de la parole, rôle facilité par sa proximité avec les régions motrices hébergeant les représentations de la bouche et de la langue (aires 4 et 6). Elle jouerait également un rôle dans le traitement syntaxique (Friederici, 2017 pour une revue) et notamment lorsque la complexité syntaxique augmente (Ben-shachar, Hendler, Kahn, Ben-bashat, & Grodzinsky, 2013; Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Grodzinsky, 2000; Röder, Stock, Neville, Bien, & Rösler, 2002; Skeide, Brauer, & Friederici, 2016; Tettamanti et al., 2002). La partie antérieure du gyrus frontal inférieur (IFG antérieur, aire 45) serait quant à elle plutôt impliquée dans les processus sémantiques (Hickok & Poeppel, 2007; Scott & Johnsrude, 2003). L'aire de Wernicke est associée au *planum temporale*, qui correspond à la partie postérieure du gyrus temporal supérieur (STG, aire 22). Cette région est accolée au cortex auditif primaire (aires 41 et 42), où s'effectuent les premiers traitements des sons entendus (langagiers ou non), et au lobe pariétal inférieur, composé au niveau caudal du gyrus angulaire (aire 39) lié notamment aux processus sémantiques et au niveau dorsal du gyrus supramarginal (aire 40) impliqué dans le traitement phonologique et articulatoire des mots.

Lors de la perception de la parole, le signal acoustique est traité par des voies parallèles spécialisées dans les différents niveaux de traitement langagier. Le cortex auditif primaire reçoit et analyse les informations acoustiques du signal de parole et les transmet à différentes

aires secondaires associatives (impliquées dans les opérations plus complexes). Ces aires sont organisées hiérarchiquement au sein du cortex temporal, dans les régions supérieures et médianes. Les traitements acoustiques de bas niveau d'un signal auditif, non spécifiques à la parole, sont réalisés dans l'aire auditive primaire, présente au sein des gyri temporaux supérieurs dorsaux bilatéraux (STG, aires 41 et 42), comprenant le gyrus de Heschl et les régions antérieures du *planum temporale*. Les traitements secondaires spécifiques à la parole, correspondant au décodage acoustico-phonétique et à la catégorisation phonétique, ont lieu dans les régions antérieures du sillon temporal supérieur (STS). Scott et collaborateurs (2000) ont ainsi observé une activation spécifique du cortex temporal antérieur gauche lors de l'écoute de parole intelligible. Les auteurs ont fait écouter des phrases enregistrées préalablement par un locuteur à des participants dont l'activité cérébrale était enregistrée à l'aide de la tomographie par émission de positons (TEP). Quatre conditions différentes de parole étaient présentées: de la parole normale, intelligible; de la parole inversée spectralement, inintelligible sans entraînement préalable intensif; de la parole vocodée à 4 bandes (i.e. réduction du contenu spectral du signal par décomposition du spectre en plusieurs bandes de fréquence et remplacement de certaines bandes par du bruit blanc modulé en amplitude), intelligible après une rapide adaptation; et de la parole vocodée inversée spectralement, totalement inintelligible. Les résultats ont montré une activation au sein de l'hémisphère gauche dans le gyrus temporal supérieur (antérieur et ventral au cortex auditif primaire) et le sillon temporal supérieur postérieur lorsque le signal de parole comprenait de l'information phonétique (parole normale, vocodée et inversée spectralement). Cependant, seule la partie antérieure du STS s'activait lorsque le signal de parole était intelligible (parole normale et parole vocodée), suggérant que cette région permettait d'instancier les représentations phonétiques de la parole. L'implication du cortex temporal antérieur pour le traitement de la parole intelligible a également été décrite dans de nombreuses autres études (Davis & Johnsrude, 2003; Davis, Ford, Kherif, & Johnsrude, 2011; Friederici, 2011; Okada et al., 2010; Peelle, Eason, Schmitter, Schwarzbauer, & Davis, 2010; Richardson, Thomson, Scott, & Goswami, 2004; Rodd, Longe, Randall, & Tyler, 2010; Scott & Johnsrude, 2003). Cependant, il faut noter que d'autres auteurs ont rapporté une activation de la partie postérieure du STG et du gyrus temporal moyen lors du décodage acoustico-phonétique et de la catégorisation des phonèmes, et non de sa partie antérieure (Hickok & Poeppel, 2007; Rodd, Davis, & Johnsrude, 2005; Rodd et al., 2010; Zatorre, Belin, & Penhune, 2002). La localisation précise de ces fonctions sur l'axe antéro-postérieur au sein du cortex temporal fait encore l'objet de discussions à l'heure actuelle. Dans une revue, Friederici (2011) propose par

exemple que les régions antérieures du STG/STS traiteraient les informations sémantiques et syntaxiques extraites lors de la compréhension de phrases, alors que les régions postérieures intègreraient différents types d'information (langagières ou non) incluant les informations sémantiques et syntaxiques de la parole traitées dans la partie antérieure.

Une spécialisation hémisphérique est également observée lors de la perception de la parole, soulignant un engagement préférentiel de l'hémisphère droit pour le traitement des indices prosodiques, suprasegmentaux (qui est également observé lors de l'écoute de stimuli musicaux), alors que l'hémisphère gauche serait spécialisé dans le traitement précis de l'information spectro-temporelle et de l'information phonétique (Davis & Johnsrude, 2003; Hyde, Peretz, & Zatorre, 2008; McGettigan & Scott, 2012; Zatorre, 2003; Zatorre et al., 2002; Zatorre, Evans, Meyer, & Gjedde, 1992). Dans l'étude précédemment présentée, Scott et collaborateurs (2000) ont ainsi observé une activation au sein du gyrus temporal supérieur ventrolatéral droit lors de l'écoute des stimuli de parole normale et inversée spectralement, dont la hauteur et l'intonation étaient préservées.

Ainsi, lors de la perception d'un signal de parole, les régions antérieures et postérieures des cortex temporaux semblent être recrutées et impliquées dans des processus spécifiques (Davis & Johnsrude, 2003; Friederici, 2011; Scott, 2000). D'autres études ont également montré un recrutement des régions frontales lors du traitement de la parole, et en particulier une implication du système moteur (Friederici, 2011; Hickok & Poeppel, 2004, 2007; Rodd et al., 2005, 2010; Scott & Johnsrude, 2003).

3. Implication du système moteur dans la perception de la parole

Suite à la découverte des neurones miroirs et à leur rôle supposé dans le développement du langage et la perception de la parole (Lieberman & Whalen, 2000; Rizzolatti & Craighero, 2004; Rizzolatti & Arbib, 1998), de nombreuses études en neuroimagerie, utilisant des techniques telles que l'EEG⁶ (Electroencéphalographie), l'IRMf⁷ (Imagerie par Résonance

⁶ L'EEG mesure de manière continue l'activité électrique (variations de potentiel électrique) produite par les neurones au repos ou lors d'une tâche à l'aide d'électrodes placées à la surface du scalp (on parle d'EEG de

Magnétique fonctionnelle) ou encore la TMS⁸ (Stimulation Magnétique Transcranienne), se sont attachées à démontrer que la perception des sons de parole active spécifiquement les centres moteurs recrutés lors de la production de ces mêmes sons (voir Friederici, 2011 et Pulvermüller & Fadiga, 2010 pour des revues).

Dans une étude en TMS chez des adultes italiens, Fadiga et ses collaborateurs (Fadiga, Craighero, Buccino, & Rizzolatti, 2002) ont stimulé (impulsion unique) la représentation motrice de la langue dans le cortex moteur primaire gauche et enregistré les potentiels évoqués moteurs (PEM) directement sur les muscles de la langue des participants alors qu'ils écoutaient des mots et pseudo-mots bisyllabiques. Les auteurs ont montré une augmentation de l'amplitude des PEM (i.e. facilitation de l'excitabilité cortico-spinale) lors de l'écoute des stimuli comprenant la double consonne « rr » (e.g., « birra » / « bière ») qui nécessite un recrutement spécifique de la langue pour être produite en italien ; cet effet était en outre plus prononcé pour les mots que pour les pseudo-mots. En revanche, lors de l'écoute des stimuli contenant la double consonne « ff » (e.g., « baffo » / « moustache »), qui n'implique pas spécifiquement cet effecteur, aucune activité spécifique n'était évoquée sur les muscles de la langue (voir aussi Roy, Craighero, Fabbri-Destro, & Fadiga, 2008). Watkins et collègues (2003) ont quant à eux mesuré les potentiels évoqués moteurs sur les muscles des lèvres et de la main droite (après stimulation des représentations corticales motrices gauche et droite du visage et de la représentation corticale motrice gauche de la main par TMS à impulsion unique) alors que les participants (1) écoutaient un texte ou des sons non verbaux et (2) observaient des mouvements verbaux des lèvres (perception visuelle de la parole) ou des

surface). L'activité électrique recueillie résulte de la sommation des potentiels postsynaptiques dans le temps et dans l'espace (au minimum 100 000 neurones dans quelques mm³ de cortex). L'EEG permet d'étudier à la milliseconde près le déroulement temporel de l'activité cérébrale. Il est également possible de mesurer l'activité électrique cérébrale avec des électrodes intra-crâniennes (i.e. implantées en profondeur ; iEEG) ; ce type d'EEG est utilisé en particulier chez les patients épileptiques pour établir le diagnostic pré-chirurgical.

⁷L'IRMf est une technique d'imagerie dite hémodynamique et indirecte qui permet de mesurer les variations d'oxygénation du sang dans les différentes régions cérébrales lors d'une activité. Lorsqu'une aire cérébrale est impliquée dans une tâche cognitive, elle consomme plus d'oxygène, ce qui se traduit par une augmentation locale du flux sanguin (i.e. apport massif d'oxygène). C'est ce que l'on appelle la réponse BOLD (*Blood Oxygenation Level Dependent*), qui est basée sur les propriétés magnétiques de l'hémoglobine. Cette technique offre une excellente résolution spatiale (i.e. elle permet de localiser très précisément les régions activées par une tâche, par rapport à une condition contrôle), mais sa résolution temporelle est limitée.

⁸ La TMS est une technique non invasive consistant à délivrer des impulsions magnétiques à la surface du crâne avec une bobine, de sorte à induire de manière focale des courants électriques dans la région corticale ciblée. L'activité de cette région peut ainsi être modulée de façon variable selon la fréquence, le nombre d'impulsions, les intervalles et la durée des stimulations (Valero-Cabré, Pascual-Leone, & Coubar, 2011 pour une revue). Concernant le cortex moteur, les patterns d'impulsions de TMS à basse fréquence ou ceux délivrés de manière continue tendent à déprimer l'activité locale, tandis que la TMS à haute fréquence et discontinue tend à la potentialiser. Les impulsions de TMS activent la voie cortico-spinale et ses circuits associés et induisent des « secousses » dans les muscles spécifiques dont la représentation est ciblée par la stimulation au niveau cortical.

mouvements des yeux. Leurs résultats ont révélé une augmentation d'amplitude des PEM sur les muscles des lèvres, uniquement après stimulation de l'aire corticale motrice gauche du visage, lors de la perception auditive et visuelle de parole. Aucun effet spécifique n'a été observé après stimulation de la représentation corticale motrice droite du visage, gauche de la main ou encore dans les conditions non verbales (sons non verbaux et mouvements des yeux ; voir aussi les résultats de Watkins & Paus, 2004). En utilisant l'IRMf cette fois, Pulvermüller et al. (2006) ont démontré que la simple perception de phonèmes activait non seulement le cortex temporal supérieur mais aussi le gyrus précentral (principalement à gauche). En outre, cette activation variait en fonction du lieu d'articulation des phonèmes et respectait l'organisation somatotopique⁹ du cortex moteur, observée lors de l'articulation silencieuse de ces mêmes phonèmes et de mouvements non verbaux des articulateurs concernés. Ainsi, alors que l'écoute des phonèmes bilabiaux /p/ et /b/ a activé les régions motrices ventrales recrutées lors de l'articulation de ces mêmes sons, correspondant à la représentation motrice des lèvres, la perception de consonnes apico-dentales (/d/ et /t/) a suscité une activation corticale encore plus ventrale, au niveau de la représentation motrice de la langue (Figure 4 ; mais voir Arsenault & Buchsbaum, 2016 pour des résultats contradictoires). Le même type de résultats a été décrit par Wilson et collaborateurs (2004) en IRMf lors de la production et de la perception de syllabes (activation des aires motrice et prémotrice BA 4 et 6 ; voir aussi Skipper, Goldin-Meadow, Nusbaum, & Small, 2007; Skipper, Nusbaum, & Small, 2005 pour la parole audio-visuelle).

⁹ Une somatotopie fonctionnelle a été décrite par Penfield et Rasmussen (1952) à la surface du cortex moteur. Tout au long de ce cortex, il existe une représentation déformée du corps humain : la stimulation électrique de différents points corticaux produit des mouvements des parties corporelles correspondantes. Chacune des régions motrices représente un territoire corporel et sa surface est d'autant plus grande que la partie du corps considérée est complexe du point de vue de sa motricité. Les zones correspondant à la main et à la bouche sont particulièrement développées. Cette représentation déformée du corps humain à la surface du cortex moteur est appelée l'*homunculus moteur* de Penfield.

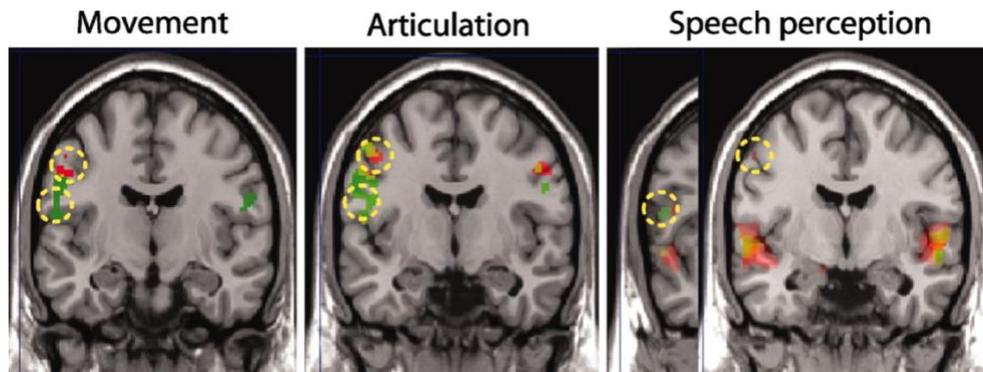


Figure 4 : Coupes frontales montrant le recouvrement d'activation somatotopique dans les régions motrices lors de l'exécution de mouvements et de la production et de la perception de phonèmes. A gauche, activité cérébrale engendrée par le mouvement des lèvres (en rouge) ou de la langue (en vert). Au centre, activité cérébrale liée à la production silencieuse de syllabes incluant une attaque consonantique bilabiale (/p/, en rouge) ou apico-dentale (/t/, en vert). A droite, activité cérébrale enregistrée lors de l'écoute de ces mêmes syllabes. Figure empruntée à Pulvermüller et al. (2006).

Dans l'ensemble, ces études révèlent donc une résonance motrice lors de la perception de la parole, étayant la théorie motrice selon laquelle les phonèmes sont reconnus et compris lorsque locuteur et auditeur partagent un répertoire articulo-moteur commun. Toutefois, elles ne permettent pas de décider si le cortex moteur joue un rôle fonctionnel dans la perception de la parole.

Des travaux en TMS ont permis de fournir des éléments de réponse à cette question, en démontrant que la stimulation du cortex prémoteur peut influencer la perception des sons de parole. L'étude *princeps* de Meister et al. (Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007) a par exemple permis de montrer que la discrimination de consonnes dans des syllabes CV (/pa/, /ta/, /ka/) masquées par du bruit blanc était significativement diminuée après stimulation du cortex prémoteur ventral gauche (impliqué dans la perception et la production de la parole) par TMS répétée¹⁰ (qui inactive temporairement l'aire stimulée et crée une « lésion virtuelle »). Aucune interférence n'était observée dans une tâche contrôle de discrimination de couleurs. D'Ausilio et collaborateurs (D'Ausilio et al., 2009) ont par la suite montré la spécificité somatotopique de ces effets dans une tâche de discrimination de phonèmes (consonnes bilabiales /p/ et /b/ ou apico-dentales /d/ et /t/ suivies de la voyelle /æ/ et

¹⁰ La TMS répétée (rTMS) consiste à délivrer un « train » de stimulations répétées à un niveau supraliminal (i.e. au-dessus du seuil moteur), ce qui supprime l'excitabilité de la région corticale ciblée (Pascual-Leone, Bartres-Faz, & Keenan, 1999 ; Pascual-Leone, Walsh, & Rothwell, 2000). En d'autres termes, une « lésion virtuelle » de cette région est créée, pouvant ou non interférer avec les processus cognitifs mis en jeu dans la tâche examinée.

masquées par du bruit blanc) où une stimulation en TMS (double impulsion excitatrice¹¹) était délivrée sur l'aire motrice primaire gauche de la langue ou des lèvres (Figure 5a). Les auteurs ont mis en évidence l'existence d'une double dissociation entre le site de la stimulation et la catégorie du stimulus. Lorsque l'aire corticale motrice de la langue était stimulée, la discrimination des consonnes dentales (qui impliquent une participation de la langue pour être produites) était plus rapide et donnait lieu à de meilleurs scores, tandis que lorsque la stimulation était délivrée sur l'aire corticale motrice de la bouche, c'est la discrimination des bilabiales (dont la production requiert un rapprochement des lèvres inférieure et supérieure) qui était facilitée (Figure 5b et 5c). Cette double dissociation fournit des éléments en faveur de la contribution du cortex moteur à la perception des sons de parole. D'autres études en TMS ont corroboré ces résultats dans des tâches d'identification ou de discrimination de phonèmes (Mottonen & Watkins, 2009) mais aussi dans des tâches phonologiques plus complexes impliquant la mémoire de travail verbale et les processus de segmentation phonologique (Romero, Walsh, & Papagno, 2006; Sato, Tremblay, & Gracco, 2009 ; voir Iacoboni, 2008; Möttönen & Watkins, 2011; Murakami, Ugawa, & Ziemann, 2013 pour des revues de littérature) ou enfin dans des tâches de compréhension (appariement mot-image avec des mots commençant par des bilabiales ou des dentales; Schomers, Kirilina, Weigand, Bajbouj, & Pulvermüller, 2015).

¹¹ La TMS à double impulsion consiste en la décharge de deux stimulations (une stimulation de conditionnement suivie d'une stimulation test) séparées par un intervalle inter-stimuli (ISI) long (> 6-7 ms) ou court (< 6-7 ms). Un ISI court va inhiber la région ciblée, alors qu'un ISI long augmentera l'excitabilité de cette région (Valero-Cabré et al., 2011).

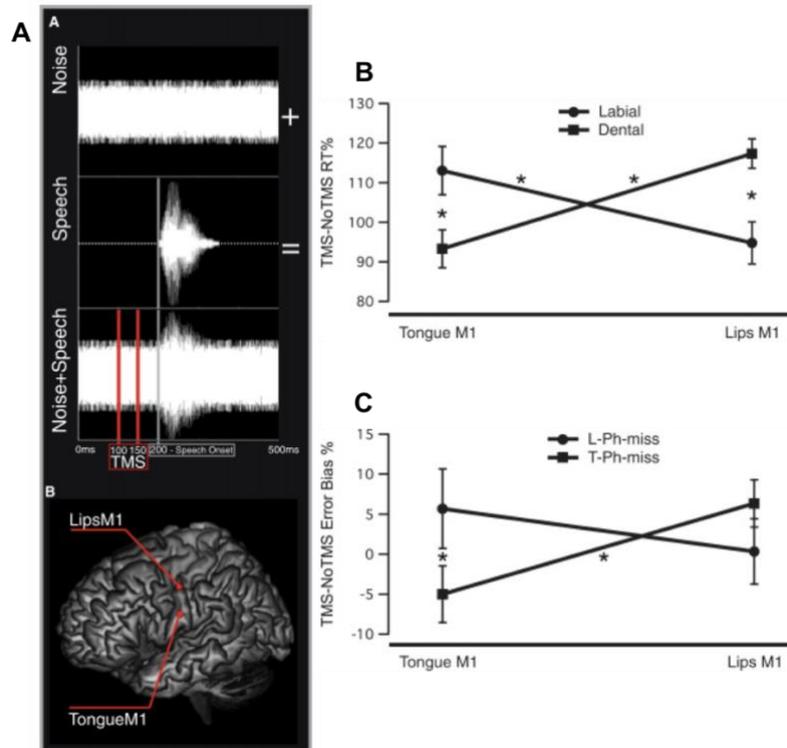


Figure 5 : Résultats de l'expérience de D'Ausilio et al (2009). (A) Stimuli expérimentaux utilisés dans l'étude (bruit, parole et parole masquée par du bruit) et régions stimulées en TMS double impulsion (régions motrices gauches des lèvres et de la langue). (B) Taux de changement (%) des temps de réponse après stimulation par TMS (i.e. ratio entre les temps de réponse après TMS et sans TMS) et (C) Taux de changement (%) des pourcentages de bonnes réponses dans une tâche de discrimination de consonnes dentales ou labiales en fonction de la région motrice stimulée (langue ou lèvres). Figures empruntées à D'Ausilio et al. (2009).

Si ces études révèlent que le cortex pré/moteur est activé lors de la perception des phonèmes et qu'il peut même l'influencer, suggérant qu'il joue un rôle causal, la question de la nécessité de cette activation motrice reste controversée (pour des discussions, Hickok, 2009, 2014; Hickok, Houde, & Rong, 2011; Pulvermüller & Fadiga, 2010; Rogalsky, Love, Driscoll, Anderson, & Hickok, 2011; Schwartz et al., 2012). La plupart des études ayant montré des liens entre perception de la parole et cortex moteur ont en effet utilisé des stimuli dégradés (i.e. parole vocodée, bruitée, compressée temporellement), suggérant que le système moteur intervienne dans la perception de la parole particulièrement dans des conditions d'écoute difficiles.

Dans une étude en IRMf combinant trois formes de parole dégradée (parole vocodée, parole segmentée en intervalles réguliers et parole masquée par du bruit blanc) comparées à

de la parole normale, Davis et Johnsrude (2003) ont observé une augmentation de l'activité cérébrale pour les trois types de dégradation au sein des régions frontales inférieures et prémotrices gauches, suggérant une implication des régions de la production de la parole dans le décodage du signal de parole dégradé. Adank et Devlin (2010) ont quant à eux étudié en IRMf l'effet du débit de parole (normal ou accéléré artificiellement par compression temporelle) sur les processus de perception de la parole chez des adultes néerlandais sains. Dans une tâche de jugement de plausibilité de phrases présentées d'abord à débit normal, puis à débit rapide, ils ont observé des modifications de l'activité neuronale spécifiques lors du traitement de la parole compressée temporellement. Au sein du cortex auditif associatif droit, le STS (sillon temporal supérieur) antérieur et postérieur s'activait fortement lors du premier bloc de stimuli (les 16 premières phrases) puis l'activation déclinait sur les trois derniers blocs sans toutefois revenir au niveau d'activation observé pour la parole produite à débit normal. Au niveau de l'hémisphère gauche, les auteurs ont relevé une activation du cortex prémoteur ventral et du STS postérieur lors de la perception du premier bloc de stimuli de parole compressée, suivie d'un déclin de cette activation pour les blocs suivants. Ces zones ne s'activaient en revanche que peu et de manière stable au cours des quatre blocs lors de la perception de parole à débit normal. Un processus d'adaptation spécifique à la perception de parole rapide (compressée) a donc été mis en évidence dans cette étude d'imagerie cérébrale (en accord avec les résultats comportementaux d'Adank & Janse, 2009). Les auteurs ont proposé que ce processus implique un changement de sensibilité aux indices acoustiques mais aussi articulatoires résultant en une « recalibration » des frontières phonémiques et en l'activation spécifique du cortex prémoteur, afin de s'adapter à l'augmentation du débit et aux phénomènes de réduction concomitants. Des arguments en faveur de l'implication du cortex moteur dans la perception de parole dégradée ont également été fournis par plusieurs autres études (Du, Buchsbaum, Grady, & Alain, 2014; Hervais-Adelman, Carlyon, Johnsrude, & Davis, 2012; Murakami, Kell, Restle, Ugawa, & Ziemann, 2015; Murakami, Restle, & Ziemann, 2011; Osnes, Hugdahl, & Specht, 2011). Osnes et collègues (2011) ont par exemple utilisé la technique de « *morphing* » pour créer un continuum allant d'un bruit blanc à une syllabe (/da/ ou /ta/) ou à un son d'instrument de musique, et demandé à des participants de déterminer quand ils percevaient les stimuli comme étant de la parole, de la musique ou du bruit. Dans un deuxième temps, leur activité cérébrale était mesurée en IRMf lors de l'écoute des mêmes stimuli. Leurs résultats en IRMf ont révélé une activation du cortex prémoteur gauche uniquement lorsque la syllabe était dégradée mais reconnaissable en tant que stimulus verbal, ce qui n'était pas observé pour les sons non verbaux. Des analyses de connectivité

effective ont par ailleurs montré une connectivité bidirectionnelle entre les régions prémotrices et temporales supérieures, et du *planum temporale* au cortex prémoteur, lors de l'écoute des sons de parole uniquement (Figure 6). Ainsi, le cortex prémoteur n'était pas sollicité lorsque des syllabes intactes étaient perçues. Ceci s'accorde avec les résultats de D'Ausilio et collègues (2012) ayant démontré une influence de la stimulation du cortex prémoteur gauche par TMS sur la discrimination de phonèmes uniquement en condition bruitée. Mukarami et collaborateurs (2011) ont également rapporté qu'une diminution de l'intelligibilité de la parole (par du bruit) augmentait l'amplitude des potentiels évoqués moteurs enregistrés sur les muscles des lèvres après stimulation par TMS sur la représentation corticale motrice de cet effecteur. Le système moteur est donc plus fortement activé lorsque la parole est difficile à identifier (voir aussi Chevillet, Jiang, Rauschecker, & Riesenhuber, 2013 pour de la catégorisation en situation d'écoute dichotique). C'est aussi le cas pour les phonèmes n'appartenant pas à notre langue maternelle comme l'a révélé une étude en IRMf de Wilson et Iacoboni (2006). Les auteurs ont en effet mis en évidence que l'écoute de phonèmes de langues très différentes (*e.g.*, birman, hindi, russe) engendrait, chez des participants anglais monolingues, une activation plus importante du cortex prémoteur ventral en regard de l'écoute de phonèmes natifs (voir aussi Callan, Jones, Callan, & Akahane-Yamada, 2004).

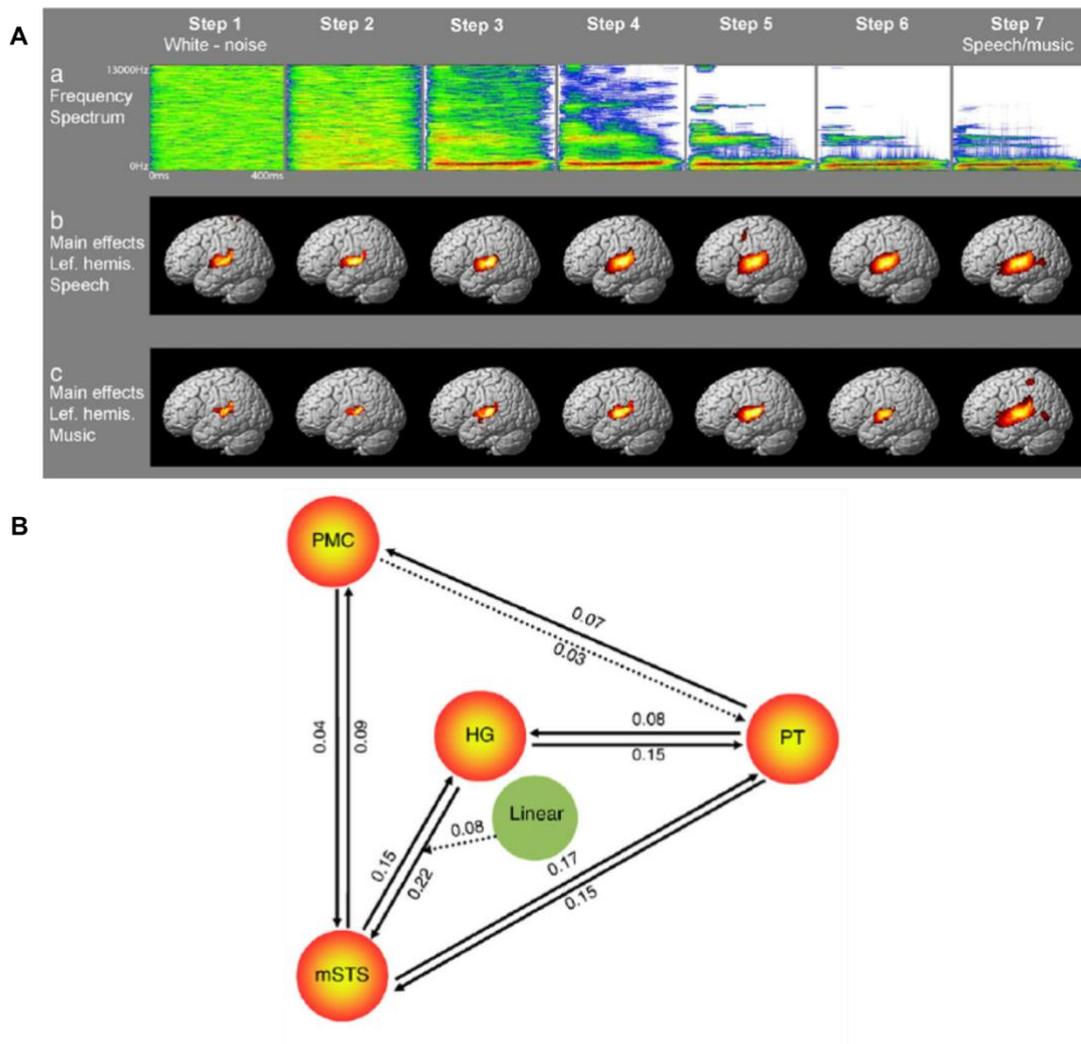


Figure 6 : Résultats de l'étude de Osnes et al. (2001). (A) En haut, spectre du continuum créé par « morphing », allant du bruit blanc à la syllabe /da/ ou à une note de musique. L'étape 5 correspond au niveau où les sujets perçoivent le stimulus comme étant de la parole bruitée. En bas et au milieu, activations corticales au sein de l'hémisphère gauche en IRMf pour chaque stimulus issu du continuum, pour de la parole (milieu) ou de la musique (bas). Le cortex prémoteur s'active à la 5^{ème} étape pour la parole, soit lorsque les sujets perçoivent le stimulus comme de la parole dégradée. (B) Modèle de causalité dynamique (*Dynamic Causal Modeling*), à partir des données obtenues en IRMf. La connectivité effective entre le gyrus de Heschl (HG), le *planum temporale* (PT), le sillon temporal supérieur médian (mSTS) et le cortex prémoteur (PMC) a été testée pour le signal de parole. Les chiffres dérivent de la moyenne Bayésienne des modèles (*Bayesian Model Averaging*). Figures empruntées à Osnes et al. (2001).

Le schéma d'activation corticale mis en évidence dans ces différentes études suggère donc que le système moteur serait préférentiellement impliqué dans la perception de la parole dans des conditions d'écoute difficiles afin d'aider les régions auditives dans les processus de

perception (Wilson, 2009). Plutôt que d'être nécessaire, le cortex pré/moteur serait donc un système complémentaire de soutien à la perception de la parole (Hickok et al., 2011). Ainsi, lorsque le signal phonétique est incomplet ou de mauvaise qualité, le système moteur impliqué dans la production de la parole serait plus spécifiquement recruté, et interagirait plus fortement avec les régions temporelles auditives, afin d'aider à la catégorisation phonémique et de ce fait à la compréhension du message.

4. Modèles neurobiologiques de la perception de la parole

En accord avec les études démontrant une implication des régions motrices lors de la perception de parole, les modèles neurobiologiques récents postulent un lien étroit entre les systèmes de perception et de production de la parole. Le système de production de la parole permettrait ainsi de contraindre le traitement phonétique de l'information de parole et d'en faciliter sa perception et sa catégorisation. Dans la suite, nous présenterons deux modèles implémentant une partie des processus de perception de la parole dans le système moteur.

Le modèle à double voie ventrale/dorsale proposé par Hickok et Poeppel (Hickok & Poeppel, 2000, 2004, 2007) postule l'existence de deux voies corticales de traitement de la parole, à l'image du traitement visuel (Mishkin, Ungerleider, & Macko, 1983). Les auteurs distinguent ainsi une voie lexico-sémantique ventrale (voie du « quoi ») orientée sur la reconnaissance de la parole et une voie d'intégration sensori-motrice dorsale (voie du « comment ») qui constituerait une interface sensori-motrice permettant d'apparier les représentations phonologiques et articulatoires (Figure 7).

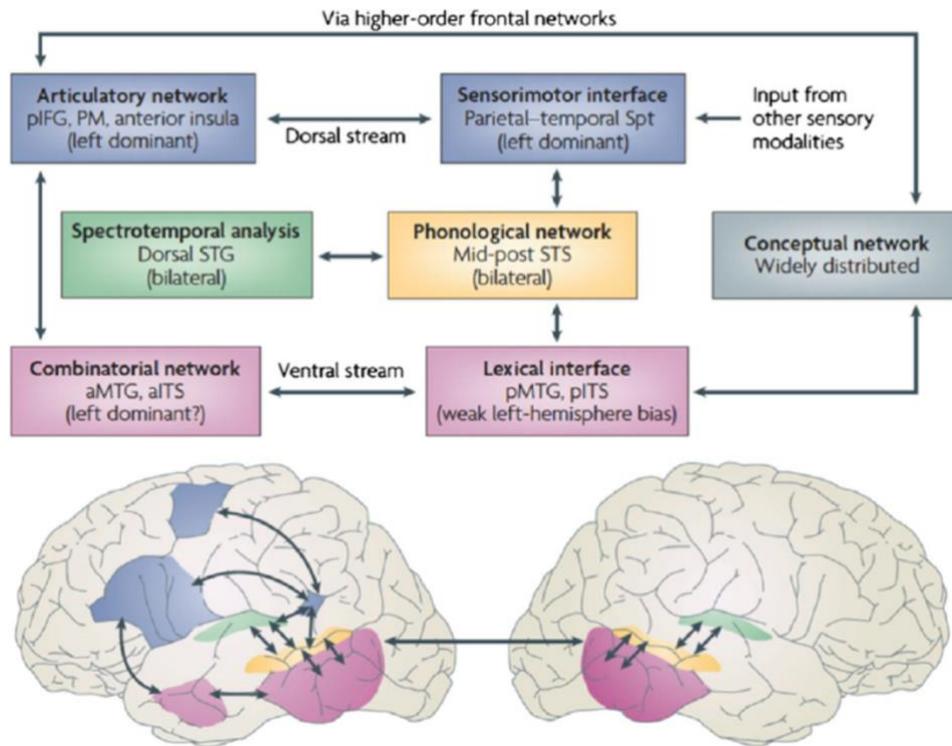


Figure 7 : Modèle à double voie ventrale/dorsale de la perception de la parole. Figure empruntée à Hickok et Poeppel (2007).

Dans ce modèle, le signal de parole est d'abord traité au sein des cortex auditifs primaires et secondaires, dans la partie dorsale moyenne des gyri temporaux supérieurs (STG), où une analyse spectro-temporelle est réalisée. Le flux sensoriel est ensuite converti en une représentation phonologique grâce au décodage acoustico-phonétique et au traitement phonologique mis en œuvre dans la partie postérieure des sillons temporaux supérieurs (STS). Puis l'information phonologique est prise en charge par les deux voies, ventrale et dorsale, avec une possible prédominance d'une voie sur l'autre selon la tâche à effectuer. La voie ventrale, bilatérale avec une prédominance hémisphérique gauche dans sa partie antérieure, permettrait la reconnaissance de la parole en tant que telle et la transformation d'une information phonologique en une information porteuse de sens (lexico-sémantique). La voie dorsale, latéralisée à gauche, serait quant à elle impliquée dans les appariements auditivo-articulatoires.

La voie ventrale comprend deux parties : la portion moyenne et inférieure du gyrus/sillon temporal postérieur, engagée dans le couplage des informations phonologiques et sémantiques (désignée comme l'interface lexicale dans le modèle), et la portion antérieure de

ces régions, plutôt latéralisée à gauche, en charge des processus combinatoires d'intégration sémantique.

La voie dorsale est également subdivisée en deux composantes, avec une interface sensori-motrice au niveau de la partie postérieure de la scissure de Sylvius (jonction temporo-pariétale : aire Spt), et un réseau articuloire au niveau du lobe frontal gauche, incluant l'aire de Broca et les régions prémotrices plus dorsales. Le réseau articuloire, abritant probablement des neurones miroirs, contribuerait à la récupération des représentations articuloires lors de la perception de la parole. Hickok et Poeppel ont proposé que la voie dorsale jouerait un rôle crucial au cours du développement de la parole chez l'enfant, et en particulier lors de la période de babillage, puisqu'elle permettrait d'ajuster les sons de parole produits par l'enfant aux prototypes de phonèmes entendus autour de lui (Doupe & Kuhl, 1999; Pulvermüller & Fadiga, 2010 pour une revue). En revanche, selon les auteurs, cette voie ne serait pas indispensable à la perception de la parole chez l'adulte même si elle continuerait de jouer un rôle, notamment dans les phénomènes de *feedback* auditif ou d'apprentissage de vocabulaire. Elle serait aussi impliquée dans les tâches méta-phonologiques désignées comme « non écologiques » (i.e. non utilisées naturellement lors de la perception de la parole), c'est-à-dire lorsque l'auditeur doit volontairement utiliser des processus moteurs pour garder en mémoire des représentations auditives (tâches phonologiques ou de mémoire de travail ; Demonet et al., 2005; Poldrack et al., 1999). Enfin, comme nous l'avons exposé précédemment, cette voie dorsale appariant les représentations auditives et motrices de la parole serait celle sollicitée lorsque les conditions ne sont pas optimales telles que dans un environnement bruité.

Une étude de Saur et al. (2008) en IRM de diffusion¹² (*Diffusion Tensor Imaging* ou DTI) a permis de révéler la réalité anatomique et fonctionnelle de ce modèle à double voie. Les auteurs ont en effet démontré que la répétition de pseudo-mots, impliquant un couplage auditivo-moteur, engageait une voie dorsale connectant le gyrus temporal supérieur postérieur à la *pars opercularis* du gyrus frontal inférieur (BA 44, aire de Broca) et au cortex prémoteur dorsal (BA 6) *via* le faisceau arqué et le faisceau longitudinal supérieur. L'écoute de phrases activait quant à elle une voie plus ventrale reliant le gyrus temporal médian postérieur et antérieur au gyrus frontal inférieur (*pars triangularis*, BA 45 et *pars orbitalis*, BA 47) *via* la

¹² L'IRM de diffusion est une technique permettant de calculer la direction de diffusion de molécules d'eau dans le cerveau et ainsi de mettre en évidence les structures fibreuses, en particulier les faisceaux de substance blanche.

capsule extrême (territoire de la substance blanche centrale) et le faisceau longitudinal inférieur.

Le modèle à double voie antérieure/postérieure est un autre modèle proposé par Rauschecker et ses collaborateurs (Rauschecker, 2011; Rauschecker & Scott, 2009) à partir d'études chez le primate non humain. Il stipule l'existence d'une voie antéro-ventrale (voie du « quoi ») impliquée dans l'identification des objets auditifs et la compréhension de la parole et incluant le cortex temporal et frontal inférieur, et d'une voie postéro-dorsale (voie du « où ») incluant le cortex prémoteur et le cortex pariétal inférieur, sous-tendant la localisation des sources auditives mais aussi l'appariement entre représentations auditives et articulatoires, en particulier lorsque la parole est difficile à comprendre. Dans leur modèle, les auteurs suggèrent en outre l'existence de modèles ou simulations internes *top-down* qui soutendraient le traitement de la parole (Figure 8).

La voie antérieure participerait aux processus de décodage acoustico-phonétique réalisés au niveau de la partie antérieure du STG et des régions frontales inférieures (BA 45) contenant les représentations phonétiques invariantes. Les informations sortantes seraient ensuite envoyées vers le cortex prémoteur ventral et l'aire de Broca (BA 44) où elles seraient appariées avec les représentations articulatoires correspondantes. Selon un mécanisme de codage prédictif (ou *forward mapping*), les commandes motrices nécessaires à la production des phonèmes seraient alors simulées et leurs conséquences sensorielles seraient ensuite transmises à l'interface sensori-motrice dans le lobule pariétal inférieur sous forme de copies d'efférence. Ces copies d'efférence dites prédictives y seraient alors comparées aux représentations phonologiques correspondant à l'information réellement perçue (voir aussi et Hickok et al., 2011). Dans la même lignée, Wilson et Iacoboni (2006) ont proposé que le cortex prémoteur génère une simulation (codage prédictif ou *forward*) de la production des phonèmes entendus, qui permettrait de prédire les conséquences acoustiques des sons de parole perçus. La comparaison entre cette simulation et l'information analysée dans le cortex temporal supérieur produirait un signal d'erreur renvoyé au cortex prémoteur de sorte à affiner la prédiction (Figure 9). Un tel mécanisme faciliterait la catégorisation phonémique en désambiguïsant l'information de parole.

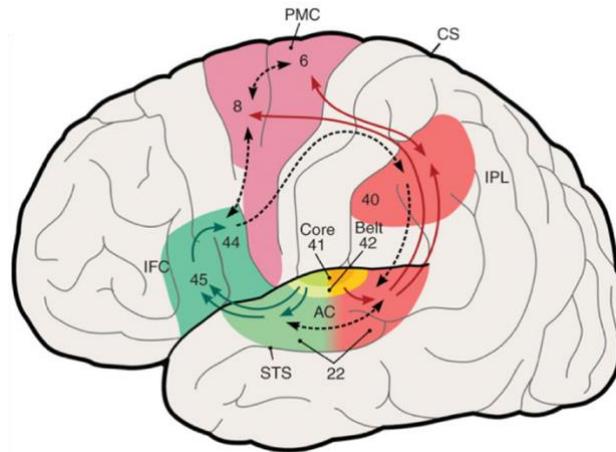


Figure 8 : Modèle à double voie antérieure/postérieure. Voies antéro-latérale (vert) et postéro-dorsale (rouge). Ces voies partent des régions auditives primaires (jaune). La voie postéro-dorsale interagit avec le cortex prémoteur (PMC) afin de comparer l'information auditive entrante avec les représentations articulatoires motrices au sein du lobule pariétal inférieur (IPL). Figure empruntée à Rauschecker & Scott (2009).

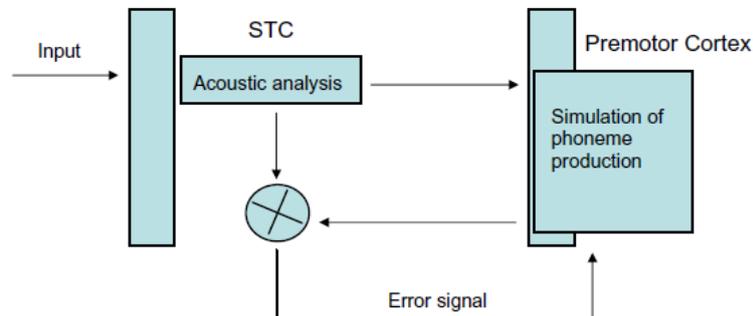


Figure 9 : D'après Wilson et Iacoboni (2006), le cortex temporal supérieur (STC) serait en charge de l'analyse acoustique du signal de parole alors que le cortex prémoteur établirait une simulation de la production des phonèmes entendus. Ce modèle interne *forward* permettrait de générer une prédiction des conséquences acoustiques de la production des phonèmes qui serait comparée à l'information traitée dans le cortex temporal supérieur. Cette comparaison produirait un signal d'erreur renvoyé au cortex prémoteur, qui corrigerait la simulation en vue de la catégorisation phonémique. Figure empruntée à Iacoboni (2008).

Dans ce deuxième chapitre, nous avons défini les processus généraux mis en jeu lors de la perception de la parole, des bases cérébrales aux derniers modèles neurobiologiques proposés pour rendre compte de la manière dont est traitée la parole dans le cerveau. Une des caractéristiques qui se dégage de ces modèles et études est le rôle primordial du système moteur dans la perception de la parole, notamment dégradée. Ainsi, les unités perceptives de la parole seraient sensori-motrices plus que purement auditives ou purement motrices. Dans le chapitre III, nous allons maintenant nous intéresser aux rythmes cérébraux et à leurs liens avec le signal de parole.

III. Les rythmes cérébraux

1. Les oscillations cérébrales

Une des propriétés fondamentales de l'activité cérébrale est la présence de fluctuations rythmiques à différentes fréquences. Ces rythmes cérébraux, ou oscillations¹³ cérébrales, résultent de l'activité électrique périodique synchrone (ou cohérente) d'une population de neurones à une certaine fréquence (Berger, 1929). Il a été proposé que les populations de neurones synchroniseraient leur activité de manière dynamique, comme une pulsation, et les oscillations reflèteraient les variations cycliques de l'excitabilité neuronale (i.e. alternance entre phases de haute excitabilité (hyperpolarisation) et de basse excitabilité (dépolariation) ; Buzsaki & Draguhn, 2004; Lakatos et al., 2005). Les oscillations ont été décrites dans le cerveau de nombreux mammifères, avec des modulations très lentes (période mesurée en minutes) jusqu'à des rythmes extrêmement rapides pouvant atteindre les 600 Hz (Buzsáki, Logothetis, & Singer, 2013).

L'activité rythmique du cerveau est observée aussi bien au repos – elle est donc intrinsèque, spontanée (i.e. indépendante de toute stimulation extérieure) – que lors d'une activité cognitive, perceptive ou sensori-motrice. Suite aux travaux *princeps* de Hans Berger en 1929, les experts de la *International Federation of Societies for Electroencephalography and Clinical Neurophysiology* (1974) ont distingué 5 rythmes cérébraux principaux en fonction de la fréquence à laquelle les neurones oscillent : les oscillations **delta** (0.5-4 Hz), **thêta** (4-8 Hz), **alpha** (8-13 Hz, où on distingue le rythme dit alpha-variant **mu**), **beta** (13-30 Hz) et **gamma** (30-90 Hz, où on distingue dans les techniques non-invasives la bande gamma basse 30-60 Hz, et haute 60-90 Hz)¹⁴. Selon l'origine anatomique (et parfois le contexte expérimental), à chaque bande de fréquence correspondent différents types d'activités ou états

¹³ Une oscillation est une variation d'une grandeur mécanique, électrique, caractérisée par un changement périodique de sens autour d'une position d'équilibre stable, passant successivement par une valeur maximale et une valeur minimale. Les oscillations peuvent être régulières (périodiques) ou décroissantes (amorties). Le cycle d'une oscillation est le temps écoulé entre deux passages successifs par la position d'équilibre. La fréquence d'une oscillation est le nombre de cycles par seconde et est exprimée en Hertz (Hz).

¹⁴ Il est à noter que les frontières entre les bandes de fréquence ont été définies quelque peu arbitrairement et peuvent donc varier d'une étude à l'autre (par exemple, la bande beta peut aller de 12 à 30 Hz ou de 15 à 30 Hz ; la frontière supérieure de la bande thêta peut également être définie à 7 ou 8 Hz). De plus, en EEG intracérébral (iEEG), on peut rapporter des activités oscillatoires allant jusqu'à 200 Hz. Cependant la dénomination de ce rythme oscillatoire est encore débattue, et notamment quant au fait de savoir s'il s'agit d'oscillations soutenues à toutes ces fréquences ou d'une activité large bande (*broadband gamma* ; Jerbi et al., 2009; Worrell et al., 2012).

cognitifs (Engel, Fries, & Singer, 2001; Tallon-baudry & Bertrand, 1999). Par exemple, les oscillations delta sont observées lors du sommeil lent profond, et les oscillations thêta lors des états de somnolence (somnolence diurne, méditation...) et de mémorisation. Les ondes alpha, lorsqu'elles sont observées au niveau occipito-pariétal, correspondent à un état éveillé diffus, yeux clos et détendu (i.e. au repos), et joueraient un rôle important dans les processus d'inhibition fonctionnelle, de somnolence et d'attention visuo-spatiale. Les oscillations dans cette même bande de fréquence mais observées au niveau du cortex sensori-moteur caractérisent notamment les activités sensori-motrices et sont appelée rythme mu. Les ondes beta présentes au niveau des régions frontales s'observent par exemple en période d'éveil actif ou lors d'un état de concentration intense ou d'un état émotionnel exagéré. Enfin, les ondes gamma sont attribuées à de nombreuses fonctions cérébrales. Elles sont particulièrement impliquées dans les phénomènes d'attention et de liage perceptif (*e.g.*, représentation des objets ; Tallon-Baudry & Bertrand, 1999) ainsi que dans la communication locale ou entre des aires cérébrales éloignées (notamment par couplage avec des bandes de fréquence plus faibles au niveau local, voir Bonnefond, Kastner, & Jensen, 2017 pour un modèle unifié liant les oscillations alpha et gamma ; Fries, 2005, 2015).

Les oscillations cérébrales peuvent être mesurées à l'aide de différentes techniques d'imagerie fonctionnelle (MEG, EEG, iEEG, IRMf, etc). Cependant, deux techniques non invasives sont particulièrement utilisées, permettant d'offrir une excellente résolution temporelle (i.e. permettant de suivre le décours temporel des activations à la milliseconde près), il s'agit de l'électroencéphalographie (EEG, définie en note de bas de page dans le chapitre II.4.) et la magnétoencéphalographie (MEG).

La magnétoencéphalographie (Figure 10), que nous avons utilisée dans ce travail de recherche, est une technique proche de l'EEG, qui enregistre les variations du champ magnétique produit par l'activité électrique des neurones au repos ou lors de la réalisation d'une tâche. Les champs magnétiques du cerveau étant infiniment plus faibles que le champ magnétique terrestre, la MEG nécessite une chambre blindée où est installé un casque contenant environ 300 capteurs spéciaux ultrasensibles et supraconducteurs (les *SQUIDS*¹⁵, 275 pour le modèle CTF que nous avons utilisé) permettant de couvrir l'ensemble du scalp d'un sujet. Les capteurs sont plongés dans de l'hélium liquide à -269°C afin de garantir leur

¹⁵*Superconducting quantum interference device.*

supraconductivité et un faible niveau de bruit. La MEG offre une excellente résolution temporelle et une meilleure résolution spatiale (quelques millimètres) que l'EEG (contrairement aux champs électriques, les flux magnétiques ne sont en effet pas déformés par les tissus traversés comme la peau, l'activité sous-jacente est donc plus facilement localisable), permettant de cartographier très précisément l'activité cérébrale. L'accès aux sources est plus fiable, *via* notamment la résolution du problème inverse qui correspond à l'estimation des décours temporels des générateurs de l'activité mesurée au niveau du scalp. C'est la raison pour laquelle nous avons choisi d'utiliser cette technique d'imagerie cérébrale dans ce travail de thèse.



Figure 10 : La magnétoencéphalographie.

2. La synchronisation oscillatoire

L'étude de la dynamique des oscillations cérébrales lors des processus de traitement de l'information a connu un essor important ces vingt dernières années et il est aujourd'hui largement admis que la synchronisation oscillatoire des populations neuronales joue un rôle important dans l'intégration sensorielle et la cognition (Engel et al., 2001; Tallon-baudry, Bertrand, & Fischer, 2001; Varela, Lachaux, Rodriguez, & Martinerie, 2001), et en particulier dans le traitement de la parole comme nous le verrons par la suite. « *Neuronal oscillations have an important role in brain operations to the extent that understanding of neuronal oscillation 'rhythms' now seems to be essential to our understanding of brain function* »

(Schroeder & Lakatos, 2009). La synchronisation des oscillations, en permettant la coopération au sein d'une même aire cérébrale ou entre différentes aires impliquées dans un même processus, constituerait ainsi le langage utilisé par les neurones pour traiter l'information plus efficacement (Fries, 2005, 2015). Autrement dit, la synchronisation sous-tendrait la communication entre les groupes de neurones, elle serait un mécanisme de codage de l'activité cérébrale. Dans cette optique, les pics/sommets rythmiques oscillatoires constitueraient des fenêtres temporelles pour la communication : « *Only coherently oscillating neuronal groups can communicate effectively, because their communication windows for input and for output are open at the same times* » (Fries, 2005).

La synchronisation mesure la relation entre la structure temporelle de deux signaux indépendamment de leur amplitude (Varela et al., 2001) : « *two signals are said to be synchronous if their rhythms coincide* ». Elle correspond à l'alignement entre les pics du cycle oscillatoire d'une population de neurones avec les pics du cycle d'une autre population : on parle alors de synchronisation de phase (la phase d'un signal comporte toutes les informations temporelles de l'activité neuronale : elle détermine le degré d'excitabilité des neurones et influence le moment où ils déchargent). En d'autres termes, la fréquence d'oscillation et la phase entre deux signaux doivent coïncider pour qu'il y ait synchronisation oscillatoire. Cette synchronisation permettrait à différentes assemblées de neurones de décharger en même temps lors de fenêtres de communication optimale, favorisant le traitement de l'information (Fries, 2015; Varela et al., 2001). En plus de la synchronisation de phase, la communication neuronale serait également sous-tendue par un processus de couplage inter-fréquences (*e.g.*, entre la phase des oscillations thêta et l'amplitude des oscillations gamma ; Canolty & Knight, 2010; Fries, 2005; Jensen & Colgin, 2007; Womelsdorf et al., 2007).

Lorsqu'on parle de synchronisation, il faut savoir que dans la littérature, on retrouve différentes formes de synchronisation : la synchronisation locale et la synchronisation à large échelle (Varela et al., 2001).

La synchronisation locale correspond à la résonance de réseaux de neurones distribués localement sur une surface égale ou inférieure à 1 cm à travers des connexions monosynaptiques présentant un délai de conduction d'environ 4-6 ms. Lorsqu'on mesure ce type de synchronisation, on mesure la puissance des potentiels de champ locaux, qui peut alors augmenter (synchronisation des réseaux de neurones locaux) ou diminuer (désynchronisation). Les oscillations rapides, telles que celles retrouvées dans la bande de

fréquence gamma, seraient particulièrement impliquées dans la synchronisation sur des échelles spatiales courtes.

La synchronisation à large échelle correspond quant à elle à la résonance de réseaux de neurones plus éloignés dans le cerveau (> 1 cm de distance, délais de transmission supérieurs à 8-10 ms). Ce type de synchronisation repose donc sur des connexions distantes, soit cortico-corticales soit thalamo-corticales mais aussi entre un signal cortical et un signal extérieur. Le rythme lent thêta se prête particulièrement à ce type de synchronisation sur de longues distances.

Dans cette thèse, nous emploierons les termes synchronisation ou couplage pour désigner des phénomènes de synchronisation à large échelle.

Enfin, différentes mesures de synchronisation peuvent être réalisées : des mesures de cohérence ou de couplage de phase (Bowyer, 2016).

La cohérence est une mesure d'interaction qui se base sur l'estimation de la dépendance linéaire entre deux signaux, dans le domaine des fréquences. Il s'agit donc de l'équivalent, dans le domaine fréquentiel, d'une mesure de corrélation dans le domaine temporel. La cohérence a des valeurs comprises entre 0 et 1. Une relation parfaitement linéaire entre deux signaux à une fréquence donnée se traduit par une valeur de cohérence égale à 1. Une absence totale de dépendance linéaire se traduit par une cohérence de 0. Le calcul de la cohérence se base sur le calcul de matrice de densité spectrale (ou inter-spectre) ; elle ne permet pas de distinguer les contributions de l'amplitude et de la phase (Buzsáki, 2006).

Le couplage de phase ou phase-locking désigne l'interaction mutuelle entre deux oscillateurs ; cette mesure est en théorie indépendante des changements d'amplitude des signaux. Elle calcule la stabilité de la différence de phase entre deux signaux sur une courte durée. Le couplage de phase est ainsi un indicateur de la relation de phase dynamique entre deux réseaux indépendamment de leur amplitude (Varela et al., 2001).

Jusqu'à présent, nous avons vu que les populations neuronales se synchronisent entre elles, sur de courtes ou de plus longues distances, pour mieux communiquer lors du traitement de l'information. Une synchronisation à large échelle « extra-corticale » peut également être observée : les oscillations cérébrales peuvent en effet aligner leur phase sur la phase d'un stimulus externe rythmique, périodique ou quasi-périodique (*e.g.*, mouvements biologiques

tels que ceux de la main, parole, musique, etc. ; voir par exemple, Gross et al., 2013 et Jerbi et al., 2007) de sorte à optimiser le traitement. Ce type de synchronisation est désigné sous le terme d' « entrainement » (*entrainment* ; Lakatos et al., 2005; Schroeder & Lakatos, 2009).

Lors de l'arrivée d'un stimulus externe possédant une structure (quasi-)rythmique, une réponse neuronale est générée sous forme de potentiels d'action dans les régions impliquées dans le traitement de ce stimulus. La qualité de cette réponse, et par conséquent la qualité du traitement de l'information, va alors dépendre de l'état oscillatoire dans lequel se trouve le système neuronal au moment où l'information est présentée, et plus précisément de la phase et de l'amplitude des oscillations (Schroeder & Lakatos, 2009). Si l'information arrive dans la phase de haute excitabilité des oscillations, le taux de décharge neuronale sera plus important et le traitement sera donc optimal. A l'inverse, si le stimulus arrive dans la phase opposée (i.e. basse excitabilité), le taux de décharge sera moindre pouvant aller jusqu'à supprimer la communication au sein des réseaux neuronaux impliqués, diminuant de ce fait l'efficacité du traitement de l'information (Figure 11; Lakatos et al., 2005; Lakatos, Chi-ming, O'Connell, Mills, & Schroeder, 2007).

Les oscillations neuronales peuvent aussi réinitialiser (*reset*) ou recalculer leur phase sur la phase du stimulus périodique ou quasi-périodique (Figure 11; Lakatos et al., 2005, 2007; Schroeder et al., 2008). Cet « *entrainment* » permettrait d'amplifier la réponse neuronale et donc de « booster » l'efficacité du traitement. Autrement dit, si une information de l'environnement présente une structure temporelle (quasi-)régulière, elle peut être traitée de manière optimale si les oscillations neuronales dans la(les) région(s) impliqué(es) dans le traitement s'alignent de sorte que cette information périodique coïncide avec les pics d'excitabilité. Une information arrivant dans les moments de haute excitabilité a en effet plus de chances d'être traitée plus rapidement et efficacement (Schroeder & Lakatos, 2009; voir aussi Peelle & Davis, 2012). Dans cette optique, l'alignement entre les oscillations neuronales et les informations (quasi-)rythmiques constituerait donc un mécanisme fondamental pour garantir une sensibilité optimale de notre organisme aux événements perceptifs (Peelle & Davis, 2012).

Le recalage de phase oscillatoire sur la phase d'un stimulus peut aussi être réalisé en amont de l'arrivée du stimulus lorsque celui-ci est régulier et donc prédictible. Les oscillations ainsi couplées aux stimuli (quasi-)rythmiques pourraient alors générer une prédiction, ou attente, sur le moment où l'information pertinente est supposée arriver (Engel,

Fries, & Singer, 2001; Morillon & Baillet, 2017). En d'autres termes, les régularités temporelles des informations extérieures pourraient être exploitées par les oscillations neuronales (de basse fréquence) de sorte à favoriser le traitement perceptif. Ce codage prédictif temporel (*predictive timing*; Arnal & Giraud, 2012) permettrait de réduire l'incertitude sur le moment où les événements sont censés se produire et par conséquent de faciliter leur détection et leur traitement. Il a en effet été proposé que l'activité neuronale oscillatoire joue un rôle dans les processus de prédiction sensorielle et d'attention sélective (Arnal & Giraud, 2012; Engel, Fries, & Singer, 2001; Fries, 2005; Morillon, Schroeder, Wyart, & Arnal, 2016; Schroeder & Lakatos, 2009). Nous reviendrons plus tard sur cette notion de prédictibilité (partie III.3.).

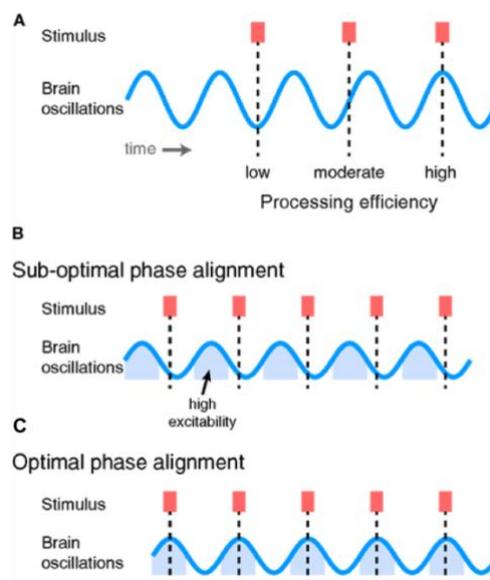


Figure 11 : (A) Représentation de l'efficacité du traitement d'un stimulus en fonction de la phase oscillatoire neuronale à laquelle il est présenté. Les informations arrivant dans une phase d'excitabilité faible seront traitées moins efficacement que si elles arrivent lorsque la phase d'excitabilité est élevée. (B et C) Lorsque le stimulus sensoriel présente une régularité temporelle, l'efficacité du traitement peut être améliorée si le système oscillatoire décale sa phase de sorte que les pics d'excitabilité neuronale soient alignés sur le rythme (i.e. la phase) du stimulus entrant. En (B), le stimulus périodique n'arrive pas dans la phase optimale d'excitabilité, alors qu'en (C), après le recalage de la phase des oscillations, il arrive dans une phase d'excitabilité maximale, permettant un traitement plus efficace de l'information. Figure empruntée à Peelle & Davis (2012).

3. Oscillations cérébrales et attention dynamique

Comme nous venons de le mentionner, les oscillations cérébrales constitueraient un mécanisme clé des processus cognitifs et notamment des processus attentionnels. L'attention peut être définie comme le processus par lequel le cerveau amplifie la représentation des informations pertinentes pour la réalisation d'une tâche au détriment des informations non pertinentes (Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). Les différents états d'excitabilité neuronale pourraient alors être utilisés comme mécanisme de sélection, où l'alignement oscillatoire (*entrainment*) sur les informations pertinentes à la tâche serait favorisé. Dans leur Théorie de l'Attention Dynamique (*Dynamic Attending Theory*, Jones, 1976; Jones & Boltz, 1989; Large & Jones, 1999; voir aussi Drake, Jones, & Baruch, 2000), Jones et ses collaborateurs définissent l'attention comme dynamique et basée sur un principe cyclique: l'attention serait elle-même un processus oscillatoire qui s'aligne sur les stimuli environnementaux rythmiques. En effet, le pic d'excitabilité des neurones module les ressources attentionnelles d'un individu, qui ne seraient donc pas distribuées de façon continue dans le temps. L'énergie attentionnelle varierait plutôt de manière cyclique, avec des maxima et des minima : les événements extérieurs synchronisés sur les pics d'énergie ou maxima seraient alors mieux traités. Selon Jones et collègues, chaque individu possède un cycle oscillatoire interne, appelé période de référence (*referent period*), qui est capable de se synchroniser sur le rythme d'un stimulus externe (niveau de référence ou *referent level*). Au sein d'un environnement rythmique (i.e. temporellement structuré), un ajustement (*attunement*) est ainsi réalisé entre la période des oscillations de l'attention et la période des événements externes. « *Dynamic attending reflects the attender's tacit use of an event's dynamic structure* » (Jones & Boltz, 1989). La régularité de ces oscillations permet alors de développer des attentes temporelles, aussi appelées prédictions, correspondant à une augmentation des ressources attentionnelles par anticipation, et donc de traiter plus efficacement l'évènement attendu. Lorsque nous attendons un évènement à un certain moment, nous sommes en effet plus attentifs à ce moment qu'à un autre, et donc plus prompts à détecter et traiter rapidement l'information que si elle est présentée à un autre moment. Nous pouvons ainsi anticiper ce qui va se produire et *quand* cela va se produire. La Théorie de l'Attention Dynamique, initialement développée pour rendre compte de la perception de séquences musicales, s'applique plus généralement au traitement de séquences structurées rythmiquement dans les modalités auditives, visuelles et motrices (Jones, Moynihan,

MacKenzie, & Puente, 2002). Dans les domaines de la parole et de la musique, un certain nombre d'études montre que les auditeurs sont capables de générer des prédictions à partir de la structure temporelle des stimuli, améliorant ainsi leurs performances. La récurrence d'éléments proéminents tels que les syllabes accentuées ou les pulsations (*beats*) permet en effet de développer des attentes sur la prochaine occurrence de ces événements (Dalla Bella, Białuńska, & Sowiński, 2013; Large & Jones, 1999). Des travaux ont montré que la structure métrique (i.e. alternance de temps forts – marqués par les syllabes accentuées – et de temps faibles – marqués par les syllabes non accentuées) de la parole dans certaines langues comme l'anglais et l'allemand, permet un meilleur traitement de l'information verbale. Zheng & Pierrehumbert (2010) ont par exemple rapporté que des participants anglophones étaient plus attentifs aux variations de durée des voyelles lorsque celles-ci occupaient une position métrique forte dans des phrases plutôt que faible. D'autres études ont décrit de meilleures performances de traitement de mots et de phrases lorsque le rythme (*timing*) était régulier plutôt qu'irrégulier (Dilley & McAuley, 2008; Quené & Port, 2005; Roncaglia-Denissen, Schmidt-Kassow, & Kotz, 2013). Un avantage de la présentation de séquences auditives rythmiques (e.g., Arnal, Doelling, & Poeppel, 2015; Bolger, Coull, & Schön, 2014; Cason & Schön, 2012; Jones, Moynihan, MacKenzie, & Puente, 2002) mais aussi de l'exécution de mouvements rythmiques (e.g., Chemin, Mouraux, & Nozaradan, 2014; Falk & Dalla Bella, 2016; Falk, Volpi-Moncorger, & Dalla Bella, 2017; Morillon, Schroeder, & Wyart, 2014; Schmidt-Kassow, Heinemann, Abel, & Kaiser, 2013) a également été rapporté sur le traitement de stimuli verbaux ou non verbaux tels que des tons et des séquences musicales. Par exemple, les attentes temporelles générées par l'écoute préalable de séquences rythmiques améliorent les capacités de segmentation phonologique de pseudo-mots (i.e. détection plus rapide d'un phonème cible s'il est présenté de façon synchrone avec le rythme inféré par la séquence, à savoir à un moment où un temps fort est attendu ; Cason & Schön, 2012). De la même manière, taper du doigt en rythme avec les syllabes accentuées d'une phrase améliore la détection d'un changement lexical dans cette phrase, par rapport à une condition où le mouvement n'est pas aligné avec ces syllabes ou si aucun mouvement n'est effectué (Falk & Dalla Bella, 2016; Falk et al., 2017). Ceci peut s'expliquer par le fait que produire un mouvement rythmique provoquerait une modulation *top-down* en provenance des régions motrices, permettant d'amplifier la réponse neuronale sur les événements saillants (se produisant aux moments attendus) dans l'information auditive entrante et affinant ainsi la sélection de cette information (Morillon & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014). Autrement dit, les prédictions temporelles générées dans le système moteur lors de la

réalisation des mouvements rythmiques optimiseraient le traitement de l'information auditive verbale ou non verbale.

La structure rythmique d'une information extérieure, et notamment de la parole, peut donc moduler les ressources attentionnelles de manière implicite et ainsi être utilisée comme un indice temporel prédictif pour faciliter la segmentation du flux continu en unités pertinentes. Une étude très élégante menée par Ghitza et Greenberg (2009) illustre parfaitement ce phénomène. Les auteurs ont présenté à des adultes anglophones plusieurs types de stimuli phrastiques différents : 1) des phrases produites naturellement ; 2) ces mêmes phrases mais compressées temporellement avec un facteur 3, c'est-à-dire que le rythme syllabique était trois fois plus rapide que le rythme original ; 3) ces phrases compressées mais segmentées en intervalles de 40 ms séparés par des périodes de silence. En fonction des conditions, la durée des silences était de 0, 20, 40, 80, 120 ou 160 ms (Figure 12A). Les silences étaient en outre insérés de façon périodique (i.e. au sein de chaque phrase, tous les silences avaient la même durée) ou non (i.e. la durée des silences au sein d'une phrase variait). La tâche des participants était d'écouter les phrases et de rappeler les mots dans l'ordre en les tapant sur un clavier. L'analyse des performances a montré que lorsque les phrases étaient compressées temporellement, la compréhension était fortement altérée (taux d'erreurs avoisinant les 50%). L'intelligibilité était également très faible, voire pire, lorsque des silences de 160 ms étaient insérés entre les segments de parole. Cependant, pour des silences allant de 20 à 120 ms, les performances s'amélioraient nettement, en particulier lorsque le rythme original de la phrase était restauré (silences de 80 ms) et que les silences étaient insérés périodiquement (taux d'erreurs de 20% ; Figure 12B). Ces résultats démontrent donc clairement que la compréhension de la parole ne dépend pas seulement de l'information acoustique contenue dans le signal de parole, mais aussi du moment où cette information est présentée. « *It is not just the information in the acoustic signal that is important, but also the timing of the information packets* » (Ghitza & Greenberg, 2009). L'intelligibilité maximale est ici observée lorsque les « paquets » d'information (segments de 40 ms dans les phrases compressées, séparés par des silences de 80 ms) sont alignés avec les segments de parole non compressés. En d'autres termes, l'information acoustique entre les versions compressée et originale est synchronisée. Ainsi, le fait de modifier la distribution temporelle de l'information de parole permet de restaurer l'intelligibilité si elle correspond à la structure initiale attendue par l'auditeur. Ghitza et Greenberg ont interprété leurs résultats comme reflétant le rôle des oscillations cérébrales endogènes (en particulier thêta, gamma et

potentiellement delta) dans le décodage de la parole, en suggérant que l'activité neuronale oscillatoire soit plus prédisposée à traiter l'information entrante sous forme de paquets d'informations contraints temporellement.

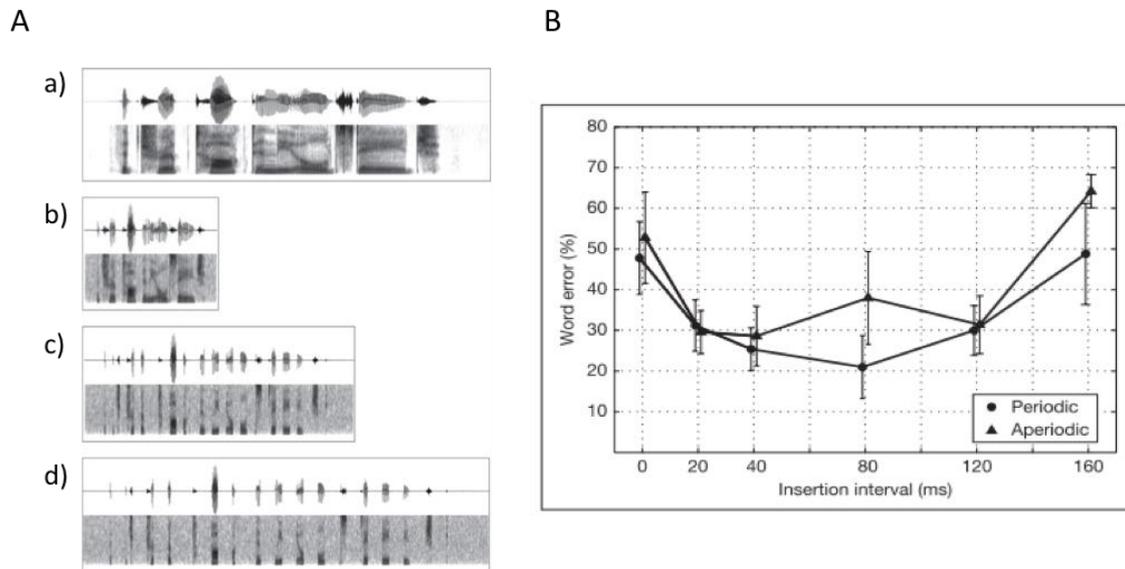


Figure 12 : (A) Exemples de stimuli dans l'étude de Ghitza et Greenberg (2009). a) Signal de parole et spectrogramme correspondant à une phrase expérimentale. b) Même phrase qu'en a) mais compressée temporellement avec un facteur 3. c) Même phrase qu'en b) mais segmentée en intervalles de 40 ms séparés par des périodes de silences de 40 ms (correspondant à un facteur de compression de 1.5 : la durée du signal correspond aux 2/3 de la durée originale). d) Même phrase qu'en c) mais avec des périodes de silence de 80 ms (la durée du signal est équivalente à celle du signal original ; la durée des segments de parole est en revanche réduite comme en c)). (B) Intelligibilité des phrases compressées temporellement en fonction de la durée des silences insérés entre les segments de parole de 40 ms. Figure empruntée à Ghitza et Greenberg (2009).

L'ensemble de ces données démontre que notre système auditif est capable, par le biais de modulations des ressources attentionnelles sous-tendues par les oscillations cérébrales, de tirer profit des régularités temporelles des informations extérieures. Les rythmes cérébraux seraient donc, de par leurs capacités de synchronisation et de prédiction, des outils de choix pour un traitement optimal du signal de parole, et notamment de son rythme. Dresser un panorama exhaustif des recherches sur les liens entre oscillations cérébrales et parole serait vain tant cette thématique suscite un engouement dans la communauté neuroscientifique : « *Speech, together with music, is arguably the most interesting stimulus to study the role of neuronal oscillations in perception* » (Morillon & Schroeder, 2015). L'objectif du chapitre

suivant sera néanmoins de présenter les principaux modèles et études expérimentales démontrant que les oscillations cérébrales jouent un rôle majeur dans le traitement de la parole.

4. Rythmes cérébraux et prédiction de la parole

4.1. Modèles oscillatoires de la parole

Comment expliquer que nous sommes capables d'identifier la parole en une fraction de seconde ? Comment pouvons-nous découper ce flux temporel continu en segments de taille équivalente à celle de nos représentations linguistiques stockées dans le lexique mental ? A partir d'un certain nombre de données issues de la psychophysique et des neurosciences, Poeppel (2003) a développé un modèle, celui de l'échantillonnage temporel asymétrique (*Asymmetric Sampling Theory* ou AST), afin de fournir des éléments de réponse à cette question. Comme nous l'avons décrit dans le chapitre I.2., le rythme du signal de parole est marqué par des indices phonologiques segmentaux (les phonèmes (voyelles et consonnes) et syllabes) et suprasegmentaux (prosodiques : durée, intensité, hauteur, accentuation) qui s'enchaînent dans le temps. Se basant sur les travaux de Rosen (1992), qui définit une organisation du signal de parole en deux échelles temporelles principales associées aux unités de base du signal acoustique, les syllabes (150-300 ms, correspondant à une fréquence d'environ 4-7 Hz, rythme thêta) et les transitions formantiques, caractérisant les phonèmes (20-40 ms, correspondant à une fréquence d'environ 25-40 Hz, rythme gamma), Poeppel propose une organisation temporelle similaire au sein du système cérébral auditif. Plus précisément, son modèle repose sur l'existence de fenêtres d'intégration temporelle de taille variable (principalement 25 ms et 200 ms), permettant de découper l'information temporelle continue en « paquets » pour être analysée. Le cerveau est composé de différents types de cellules nerveuses qui présentent une activité oscillatoire spontanée plus ou moins rapide. En particulier, au sein du système auditif, on trouve des grandes cellules pyramidales présentant une activité oscillatoire spontanée rapide d'environ 40 Hz (constante d'intégration de ~ 25 ms ; ondes gamma), et de plus petites cellules à activité lente (constante d'intégration de ~200 ms ; 4-7 Hz, ondes thêta ; Hutsler & Galuske, 2003). La proportion de ces cellules au sein des cortex auditifs droit et gauche est en outre asymétrique : le cortex auditif droit est composé majoritairement de cellules de petite taille oscillant au rythme thêta, alors que le gauche contient une plus grande proportion de grandes cellules oscillant dans la bande de fréquence

gamma. Le modèle AST suggère ainsi l'existence d'un couplage étroit entre l'activité rythmique des neurones des cortex auditifs gauche et droit et le rythme du signal de parole. Plus précisément, le cortex auditif gauche offrirait plutôt un mode de découpage de l'information verbale en unités de petite taille, les phonèmes, alors que le cortex auditif droit échantillonnerait le signal en plus grandes unités et participerait donc au traitement de l'information syllabique, caractérisant notamment le débit et donc le rythme de la parole¹⁶ (Figure 13). Les informations phonémique et syllabique, traitées en parallèle par les cortex auditifs gauche et droit respectivement, auraient alors le bon format pour être associées aux représentations linguistiques stockées en mémoire.

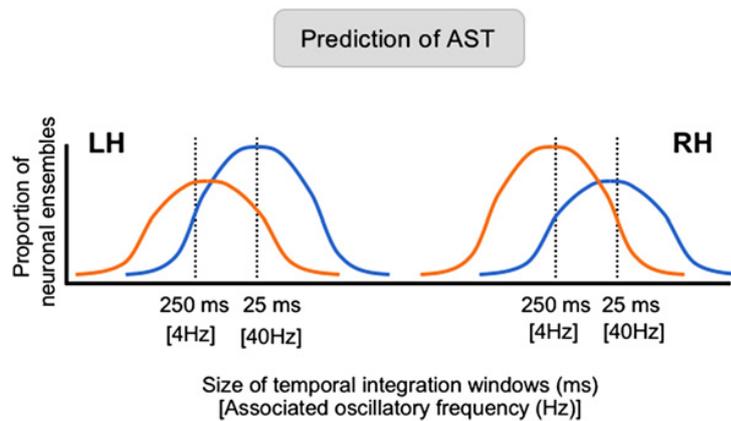


Figure 13 : Prédiction du modèle AST de Poeppel (2003). Le cortex auditif gauche, contenant plus de grandes cellules pyramidales ayant une activité oscillatoire rapide (40 Hz), serait plus à même d'échantillonner l'information de parole en unités phonémiques (25 ms). Le cortex auditif droit, contenant une plus grande proportion de plus petites cellules oscillant plus lentement (4 Hz), découperait l'information en unités de la taille des syllabes (250 ms). Figure empruntée à Giraud et al. (2007).

Giraud et Poeppel (2012) ont par la suite étendu ce modèle AST en proposant que les oscillations gamma basses (25-35 Hz), thêta (4-8 Hz) et delta (1-3 Hz) jouent un rôle dans les traitements phonémique, syllabique et phrastique (au sens de groupes de mots ou groupes rythmiques) respectivement lors de la perception de la parole. Les auteurs ont plus

¹⁶ Le modèle d'échantillonnage asymétrique de Poeppel (2003) partage certaines caractéristiques du modèle proposé par Zatorre et collègues (Zatorre, Belin, & Penhune, 2002) pour rendre compte du traitement de la musique et de la parole. Brièvement, ce modèle, corroboré par un certain nombre d'études, suggère également une spécialisation fonctionnelle asymétrique du système auditif : les régions auditives gauches présenteraient une meilleure résolution temporelle alors que la résolution spectrale serait plus élevée dans les régions droites.

particulièrement suggéré l'existence d'une cascade de processus qui permettrait de transformer le signal continu de parole en unités distinctes et qui serait basée sur un couplage étroit ou *nesting* (i.e. nichage) entre les oscillations thêta et gamma (Figure 14). Après un encodage des propriétés spectro-temporelles du signal de parole dans le cortex auditif primaire, les informations saillantes (*edges* ou crêtes au sein de l'enveloppe d'amplitude) induiraient un recalage de la phase des oscillations intrinsèques thêta et probablement gamma (étape 1). Les oscillations thêta s'aligneraient spécifiquement sur les modulations lentes de l'enveloppe d'amplitude (qui, comme nous l'avons vu dans la partie I.3., caractérisent le rythme et le débit de parole, et où l'information syllabique domine ; étape 2). Les oscillations thêta et gamma, faiblement couplées au repos et traitant en parallèle l'information verbale, deviendraient alors fortement couplées : la phase des oscillations thêta modulerait l'amplitude (et potentiellement la phase) des oscillations gamma (*theta-gamma nesting* ; étape 3). La puissance des oscillations gamma régulerait ensuite la décharge neuronale du cortex auditif primaire vers les aires de plus haut niveau de manière à créer des paquets de trains de décharge et des cycles d'excitabilité correspondant aux échelles phonémique et syllabique (étape 4). Enfin, l'excitabilité neuronale serait modulée pour être alignée avec la structure du signal de parole (étape 5). Les auteurs proposent donc que la construction de représentations intelligibles de la parole repose à la fois sur (1) le couplage de phase (*phase-locking*) entre l'information entrante et le cortex auditif sur au moins deux échelles temporelles, et (2) le couplage hiérarchique entre les oscillations corticales thêta et gamma qui permettrait aux traitements phonémique et syllabique d'être réalisés en parallèle tout en restant hiérarchisés. Ainsi, le signal de parole réinitialiserait (*reset*) l'activité oscillatoire endogène du cortex auditif, principalement dans les bandes de fréquence correspondant aux taux d'échantillonnage optimaux pour les découpages phonémique et syllabique. En retour, les oscillations corticales permettraient d'organiser temporellement le signal continu de parole pour faciliter sa reconnaissance.

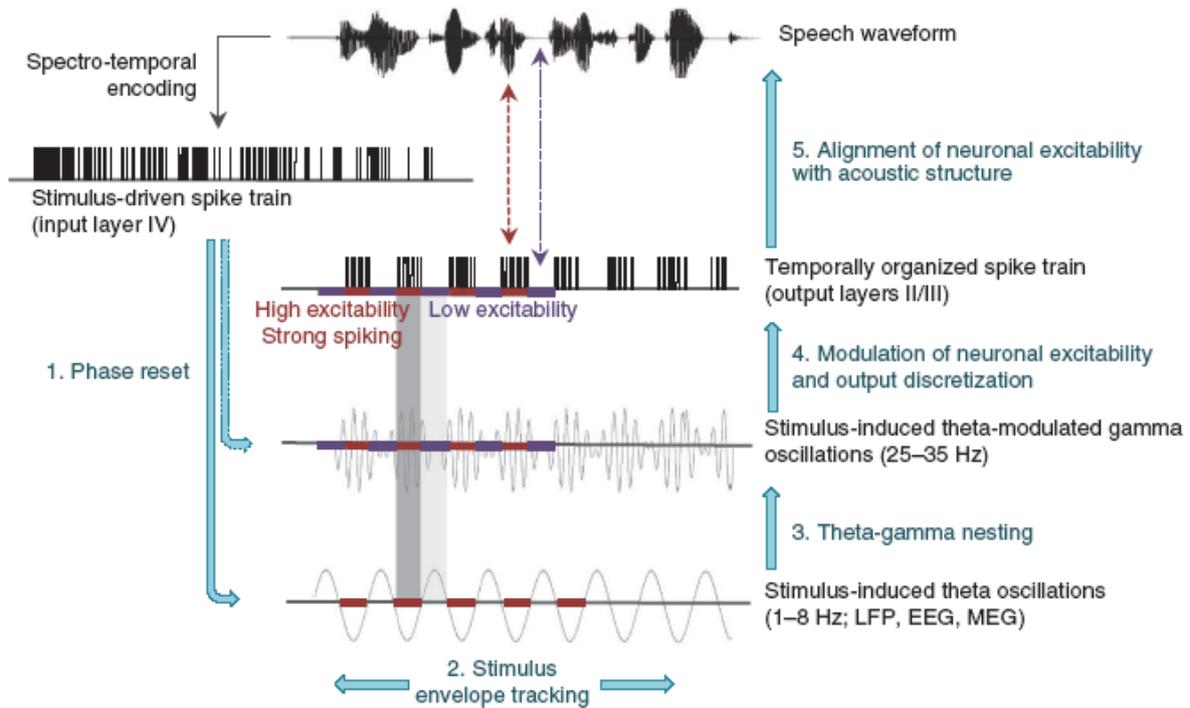


Figure 14: Modèle de Giraud et Poeppel (2012) permettant d'expliquer comment les oscillations corticales échantillonnent et s'alignent sur le signal continu de parole. Les propriétés spectro-temporelles de la parole seraient encodées dans le cortex auditif primaire (une décharge neuronale dans la couche corticale IV est représentée). La plupart des neurones synchroniseraient leur phase sur les modulations d'amplitude de la parole. Un recalage de la phase des oscillations thêta dans les couches superficielles se produirait ensuite sur les événements saillants dans le signal (étape 1). Ces oscillations θ s'aligneraient (*track*) alors sur l'enveloppe temporelle de la parole (étape 2). Par le biais d'un couplage étroit (*nesting*), les oscillations thêta moduleraient l'amplitude des oscillations gamma (étape 3). La puissance des oscillations gamma contrôlerait l'excitabilité des neurones permettant la transmission de l'information du cortex auditif primaire aux aires de plus haut niveau (étape 4). L'excitabilité neuronale s'alignerait enfin sur la structure du signal de parole : la puissance gamma est plus forte dans les moments de faible énergie dans le signal (étape 5). Figure empruntée à Giraud & Poeppel (2012).

Si le modèle AST (Poeppel, 2003) et celui de Giraud et Poeppel (2012) se focalisent sur les bandes de fréquence thêta et gamma, Ghitza (2011), dans son modèle baptisé « Tempo », introduit les oscillations beta (15-30 Hz) au sein de la cascade d'oscillations jouant un rôle fondamental dans l'intelligibilité de la parole. Le modèle Tempo stipule l'existence de deux modes parallèles de traitement de l'information verbale : le découpage (*parsing*) et le décodage. En accord avec le modèle précédent, Ghitza (2011) propose que pour le découpage, les oscillations thêta s'aligneraient sur les modulations rythmiques de l'enveloppe temporelle,

conduisant à un découpage syllabique du signal. Les oscillations thêta moduleraient alors les oscillations beta (i.e. la fréquence beta serait un multiple de la fréquence thêta, et la phase beta serait réinitialisée au début d'un cycle thêta), qui auraient pour fonction de réaliser un découpage plus fin de l'information (segments de 50 ms, correspondant à la taille des dyades¹⁷). Enfin, un troisième oscillateur dans la bande gamma (dont la fréquence serait un multiple de la fréquence beta de sorte à être toujours aligné avec les oscillations beta) serait en charge du traitement des transitions spectro-temporelles rapides associées aux dyades. En parallèle, pour assurer le décodage, les différentes informations extraites par la procédure de découpage seraient appariées à des représentations en mémoire sur deux échelles temporelles (syllabique, ~200 ms et dyadique, ~50 ms). Ainsi, ce modèle propose également un couplage (*nesting*) étroit entre les oscillations des différentes bandes de fréquence pertinentes pour le traitement de la parole : « *the theta oscillator is the master in the cascaded array, and the other oscillators entrain to the theta* » (Ghitza, 2011 ; voir aussi Ghitza, 2017 pour un rôle des oscillations delta dans la segmentation prosodique). Ghitza souligne également qu'une propriété cruciale de cette cascade oscillatoire est qu'elle reste synchronisée à l'information quasi-rythmique de parole, quand bien même le rythme change (un peu). Autrement dit, la durée des cycles oscillatoires s'adapterait à l'information entrante : l'intelligibilité de la parole resterait haute tant que les oscillations pourraient synchroniser leur phase avec les modulations du signal entrant ; une absence de synchronisation (parce-que les oscillateurs atteindraient la limite de leur bande de fréquence) conduirait à une chute drastique d'intelligibilité.

En résumé, les modèles, plus convergents que divergents, proposent l'existence d'une cascade oscillatoire au sein du cortex auditif, dans laquelle les oscillations thêta se caleraient sur le rythme de la parole et moduleraient les oscillations gamma et potentiellement beta. Ceci permettrait de hiérarchiser les traitements syllabique et phonémique (qui pourraient cependant rester simultanés). Un tel alignement de l'activité oscillatoire des régions auditives avec l'information acoustique contenue dans le signal de parole permettrait d'augmenter la sensibilité aux indices acoustiques pertinents, assurant un traitement rapide et efficace de la parole (Figure 15 ; Peelle & Davis, 2012).

¹⁷ Une dyade est le reflet acoustique du geste dynamique des articulateurs effectué pour passer d'un phone (i.e. réalisation concrète d'un phonème, variante d'un phonème) à un autre. Ce mouvement est souvent associé à une « crête » (*acoustic edge*) dans le signal acoustique.

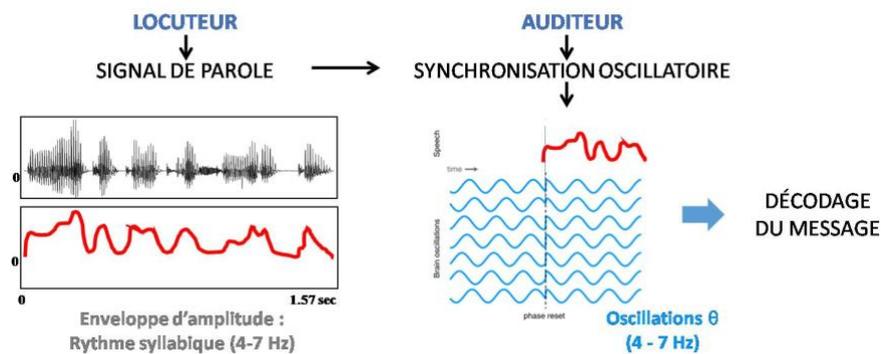


Figure 15 : La perception de la parole serait permise, au moins en partie, par la synchronisation des oscillations corticales sur le signal de parole, en particulier par un alignement des oscillations lentes du cortex auditif dans la bande de fréquence thêta sur les modulations lentes de l'enveloppe d'amplitude du signal de parole. Figure inspirée de Peelle & Davis (2012).

4.2. Synchronisation oscillatoire sur le signal de parole

Comme mentionné précédemment, l'étude des oscillations cérébrales en lien avec la parole suscite un vif intérêt depuis quelques années. Nous présenterons ici des arguments expérimentaux en faveur des modèles oscillatoires décrits ci-dessus.

Avant de s'intéresser à la perception de la parole, intéressons-nous d'abord à la dynamique des oscillations cérébrales dans les bandes de fréquence thêta et gamma au repos. Si, comme le suggère Poeppel (2003), il existe une asymétrie fonctionnelle des oscillations du cortex auditif, existe-t-elle déjà de manière intrinsèque ? Dans une étude *princeps* combinant des enregistrements en EEG et en IRMf, Giraud et collaborateurs (2007) ont examiné les oscillations cérébrales dans les bandes thêta et gamma chez des participants adultes au repos (i.e. en l'absence de toute tâche expérimentale). Ils ont montré que l'activité hémodynamique au sein du cortex temporal supérieur covariait avec l'activité oscillatoire dans la bande de fréquence thêta (3-6 Hz). Ce rythme était prédominant à droite au niveau du gyrus de Heschl incluant le cortex auditif primaire, mais, contrairement aux prédictions du modèle AST (Poeppel, 2003), il était également observé à gauche dans les régions temporales plus antérieures. Les auteurs ont également rapporté une dominance hémisphérique gauche des oscillations gamma (28-40 Hz) au sein du cortex temporal supérieur (partie latérale du gyrus de Heschl ; Figure 16A). De manière intéressante, ce pattern ne se limitait pas au cortex auditif. La puissance des oscillations thêta était en effet également corrélée à l'activité hémodynamique d'une partie du cortex moteur contrôlant les mouvements de la bouche ;

l'activité oscillatoire gamma covariait quant à elle avec l'activité synaptique d'une région motrice un peu plus ventrale contrôlant les mouvements de la langue. Aucune latéralisation hémisphérique n'a été observée dans ces régions motrices (Figure 16B). Ainsi, les oscillations thêta, prédisposées à segmenter la parole en unités syllabiques, s'expriment naturellement dans les cortex auditif (droit principalement) et moteur de la bouche (i.e. mouvements de la mâchoire pour produire les syllabes), alors que les oscillations gamma offrant un mode de découpage phonémique sont associées au cortex auditif gauche et au cortex moteur de la langue (i.e. déplacement de la langue pour produire les phonèmes consonantiques ; voir aussi Morillon, Lehongre, Frackowiak, Ducorps, & Kleinschmidt, 2010). Ces données s'accordent donc avec l'idée de liens étroits entre perception et production de la parole, et en particulier avec la théorie du cadre et du contenu de MacNeilage et Davis (2001 ; voir chapitre I.3.).

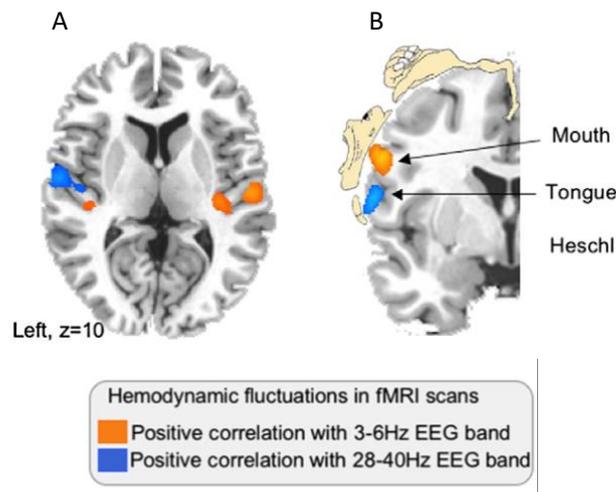


Figure 16 : Corrélation entre l'activité hémodynamique mesurée en IRMf et l'activité oscillatoire mesurée en EEG dans les bandes de fréquence thêta (3-6 Hz) et gamma (28-40 Hz) au repos, (A) dans le cortex auditif et (B) dans le cortex moteur. Les résultats sont présentés sur des coupes transversales du cerveau. Figure empruntée à Giraud et al., (2007).

Des résultats comparables, révélant la spécificité oscillatoire endogène du cortex auditif pour les fenêtres d'intégration temporelle pertinentes pour le traitement du signal de parole, ont par la suite été rapportés par Morillon et collaborateurs (2012) au moyen d'enregistrements intra-crâniens chez une patiente épileptique. Les auteurs ont en effet retrouvé une dominance de l'activité oscillatoire lente thêta (5-6 Hz) dans le cortex auditif primaire droit et une dominance de l'activité gamma (25-45 Hz) dans le cortex auditif

primaire gauche, à la fois au repos et lors du traitement de syllabes (nous y reviendrons plus tard). En outre, corroborant les résultats de Giraud et al. (2007), ils ont observé que l'activité thêta (de fréquence légèrement plus élevée, 7-8 Hz) s'exprimait également dans les cortex auditifs primaire et associatif gauches (cortex temporal antérieur, BA 22), mais aussi que la phase de cette activité thêta dans l'aire BA 22 était corrélée à la puissance des oscillations gamma hautes induites. Ainsi, l'activité thêta est également présente dans une région impliquée dans des traitements linguistiques plus avancés. Pour interpréter ce dernier résultat, Morillon et ses collègues (2012) ont proposé que l'activité oscillatoire thêta dans les cortex auditifs droit et gauche jouerait des rôles distincts (voir aussi Giraud et al., 2007; Giraud & Poeppel, 2012). Alors que les oscillations thêta à droite permettraient un découpage de l'information de parole selon le rythme syllabique, leurs équivalents dans l'hémisphère gauche pourraient être impliqués dans le couplage avec les oscillations gamma. Plus précisément, c'est au sein de l'hémisphère gauche que les segments de taille phonémique, échantillonnés dans la bande gamma, pourraient être intégrés dans des segments plus longs de taille syllabique extraits par la bande thêta. Un tel couplage thêta-gamma au sein des régions auditives gauches permettrait un traitement de l'information linguistique efficace et moins coûteux au niveau cognitif (comparé à un traitement de chaque type d'unité par un seul hémisphère). Dans cette optique, l'activité thêta latéralisée à gauche occuperait une fonction d'intégration de haut niveau plutôt qu'une simple fonction de découpage de l'information : elle permettrait d'accéder aux représentations phonologiques (plus abstraites) du signal de parole à partir de l'extraction et de la segmentation des indices acoustiques (Giraud & Poeppel, 2012).

Les études décrites ci-dessus font état de la dynamique oscillatoire dans les bandes thêta et gamma au repos. Qu'en est-il du traitement de la parole ? Un nombre florissant de travaux montre que les oscillations corticales dans ces mêmes bandes de fréquence se synchronisent de manière spécifique sur le signal verbal, révélant le caractère multi-temporel du traitement de la parole (Giraud & Poeppel, 2012; Howard & Poeppel, 2010, 2012; Luo & Poeppel, 2007; Morillon et al., 2012; Ghitza, Giraud, & Poeppel, 2012 pour une revue). L'*entraînement* des oscillations thêta sur les fluctuations quasi-rythmiques de l'enveloppe d'amplitude du signal de parole, qui véhiculent entre autres l'information syllabique, a fait l'objet d'un nombre particulièrement important de recherches (*e.g.*, Cogan & Poeppel, 2011; Ding & Simon, 2013; Doelling, Arnal, Ghitza, & Poeppel, 2014; Gross et al., 2013; Howard & Poeppel, 2010;

Kerlin, Shahin, & Miller, 2010; Peelle, Gross, & Davis, 2013; Vander Ghinst et al., 2016; Zion Golumbic et al., 2013). Il a en outre été suggéré que ce couplage joue un rôle clé dans l'intelligibilité du signal, même si la question de la nécessité de la compréhension pour l'alignement cortico-acoustique reste débattue à l'heure actuelle (Ghitza & Greenberg, 2009; Peelle & Davis, 2012; Zion Golumbic, Poeppel, & Schroeder, 2012; Zoefel & Vanrullen, 2015). Ahissar et collaborateurs (2001) ont été les premiers à démontrer, à l'aide de la MEG chez des adultes anglophones, que l'activité du cortex auditif primaire gauche était couplée à l'enveloppe temporelle du signal de parole. Néanmoins, ceci n'était observé que pour des phrases intelligibles qui étaient compressées temporellement de façon modérée (compression à 75% de la durée originale du signal, correspondant à des fluctuations de l'enveloppe aux alentours de 4 Hz, et compression à 50% avec une fréquence d'amplitude d'environ 6 Hz). Lorsque les phrases étaient fortement accélérées et donc inintelligibles (compression à 35% correspondant à une fréquence de modulations d'environ 9 Hz, et à 20% correspondant environ à 15 Hz), la synchronisation entre les deux signaux n'était plus observée. De plus, les auteurs ont mis en évidence que le couplage de phase entre l'activité corticale et l'enveloppe prédisait les capacités de compréhension des phrases des participants (i.e. tâche de jugement de phrases). Lorsque les performances des participants chutaient avec la diminution d'intelligibilité des phrases compressées, le couplage était également plus faible. Ceci était principalement lié au fait que les modulations corticales évoquées, dont la fréquence était proche de celle de l'enveloppe de la parole à un débit normal (< 10 Hz), étaient maintenues à cette fréquence même lorsque les fluctuations de l'enveloppe étaient plus rapides suite à la compression temporelle. Les auteurs ont alors suggéré que la correspondance entre le débit de parole et les capacités oscillatoires des régions auditives soit un prérequis pour une compréhension efficace : « *the frequency tuning of cortical neurons determines the upper limit of speech modulation rate that can be followed by an individual, and comprehension of a sentence requires an online phase-locking to the temporal envelope of that sentence.* ». Plus précisément, ils ont proposé qu'avec des phrases courtes telles que celles utilisées dans leur étude, le cerveau ne disposait pas de suffisamment de temps pour adapter sa fréquence de réponse à celle des stimuli, et qu'il était donc nécessaire que les modulations de l'information verbale s'inscrivent dans la bande de fréquence dans laquelle le cortex auditif oscille *a priori*. Dans le cas d'un débit de parole apparié à la fréquence oscillatoire corticale, le couplage cortico-acoustique serait fort ; en revanche, un débit plus rapide (par accélération artificielle ici) conduirait à une perte de couplage et donc à une diminution d'intelligibilité. Des résultats comparables ont été décrits quelques années plus tard par Luo et Poeppel (2007). Les auteurs

ont demandé à des adultes anglophones d'écouter des paires de phrases, la première produite naturellement dans sa version intelligible, et la seconde inintelligible (i.e. chimère parole-bruit¹⁸), et de décider s'il s'agissait de la même phrase ou non. Les auteurs ont d'abord observé que la phase des oscillations thêta (4-8 Hz ; « *theta phase pattern* ») dans le cortex auditif, avec une asymétrie en faveur de l'hémisphère droit, permettait de discriminer les phrases (i.e. la cohérence inter-essais était plus forte pour des groupes de phrases identiques vs différentes). Ceci n'était pas observé lorsque la puissance thêta était prise en compte ; de la même manière, la puissance thêta était identique avant et pendant la présentation des phrases, confirmant que c'est bien la phase des oscillations intrinsèques thêta qui était modulée par le signal de parole (par un recalage de phase) et qui permettait de discriminer les phrases (mais voir Howard & Poeppel, 2012 pour des résultats montrant également des modulations d'amplitude et donc une activité oscillatoire supplémentaire en réponse à la parole). Ensuite, en accord avec les résultats d'Ahissar et al. (2001), les auteurs ont rapporté une corrélation entre le pattern de phase thêta et le degré d'intelligibilité des phrases : la phase des oscillations était moins à même de discriminer les phrases (i.e. le couplage avec la parole était moins robuste) à mesure que les stimuli devenaient moins intelligibles. Pour interpréter leurs résultats, Luo et Poeppel ont conclu que les oscillations thêta constitueraient le code neuronal de la parole : la parole continue serait traitée par une fenêtre d'intégration temporelle endogène de ~200 ms, correspondant à la période de la phase thêta, qui serait réinitialisée (*reset*) et « glisserait » (*slide*) de manière dynamique sur le signal verbal pour le segmenter.

Ces études, si elles ont le mérite de mettre en évidence la relation étroite entre oscillations corticales, enveloppe de la parole et intelligibilité, ne permettent cependant pas d'en déterminer la causalité : est-ce parce-que les oscillations ne peuvent pas se synchroniser sur la parole que celle-ci est inintelligible ou inversement, est-ce parce-que la parole n'est pas intelligible que le couplage cortico-acoustique échoue ? Si cette question reste encore ouverte à ce jour, des travaux ultérieurs ont révélé que la compréhension n'était pas une condition *sine qua non* au couplage cerveau/parole. La synchronisation des oscillations θ sur les modulations lentes d'enveloppe d'amplitude serait effective que le signal de parole soit intelligible ou non. Dans une expérience similaire à celle précédemment décrite, Howard & Poeppel (2010) ont observé que la phase θ permettait également de discriminer des phrases présentées à l'envers,

¹⁸ Les chimères parole-bruit (*speech-noise chimeras*) sont formés en couplant un signal verbal avec du bruit ayant le même spectre de puissance. Ces chimères sont synthétisées avec un certain nombre de bandes de fréquence ; l'information contenue dans l'enveloppe d'amplitude ou la structure temporelle fine est également supprimée. Ces manipulations rendent le signal de parole plus ou moins intelligible (Smith, Delgutte, & Oxenham, 2002).

c'est-à-dire dont l'enveloppe temporelle était préservée (mais inversée temporellement) mais qui étaient totalement incompréhensibles. Les auteurs ont ainsi proposé que le couplage θ sur le signal de parole reflète le traitement des propriétés acoustiques de la parole nécessaires à l'intelligibilité (i.e. modulations lentes d'enveloppe d'amplitude) plutôt que les traitements linguistiques plus avancés liés à la compréhension *per se* (lexical, sémantique et syntaxique). Dans cette optique, les oscillations θ assureraient un suivi des fluctuations du signal de parole indépendamment du fait que l'auditeur le comprenne. De tels résultats ont été corroborés, au moins en partie, par un grand nombre de travaux (Gross et al., 2013; Hertrich, Dietrich, Trouvain, Moos, & Ackermann, 2012; Luo & Poeppel, 2012; Mai, Minett, & Wang, 2016; Nourski et al., 2009; Park, Ince, Schyns, Thut, & Gross, 2015; Peelle et al., 2013; Pefkou, Arnal, Fontolan, & Giraud, 2017). Dans une étude en MEG, Peelle et collaborateurs (2013) ont demandé à des participants anglophones de répéter des phrases en parole vocodée¹⁹ (l'enveloppe temporelle est en revanche préservée), dont l'intelligibilité variait en fonction du nombre de bandes de fréquence préservées (i.e. parole inintelligible pour 1 bande vs parole modérément intelligible pour 4 bandes vs parole totalement intelligible pour 16 bandes). Leurs résultats ont révélé une augmentation de cohérence entre les oscillations cérébrales dans la bande de fréquence θ (4-7 Hz) et l'enveloppe d'amplitude du signal de parole, même lorsque celle-ci n'était pas intelligible (parole vocodée 1 bande et parole « retournée spectralement » 4 bandes). Ceci était observé bilatéralement dans les régions temporales supérieure et médiane, le gyrus frontal inférieur et le cortex moteur. Toutefois, une analyse centrée sur le cortex auditif et comparant les conditions 1 bande et 16 bandes a mis en évidence que la synchronisation était plus forte et plus latéralisée dans l'hémisphère gauche lorsque la parole vocodée était intelligible (16 bandes ; Figure 17A). De la même manière, les oscillations du cortex auditif gauche (gyrus temporal médian) présentaient une plus forte cohérence avec la parole vocodée à 4 bandes (intelligible) qu'avec la parole vocodée « retournée » à 4 bandes (inintelligible) qui présentent pourtant la même complexité spectrale (Figure 17B). Selon cette étude, le couplage cortico-acoustique pourrait donc se produire pour de la parole non intelligible ; néanmoins, il ne dépendrait pas uniquement des indices acoustiques rythmiques présents dans l'enveloppe d'amplitude mais bénéficierait de l'information linguistique et serait alors plus latéralisé dans l'hémisphère gauche (contrairement aux prédictions du modèle AST de Poeppel, 2003).

¹⁹ Pour rappel, la parole vocodée consiste en une réduction du contenu spectral du signal par décomposition du spectre en plusieurs bandes de fréquence et remplacement de certaines bandes par du bruit blanc modulé en amplitude.

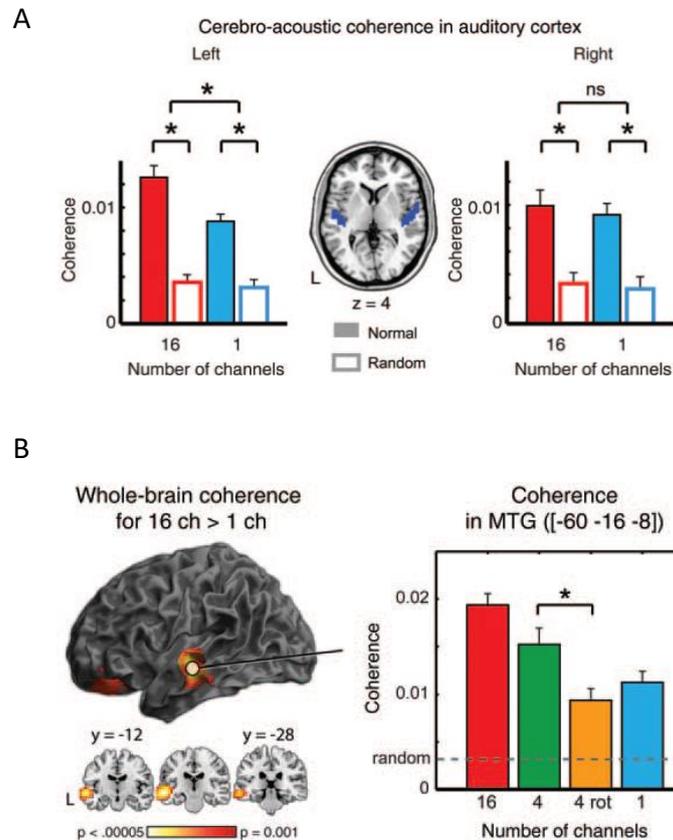


Figure 17: (A) Cohérence « cérébro-acoustique » dans le cortex auditif dans la bande de fréquence thêta (4-7 Hz) pour les conditions de parole vocodée intelligible (16 bandes) et inintelligible (1 bande). L'interaction montre une cohérence plus forte dans l'hémisphère gauche pour la condition intelligible. (B) Les oscillations du gyrus temporal médian gauche présentent une cohérence avec l'enveloppe d'amplitude de parole plus forte dans la condition intelligible (16 bandes) vs non intelligible (1 bande) ; une différence est également observée entre les conditions 4 bandes intelligible et inintelligible (parole « retournée » ou « rotated ») qui possèdent le même niveau de détail spectral.

Figure empruntée à Peelle et al. (2013).

Jusqu'à présent, les études présentées se sont principalement intéressées aux oscillations thêta et à leur synchronisation sur l'enveloppe temporelle de la parole. D'autres études ont plus directement testé les prédictions du modèle AST (Giraud & Poeppel, 2012; Poeppel, 2003) en examinant les interactions entre les oscillations dans les bandes thêta, gamma (et delta), et le signal de parole (Gross et al., 2013; Morillon et al., 2012; Park et al., 2015a; Zion-Golumbic et al., 2013). Dans leur étude chez une patiente épileptique francophone (précédemment décrite), Morillon et al. (2012) ont mis en évidence, lors de

l'écoute de syllabes isolées (/ba/ et /pa/), des réponses gamma basses (25-45 Hz) évoquées²⁰ plus fortes à gauche dans le cortex auditif primaire (et secondaire dans une moindre mesure), et des réponses gamma hautes (> 50 Hz) induites bilatérales dans le cortex auditif primaire mais latéralisées à gauche dans le cortex associatif (BA 22 ; Figure 18A). Ainsi, l'asymétrie en faveur de l'hémisphère gauche s'accroît à mesure que le traitement était réalisé dans les régions occupant une place plus haute dans la hiérarchie. Au niveau de la bande thêta, les auteurs ont observé une activité thêta haute (~7.5 Hz) évoquée plus importante dans les cortex auditifs primaire, secondaire et associatif gauches, alors que l'activité thêta basse (~5.5 Hz) était latéralisée à droite et maximale dans le cortex auditif secondaire. Par ailleurs, une analyse du couplage inter-fréquences dans le cortex associatif gauche (BA 22) a révélé une co-modulation entre la phase des oscillations thêta hautes (~7.5 Hz) et l'amplitude des oscillations gamma hautes (*nesting* thêta-gamma ; Figure 18B). Bien que la généralisation de ces résultats obtenus chez une patiente avec des électrodes implantées soit soumise à précaution, cette étude démontre donc que les réponses oscillatoires évoquées de basse fréquence (gamma basses à gauche et thêta basses à droite) assurent un suivi (*tracking*) des indices acoustiques du signal de parole (syllabes isolées ici). Au sein de l'hémisphère gauche, l'activité thêta haute pourrait quant à elle, comme évoqué précédemment, structurer temporellement l'activité gamma haute et ainsi assurer la transition du traitement des traits acoustiques à un traitement phonologique plus abstrait (Giraud & Poeppel, 2012).

²⁰ Les réponses évoquées sont reflétées dans la puissance/amplitude et la phase (*phase-locking value*) des oscillations, alors que les réponses induites ne sont reflétées que dans la puissance.

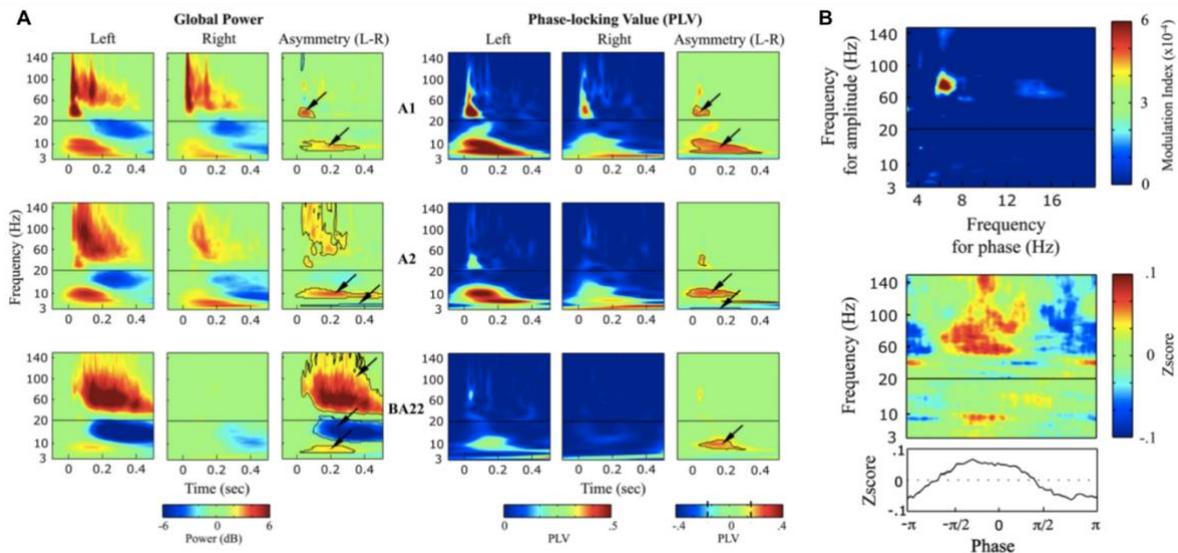


Figure 18 : (A) Cartes temps-fréquence montrant les modulations de puissance (à gauche) et valeurs de *phase-locking* (PLV, reflétant l'activité évoquée, à droite) lors de l'écoute de syllabes au sein des régions auditives (A1, A2 et BA22) gauches et droites. La composante asymétrique oscillatoire (gauche-droite) est également représentée. (B) Co-modulation entre la phase des oscillations thêta hautes (~7.5 Hz) et l'amplitude des oscillations gamma hautes dans l'aire BA22. En haut, comodulogramme entre la phase des basses fréquences (3-20 Hz) et la puissance (3-150 Hz). Un index de modulation positif indique un couplage significatif. En bas, distribution de la puissance (3-150 Hz) selon la phase thêta (7.5 Hz). Figure empruntée à Morillon et al. (2012).

Dans une étude en MEG utilisant de la parole continue (7 minutes) chez des adultes anglophones, Gross et collaborateurs (2013) ont quant à eux rapporté que la phase des oscillations delta (1-3 Hz) et thêta (3-7 Hz) dans le cortex auditif et les régions fronto-pariétales était plus fortement couplée à l'enveloppe d'amplitude de la parole lorsque le signal était intact plutôt que présenté à l'envers et donc inintelligible. Ce couplage était en outre plus marqué à droite, au niveau temporel pour la bande thêta et au niveau fronto-pariétal pour la bande delta. Les fluctuations lentes de l'enveloppe (3-7 Hz) modulaient également l'amplitude des oscillations gamma (35-45 Hz) au sein du cortex auditif, frontal et pariétal, là encore de façon plus importante pour la parole intelligible vs inintelligible, mais cette fois avec une asymétrie en faveur de l'hémisphère gauche (Figure 19A). Ces données confirment donc l'existence de deux mécanismes de « suivi » de la parole, sensibles à l'intelligibilité du signal : les modulations quasi-rythmiques de la parole entraînent (i.e. alignent) (1) la phase des oscillations delta et thêta dans les régions auditives droites, et (2) la dynamique d'amplitude des oscillations gamma dans les régions auditives gauches. Les auteurs ont

également mis en évidence un couplage (*nesting*) entre la phase thêta et l'amplitude gamma plus fort dans la condition de parole intacte par rapport à la condition de parole inversée, au sein des régions temporales auditives bilatérales et des régions frontales et pariétales gauches (Figure 19B). Un couplage, quoique moins fort, était également présent entre la phase delta et l'amplitude thêta dans les régions temporales majoritairement gauches. Enfin, les auteurs ont cherché à savoir si la phase des oscillations thêta était réinitialisée par les attaques des syllabes dans le signal continu (qui se manifestent par des « crêtes » (*edges*) dans l'enveloppe d'amplitude). Leurs résultats ont montré une augmentation de l'alignement entre les signaux MEG et acoustique dans les cortex auditifs droit et gauche, qui n'était pas seulement transitoire (i.e. ne se limitait pas au début de chaque crête ou *edge onset*) mais perdurait dans le temps (voir aussi Doelling, Arnal, Ghitza, & Poeppel, 2014 pour un alignement delta-thêta sur les attaques syllabiques). Cet alignement sur la parole s'accompagnait d'une diminution de couplage de phase thêta (découplage transitoire) entre les deux hémisphères dans les régions correspondantes. Dans l'ensemble, ces données s'accordent donc avec le modèle AST (Poeppel, 2003) et avec l'idée d'un couplage hiérarchique des oscillations dans les différentes bandes de fréquence lors du traitement de la parole, qui permettrait de traiter les informations phonémique, syllabique et phrastique en parallèle tout en préservant leurs relations mutuelles (Ghitza, 2011; Giraud & Poeppel, 2012). Les résultats de Gross et al. corroborent également ceux de Peelle et collègues (2013) quant à une influence du contenu linguistique sur l'*entrainement* oscillatoire sur l'enveloppe de la parole. Les auteurs ont attribué cet effet à des processus *top-down* permettant de générer des prédictions sur les événements saillants et les mots contenus dans le signal.

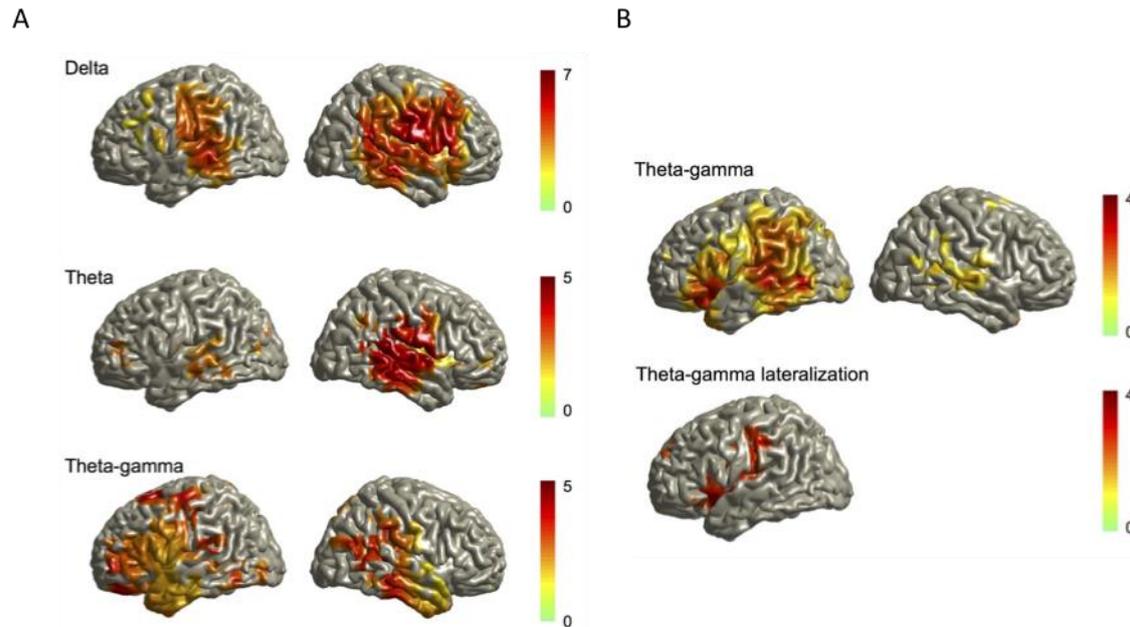


Figure 19 : (A) Alignement (information mutuelle) entre la phase des oscillations corticales delta (1-3 Hz), thêta (3-7 Hz) et gamma (35-45 Hz) et la phase (thêta) du signal de parole (enveloppe d'amplitude) dans les deux hémisphères cérébraux, pour la condition intelligible (parole à l'endroit) vs inintelligible (parole inversée). (B) Couplage entre la phase des oscillations corticales thêta et l'amplitude des oscillations corticales gamma pour la parole intelligible vs inintelligible. Le pattern de latéralisation de ce couplage est également représenté. Figure empruntée à Gross et al. (2013).

Des influences *top-down* sur l'alignement entre rythmes cérébraux et linguistiques provenant de régions corticales de plus haut niveau ont par ailleurs été décrites. Dans le même paradigme que celui utilisé par Gross et al. (2013), Park et collaborateurs (2015) ont par exemple montré que l'activité des régions frontales motrices, principalement gauches, modulait la phase des oscillations delta et thêta du cortex auditif gauche lors de l'écoute de parole intelligible (présentée à l'endroit) vs inintelligible (à l'envers). En outre, plus ces influences *top-down* étaient fortes, meilleur était l'alignement cerveau/parole : le couplage cortico-acoustique dans les bandes delta et thêta était en effet corrélé de manière positive à l'activité oscillatoire des gyri précentral et frontal gauches et des régions temporales postérieures. Keitel et collègues (2017) ont complété ces résultats en montrant que ces influences cortico-corticales lors de la perception de la parole se produisaient dans des bandes de fréquence différentes. Ils ont ainsi décrit une modulation du couplage des oscillations delta sur l'enveloppe d'amplitude de la parole par trois réseaux : 1) la puissance des oscillations beta dans les régions orbitofrontales (bilatérales avec une prédominance gauche) modulait le

couplage cortico-acoustique dans le gyrus temporal supérieur gauche, 2) la puissance des oscillations alpha dans les régions centrales (en général bilatérales) modulait négativement le couplage au sein du gyrus de Heschl et du gyrus temporal supérieur antérieur gauche (i.e. le couplage cortico-acoustique augmentait lorsque la puissance alpha au niveau central diminuait) et 3) la puissance des oscillations thêta dans les régions pariétales modulait négativement le couplage dans le gyrus temporal postérieur droit (i.e. le couplage cortico-acoustique augmentait lorsque la puissance thêta au niveau pariétal diminuait). Dans les deux études, les auteurs ont suggéré que ces mécanismes permettent de faciliter la segmentation du signal de parole en générant des prédictions *top-down*. En particulier, l'influence des régions frontales motrices sur le cortex auditif s'accorde avec les modèles à double voie (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009) et l'idée que les régions de production de la parole sont capables d'établir des prédictions sur le signal perçu de sorte à favoriser la catégorisation phonémique et ainsi la compréhension (Iacoboni, 2008 ; cf. chapitre II.5.). Le cortex moteur pourrait donc contrôler (*monitor*) le rythme de l'information auditive verbale (liée aux mouvements rythmiques articulatoires nécessaires à la production) afin de mieux anticiper l'occurrence des événements acoustiques pertinents pour la reconnaissance (Arnal, Doelling, & Poeppel, 2015; Morillon & Schroeder, 2015; Morillon & Baillet, 2017; Morillon, Schroeder, & Wyart, 2014; Strauß & Schwartz, 2017). Les oscillations alpha et beta au sein des régions motrices seraient des candidats idéaux pour remplir cette fonction de codage prédictif (Arnal, Doelling, & Poeppel, 2015; Arnal & Giraud, 2012). En accord avec cette interprétation, une étude en MEG menée par Alho et collaborateurs (2014) a par exemple mis en évidence une connectivité renforcée (synchronisation de phase), au sein de l'hémisphère gauche, entre le cortex prémoteur (ventral et dorsal) et la jonction temporo-pariétale, ainsi qu'entre le cortex auditif et le cortex prémoteur ventral dans la bande de fréquence beta (~20 Hz) dans une tâche où des adultes devaient identifier des syllabes bruitées. Cette synchronisation cortico-corticale était en outre corrélée aux performances des participants.

Enfin, d'autres types d'influences *top-down* sur l'alignement cerveau/parole dans les régions auditives ont été démontrés, que les modulations proviennent de processus attentionnels (e.g., Ding & Simon, 2012; Kerlin, Shahin, & Miller, 2010; Vander Ghinst et al., 2016; Zion Golumbic, Ding, et al., 2013; Zion Golumbic, Cogan, Schroeder, & Poeppel, 2013) ou de l'information visuelle (e.g., Luo, Liu, & Poeppel, 2010; Zion Golumbic, Cogan, Schroeder, & Poeppel, 2013). Dans une situation de type *cocktail party* où des phrases

différentes étaient présentées dans chaque oreille, Kerlin et collègues (2010) ont par exemple montré que la réponse oscillatoire θ sur le signal de parole était plus forte pour les phrases auxquelles l'auditeur devait prêter attention. Ce gain attentionnel était par ailleurs corrélé au pattern de latéralisation des oscillations alpha dans les régions occipito-pariétales (i.e. suppression alpha du côté controlatéral au côté où l'attention était dirigée, et augmentation alpha du côté ipsilatéral). D'autres études ont révélé que, de la même manière, la réponse oscillatoire auditive à la parole est amplifiée lorsque le signal est complété par une information visuelle congruente (Crosse, Butler, & Lalor, 2015; Luo et al., 2010; Park, Kayser, Thut, & Gross, 2016; Zion Golumbic, Ding, et al., 2013; Zion Golumbic, Cogan, et al., 2013). Autrement dit, le fait de voir les gestes articulatoires du locuteur aide les oscillations corticales auditives à se synchroniser sur les modulations lentes de l'enveloppe temporelle, améliorant la sensibilité aux événements saillants pour la segmentation (voir Peelle & Sommers, 2015 pour une revue). Luo et collègues (2010) ont présenté à des participants anglophones des extraits de films contenant des conversations dans lesquelles les informations auditive et visuelle étaient congruentes ou non. Leurs résultats ont révélé que la phase des oscillations dans les bandes delta et θ au sein des régions auditives et visuelles était modulée par les stimuli audio-visuels. De manière intéressante, les réponses dans ces régions étaient plus cohérentes entre les essais dans la condition congruente *vs* incongrue : dans les régions auditives par exemple, les modulations de phase pour les vidéos ayant la même information auditive étaient plus comparables lorsque l'information visuelle était congruente. En utilisant de la parole audio-visuelle continue également, Park et collègues (2016) ont quant à eux mis en évidence un *entrainement* des oscillations (1-7 Hz) du cortex visuel bilatéral sur les mouvements rythmiques des lèvres du locuteur, qui était plus fort (à gauche) lorsque l'attention des participants était dirigée sur l'information visuelle (cet alignement était en outre indépendant de l'alignement du cortex auditif sur le signal auditif de parole). L'activité des cortex moteur et prémoteur gauches (BA 4/6) était également synchronisée avec les gestes articulatoires perçus et modulée par l'attention ; en outre, elle prédisait de manière significative les performances de compréhension des participants. Ce dernier résultat souligne là encore le rôle prédictif des régions motrices qui permettraient de faciliter le traitement de la parole en anticipant l'occurrence des événements acoustiques pertinents pour la reconnaissance (Arnal, Doelling, & Poeppel, 2015; Morillon & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014; Strauß & Schwartz, 2017). Plus généralement, le bénéfice de l'information visuelle sur la perception de la parole pourrait s'expliquer par le fait que cette information visuelle, issue des gestes faciaux et manuels, coïncide avec le rythme

syllabique (thêta, pour les mouvements articulatoires de la bouche et de la mâchoire) et prosodique (delta, pour les mouvements de la tête ou des sourcils par exemple) de la parole produite et précède l'information auditive (~100 à 300 ms ; Chandrasekaran, Trubanova, Stillitano, Caplier, & Ghazanfar, 2009; Hadar, Steiner, Grant, & Rose, 1983). En accord avec la théorie de l'attention dynamique (Large & Jones, 1999), voir le visage du locuteur pourrait alors fournir des indices majeurs permettant de prédire *on-line* le signal verbal auditif à venir (Schroeder et al., 2008; Schwartz, Berthommier, & Savariaux, 2004; van Wassenhove, Grant, & Poeppel, 2005). Schroeder et collaborateurs (2008; voir aussi Zion Golumbic, Cogan, Schroeder, & Poeppel, 2013) ont notamment proposé que le rythme delta/thêta de l'*input* visuel puisse moduler la phase des oscillations du cortex auditif de l'auditeur dans les bandes de fréquence correspondantes. De cette manière, les événements saillants (syllabes) dans le signal auditif de parole pourraient être alignés avec les phases de haute excitabilité des oscillations thêta du cortex auditif, résultant en une amplification de la réponse oscillatoire. Un tel processus serait particulièrement important lorsque le signal de parole est dégradé par du bruit (McGettigan et al., 2012) ou dans des situations de *cocktail party* (Vander Ghinst et al., 2016; Zion Golumbic, Cogan, et al., 2013).

Tout au long du chapitre III.4., nous avons tenté de donner au lecteur un large aperçu des recherches examinant le rôle des oscillations cérébrales dans la perception de la parole. Si l'enthousiasme autour de cette thématique est toujours vif à l'heure actuelle, et les questions toujours plus nombreuses, il est relativement frappant de constater que la communauté s'est principalement intéressée à la manière dont le cerveau adulte se synchronise sur le signal de parole, laissant l'aspect développemental relativement peu investigué. Comme nous le verrons dans la suite, les aspects rythmiques de la parole constituent l'une des pierres angulaires du développement du langage oral, suggérant que les oscillations cérébrales puissent jouer un rôle important, équivalent voire plus important que celui observé chez l'adulte, dans les processus de perception de la parole. Le prochain chapitre aura pour objectif de présenter les quelques études ayant examiné la dynamique des oscillations cérébrales chez l'enfant présentant un développement typique, au repos ou lors du traitement de stimuli auditifs rythmiques.

IV. Les rythmes chez l'enfant

1. Rythme de la parole et développement du langage

La dynamique temporelle de la parole et sa structure prosodique sont particulièrement attirantes pour le nourrisson et il est aujourd'hui admis que l'information véhiculée par la structure quasi-rythmique de la parole joue un rôle fondamental dans l'acquisition et le développement du langage en aidant l'enfant à organiser les informations de parole. Dès la naissance, l'enfant est sensible au rythme de la parole et aux informations prosodiques (fréquence fondamentale, durée, intensité), qui constitueraient des indices majeurs lui permettant d'initialiser la discrimination et la segmentation du signal de parole mais aussi d'appréhender le traitement morphosyntaxique (Christophe, Gout, Peperkamp, & Morgan, 2003; Jusczyk, 1997; Jusczyk, 1995; Mehler, Dupoux, Nazzi, & Dehaene-Lambertz, 1996; Mehler et al., 1988; Morgan & Demuth, 1996; Nazzi & Ramus, 2003; Port, 2003). Dans une étude *princeps* utilisant la méthode de succion de haute amplitude (ou succion non nutritive), Mehler et collaborateurs (1988) ont démontré que dès les premiers jours de vie, des nouveau-nés français étaient capables de discriminer des phrases russes de phrases françaises, et qu'ils montraient une préférence pour leur langue maternelle comme révélé par une amplitude de succion plus forte à l'écoute du français. Dans une autre expérience, les auteurs ont mis en évidence que si les signaux de parole dans les deux langues étaient inversés temporellement (modifiant la structure temporelle caractéristique du signal), alors la discrimination n'était plus possible. A l'inverse, un filtrage du signal supprimant la majeure partie de l'information phonétique mais conservant l'information prosodique n'avait aucune incidence sur les capacités de discrimination des langues. Ces résultats suggèrent donc que les nouveau-nés exploitent très tôt les propriétés temporelles dynamiques et notamment prosodiques de la parole pour traiter ce signal si particulier (voir aussi Ramus, Hauser, Miller, Mauris & Mehler, 2000). Nazzi et collègues (1998) ont par la suite montré que la familiarité avec la langue n'était pas un prérequis nécessaire à la discrimination. Ainsi, des nouveau-nés français (dont la langue est syllabique) étaient capables de discriminer deux langues inconnues appartenant à deux classes rythmiques non natives différentes (anglais britannique, langue accentuelle vs japonais, langue moraïque) tandis que lorsqu'il s'agissait de deux langues appartenant à la même classe rythmique non native (accentuelle : anglais britannique vs néerlandais), aucune discrimination n'était mise en évidence (voir aussi Bosch & Sebastian-

Galles, 1997; Nazzi, Jusczyk, & Johnson, 2000 pour des études chez des bébés plus âgés de 4-5 mois montrant que la discrimination entre des langues d'une même classe rythmique est limitée à la classe à laquelle la langue maternelle appartient). D'autres études ont révélé le rôle joué par l'information prosodique dans la segmentation de la parole en syllabes, mots et propositions chez les bébés. Mandel et collaborateurs (1994) ont par exemple décrit que des bébés âgés de deux mois détectaient mieux des changements de phonèmes dans des phrases produites avec une intonation naturelle plutôt que dans des listes de mots. Jusczyk et collègues (1993) ont quant à eux montré que les bébés américains de 6 mois distinguaient des mots de leur langue de ceux d'une langue étrangère (norvégien) présentant une organisation prosodique distincte.

La prosodie attirerait ainsi l'attention des bébés en valorisant la structure phonétique et rythmique des unités au sein du signal de parole, leur permettant de structurer et de segmenter le flux continu plus facilement. Pour reprendre les propos de Mandel et collègues (1994), la prosodie serait une « glu perceptive » pour les séquences de parole. L'intonation est d'ailleurs l'un des paramètres que les adultes modifient particulièrement lorsqu'ils s'adressent à un très jeune enfant ; il s'agit du « parler bébé » (*motherese*) ou « langage adressé à l'enfant » (*child directed speech*) caractérisé par une exagération des contours prosodiques (i.e. plus grande musicalité) et du rythme, avec une hauteur (F0) plus élevée, un débit plus lent, une articulation soignée, des expressions plus courtes et des répétitions plus nombreuses (Khomsî, 1984; Snow & Ferguson, 1977). Les comptines enfantines sont également des productions linguistiques caractérisées par une forte accentuation syllabique (Bailly, 1997; Stern, 1974). Ces changements prosodiques s'accompagnent en outre souvent d'expressions faciales exagérées (e.g., grands sourires, haussement des sourcils) et de mouvements rythmiques du corps ou d'ajustements posturaux qui focalisent l'attention du bébé et renforcent son intérêt pour la communication. Plusieurs études ont d'ailleurs décrit une préférence des bébés pour le « parler bébé » en comparaison d'un discours adulte adressé à un autre adulte (Fernald, 1985; Pegg, Werker, & McLeod, 1992; Werker, Pegg, & McLeod, 1994).

D'autres études révèlent également le rôle fondamental des indices acoustiques rythmiques du signal de parole, et notamment de la durée syllabique et du temps de montée de l'enveloppe d'amplitude (*rise time*, reflétant les attaques des syllabes successives ; cf. partie I.3.), dans le développement de la phonologie et de la morphosyntaxe (Goswami, 2011; Nazzi & Ramus, 2003; Richards & Goswami, 2015). Comme nous l'avons vu dans la partie I.1., les

bébés, très tôt, sont particulièrement sensibles aux règles régissant la structure syllabique de leur langue (Bertoncini, Bijeljac-Babic, Jusczyk, Kennedy, & Mehler, 1988; Bertoncini & Mehler, 1981; Bijeljac-Babic, Bertoncini, & Mehler, 1993). Christophe et collègues (Christophe, Dupoux, Bertoncini, & Mehler, 1994) ont également montré que les nouveau-nés français de 3 jours étaient capables de détecter des différences de durée de syllabes marquant des frontières de mots (i.e. ils discriminent /mati/ issu de « mathématicien » de /mati/ issu des deux mots « schéma tigré »). Avant d'acquérir une conscience phonémique (i.e. capacité à reconnaître et manipuler les unités phonologiques ; généralement au moment de l'apprentissage de la lecture et de l'écriture, vers 5-6 ans), les enfants vont ainsi d'abord se baser sur des indices acoustiques leur permettant d'extraire les syllabes (conscience syllabique vers 3-4 ans) puis la structure de ces syllabes (attaque-rime ; vers 4-5 ans), et ce dans plusieurs langues de structure phonologique différente (Ziegler & Goswami, 2005).

La sensibilité à l'information syllabique est par ailleurs renforcée par la dimension multimodale de la parole : à l'information auditive quasi-rythmique s'ajoute de l'information visuelle caractérisée principalement par les mouvements articulatoires du locuteur. Ces mouvements correspondent à des oscillations de la mandibule (i.e. cycles d'ouverture et de fermeture) à une fréquence avoisinant les 4-7 Hz et sont à l'origine du cadre syllabique (Chandrasekaran et al., 2009; MacNeilage, 1998; MacNeilage & Davis, 2001). Dès leur naissance, les enfants sont ainsi en permanence stimulés par des informations verbales rythmiques auditives et visuo-motrices dont la congruence va leur permettre de segmenter le signal de parole plus facilement. En témoignent les résultats de Dodd (1979) montrant que les bébés de 4 mois s'agitent (i.e. sont mécontents) lorsque l'information verbale visuelle est présentée en décalage par rapport à l'information verbale auditive. Autrement dit, les bébés détectent, et sont perturbés, lorsque les mouvements rythmiques bucco-faciaux véhiculant l'information syllabique (« prosodie visuelle ») ne sont pas synchronisés avec l'information prosodique (débit syllabique) dans le flux auditif. Le rythme syllabique se retrouve très tôt dans la production du jeune enfant. Vers l'âge de 6-7 mois, les bébés commencent eux-mêmes à réaliser des mouvements cycliques d'ouverture et de fermeture de la mandibule, il s'agit de l'étape du babillage où émergent les premières formes syllabiques (Kern, 2001). Après cette étape assez rudimentaire, le babillage canonique se met ensuite en place entre 7 et 10 mois (Boisson-Bardies, 1996) : les syllabes ont une structure simple, principalement consonne-voyelle (CV) et sont le plus souvent redupliquées (« bababa ») et perceptuellement isochrones (Davis & MacNeilage, 1995). L'apparition de ces premières syllabes coïncide d'ailleurs avec

la production de mouvements rythmiques des bras et des mains (Iverson, 2010; voir Iverson & Thelen, 1999 pour un couplage entre les gestes rythmiques manuels et articulatoires). Le babillage canonique sera ensuite progressivement remplacé par un babillage dit diversifié où les syllabes diffèrent les unes des autres de par les consonnes ou les voyelles qui les composent. Chaque cycle d'ouverture et de fermeture de la mandibule, dont la fréquence d'oscillation se situe autour de 3 Hz chez le bébé (Canault, Laboissière, Perrier, & Sock, 2008; Canault, 2007; voir aussi Petitto, Holowka, Sergio, Levy, & Ostry, 2004 pour une étude chez des enfants nés de parents sourds-muets babillant avec les mains à un rythme de 3 Hz), constitue donc une forme syllabique et petit à petit, l'enfant va moduler ses productions à partir de ce qu'il perçoit de son entourage pour former les différentes unités phonologiques de sa langue (Green, Moore, Higashikawa, & Steeve, 2000; MacNeilage, 1998; MacNeilage & Davis, 2001; Strauß & Schwartz, 2017). Le babillage serait ainsi permis par l'interaction entre les systèmes de production et de perception de la parole, qui impliquerait la voie dorsale sensori-motrice du modèle à double voie (Hickok & Poeppel, 2007; cf. partie II.5.), permettant d'ajuster les sons de parole produits par l'enfant aux prototypes de phonèmes entendus autour de lui (Doupe & Kuhl, 1999; Kuhl, 2004; Pulvermüller & Fadiga, 2010).

En résumé, dès son plus jeune âge, l'enfant est donc exposé à une information acoustique et visuelle de parole dont les modulations rythmiques se situent dans la bande de fréquence thêta 4-7 Hz. Au stade du babillage, le bébé va lui-même commencer à produire des syllabes et à ajuster ses productions sur le signal verbal perçu, adoptant également un rythme mandibulaire autour de 3 Hz. Le rythme thêta constitue donc un rythme majeur dans le développement du langage oral ; peut-on alors retrouver une sensibilité particulière pour traiter ce rythme dans le cerveau en développement ?

2. Dynamique cérébrale oscillatoire chez l'enfant

2.1. Dynamique cérébrale oscillatoire chez l'enfant au repos

Comme proposé par le modèle AST (Poeppel, 2003) et montré par les études sur la dynamique corticale oscillatoire de repos chez l'adulte (Giraud et al., 2007; Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012; cf. chapitre III.4.2.), une prédisposition oscillatoire semble exister chez l'adulte au sein des hémisphères droit et gauche pour traiter les différentes échelles temporelles de la parole (rythmes syllabique thêta et phonémique

gamma respectivement). Au vu du rôle primordial de la structure rythmique de la parole dans l'acquisition et le développement du langage oral, il est raisonnable de penser que le système cérébral oscillatoire de l'enfant doit également présenter une telle prédisposition. Étonnamment, peu de travaux à notre connaissance ont étudié la dynamique des oscillations corticales dans les bandes de fréquence pertinentes pour la parole chez l'enfant au repos (voir Uhlhaas, Roux, Rodriguez, Rotarska-Jagiela, & Singer, 2010 pour une revue). Des études transversales menées de l'enfance à l'âge adulte ont montré une diminution exponentielle de la puissance des oscillations de basse fréquence delta, thêta et alpha basses (8-10 Hz) avec l'âge, alors que les oscillations de plus haute fréquence (alpha hautes, 10-13 Hz) et beta suivaient des dynamiques différentes, avec une augmentation de puissance dans l'enfance suivie d'une diminution (alpha en particulier) au cours de l'adolescence (Gasser, Verleger, Bacher, & Sroka, 1988; Lüchinger, Michels, Martin, & Brandeis, 2011, 2012; Michels et al., 2013). Whitford et collègues (2007) ont ainsi décrit une diminution de l'amplitude des oscillations cérébrales entre l'âge de 10 et 30 ans, particulièrement dans les bandes delta et thêta. Michels et collaborateurs (2013) ont également rapporté que la puissance des oscillations delta, thêta et alpha (8-13 Hz) était plus forte chez des enfants âgés de 10 ans en moyenne que chez des jeunes adultes (âge moyen 25 ans). Le même type de résultats a été observé au cours du sommeil lent (Campbell & Feinberg, 2009) : des modifications de l'activité oscillatoire se produisent en effet entre 11 et 16.5 ans avec une diminution de la puissance des oscillations delta et thêta, attribuée à la maturation synaptique observée à cette période (Feinberg & Campbell, 2010). Les oscillations intrinsèques gamma semblent quant à elles suivre une dynamique quelque peu différente, subissant des modifications plus tôt dans l'enfance, même s'il faut noter que les investigations dans cette bande de fréquence chez l'enfant au repos restent encore plus rares. Une étude de Takano et Ogawa (1998) chez des enfants de 3 à 12 ans a ainsi permis de montrer une augmentation non linéaire de la puissance gamma (35-45 Hz), en particulier au niveau frontal, avec l'âge avec un pic autour de 4-5 ans. Deux autres études menées chez des enfants plus jeunes ont par ailleurs révélé une corrélation entre l'activité oscillatoire spontanée gamma dans les régions frontales dans les trois premières années de vie et le développement du langage et des fonctions cognitives (Benasich, Gou, Choudhury, & Harris, 2008; Gou, Choudhury, & Benasich, 2011), suggérant un rôle fonctionnel des oscillations gamma basses dans la maturation des fonctions cognitives. Benasich et collaborateurs (2008) ont mis en évidence, chez des enfants âgés de 16, 24 et 36 mois, que la puissance des oscillations gamma (31-50 Hz) de repos dans les régions frontales et préfrontales était corrélée positivement aux performances langagières

expressives (16 et 24 mois) et réceptives (24 et 36 mois) des enfants ainsi qu'à leurs capacités attentionnelles (24 mois). La puissance gamma (31-50 Hz) était en outre moins forte chez les enfants présentant un risque de développer un trouble d'apprentissage du langage (historique familial) que chez les enfants contrôles appariés. Les auteurs ont suggéré que les variations dans la maturation cérébrale puissent moduler l'émergence de l'activité oscillatoire intrinsèque de haute fréquence, qui pourrait elle-même permettre le bon développement des fonctions cognitives et langagières. Dans une étude longitudinale chez les mêmes enfants, la même équipe (Gou, Choudhury, & Benasich, 2011) a par la suite démontré que les oscillations spontanées gamma aux âges de 16, 24 et 36 mois prédisaient les performances phonologiques (i.e. répétition de non-mots) et morpho-syntaxiques (i.e. langage réceptif et expressif ; corrélation également avec les oscillations à 16 mois) des enfants quelques années plus tard, à l'âge de 4-5 ans.

Une seule étude à notre connaissance s'est plus particulièrement intéressée à la fois aux oscillations thêta et gamma, pertinentes pour le traitement de la parole, chez les enfants au repos. Thompson et collaborateurs (2016) ont examiné en EEG l'activité oscillatoire endogène de 2 à 55 Hz chez des enfants âgés entre 3 et 5 ans. Leurs résultats ont révélé que les oscillations gamma (20-50 Hz) étaient plus fortement latéralisées à gauche comparées aux oscillations thêta (3-7 Hz) qui étaient distribuées symétriquement. Ce pattern ne changeait pas entre 3 et 5 ans ni en fonction du sexe des enfants. Les auteurs ont par ailleurs rapporté, en accord avec les deux études précédemment décrites, que les enfants qui percevaient le mieux la parole présentée dans du bruit étaient aussi ceux ayant une activité oscillatoire gamma endogène plus fortement latéralisée à gauche. Ces résultats s'accordent donc avec les prédictions du modèle AST (Poeppel, 2003), au moins concernant la bande de fréquence gamma. Selon les auteurs, l'absence de latéralisation thêta à droite entre 3 et 5 ans pourrait être liée au développement plus tardif du cortex auditif droit en regard du gauche (Paus et al., 1999). Rappelons néanmoins que des études chez l'adulte ont révélé la présence d'oscillations spontanées thêta également dans le cortex temporal gauche (Giraud et al., 2007; Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012). D'autres études développementales sur des tranches d'âge plus grandes seraient nécessaires afin d'examiner plus précisément la mise en place de la latéralisation des oscillations thêta dans l'hémisphère droit au repos.

2.2. Dynamique cérébrale oscillatoire chez l'enfant lors d'une stimulation auditive

Les études qui se sont intéressées à la synchronisation cortico-acoustique chez les enfants ont pour la plupart utilisé des stimuli non langagiers et toutes n'ont pas examiné la latéralisation hémisphérique de l'activité oscillatoire. Telkemeyer et collaborateurs (2009) ont mesuré l'activité cérébrale de nouveau-nés de 3 jours, en combinant l'EEG et la spectroscopie proche infrarouge fonctionnelle²¹ (fNIRS pour *functional Near-Infrared Spectroscopy*), lors de l'écoute de bruits complexes de structure temporelle variée (i.e. concaténation de segments de durées différentes – 12, 25, 160 et 300 ms – correspondant à des fréquences de 3, 6, 40 et 83 Hz respectivement). Leurs résultats ont montré que les modulations rapides (segments de 25 ms, 40 Hz) donnaient lieu à une réponse hémodynamique plus forte dans les régions temporales inférieure et postérieure bilatérales (*vs* 12 ms) et dans le cortex temporal inférieur gauche (*vs* 160/300 ms). La réponse vasculaire la plus faible était observée pour les modulations avec des segments de 12 ms et aucune différence n'était observée entre les conditions de modulations lentes 160 et 300 ms. Les auteurs ont en outre décrit une réponse plus importante du cortex temporo-pariétal droit pour ces deux dernières conditions. Les résultats en EEG n'ont en revanche pas révélé de différences entre les conditions, ce que les auteurs attribuent plutôt à des considérations méthodologiques. Dans une étude publiée deux ans plus tard (Telkemeyer et al., 2011) et utilisant le même protocole, les mêmes auteurs ont confirmé l'activation bilatérale du cortex temporal (inférieur gauche et temporo-pariétal droit) lors du traitement des modulations acoustiques rapides (segments de 12 et 25 ms) chez des bébé de 6 mois ; l'activation pour les modulations lentes (segments de 160 et 300 ms) était en revanche latéralisée à droite dans les régions frontale inférieure et temporo-pariétale, en accord avec les résultats chez les nouveau-nés (Figure 20A ; voir aussi Homae, Watanabe, Nakano, Asakawa, & Taga, 2006 et Perani et al., 2010 pour des activations à droite lors du traitement d'informations prosodiques et musicales (*pitch*) chez les bébés). Des analyses temps-fréquence ont par ailleurs révélé une désynchronisation oscillatoire thêta (4-8 Hz) pour les modulations lentes, qui était moins prononcée pour les modulations rapides (Figure 20B).

²¹ La NIRS est une technique non invasive d'imagerie optique qui mesure de manière indirecte l'activité du cerveau au moyen de capteurs optiques placés sur le cuir chevelu. Comme l'IRMf, la NIRS permet de mesurer les changements hémodynamiques associés à l'activité neuronale (Villringer, Planck, Hock, Schleinkofer, & Dirnagl, 1993) mais contrairement à l'IRMf, elle est sensible aux deux états de la molécule d'hémoglobine (oxygénée et désoxygénée/réduite). En émettant un faible faisceau de lumière dans le spectre du proche infrarouge sur le cuir chevelu d'un individu, il est ainsi possible de mesurer les changements de concentration en hémoglobine oxygénée et désoxygénée dans les régions activées par une tâche.

Dans l'ensemble, ces études démontrent donc que dès la naissance, le cerveau est déjà particulièrement sensible à la structure temporelle des stimuli auditifs, confortant l'idée que les caractéristiques acoustiques de la parole influencent la latéralisation hémisphérique dès les premières étapes d'acquisition du langage (Telkemeyer et al., 2011). Une étude de Vanvooren et collègues (Vanvooren, Poelmans, Hofmann, Ghesquière, & Wouters, 2014) chez des enfants âgés de 5 ans a corroboré ce pattern de latéralisation pour les rythmes lent (~syllabique) et rapide (~phonémique) en montrant des réponses corticales évoquées plus fortes à droite pour des bruits dont la fréquence de modulation était de 4 Hz et bilatérales pour des modulations de 20 Hz.

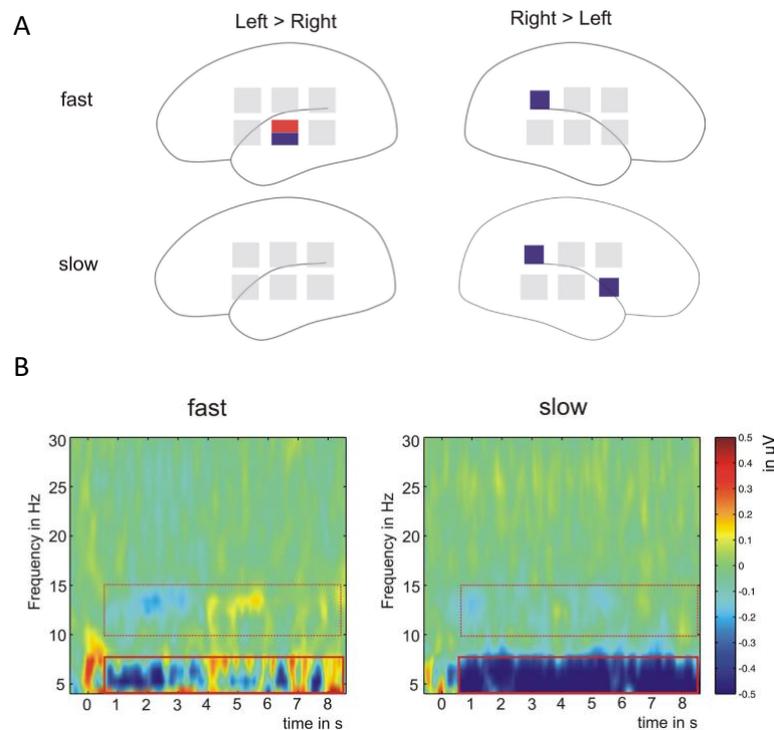


Figure 20 : (A) Réponses hémodynamiques mesurées par la spectroscopie proche infrarouge (NIRS) dans les hémisphères gauche et droit lors du traitement de modulations acoustiques rapides (segments de 12 et 25 ms) et lentes (segments de 160 et 300 ms) chez des bébés de 6 mois. Les carrés rouges indiquent une réponse significative de l'oxy-hémoglobine, les carrés bleus de la désoxy-hémoglobine. (B) Cartes temps-fréquence pour les conditions de modulations rapides (à gauche) et lentes (à droite). Les rectangles rouges en traits pleins montrent les différences significatives entre les conditions dans la bande thêta (4-8 Hz). Figures empruntées à Telkemeyer et al. (2011).

Enfin, dans l'une des rares études transversales, Shahin et collaborateurs (2010) ont examiné la trajectoire développementale des réponses oscillatoires mesurées en EEG lors de l'écoute de notes de musique (violon et piano) et de tons purs chez des participants âgés de 4 à 25 ans. Ils ont observé une augmentation de la synchronisation oscillatoire (*phase locking*) dans les bandes de fréquence thêta (4-8 Hz), alpha (8-14 Hz), beta (14-25 Hz) et gamma (25-70 Hz) avec l'âge ; ce couplage était en outre plus fort pour les stimuli musicaux, plus complexes spectralement, que pour les tons purs. Cependant, la trajectoire suivie par les basses et les hautes fréquences était différente, suggérant des mécanismes sous-jacents distincts. Alors que le couplage des oscillations thêta, alpha et beta était déjà présent vers 6-7 ans, avec une augmentation franche sur la période 8-13 ans, la synchronisation des oscillations gamma sur les stimuli était plus marquée à partir de 14 ans. Par ailleurs, les résultats ont montré que les basses fréquences (< 25 Hz) étaient sensibles à l'information temporelle, et notamment au *rise time* des stimuli, dès l'âge de 7 ans : le couplage de phase était différent (en termes de latence) entre le ton pur et la note de violon d'une part, et les notes de violon et de piano d'autre part, qui présentent des *onset* différents (~10 ms pour le ton pur et le piano vs ~50 ms pour le violon). Autrement dit, les oscillations thêta, alpha et beta étaient modulées par les caractéristiques temporelles des stimuli acoustiques. A l'inverse, le couplage des oscillations gamma n'était pas affecté par ces informations temporelles, suggérant qu'il soit lié à la complexité spectrale uniquement. Les auteurs ont proposé que la tranche d'âge 10-13 ans (voire 8-13 ans) constitue une période charnière de transition où l'activité oscillatoire de basse fréquence subirait une forte augmentation, en lien avec la maturation des connexions synaptiques dans les couches superficielles du cortex auditif observée entre 5 et 12 ans. L'activité oscillatoire gamma, particulièrement présente après cette période, pourrait alors reposer sur des mécanismes dépendants de cette maturation.

On retrouve donc déjà chez les enfants un encodage des caractéristiques spectrales et temporelles du signal auditif dans les bandes de fréquence sensibles aux différents rythmes de la parole, au moins chez l'adulte. Comme nous l'avons indiqué précédemment, ces études ont toutefois utilisé des stimuli non verbaux et il était important d'examiner la synchronisation des oscillations corticales sur le signal de parole perçu par l'enfant. Là encore, une revue de la littérature permet de constater un nombre incroyablement limité d'études à ce sujet. Dans une étude en EEG, Abrams et collaborateurs (2008) se sont intéressés au traitement cortical de l'information rythmique syllabique chez des enfants âgés de 9 à 13 ans. Alors que les enfants

regardaient un film, une même phrase leur était présentée dans trois modes différents : de la parole claire (« *clear speech* », caractérisée par une excellente diction et donc une intelligibilité parfaite), de la parole conversationnelle (i.e. informelle, caractérisée par une moins bonne articulation, une intensité plus faible et donc une intelligibilité sensiblement moindre) et de la parole compressée temporellement à 50% de la durée originale (réduisant l'intelligibilité). Ces trois stimuli différaient donc notamment de par leur enveloppe temporelle, où les modulations d'amplitude étaient plus marquées dans la parole claire vs conversationnelle et plus rapides dans la parole compressée. Les résultats ont révélé un couplage entre l'enveloppe d'amplitude du signal de parole et les réponses corticales mesurées au niveau des régions temporales, quel que soit le mode de production de la phrase (c'est-à-dire que les modulations de l'enveloppe soient grandes dans la parole claire intelligible ou plus rapides dans la parole compressée ; Figure 21A). De manière intéressante, les auteurs ont en outre observé que la réponse évoquée dans le cortex auditif droit « suivait » plus finement le contour de l'enveloppe de la phrase que celle du cortex auditif gauche (Figure 21B ; voir aussi Abrams, Nicol, Zecker, & Kraus, 2009). Cette dominance hémisphérique droite pour traiter les fluctuations (quasi-rythmiques) lentes de la parole, correspondant au rythme syllabique, s'accorde donc avec le modèle AST de Poeppel (2003). Une autre étude en EEG menée par Power et collègues (2012) chez des enfants âgés de 13 ans en moyenne a montré que les oscillations des cortex auditif droit et gauche étaient capables de se synchroniser sur un signal verbal rythmique. Le stimulus utilisé était la syllabe /ba/ répétée à une fréquence de 2 Hz (bande delta), qui était présentée en modalité soit auditive uniquement, soit visuelle uniquement (i.e. vision du locuteur prononçant la séquence en absence d'information auditive), soit audio-visuelle (i.e. vision du locuteur prononçant la séquence avec l'information auditive). Les enfants devaient écouter/regarder les séquences et appuyer sur un bouton dès qu'une syllabe perturbait le rythme isochrone (i.e. la syllabe était présentée avec un certain retard par rapport à ce qui était attendu). Les auteurs ont mis en évidence une synchronisation (*entrainment*) des oscillations delta (1-3 Hz) sur les stimuli dans les trois conditions. Une synchronisation thêta (3-5 Hz) était également observée lorsque l'information auditive était présente, à savoir dans les conditions auditive et audio-visuelle. Ce couplage thêta était en outre corrélé aux performances langagières des enfants dans des tâches de lecture et de dénomination rapide d'objets (phonologie), confortant l'idée d'un rôle de l'activité oscillatoire thêta dans le découpage et l'encodage syllabique (que ce soit à l'oral ou à l'écrit).

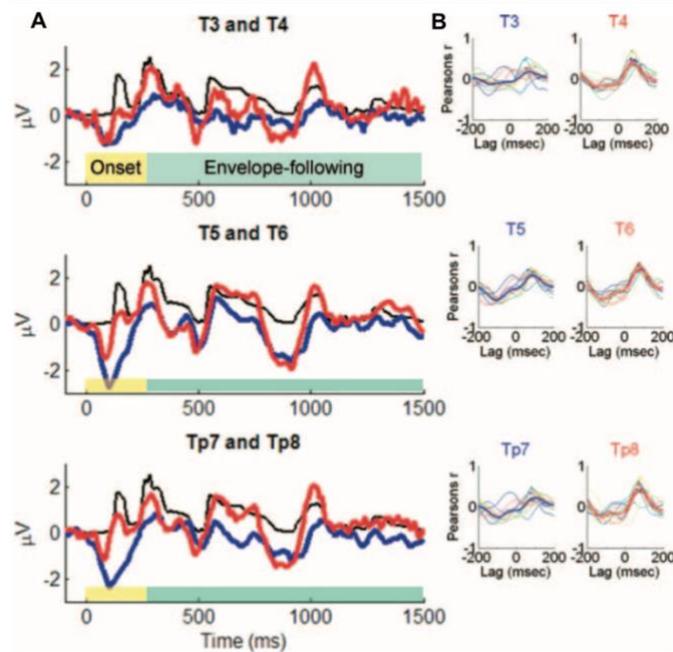


Figure 21 : (A) Réponses corticales au niveau de 3 paires d'électrodes temporales et enveloppe temporelle de la parole claire intelligible. La ligne noire correspond à l'enveloppe de parole, la ligne rouge à la réponse corticale dans l'hémisphère droit et la ligne bleue représente l'activité corticale à gauche. (B) Corrélogrammes entre l'enveloppe temporelle de la parole claire intelligible et les réponses corticales individuelles des sujets pour chaque paire d'électrodes. Figure empruntée à Abrams et al. (2008).

Pour résumer, les données disponibles à ce jour suggèrent donc que les oscillations cérébrales des enfants sont capables de se synchroniser sur un stimuli auditif rythmique langagier ou non, dès la naissance, et suivant un pattern de latéralisation hémisphérique globalement similaire à celui décrit chez l'adulte (Giraud & Poeppel, 2012; Hickok & Poeppel, 2007). Une telle dynamique oscillatoire pourrait refléter la sensibilité des enfants aux informations rythmiques prosodique et syllabique du signal de parole, qui constitue un fondement de l'acquisition et du développement du langage. Dans cette optique, on peut raisonnablement penser qu'un dysfonctionnement oscillatoire puisse altérer le traitement du rythme de la parole et par conséquent sa segmentation en unités linguistiques pertinentes, affectant les capacités langagières. Sans une bonne segmentation du signal de parole, le développement de la phonologie peut en effet être déficitaire et conduire par la suite à différents troubles du langage oral et écrit. Dans le cadre de cette thèse, nous nous sommes particulièrement intéressés à un trouble du langage oral lié à un déficit phonologique, la

dysphasie. Dans la suite de cette Introduction, nous en décrivons les caractéristiques et présenterons des preuves expérimentales d'un déficit de traitement du rythme dans ce trouble neurodéveloppemental.

V. La dysphasie

1. Définition

La dysphasie est un trouble du langage oral chez l'enfant. Elle appartient aux troubles développementaux spécifiques des apprentissages, également appelés troubles « DYS » et touchant près de 40 % des enfants. Ces troubles sont définis comme un ensemble hétérogène de troubles causés par une dysfonction, détectée ou non, du système nerveux central au cours du développement, mais n'ayant pas pour origine un handicap visuel, auditif ou moteur, une arriération mentale, un trouble affectif ou un milieu défavorisé (Mazeau & Pouhet, 2014).

L'étude des troubles du langage oral chez l'enfant a connu un essor considérable à partir de la deuxième moitié du XX^{ème} siècle. Des travaux en psychologie, linguistique, neurosciences et génétique se sont développés et de nombreuses terminologies ont été proposées afin de labelliser ces troubles. Dans les années 50, Ajuriaguerra et collègues ont défini sous le terme de « audimuté » un groupe de patients présentant des difficultés pour s'exprimer et comprendre (Ajuriaguerra, Borel-Maisonny, Diatkine, Narlian, & Stambak, 1958). Quelques années plus tard, les termes « dysphasie » (Ajuriaguerra et al., 1962), « aphasie congénitale » (Benton, 1964) ou « trouble développemental du langage » ont été proposés, accentuant l'aspect développemental du trouble (Rapin & Allen, 1983). Dans la littérature scientifique internationale, des terminologies mettant en avant la sévérité et la spécificité des troubles du langage sont majoritairement employées. On retrouve alors principalement « *specific language impairment* » (SLI ; Bishop, 1992) ou « trouble spécifique du langage » (TSL) mais aussi « *developmental language disorder/impairment* » (DLD ; International Classification of Diseases, ICD-10 ; World Health Organization, 2005), « *primary language impairment* » ou encore « *developmental dysphasia* » (principalement en Europe : Hüneke & Lascelles, 2014; Reilly, Tomblin, et al., 2014; Reilly, Bishop, & Tomblin, 2014; Bishop, 2017). Dans la dernière version du DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association, 2013*) la dysphasie est classée dans la rubrique des « troubles de la communication » sous le terme de « trouble du langage » (« *Language Disorder* »), reflétant le déficit persistant d'acquisition et d'utilisation du langage dans ses différentes modalités (parlé, écrit, langue des signes ou autre) résultant d'un trouble de la compréhension ou de l'expression.

Cependant, aucune de ces terminologies ne fait consensus auprès de la communauté clinique et scientifique, et la terminologie concernant les troubles du langage oral est toujours au centre d'un débat international très vif (Bishop, 2014, 2017; Reilly, Tomblin, et al., 2014; Reilly, Bishop, et al., 2014). SLI est le terme le plus couramment retrouvé dans la littérature scientifique internationale et la recherche académique. Cependant, certains chercheurs et les cliniciens considèrent qu'il n'est pas adapté pour définir un enfant présentant des troubles du langage oral. Le terme « spécifique » implique en effet que le trouble du langage se produit en l'absence d'autres déficits du développement. Or, l'ensemble des cliniciens s'accordent sur le fait qu'un trouble du langage oral apparaît rarement seul ; des troubles associés sont pratiquement toujours observés tels que des difficultés d'apprentissage de la lecture, des troubles de la coordination motrice, des troubles de l'attention, etc (Leonard, 2014; Reilly, Tomblin, et al., 2014; Rice, 2015). Nous détaillerons plus précisément la comorbidité de la dysphasie plus tard. Par conséquent, suivant les recommandations de *l'American Speech and Hearing Association* (ASHA), le terme SLI a été modifié par « *language disorder* » (LD) dans la dernière version du DSM-5. Toutefois, Bishop (2014) a très justement remarqué que l'acronyme de ce terme pouvait être confondu avec les troubles des apprentissages définis sous le terme « *Learning Difficulty* » et qu'une recherche dans un moteur de recherche pourrait donner de nombreux résultats sans qu'ils ne soient forcément liés au trouble du langage des enfants. Elle a alors proposé le terme « *Primary Language Impairment* » afin de préciser que le trouble développemental du langage n'est pas secondaire à un autre trouble, mais bien le trouble principal. Certains auteurs relèvent cependant qu'il n'est pas toujours facile de juger quelle condition est primaire ou secondaire lorsqu'un enfant présente des troubles associés (Conti-Ramsden, 2014). Tout récemment, le terme « *Developmental Language Disorder* » (DLD) a été choisi par un groupe d'experts de l'ICD-11 (Bishop, 2017). Ce terme, peu controversé encore, souligne bien la nature congénitale du trouble du langage. Toutefois, il faut noter que le terme « *developmental* » peut être mal interprété par les patients et leurs familles, le considérant comme inapproprié lorsque le patient grandit et devient adulte (Clark & Carter, 2014) ; c'est notamment ce qui est généralement observé avec la dyslexie développementale.

En France, le terme « dysphasie » reste encore prédominant dans le milieu clinique à l'heure actuelle, c'est pourquoi nous adopterons cette terminologie dans la suite de cette thèse.

La dysphasie est définie comme un trouble neurodéveloppemental sévère et durable de l'élaboration du langage oral, non lié à un déficit auditif, à un déficit cognitif (principalement un retard mental), à un trouble envahissant du développement (spectre autistique), à un trouble neurologique (lésion acquise, malformation des organes phonateurs, déficit sensoriel), ou encore à une carence affective ou éducative (Gérard, 1993; Maillart & Schelstraete, 2012). Elle peut toucher l'ensemble des composantes langagières de l'expression et/ou de la compréhension du langage de manière plus ou moins sévère : le son et son organisation dans la langue (phonologie), le vocabulaire (lexique), l'organisation des mots dans les phrases et les marques grammaticales (morphosyntaxe), la mélodie du discours (prosodie), le sens des sons (sémantique) et leur signification selon le contexte (pragmatique). De plus, l'atteinte langagière sera toujours supérieure aux autres fonctions cognitives : les enfants dysphasiques sont caractérisés par un quotient intellectuel verbal nettement inférieur au quotient intellectuel non-verbal (Parijsse & Maillart, 2010).

Ainsi, la dysphasie est caractérisée par la diversité des difficultés langagières observées et elle est considérée comme un trouble spécifique dans le sens où aucune autre cause ne peut expliquer les performances langagières inférieures aux normes établies pour l'âge (Parijsse & Maillart, 2010). Dans la littérature francophone, une distinction est faite entre la dysphasie et le simple retard de langage, ce qui n'est pas le cas dans la littérature internationale (DSM-5, 2013; de Weck & Marro, 2010; Schelstraete & Collette, 2012). Les enfants présentant un retard simple de langage vont acquérir et développer les différentes composantes langagières de façon plus lente par rapport à la courbe développementale générale. Ce retard est normalement rattrapé à la fin de l'école maternelle, soit vers 5-6 ans (voir Desmarais, Sylvestre, Meyer, Bairati, & Rouleau, 2008, pour une revue de la littérature). Chez les enfants dysphasiques, les troubles langagiers vont au contraire persister au-delà de 5-6 ans, malgré la mise en place d'une rééducation orthophonique (si elle est mise en place), et seront caractérisés par des productions langagières déviantes qui ne sont pas caractéristiques de l'évolution normale du langage (Gérard, 1993). Le diagnostic de dysphasie sera donc généralement posé à cet âge et visera à valider le déficit dans un ou plusieurs domaines langagiers d'une part, et le maintien d'un raisonnement intellectuel perceptif dans la norme d'autre part.

La dysphasie touche 3 à 7 % des enfants d'âge scolaire (Tomblin et al., 1997), avec une prédominance chez les garçons, spécifiant le caractère familial de ce trouble. Elle est souvent associée à d'autres troubles DYS et notamment dans 45 % des cas à une dyslexie (trouble de

l'acquisition et de l'automatisation de la lecture ; Snowling, Bishop, & Stothard, 2000). On retrouve aussi régulièrement des troubles de l'attention avec ou sans hyperactivité (TDAH) ou encore des troubles de la motricité fine et globale (dyspraxie). Enfin, les enfants dysphasiques peuvent présenter des difficultés à s'orienter dans l'espace et le temps ; des troubles du comportement peuvent également se développer suite aux difficultés de communication de ces enfants (Baker & Cantwell, 1991; Hill, 2001).

2. Classification des dysphasies

Face à la grande hétérogénéité du tableau clinique de la dysphasie (selon la modalité et la fonction langagière atteinte), un certain nombre de classifications ont été proposées afin de répertorier l'ensemble des formes possibles. Cependant, comme nous l'avons vu, aucune ne fait réellement consensus et ce n'est que très récemment qu'un terme (*Developmental Language Disorder*) qui pourrait contenter la communauté internationale a été proposé (Bishop, 2014, 2017; Reilly, Tomblin, et al., 2014; Reilly, Bishop, et al., 2014). Dans le cadre de ce travail de thèse, nous allons principalement détailler la classification proposée par Rapin et Allen (1983) qui est la classification majoritairement utilisée par les cliniciens en France. Nous présenterons ensuite son adaptation réalisée par Gérard (1993).

La classification de Rapin et Allen (1983) s'appuie sur une approche neurolinguistique distinguant 3 groupes de troubles, les troubles prédominants au niveau expressif, réceptif ou touchant les processus de traitement central et de la formulation. A partir de ces trois groupes, 6 profils cliniques sont proposés :

- **La dyspraxie verbale.** Le langage expressif est absent ou très réduit suite à un déficit d'organisation, de programmation et de production des gestes articulatoires associés aux sons de la parole et aux séquences verbales, en l'absence de trouble neuromusculaire tel qu'une dysarthrie. Les représentations phonologiques sont correctement encodées mais le trouble praxique rend difficile leur traduction en représentations phonétiques. Ces enfants présentent par exemple des difficultés pour répéter des mots de trois syllabes. Leur expression n'est pas fluide avec un vocabulaire pauvre, des mots tronqués, isolés et sans structure grammaticale. Elle est le plus

souvent de style télégraphique. Un mode de communication signé est souvent observé. La compréhension verbale n'est quant à elle que très légèrement atteinte²².

- **Le trouble de la programmation ou de la production phonologique.** L'expression est abondante. Les enfants sont bavards mais ils présentent des troubles de l'articulation. Ils sont capables de répéter des phonèmes isolés mais ne peuvent les produire spontanément. La syntaxe semble être présente mais n'est pas respectée, rendant le discours inintelligible. La compréhension est affaiblie mais toujours meilleure que la production.
- **Le syndrome phonologico-syntaxique.** Il s'agit de la forme la plus fréquente de dysphasie. Le trouble touche de façon sévère la phonologie et l'expression syntaxique, réduisant fortement les capacités d'expression et de communication des enfants. Le discours présente une hypospontanéité et l'expression est marquée par un vocabulaire pauvre, une articulation déviante et une syntaxe réduite à des phrases courtes et dénuées de mots de fonction, donnant au discours un style télégraphique. Un fort agrammatisme est relevé. L'enfant imite le discours par l'emploi d'un jargon. La compréhension est au contraire peu voire pas touchée. Un trouble du décodage serait à l'origine de ces difficultés.
- **L'agnosie auditivo-verbale.** Dans ce type de dysphasie, une perturbation du système auditif, sans atteinte toutefois du système périphérique, empêche l'enfant d'analyser les sons de la parole. Il est alors en échec dans des tâches de discrimination et d'identification de phonèmes. Le sens des sons ne peut être extrait, entraînant un trouble important de la compréhension du langage, alors que la compréhension non verbale est satisfaisante. L'enfant se comporte alors comme s'il était sourd avec une expression orale inexistante ou rare et déformée sur le plan de l'articulation et de la prosodie. Une communication par un langage signé est souvent observée chez ces enfants.
- **Le syndrome lexical sémantique.** Il s'agit d'un trouble de l'évocation des mots (dysnomie). L'accès lexical est limité, la construction syntaxique immature mais non aberrante. Les enfants ont des difficultés à apprendre le nom des lettres et à retenir des

²² La dyspraxie verbale est une entité nosographique extrêmement rare (prévalence estimée à 0.125% de la population d'âge scolaire, Delaney et Kent 2004) et non reconnue de tous. Pour certains auteurs, il ne s'agirait que d'une variante d'une forme de dysphasie, notamment de la dysphasie de production phonologique d'après Mazeau (2005). D'autres auteurs tels que Kent (2000) et Dewey (1993) proposent que la dyspraxie verbale soit liée à un trouble plus général du développement des praxies, dont le pendant chez l'adulte serait l'apraxie verbale (atteinte acquise de la programmation des mouvements orofaciaux).

mots nouveaux. Le langage spontané produit en réponse à des questions orales est marqué par des hésitations, des faux départs, une lenteur de réponse, des essais d'autocorrection, l'utilisation de mots de remplissage et des paraphrasies sémantiques ou phonologiques. La dénomination d'objets est bonne par rapport à l'évocation sur demande, en raison d'un trouble d'accès lexical. L'articulation, la fluidité et l'intelligibilité du discours sont préservées.

- **Le syndrome sémantique-pragmatique.** Il touche essentiellement la compréhension, l'expression étant de bonne qualité. Le langage peut se développer normalement ou apparaître tardivement et entraîner un comportement logorrhéique. Une bonne mémoire auditive permet de reproduire des phrases correctes mais avec un caractère stéréotypé. La compréhension de mots isolés est bonne, mais elle est déficitaire lors de questions ouvertes ou de situations inhabituelles. L'interprétation des métaphores et des contenus abstraits est également problématique. Les enfants présentent une faible intention de communiquer ainsi qu'un non-respect des règles de communication (contact visuel pauvre, non-respect du tour de parole), ce qui distingue ce syndrome des autres types de dysphasie.

Gérard (1993) a ensuite adapté cette classification en se basant sur le modèle neuropsychologique de Crosson (1985 ; Figure 22) qui décrivait alors les relations réciproques entre trois centres corticaux de l'hémisphère gauche : (1) les centres antérieurs qui programment l'encodage linguistique (production de la parole) à l'aide de deux modules : le centre de formulation définissant le contenu sémantique et syntaxique, et le centre de programmation qui régit les opérations nécessaires à l'actualisation de ce contenu ; (2) les centres postérieurs qui décodent et donnent un sens aux phonèmes (perception et compréhension du message linguistique); et (3) les centres sous-corticaux qui effectuent un contrôle sur les deux autres centres et permettent leur bonne cohérence.

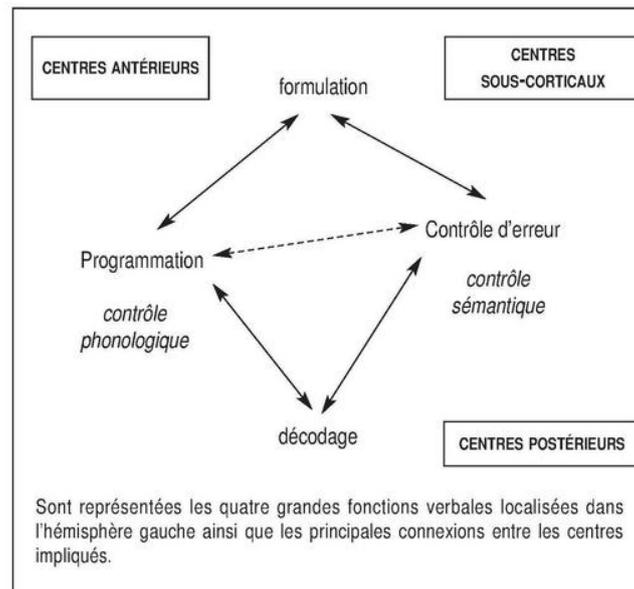


Figure 22 : Modèle de Crosson (1985) selon Gérard (1993). Figure empruntée à Gérard (1993).

Selon Gérard (1993), un ou plusieurs dysfonctionnements de ces centres seraient à l'origine de chaque profil de dysphasie. Ainsi, l'auteur distingue 5 formes de dysphasie à partir de la classification de Rapin et Allen (1983) :

- Le syndrome phonologico-syntaxique qui serait lié à une défaillance de la jonction formulation-programmation.
- Le trouble de production phonologique résultant d'un déficit au niveau du contrôle de la production phonologique, c'est-à-dire de la mise en place des organes bucco-phonatoires pour produire un son spécifique.
- La dysphasie réceptive qui serait due à un trouble du décodage de l'information.
- La dysphasie mnésique ou lexicale syntaxique, liée à un trouble du système de contrôle sémantique.
- La dysphasie sémantique-pragmatique résultant d'un déficit de la fonction de formulation.

Ces différentes classifications permettent de mieux diagnostiquer un enfant dysphasique et ainsi de lui proposer un suivi orthophonique et une remédiation adaptée à son trouble. Cependant, le profil d'un enfant dysphasique n'est pas stable dans le temps et peut varier d'une forme de dysphasie à une autre au cours des années (Botting & Conti-Ramsden, 2004). Cette dynamique du profil dysphasique est liée au développement en cours du langage chez

les enfants (avec notamment l'entrée dans l'écrit), mais aussi aux stratégies de compensation qu'ils sont susceptibles de mettre en place pour combler leurs difficultés. Il est donc important de suivre cette évolution et d'adapter la rééducation en permanence.

3. Etiologie des dysphasies

L'origine de la dysphasie pose aujourd'hui encore question et n'a pas trouvé de véritable consensus, probablement parce qu'il existe des origines multiples, et non une origine unique, à ce trouble développemental. Face à l'hétérogénéité des déficits, plusieurs théories ou modèles explicatifs ont été proposés, des théories linguistiques expliquant les troubles langagiers dans la dysphasie comme résultant d'un déficit de la connaissance linguistique (i.e. délai de maturation ou représentations langagières déficitaires) aux théories basées sur un déficit dans des traitements cognitifs spécifiquement linguistiques ou non (Maillart & Schelstraete, 2012), en passant par des hypothèses génétiques (Bishop, 2006) ou relevant d'anomalies neurologiques (Chiron et al., 1999). Toutefois, deux courants semblent dominer à l'heure actuelle, à savoir ceux basés sur des facteurs neurobiologiques et génétiques comme origines possibles de la dysphasie.

3.1. Les facteurs neurobiologiques

Les études en imagerie cérébrale ont permis de relever un certain nombre de particularités neurologiques chez les enfants présentant une dysphasie.

Au niveau fonctionnel d'abord, des anomalies au sein de l'ensemble du système langagier semblent présentes chez ces enfants par rapport à la population générale (Badcock, Bishop, Hardiman, Barry, & Watkins, 2012; Chiron et al., 1999; Cohen, Campbell, & Yaghami, 1989; de Guibert et al., 2011; Dibbets, Bakker, & Jolles, 2006; Ellis-Weismer, Plante, Jones, & Tomblin, 2005; Hugdahl et al., 2004; Hwang et al., 2006; Leonard, Eckert, Given, Virginia, & Eden, 2006; Ors et al., 2005; Tzourio, Heim, Zilbovicius, Gerard, & Mazoyer, 1994; Whitehouse & Bishop, 2008). Dans une étude en IRMf chez des enfants francophones âgés de 7 à 18 ans, de Guibert et collaborateurs (2011) ont étudié la réponse neuronale d'un groupe d'enfants dysphasiques comparé à un groupe d'enfants contrôles lors de différentes tâches de langage (deux tâches lexico-sémantiques auditives et deux tâches phonologiques visuelles). Ils ont observé une absence de latéralisation hémisphérique gauche

dans les régions périsylviennes « classiques » du langage (*pars opercularis* et *triangularis* du gyrus frontal inférieur, gyrus supramarginal et gyrus temporal supérieur ou STG) chez les dysphasiques (Figure 23A). Cette symétrie hémisphérique au sein des régions périsylviennes a également été observée par Badcock et collaborateurs (2012). Ces auteurs ont comparé des enfants dysphasiques à des enfants contrôles lors d'une tâche auditive de définition de mots en IRMf. Les enfants entendaient la définition d'un mot et devaient le nommer silencieusement. Une condition contrôle leur était présentée correspondant à une définition en parole inversée (rendant la définition inintelligible). Les enfants dysphasiques présentaient une activité significativement réduite au sein du gyrus frontal inférieur gauche, du STG gauche et du putamen droit lors de la condition de parole normale. Les auteurs ont également mesuré l'indice de latéralisation de chaque groupe de participants pour chaque condition de parole ; ils ont observé une latéralisation moins marquée chez les dysphasiques lors de la perception de parole avec même une tendance à une latéralisation plus forte à droite.

de Guibert et collaborateurs (2011) ont également révélé une hypoactivation de l'aire de Wernicke au carrefour des régions temporales supérieure postérieure et supramarginale gauches chez ces enfants lors d'une tâche de définition de mots (i.e. nommer silencieusement un mot après écoute de la définition du concept correspondant, par exemple « éléphant » pour « un gros animal avec une trompe » ; Figure 23B). Cette hypoactivation était particulièrement forte dans la région pariéto-temporale gauche Spt qui jouerait un rôle d'interface sensori-motrice au sein de la voie dorsale du modèle à double voie de perception de la parole (Hickok & Poeppel, 2007; cf. partie II.4.) et qui serait notamment impliquée dans la perception de la parole complexe/dégradée et dans les processus phonologiques en perception et production (Buchsbaum, Hickok, & Humphries, 2001; Price, 2010). Une lésion de cette zone a d'ailleurs été rapportée dans des cas d'aphasie de conduction principalement caractérisée par des troubles de la répétition (omissions, paraphasies phonémiques) avec une compréhension préservée (Anderson et al., 1999; Buchsbaum et al., 2011; Dronkers, 2000; Hillis, 2007). Hugdahl et collègues (2004) ont également décrit une hypoactivation des régions temporales gauches lors de l'écoute de sons de parole (voyelles isolées, pseudomots et mots) chez les individus d'une même famille présentant une dyspraxie verbale par rapport à des contrôles. Cette hypoactivation était particulièrement marquée au sein du gyrus temporal médian et du sillon temporal supérieur dans le groupe présentant une dyspraxie verbale par rapport au groupe contrôle. Des résultats similaires ont été observés dans des études utilisant la tomographie par émission monophotonique (SPECT), qui est une technique d'imagerie

cérébrale permettant de mesurer le flux sanguin cérébral au repos ou lors d'une tâche cognitive (Chiron et al., 1999; Hwang et al., 2006; Lou, Henriksen, & Bruhn, 1984; Tzourio-Mazoyer et al., 1994). Ces études réalisées sur un groupe d'enfants dysphasiques comparés à des enfants contrôles ont ainsi montré une hypoperfusion des régions périssylviennes antérieures chez les dysphasiques au repos (Lou et al., 1984) et des régions pariétales inférieures droites au repos (Hwang et al., 2006) et lors d'une tâche de discrimination de phonèmes (Tzourio-Mazoyer et al., 1994) ainsi qu'au sein de la région de Broca lors d'une tâche d'écoute dichotique (écoute de paires de nombres, Chiron et al., 1999).

Enfin, de Guibert et al. (2011) ont décrit une hyperactivation à droite au niveau de l'insula antérieure, du gyrus frontal inférieur et de la tête du noyau caudé lors d'une tâche phonologique dans laquelle les enfants dysphasiques devaient nommer silencieusement un triplet d'images dont les noms différaient sur un trait phonologique au niveau du premier phonème (e.g., /pul/-/bul/-/mul/ ; Figure 23C). Ces régions cérébrales appartiennent au système langagier et jouent notamment un rôle dans les aspects moteurs de la parole (coordination et contrôle moteur de l'articulation et du tractus vocal ; Keller, Crow, Foundas, Amunts, & Roberts, 2009; Price, 2010). L'insula antérieure et le gyrus frontal inférieur joueraient également un rôle dans les aspects suprasegmentaux de la parole (prosodie) avec une prédominance au sein de l'hémisphère droit (Ackermann & Riecker, 2010; Bohland & Guenther, 2006; Brown, Ingham, Ingham, Laird, & Fox, 2005). L'hyperactivation de ces régions dans la dysphasie pourrait alors refléter des mécanismes compensatoires prosodiques et/ou articulatoires (pour l'insula et le gyrus frontal inférieur) et d'initiation des processus phonologiques (pour le noyau caudé) dans le but de pallier le déficit phonologique, en particulier dans un contexte de conflit phonologique (par exemple, la différence entre les paires minimales pain / bain / main ; de Guibert et al., 2011 ; voir aussi Hugdahl et al., 2004; Vargha-Khadem, Gadian, Copp, & Mishkin, 2005). Ors et collègues (2005) ont quant à eux observé à l'aide de la SPECT chez des enfants dysphasiques au repos, une hypoperfusion des régions pariétales droites et une hyperperfusion des régions temporales droites. La distribution du flux sanguin cérébral semble donc symétrique au sein des régions temporales chez les enfants dysphasiques, alors qu'elle est plus forte à gauche chez le groupe contrôle.

L'ensemble de ces études suggère donc une activité cérébrale atypique chez les individus présentant une dysphasie avec notamment une baisse de la supériorité de l'hémisphère gauche dans les processus langagiers. Cette latéralisation atypique constituerait ainsi un marqueur biologique spécifique de la dysphasie (Whitehouse & Bishop, 2008).

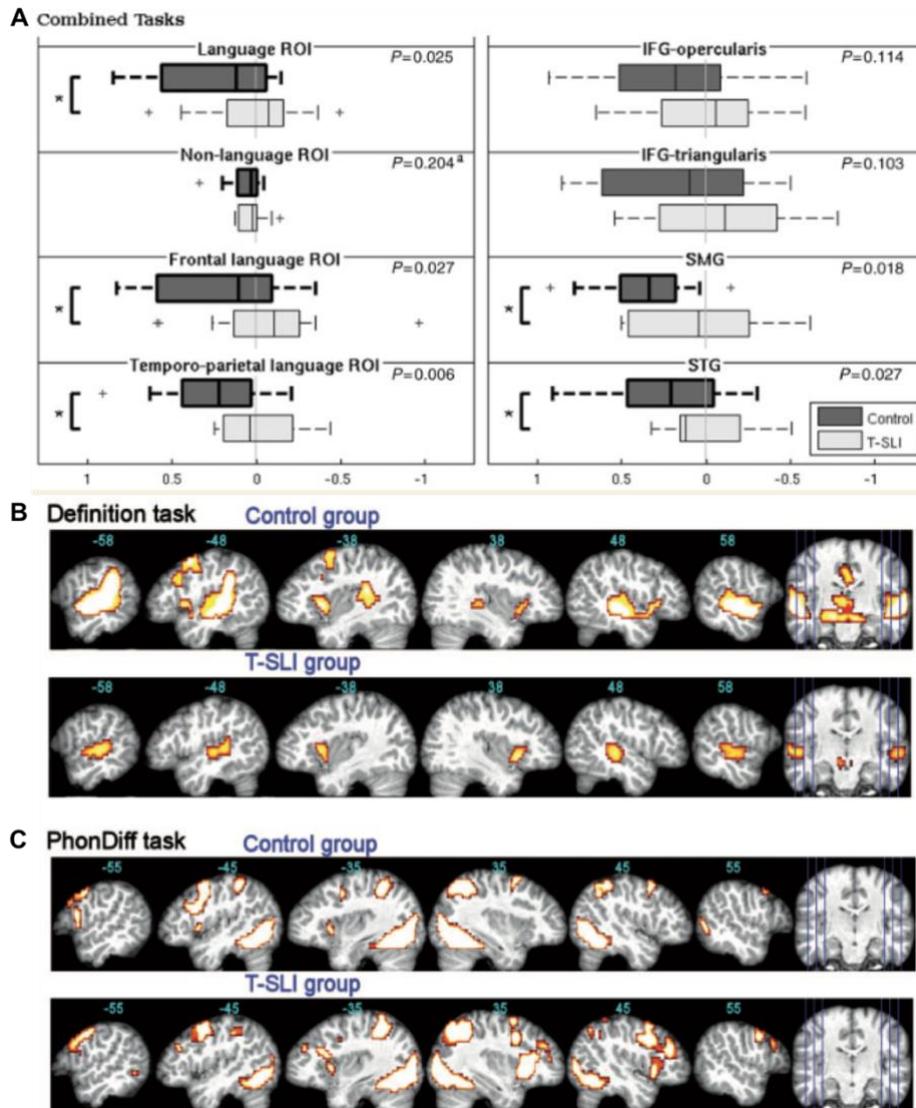


Figure 23 : (A) Indices de latéralisation (positif = gauche, négatif = droite) du groupe contrôle et du groupe dysphasique pour chaque région d'intérêt (ROI) (* signifie une latéralisation significative au sein du groupe, ou entre les groupes). (B) Cartes fonctionnelles des sujets contrôles (haut) et dysphasiques (bas) pour la tâche de définition de mots. (C) Cartes fonctionnelles des sujets contrôles (haut) et dysphasiques (bas) pour la tâche phonologique. Figures empruntées à Guibert et al. (2011).

Au niveau morphologique, des études ont montré une réduction du volume de certaines des régions identifiées dans les études fonctionnelles (région périsylvienne temporo-pariétale, *planum temporale* gauche et noyau caudé bilatéral) chez les enfants dysphasiques (Gauger, Lombardino, & Leonard, 1997; Jernigan, Hessenlink, Sowell, & Tallal, 1991; Leonard et al., 2002; Soriano-Mas et al., 2009; Watkins et al., 2002b). Dans l'étude évoquée précédemment, Badcock et collaborateurs (2012) ont examiné la structure morphologique des régions cérébrales impliquées dans la parole chez les enfants dysphasiques. Ils ont observé une forte

concentration de substance grise au sein du gyrus frontal inférieur gauche et une concentration plus faible de substance grise au sein des STS bilatéraux et du noyau caudé droit (non significatif à gauche, mais volume réduit). Ainsi les enfants dysphasiques présentent des anomalies structurelles et fonctionnelles : le cortex frontal inférieur gauche possède plus de substance grise mais est moins activé lors d'une tâche langagière, alors que les cortex temporaux postérieurs (STG et STS bilatéraux) possèdent moins de substance grise et sont moins activés. Les auteurs ont par ailleurs décrit une corrélation négative entre le volume du noyau caudé droit et les performances des enfants dysphasiques dans une tâche de répétition de non-mots. Ces résultats sont similaires à ceux rapportés par Watkins et collègues (2002b) chez les membres de la famille KE. Les membres de cette famille présentaient des troubles sévères d'articulation, caractérisés par des difficultés de coordination des muscles oro-faciaux, et des déficits linguistiques (trouble de la compréhension de phrases à la structure syntaxique complexe, troubles lexicaux et grammaticaux ; Vargha-Khadem et al., 1998; Watkins, Dronkers, & Vargha-Khadem, 2002a). Watkins et collègues (2002b) ont principalement observé une réduction du volume du noyau caudé bilatéralement qui était corrélée négativement avec les performances des dysphasiques dans une tâche de répétition de mots (noyau caudé droit) et positivement avec les troubles du contrôle oro-moteur et de l'articulation observés chez ces individus. Une réduction du volume de substance grise des régions temporales postérieures et frontales inférieures gauches a également été observée au sein de cette famille (Belton, Salmond, Watkins, Vargah-Khadem, & Gadian, 2003; Watkins et al., 2002b). Soriano-Mas et collègues (2009) ont quant à eux examiné la courbe développementale de la composition structurelle du cortex en analysant les IRM de 36 enfants âgés de 5 à 17 ans présentant une dysphasie, en comparaison de participants contrôles. Les résultats ont montré un volume global plus important de substance blanche et grise chez les enfants dysphasiques par rapport aux contrôles, avec une augmentation plus grande chez les enfants les plus jeunes. L'augmentation de volume de la substance grise était décrite au sein de deux régions principales chez les enfants dysphasiques : la région périsylvienne droite et le gyrus occipital médian gauche. D'autres régions montrant un pattern similaire ont été retrouvées chez les enfants dysphasiques les plus jeunes, incluant notamment le cortex temporal postérieur, le noyau caudé, le cortex moteur et le précuneus. Aucune différence significative entre les enfants dysphasiques plus âgés et leurs contrôles n'a été observée. Ces différences structurelles spécifiques aux jeunes enfants dysphasiques confirment le caractère neurodéveloppemental de la dysphasie.

La dysphasie semble donc caractérisée par différentes anomalies structurelles. Toutefois, le nombre et les régions altérées ne sont pas toujours les mêmes au sein des différentes études et certaines ne montrent même aucune différence morphologique entre les personnes présentant une dysphasie et la population générale (Gauger et al., 1997; Jäncke, Siegenthaler, Preis, & Steinmetz, 2007; Preis, Jäncke, Schittler, Huanq, & Steinmetz, 1998; Vargha-Khadem et al., 1998). Le caractère variable des résultats de ces études s'accorde avec la grande hétérogénéité des troubles dysphasiques observée en clinique et reflétée dans les classifications (Gérard, 1993; Mazeau, 2005; Rapin & Allen, 1983).

3.2. Les facteurs génétiques

Un grand nombre d'études se sont intéressées à la part de la génétique dans la dysphasie. On observe que les garçons sont plus touchés que les filles (Leonard, 1988). De plus, les risques de partager une dysphasie sont supérieurs chez les jumeaux monozygotes, possédant le même patrimoine génétique, que chez les dizygotes (Bishop, North, & Donlan, 1995). Enfin, l'étude de la célèbre famille KE précise le caractère familial de ce trouble (Hurst, Baraitser, Auger, Graham, & Norell, 1990; Vargha-Khadem et al., 2005). L'étude du patrimoine génétique des membres de cette famille a en effet révélé la présence d'une mutation du gène FOXP2, situé sur le chromosome 7, partagée par tous les membres atteints. Ce gène est principalement exprimé au sein des populations neuronales situées dans les ganglions de la base et notamment du noyau caudé, mais aussi au niveau du cortex, du cervelet et du thalamus. Ces différentes régions sont par ailleurs impliquées dans les boucles cortico-striato-thalamo-corticales jouant un rôle dans la planification motrice et l'apprentissage. Il a alors été rapidement proposé que ce gène joue un rôle important dans l'acquisition et le développement du langage et que sa mutation puisse être à l'origine, au moins en partie, des anomalies neurologiques observées dans la dysphasie (Bishop, 2013; Liégeois et al., 2003; Vargha-Khadem et al., 2005; Vargha-Khadem et al., 1998; Watkins et al., 2002a). Cependant, le tableau n'est pas si clair puisque cette mutation n'est pas retrouvée chez tous les dysphasiques et ne peut donc expliquer à elle seule l'apparition des troubles (Marcus & Fisher, 2003). D'autres facteurs génétiques ont été envisagés dans la dysphasie et sont encore à l'étude actuellement, mais aucun ne semble être confirmé jusqu'à présent.

3.3. Les autres facteurs/causes

Un certain nombre d'autres causes potentielles de la dysphasie ont été proposées. Notre but ici n'étant pas d'en établir une liste exhaustive, nous en mentionnerons uniquement certaines. Parmi elles, on retrouve notamment un trouble des traitement généraux (Gathercole & Baddeley, 1990, 1993; Ullman & Pierpont, 2005), associé ou non à une limitation des capacités de traitement cognitif (Montgomery, 2002, 2005; Stark & Montgomery, 1995), ou encore un trouble du traitement temporel (Tallal, 2003; Tallal, Stark, & Curtiss, 1976).

Gathercole et Baddeley (1990) ont évoqué l'existence d'un déficit de stockage des informations phonologiques en mémoire de travail chez les enfants dysphasiques, résultant soit d'une analyse imparfaite des segments de parole, formant alors des représentations phonologiques insuffisamment spécifiques, soit d'une trace mnésique s'effaçant trop rapidement (Ellis-Weismer, Evans, & Hesketh, 1999; Ullman & Pierpont, 2005). Des données en IRMf ont confirmé que le réseau cérébral sous-tendant la mémoire de travail était moins fonctionnel chez les enfants dysphasiques dans des tâches de reconnaissance et d'encodage et fortement corrélé aux performances des sujets dans des tâches de répétition de non-mots (Badcock et al., 2012; Ellis-Weismer et al., 2005; Hugdahl et al., 2004; Watkins et al., 2002b). La boucle phonologique de la mémoire de travail jouant un rôle dans la récupération des patrons articulatoires au niveau de la voie dorsale du modèle de Hickok et Poeppel (2007), son dysfonctionnement dans la dysphasie supposerait alors l'existence d'un trouble au niveau de la voie dorsale de la perception de la parole.

Ce déficit de la mémoire de travail serait par ailleurs corrélé aux capacités limitées de traitement des enfants dysphasiques (Montgomery & Windsor, 2014). Dans une expérience testant l'interaction entre le rythme de présentation d'un stimulus et les capacités de mémoire de travail d'enfants dysphasiques appariés à des enfants contrôles, Montgomery (2003) a montré que les enfants dysphasiques avaient des performances plus faibles que leurs pairs dans une tâche de compréhension de phrases (pointer l'image correspondant à la phrase entendue). Ces enfants étaient par ailleurs meilleurs lorsque les phrases étaient courtes, et lorsqu'elles étaient présentées à un rythme plus lent (augmentation de 25% de la durée originale du signal). Les auteurs ont alors suggéré que ces enfants aient besoin d'un délai plus long pour effectuer correctement le traitement cognitif du langage parlé. Montgomery et ses collègues ont également montré que les enfants dysphasiques étaient plus lents que des enfants de la population générale du même âge à détecter des mots cibles dans des phrases, suggérant une vitesse de traitement limitée (Montgomery, 2002; Montgomery, 2005; Stark &

Montgomery, 1995). La vitesse de traitement correspond ici à la quantité d'information qui peut être traitée par un individu sur une certaine durée. Cependant, cette limitation des capacités de traitement cognitif ne semble pas spécifique au langage et est également observée dans des tâches non langagières de rapidité de réponse comparant des enfants dysphasiques à des enfants contrôles (Miller, Leonard, & Tomblin, 2001), suggérant une lenteur générale de traitement chez les enfants dysphasiques (Kail, 1994).

Une autre hypothèse très influente dans la recherche sur la dysphasie, celle d'un déficit du traitement temporel, a été proposée par Tallal et ses collègues (Tallal et al., 1996; Tallal & Piercy, 1973a, 1973b, 1975). Les auteurs ont décrit des difficultés chez les enfants dysphasiques pour traiter les stimuli auditifs brefs, verbaux ou non, perçus en succession rapide ; en revanche, si la vitesse de présentation est ralentie ou la durée des sons allongée, leurs performances s'améliorent. Ceci a conduit Tallal à proposer l'existence d'un trouble du traitement des changements acoustiques rapides qui serait à l'origine d'un déficit de segmentation du signal de parole en unités correspondant aux représentations phonologiques. Le déficit surviendrait au niveau précoce du traitement auditif et serait lié à une immaturité du système auditif ou à une organisation neurologique atypique. Toutefois, cette hypothèse a été et reste controversée, le trouble de traitement temporel n'étant pas présent chez tous les dysphasiques et certains présentant également des difficultés à traiter les stimuli longs (McArthur & Bishop, 2004; Mody, Studdert-Kennedy, & Brady, 1997).

En résumé, la multiplicité des hypothèses explicatives de la dysphasie fait écho à la grande hétérogénéité des troubles dysphasiques. Si la composante génétique et héréditaire semble avérée, elle ne peut à elle seule, tout comme les autres facteurs, expliquer l'apparition des différents déficits rapportés chez les enfants. Comme suggéré par certains auteurs, un tableau plus complexe doit être envisagé : la dysphasie pourrait ainsi très probablement résulter de l'interaction d'un ensemble de facteurs, génétiques mais aussi environnementaux, affectant le développement anatomo-fonctionnel et l'organisation des réseaux langagiers, favorisant ainsi l'apparition du trouble du développement du langage oral chez l'enfant (Bishop, 2006; Fisher, Lai, & Monaco, 2003; Friederici, 2006; Gérard, 1993; Krishnan, Watkins, & Bishop, 2016).

4. Rythmes et dysphasie

4.1. Troubles de discrimination des indices rythmiques de la parole

Nous avons vu précédemment (partie I.1.) qu'une perception efficace des indices rythmiques de la parole est primordiale pour le bon développement du langage et notamment la formation des représentations phonologiques. Un certain nombre d'études ont montré que les enfants présentant une dysphasie montraient des difficultés à percevoir et à traiter l'information rythmique contenue dans un signal auditif.

Dans les années 70, dans le cadre de son hypothèse d'un déficit de traitement temporel susmentionnée, Tallal suggérait déjà la présence d'un trouble du traitement du rythme dans la dysphasie (Tallal et al., 1996; Tallal & Piercy, 1973a, 1973b, 1975). L'auteur a testé les capacités de discrimination d'enfants dysphasiques anglophones âgés de 6 à 9 ans appariés à des enfants contrôles lors de l'écoute de stimuli acoustiques complexes non verbaux de fréquences comprises dans la gamme de fréquences de la parole. Elle a alors montré que ces enfants présentaient des performances plus faibles que leurs pairs lorsque la durée des sons était courte (< 250 ms). Ils étaient également sensibles à l'intervalle entre deux sons (intervalle inter-stimuli, ISI) se révélant incapables de les différencier dans le cas d'un intervalle trop bref (< 150 ms). Cependant, lorsque la durée du son et/ou l'intervalle inter-stimuli (son > 250 ms, ISI > 150 ms) étaient allongés, les performances des enfants dysphasiques devenaient semblables à celles des enfants contrôles (Tallal & Piercy, 1973a, 1973b). Le même type de résultats a été rapporté avec des stimuli langagiers (i.e. groupes syllabiques). Les enfants dysphasiques présentaient en effet des difficultés particulières à discriminer les syllabes /ba/ et /da/ caractérisées par des transitions formantiques très brèves, alors que si la durée de ces syllabes était allongée, la distinction entre les deux syllabes était possible (Tallal & Piercy, 1975). De ces différentes études, les auteurs ont conclu que les difficultés des enfants dysphasiques à traiter les sons brefs pourraient expliquer leurs troubles phonologiques et donc se répercuter sur le traitement général du signal de parole. En effet, dans le signal de parole, les informations acoustiques rapides de l'ordre de quelques dizaines de millisecondes correspondent aux contrastes phonétiques, critiques pour le traitement et l'identification des sons de la parole. Ainsi, si les enfants présentant une dysphasie ne sont pas capables d'extraire les différents phonèmes de la parole, ils ne peuvent pas former de catégories phonologiques stables, ce qui va entraver la perception et la production du langage.

Si Tallal s'est concentrée sur les indices de durée, d'autres se sont plus récemment intéressés à d'autres indices suprasegmentaux importants pour le rythme de la parole et son accentuation, à savoir le *rise time* (temps de montée de l'enveloppe d'amplitude du signal de parole ; cf. chapitre I.3.). Corriveau et collaborateurs (2007) ont étudié les capacités de perception auditive du rythme chez des enfants anglophones dysphasiques appariés à des enfants présentant un développement langagier typique du même âge (7-11 ans). Trois tâches de discrimination auditive non verbale leur étaient proposées : les deux premières portaient sur des stimuli non verbaux variant sur le *rise time* et la durée des sons ; la troisième tâche impliquait des stimuli différant sur un indice non rythmique, l'intensité. Pour chaque tâche, l'enfant écoutait deux dinosaures émettre un son non verbal et devait choisir le dinosaure ayant produit le son cible défini au préalable. Les enfants devaient ainsi choisir 1) le son ayant le *rise time* le plus court (variant de 15 à 300 ms), 2) le son le plus long (durée variant de 400 à 640 ms) et 3) le son le plus fort (intensité variant de 73 à 81.1 dB SPL). Les résultats ont révélé que les enfants dysphasiques présentaient des performances significativement plus faibles que leurs pairs dans les deux tâches où les indices rythmiques étaient manipulés, alors qu'aucune différence significative entre les groupes n'était observée dans la tâche portant sur l'intensité des sons. Les auteurs ont ensuite examiné les corrélations entre les performances des enfants dans ces différentes tâches et leurs capacités langagières. Ils ont alors montré que les performances dans les tâches de jugement de durée et de *rise time* étaient de bons prédicteurs des capacités langagières des sujets, en particulier concernant le vocabulaire, la conscience phonologique, la mémoire de travail et la lecture.

D'autres études ont par la suite conforté l'existence d'un déficit de perception du *rise time* chez les enfants présentant des troubles développementaux du langage. Beattie et Manis (2012) ont examiné la sensibilité au *rise time* et à l'intensité de stimuli auditifs non langagiers dans deux tâches auditives identiques à celles utilisées par Corriveau et al (2007), chez trois groupes d'enfants anglophones âgés de 8.4 à 16 ans : un groupe contrôle sans troubles du langage, un groupe présentant une dyslexie pure et un groupe présentant une dyslexie associée à une dysphasie. Ils ont notamment démontré que les enfants dyslexiques, ayant ou non une dysphasie associée, avaient des performances amoindries pour percevoir le *rise time* par rapport aux enfants contrôles. Ce déficit de traitement du *rise time* dans la dyslexie a été décrit par ailleurs (Goswami, 2011; Pasquini, Corriveau, & Goswami, 2007; Thomson & Goswami, 2008; Woodruff Carr, White-Schwoch, Tierney, Strait, & Kraus, 2014), suggérant une difficulté générale de traitement de cet indice rythmique dans les troubles

développementaux du langage. Richards et Goswami (2015) ont récemment reproduit l'étude de Corriveau et al. (2007) chez un groupe d'enfants anglophones dysphasiques et contrôles âgés de 8 à 12 ans, en incluant une tâche psychoacoustique supplémentaire de discrimination fréquentielle des stimuli non verbaux (variation de la hauteur du stimulus) et deux tâches prosodiques mesurant la sensibilité à l'accentuation. La première tâche (tâche « DeeDee ») teste la sensibilité prosodique et consiste pour les enfants à écouter des noms de personnages accentués correctement ou non sous la forme d'une série de syllabes accentuées (DEE) ou non accentuées (dee). La deuxième tâche est une tâche de perception de l'accent lexical dans laquelle les enfants doivent juger de la bonne accentuation de mots plurisyllabiques. Les résultats s'accordent avec ceux des études précédentes puisqu'ils ont révélé que le groupe d'enfants dysphasiques présentait des difficultés pour percevoir le *rise time* et la durée du signal auditif non verbal, alors que leur sensibilité aux variations d'intensité et de fréquence n'était pas altérée. De plus, ces résultats ont montré des performances limitées dans les deux tâches prosodiques chez les enfants dysphasiques comparés à leurs pairs. Différentes analyses de régressions multiples ont été réalisées avec en variable indépendante chaque mesure psychoacoustique (*rise time*, intensité, fréquence et durée) et en variable dépendante les performances des enfants dans les deux tâches prosodiques. Le QI des enfants était par ailleurs contrôlé. Les auteurs ont observé que les différences individuelles de sensibilité au *rise time* prédisaient la variabilité des performances observées dans le traitement de l'accent lexical chez les enfants dysphasiques. Richards et Goswami en ont conclu que le traitement du *rise time* serait un élément essentiel au bon développement des représentations prosodiques et phonologiques (voir aussi Leong & Goswami, 2014b pour le déficit de traitement prosodique et syllabique dans la dyslexie). Cette interprétation s'inscrit dans le cadre du modèle d'échantillonnage temporel (*temporal sampling framework*, TSF ; Figure 24) proposé initialement pour la dyslexie par Goswami (2011), qui suggère qu'un trouble de la perception de la structure rythmique de l'enveloppe du signal de parole (correspondant aux modulations lentes d'amplitude et incluant le *rise time* et le rythme syllabique) puisse compromettre la segmentation en syllabes et en mots et par conséquent la formation de représentations phonologiques stables. Cette théorie s'inspire ainsi du modèle AST de l'échantillonnage temporel asymétrique de Poeppel (2003, défini dans la partie III.4.1.) en proposant un dysfonctionnement de la synchronisation (*entrainment*) des oscillations thêta (et potentiellement delta) sur les modulations lentes du signal de parole (véhiculant l'information syllabique et prosodique) au sein des cortex auditifs. Ce déficit oscillatoire thêta pourrait ensuite entraîner un échantillonnage imparfait des indices phonémiques dans la bande de

fréquence gamma (i.e. *theta-gamma nesting*) et sous-tendre le développement phonologique atypique observé dans la dyslexie.

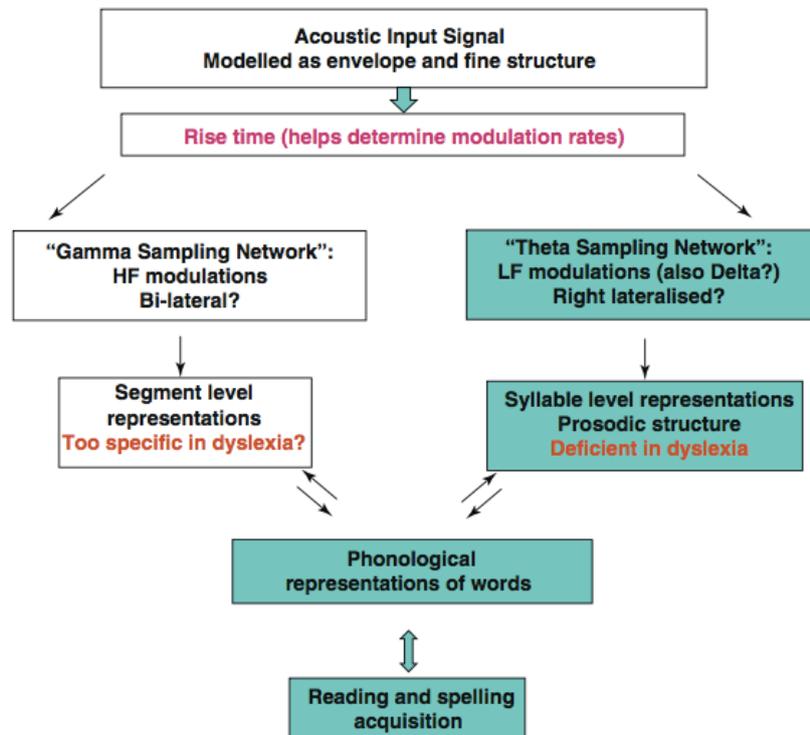


Figure 24 : Modèle d'échantillonnage temporel proposé par Goswami (2011). Figure empruntée à Goswami (2011).

Ainsi, l'ensemble de ces études révèle l'importance d'une perception efficace des indices acoustiques présents dans l'enveloppe d'amplitude du signal de parole dans les processus phonologiques et prosodiques. Cependant aucune des études mentionnées jusqu'à présent n'a examiné les capacités des enfants dysphasiques pour percevoir le *rise time* dans la parole. C'est ce qu'ont proposé Cumming et ses collaborateurs (2015a) chez des enfants anglophones (d'âge moyen 9.6 ans) présentant une dysphasie associée ou non à des troubles phonologiques et de la lecture dans les mêmes tâches psychoacoustiques et prosodiques que celles utilisées par Richards et Goswami (2015) mais en ajoutant des stimuli langagiers (i.e. syllabe /ba/). Ils ont alors montré que les enfants dysphasiques dont la phonologie était fortement altérée présentaient des difficultés à percevoir le *rise time* et la durée des stimuli verbaux par rapport à leurs pairs, alors que pour les stimuli non-verbaux, la différence entre les deux groupes n'était pas significative. De plus, ces enfants étaient moins bons dans les deux tâches prosodiques et leurs performances individuelles dans ces tâches étaient

principalement expliquées par leurs performances pour traiter le *rise time* et la durée. Afin d'expliquer leurs résultats, et dans la lignée de la théorie de l'échantillonnage temporel de Goswami (2011), Cumming et collègues (2015a) ont avancé l'hypothèse « *prosodic phrasing hypothesis* » selon laquelle les troubles morpho-syntaxiques souvent observés chez les enfants dysphasiques résulteraient en partie de leur insensibilité aux indices de *rise time* et de durée. Cette sensibilité réduite affecterait l'extraction des indices rythmiques de basse fréquence contenus dans l'enveloppe temporelle du signal de parole, entravant la segmentation des syllabes et des mots, et par conséquent le traitement de la structure phonologique à la base des représentations syntaxiques (Cumming, Wilson, & Goswami, 2015a).

Depuis quelques années, des preuves croissantes de l'existence d'un trouble de la perception des indices acoustiques importants pour le rythme de la parole ont donc été apportées chez les enfants dysphasiques. De façon intéressante, un certain nombre d'études a montré que ce trouble du rythme pouvait se manifester non seulement au niveau perceptif mais aussi au niveau expressif, moteur. Les dysphasiques ainsi que les dyslexiques, sans trouble de la dextérité motrice, présentent en effet des difficultés à s'ajuster au rythme lent d'un signal auditif lors de tâches de *tapping* (i.e. battre le rythme avec son doigt; Corriveau & Goswami, 2009; Cumming, Wilson, Leong, Colling, & Goswami, 2015b; Goswami, 2012; Hill, 2001; Thomson, Fryer, & Maltby, 2006; Thomson & Goswami, 2008). Corriveau et Goswami (2009) ont ainsi montré que des enfants dysphasiques anglophones âgés entre 7 et 11 ans n'étaient pas capables de battre du doigt correctement en rythme avec un métronome, en particulier lorsque le rythme était lent (1.5 et 2 Hz, correspondant à la « fréquence prosodique ») et cela en l'absence de trouble moteur (mesuré à l'aide d'une tâche standardisée, la *Pegboard task*). Plus précisément, les enfants avaient tendance à anticiper le battement (i.e. ils étaient en avance sur le signal perçu) et présentaient une plus forte variabilité entre eux que les enfants contrôles. En outre, leurs performances dans cette tâche de *tapping* expliquaient une part importante de la variance unique de leurs compétences linguistiques et plus particulièrement en phonologie et en littéracie (voir aussi Woodruff Carr et al., 2014).

D'autres études ont utilisé des séquences musicales possédant une métrique régulière ou des comptines enfantines en guise de stimuli et ont montré des résultats similaires (Cumming, Wilson, Leong, et al., 2015b; Leong & Goswami, 2014a). Cumming et collaborateurs (2015b) ont comparé les performances d'enfants anglophones d'âge moyen 9 ans et 6 mois présentant

une dysphasie pure ou associée à des troubles phonologiques et de lecture à un groupe d'enfants contrôles. Ils ont mesuré leurs performances dans différentes tâches phonologiques, trois tâches auditives (*rise time*, durée et fréquence) utilisant des stimuli de parole ou non et des tâches incluant des stimuli musicaux ou de parole (syllabe /ba/) : 1) une tâche de jugement du rythme musical (les enfants devaient décider si deux séquences musicales présentaient un rythme identique ou non), 2) une tâche où les enfants devaient battre du doigt en rythme avec les stimuli musicaux (« *tapping* ») et 3) trois tâches évaluant les capacités de perception du rythme et/ou de la hauteur des stimuli musicaux (mélodies de 5/6 notes) ou verbaux (phrases de 5/6 syllabes). Dans ces dernières tâches, les enfants entendaient un stimulus cible puis deux autres stimuli : un identique au stimulus cible et l'autre présentant a) soit une variation du rythme (modification de l'accentuation sans toucher à la F0), b) soit une variation de la hauteur (variation de la mélodie sans toucher au rythme), c) soit une variation du rythme et de la hauteur. Les enfants entendaient ensuite une nouvelle fois le stimulus cible et devaient juger quel stimulus parmi les deux entendus correspondait à la cible. Les résultats ont révélé que 1) les deux groupes de dysphasiques avaient des performances comparables au groupe contrôle dans la tâche de jugement du rythme musical ; 2) en revanche, dans la tâche de « *tapping* », l'ensemble des enfants dysphasiques présentaient une plus faible synchronisation motrice sur le stimulus auditif que les enfants contrôles (ces résultats ne dépassaient toutefois le seuil de significativité que chez les enfants présentant une dysphasie pure). 3) Enfin, dans les tâches examinant les capacités de perception du rythme et/ou de la hauteur des stimuli musicaux et verbaux, les auteurs ont observé une difficulté de perception de ces deux variables chez les enfants présentant une dysphasie associée à un trouble phonologique par rapport aux enfants contrôles. Ces résultats n'étaient pas observés chez les enfants présentant une dysphasie pure. Il faut cependant noter que l'ensemble des enfants dysphasiques ou contrôles semblaient trouver plus facile de comparer deux stimuli lorsqu'ils présentaient à la fois les indices de rythme et de hauteur, plutôt qu'un seul type d'indice. Enfin, à l'aide d'analyses de régressions multiples, les auteurs ont rapporté que les différences individuelles dans les différentes tâches rythmiques de parole étaient de bons prédicteurs du développement du langage réceptif et expressif ainsi que de la phonologie et de la littéracie. La sensibilité à la hauteur des stimuli prédisait quant à elle essentiellement les performances phonologiques et de littéracie des enfants. Dans cette étude, la tâche de « *tapping* » ne semblait au contraire pas être un bon prédicteur des performances linguistiques et phonologiques des enfants dysphasiques, ce qui était pourtant observé dans les études précédemment décrites (Corriveau & Goswami, 2009; Thomson et al., 2006; Thomson & Goswami, 2008). Ainsi, l'ensemble des

enfants dysphasiques semblent avoir des difficultés pour traiter les indices rythmiques, que ce soit dans la musique ou la parole ; cette étude révèle cependant que certaines mesures sont plus sensibles que d'autres avec des différences selon les troubles linguistiques des enfants.

Leong et Goswami (2014a) ont quant à elles étudié les capacités d'adultes dyslexiques de langue maternelle anglaise à percevoir et produire le rythme de la parole. Pour cela, elles ont utilisé des comptines enfantines, caractérisées par leur forte rythmicité. Dans la tâche de perception, les sujets avaient pour tâche de battre le rythme sur celui de la comptine (le rythme syllabique était de 4 Hz et les syllabes accentuées apparaissaient à une fréquence de 2 Hz) ; dans l'expérience de production, ils devaient reproduire oralement ces comptines en rythme avec un métronome. Différentes mesures rythmiques ont été réalisées : la périodicité des réponses (calcul de l'intervalle entre deux battements ou deux syllabes accentuées), la précision (mesure de la distance entre le battement/syllabe produit et l'*onset* de la voyelle de la syllabe accentuée) et la synchronisation oscillatoire (i.e. couplage de phase) de la réponse rythmique des sujets sur les différentes échelles temporelles de l'enveloppe d'amplitude du signal de parole (accentuation, syllabes et phonèmes). Les résultats ont révélé un dysfonctionnement du traitement du rythme syllabique dans la dyslexie. Dans la tâche de perception, alors que les dyslexiques et les contrôles montraient un *entraînement* de leurs battements sur le rythme syllabique, les auteurs ont observé que les sujets dyslexiques étaient en avance par rapport aux contrôles (aucune différence inter-groupes n'a été observée pour les mesures de périodicité et de précision). Ainsi, les dyslexiques synchronisaient systématiquement leurs battements sur la phase oscillatoire précédente par rapport aux contrôles. Contrairement aux contrôles, les dyslexiques montraient également une synchronisation sur le rythme des syllabes accentuées, suggérant qu'ils exploitent l'information rythmique syllabique et accentuelle afin d'identifier le rythme d'une phrase, alors que les contrôles n'utiliseraient que le rythme syllabique. De plus, des analyses de régression ont montré que ces performances étaient liées à la conscience phonologique et aux capacités orthographiques des sujets (les résultats étant principalement portés par le groupe dyslexique). Concernant la tâche de production, une altération de la production des syllabes a été observée chez les dyslexiques (raccourcissement de la durée) et la synchronisation au niveau des rythmes syllabique et phonémique était également altérée. Ces résultats supportent ainsi l'idée d'un trouble de synchronisation sur les rythmes lents de la parole (rythmes syllabique et prosodique) dans la dyslexie, contribuant probablement au développement de représentations phonologiques atypiques, caractéristiques de ce trouble du langage.

Enfin, il faut noter que des troubles de la discrimination du rythme, au niveau perceptif aussi bien qu'expressif, ont également été observés chez des enfants et adultes anglophones présentant une dyspraxie verbale (qui, pour rappel, est un sous-type de dysphasie selon la classification de Rapin et Allen, 1983; Alcock, Passingham, Watkins, & Vargha-Khadem, 2000; Peter & Stoel-Gammon, 2008). Alcock et collègues (2000) ont testé les performances d'adultes présentant une dyspraxie verbale, dont 9 issus de la même famille (famille KE), comparés à un groupe contrôle, dans différentes tâches de perception de stimuli musicaux (discrimination de la hauteur de notes, de la mélodie de chansons ou de séquences musicales familières et du rythme de séquences musicales) et de production (répétition de notes dont la hauteur varie, production de mélodie et battre le rythme de différentes notes perçues ou les reproduire oralement). Ils ont rapporté une différence entre les deux groupes de sujets uniquement dans les tâches de discrimination rythmique (au niveau perceptif et productif), ce qui rejoint les résultats des études précédentes chez les dysphasiques et les dyslexiques.

Les tâches de synchronisation motrice sur un stimulus auditif rythmique impliquent un couplage auditivo-moteur fin dans lequel les informations temporelles précises du stimulus (soit le début du son) sont extraites par le système auditif et intégrées au niveau des régions motrices (Tierney & Kraus, 2014). La synchronisation nécessite en outre de contrôler et d'ajuster en permanence et de façon appropriée les commandes motrices pour respecter le rythme donné par le signal auditif. Comme déjà mentionné dans le chapitre IV.1., le couplage auditivo-moteur joue également un rôle clé dans l'acquisition du langage et notamment dans le développement des compétences phonologiques des enfants (Goswami, 2011; Tierney & Kraus, 2014). Au vu de ces similitudes, il a alors été suggéré qu'un mécanisme neuronal sous-tendant la perception et la production du *timing* des informations rythmiques (notamment les rythmes syllabique et prosodique), et impliquant les régions cérébrales auditives et motrices (Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007), puisse être compromis dans la dysphasie (et dans la dyslexie) et affecter à la fois le développement du langage et de la motricité (Corriveau & Goswami, 2009; Leong & Goswami, 2014a). Ceci pourrait rendre compte, au moins en partie, de la comorbidité des troubles langagiers et moteurs fréquemment décrite chez les dysphasiques, en particulier lorsque les troubles langagiers se manifestent sur le plan expressif (Bishop, 2002; Hill, 2001).

D'autres éléments en faveur de liens étroits entre les versants perceptif et expressif chez les dysphasiques proviennent d'études comportementales montrant des corrélations entre les

performances d'enfants présentant une dyspraxie verbale dans des tâches de production et de perception de la parole. Les travaux de Groenen et al. (1996) ont par exemple montré que les enfants dyspraxiques (sur le plan verbal) de langue maternelle allemande avaient des performances inférieures aux enfants contrôles lorsqu'ils devaient discriminer la frontière entre les phonèmes /d/ et /b/, qui diffèrent de par leur lieu d'articulation, et que ces performances étaient corrélées au nombre de phonèmes que les enfants dyspraxiques substituaient dans une tâche de production. Le même type de résultats a été observé par Nijland (2009) chez des enfants dyspraxiques germanophones dans des tâches de jugement de rimes, de discrimination de phonèmes et de production de mots et de non-mots. Enfin, comme mentionné précédemment dans la partie V.3. de l'introduction, des anomalies structurelles et fonctionnelles ont été décrites dans la dysphasie (incluant la dyspraxie verbale) dans les régions motrices corticales et sous-corticales telles que le cervelet, le noyau caudé et l'aire motrice supplémentaire (Belton et al., 2003; de Guibert et al., 2011; Vargha-Khadem et al., 2005; Watkins et al., 2002b). Ces régions sont connues pour être impliquées dans la planification motrice et la production de la parole mais aussi dans les processus temporels aux niveaux moteur (Grahn & Brett, 2007) et linguistique (Kotz & Schmidt-Kassow, 2015; Kotz, Schwartze, & Schmidt-Kassow, 2009).

Ainsi, des troubles du traitement du rythme en perception et production sont décrits dans la dysphasie et la dyslexie. Un déficit de la synchronisation motrice sur des stimuli auditifs rythmiques verbaux ou non a été observé dans les bandes de fréquence correspondant à l'information rythmique syllabique et prosodique (2 à 5 Hz). Il a alors été suggéré que ce déficit soit lié à une dynamique oscillatoire corticale atypique dans les régions cérébrales sous-tendant la perception et la production de la parole chez les individus présentant une dysphasie ou une dyslexie.

4.2. Oscillations corticales et dysphasie

Les modèles proposés par le groupe de Goswami (2011; *Temporal Sampling Framework*, cf. Figure 24) et de Cumming (2015a; *prosodic phrasing hypothesis*) supposent que les déficits de traitement du rythme dans les troubles développementaux du langage seraient liés à un dysfonctionnement de l'échantillonnage neuronal dans une ou plusieurs bande(s) de fréquence corticale primordiale(s) pour le traitement de la parole. Comme

mentionné précédemment, ce déficit se manifesterait plus précisément dans les basses fréquences des régions auditives droites permettant de segmenter le signal de parole en unités syllabiques (bande thêta 4-7 Hz) et prosodiques (bande delta, 1-4 Hz). Pour Giraud et Poeppel (2012) en revanche, un fonctionnement oscillatoire atypique chez ces enfants se produirait plutôt au niveau des oscillations gamma basses (25-45 Hz) au sein des régions temporales gauches. Un tel dysfonctionnement entrainerait un sur-échantillonnage ou un sous-échantillonnage de l'information phonémique du signal de parole aboutissant à la formation de représentations phonologiques atypiques.

Plusieurs études en imagerie cérébrale ont révélé la présence de telles anomalies oscillatoires dans les réseaux corticaux langagiers chez les dyslexiques (Hämäläinen, Rupp, Soltész, Szücs, & Goswami, 2012; Lehongre, Morillon, Giraud, & Ramus, 2013; Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011; Poelmans et al., 2012; Vanvooren et al., 2014). Dans une étude en MEG, Lehongre et ses collègues (2011) ont mesuré les réponses corticales auditives (*auditory steady-state responses*, ASSRs) induites par la répétition d'un bruit blanc modulé en amplitude chez des adultes francophones dyslexiques et normo-lecteurs. Au sein du stimulus, les modulations d'amplitude augmentaient de façon linéaire de 10 à 80 Hz. Chez les adultes contrôles, les auteurs ont mis en évidence un entraînement cortical sur les modulations acoustiques latéralisé à gauche au sein du *planum temporale* dans la bande de fréquence gamma basse (25-35 Hz), en accord avec le modèle AST de Poeppel (2003). Dans la bande gamma plus haute (55-80 Hz), cette asymétrie était inversée et les réponses corticales étaient cette fois plus importantes dans le cortex temporal droit. Chez les adultes dyslexiques en revanche, la latéralisation à gauche dans la bande 25-35 Hz était absente au niveau du *planum temporale* (voire les réponses étaient plus fortes à droite) ; les réponses ASSR étaient également réduites dans les régions articulaires gauches et les cortex somatosensoriel et supramarginal gauches, impliquées dans les processus phonologiques. En outre, lorsque la bande gamma haute (55-80 Hz) était considérée, les réponses corticales auditives étaient plus fortes chez les dyslexiques que chez les normo-lecteurs, et ce dans les deux hémisphères. Lehongre et collaborateurs ont ensuite examiné les corrélations entre les réponses corticales et les performances des deux groupes de participants dans différentes tâches langagières. Leurs résultats ont montré une corrélation significative entre les réponses ASSR dans la bande gamma basse au sein des cortex auditifs gauche et droit et la vitesse de lecture uniquement chez les participants normo-lecteurs. Les performances dans des tâches évaluant la conscience phonologique (e.g., répétition de non-mots) étaient également corrélées avec les réponses du

planum temporale gauche chez les normo-lecteurs, mais droit chez les dyslexiques. En résumé, ces données chez les dyslexiques révèlent donc d'une part une sensibilité moindre du cortex auditif gauche aux modulations correspondant au rythme phonémique de la parole (25-35 Hz), et d'autre part une sensibilité accrue au niveau bilatéral aux modulations plus rapides (55-80 Hz). Les auteurs ont alors suggéré que la latéralisation atypique des oscillations gamma (basses) chez cette population entraîne des difficultés à extraire et encoder les indices phonémiques dans l'hémisphère gauche, conduisant à des représentations phonémiques imparfaites, et par conséquent aux déficits de traitement phonologique et de correspondances grapho-phonémiques caractéristiques de la dyslexie. Il faut également noter que l'hypersensibilité aux modulations plus rapides et donc le sur-échantillonnage du signal auditif en unités de taille infra-phonémique pourraient rendre compte du mode de perception allophonique de la parole (i.e. perception des différences intra-catégorielles plutôt que catégorielles) rapporté dans la dyslexie dans d'autres études (e.g., Noordenbos, Segers, Serniclaes, Mitterer, & Verhoeven, 2012; Serniclaes, Sprenger-Charolles, Carré, & Demonet, 2001; Serniclaes, Van Heghe, Mousty, Carré, & Sprenger-Charolles, 2004). L'étude en EEG de Poelmans et al. (2012) chez des participants néerlandais a conforté ces résultats d'un pattern oscillatoire atypique chez les dyslexiques, les auteurs ayant rapporté des différences entre dyslexiques et normo-lecteurs au niveau des oscillations à 20 Hz dans l'hémisphère gauche.

Afin de répliquer leurs résultats et d'élargir leurs investigations à d'autres bandes de fréquence, Lehongre et collaborateurs (2013) ont associé EEG et IRMf et mesuré la corrélation entre la puissance du signal EEG et du signal BOLD (*Blood Oxygenation Level Dependent*) tel que mesuré en IRMf dans les bandes de fréquence delta, thêta et gamma chez des sujets francophones dyslexiques ou normo-lecteurs lors de l'écoute de stimuli auditifs verbaux (documentaire audiovisuel comprenant toute la gamme des modulations d'amplitude pertinentes pour la parole). Ils ont démontré que les oscillations thêta dominaient dans le cortex auditif droit (gyrus de Heschl et *planum temporale*) chez les normo-lecteurs (cf. Giraud et al., 2007; Morillon et al., 2010; Poeppel, 2003), alors que les oscillations delta et gamma semblaient être réparties bilatéralement (Figure 25A). Chez les dyslexiques, aucune spécialisation hémisphérique en fonction des bandes de fréquence n'était observée (Figure 25B). Toutefois, la corrélation entre les signaux EEG et BOLD était plus forte dans l'hémisphère droit mais sans distinction entre les trois bandes de fréquence étudiées. Une comparaison directe entre les groupes a ensuite révélé une latéralisation plus forte des

oscillations gamma dans l'hémisphère gauche chez les normo-lecteurs que chez les dyslexiques ; la différence inter-groupes n'était pas significative dans l'hémisphère droit pour les trois bandes de fréquence (Figure 25, gamma (C), thêta (D) et delta (E)). Ces résultats s'accordent avec les résultats précédents de la même équipe indiquant un dysfonctionnement oscillatoire gamma au sein de l'hémisphère gauche dans la dyslexie. Néanmoins, aucun élément en faveur d'un pattern oscillatoire atypique dans les bandes de fréquence delta et thêta, tel que proposé par Goswami (2011), n'a été mis en évidence ici.

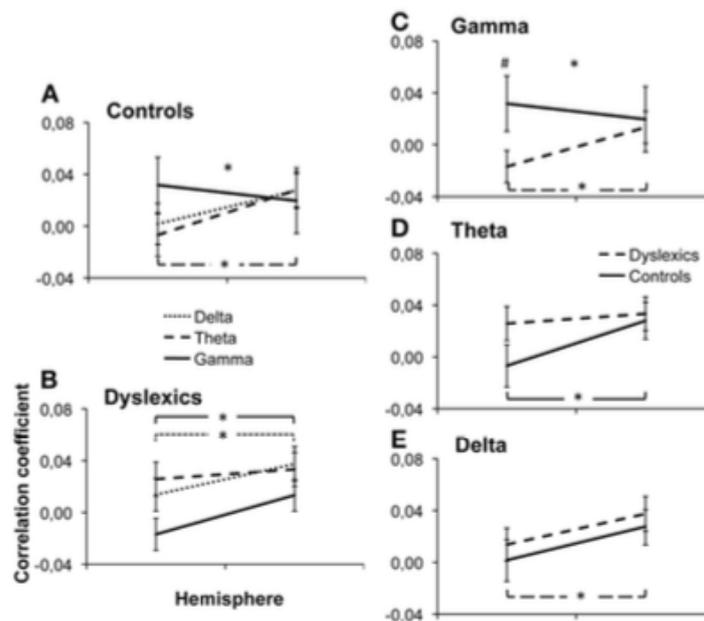


Figure 25 : Interaction Groupe × Fréquence × Hémisphère pour chaque groupe de participants dans chaque bande de fréquence (delta, thêta, gamma) (A) chez les contrôles et (B) chez les dyslexiques.

Interaction Groupe × Fréquence (C) dans la bande gamma, (D) dans la bande thêta et (E) dans la bande delta. Figure empruntée à Lehongre et al. (2013).

D'autres études ont montré une cohérence réduite entre l'enveloppe temporelle du signal de parole et les oscillations corticales delta chez les adultes et les enfants dyslexiques (Hämäläinen et al., 2012; Lizarazu et al., 2015; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; Power, Mead, Barnes, & Goswami, 2013; Soltész, Szűcs, Leong, White, & Goswami, 2013; voir aussi Abrams et al., 2009 pour une comparaison entre bons et mauvais lecteurs). Molinaro et collaborateurs (2016) ont examiné en MEG la dynamique oscillatoire chez des enfants et des adultes dyslexiques hispanophones lors de l'écoute de phrases produites naturellement à un débit normal. Les auteurs ont mesuré la cohérence entre les

modulations lentes du signal de parole et l'activité dans différentes régions cérébrales dans les bandes de fréquence inférieures à 40 Hz. Chez le groupe contrôle (Figure 26A), les auteurs ont relevé deux pics de cohérence, un premier dans la bande delta (0.5-1 Hz) et un second en thêta (5.8-6.3 Hz), au sein des cortex auditifs primaires droit et gauche, des régions temporales supérieures et médiales droites, du gyrus temporal supérieur gauche et du gyrus frontal inférieur gauche. Chez le groupe de dyslexiques (Figure 26B), ils ont observé une diminution de la cohérence dans la bande delta par rapport au groupe contrôle, au niveau du cortex auditif droit (incluant les régions temporales supérieures) et de la *pars opercularis* du gyrus frontal inférieur gauche. Cette réduction de la cohérence en delta était présente chez les enfants et les adultes. Aucune différence entre les groupes n'a cependant été rapportée dans la bande thêta (contrairement aux prédictions du modèle TSF de Goswami, 2011). Par ailleurs, des analyses supplémentaires ont révélé, chez les dyslexiques, une connectivité réduite entre les régions auditives droites et le gyrus frontal inférieur gauche par rapport aux sujets normo-lecteurs, quel que soit leur âge. Enfin, les auteurs ont rapporté une corrélation positive entre la cohérence cortico-acoustique au sein du gyrus frontal inférieur et les performances en littéracie des dyslexiques et, la cohérence dans le cortex auditif droit et les capacités de conscience phonologique. Ces résultats s'accordent globalement avec le modèle d'échantillonnage temporel proposé par Goswami (2011), et suggèrent que les troubles phonologiques observés dans la dyslexie puissent être liés à un dysfonctionnement de la dynamique des oscillations cérébrales dans la bande delta au sein des régions auditives droites et frontale inférieure gauche lors du traitement de l'information prosodique du signal de parole. De plus, cette étude montre que dans la dyslexie, la communication entre le cortex auditif droit et le gyrus frontal inférieur gauche semble être diminuée par rapport à ce qui est observé chez les sujets contrôles, suggérant un dysfonctionnement de la voie dorsale. Dans une étude antérieure en MEG de ce groupe de chercheurs sur la même population de sujets, mais utilisant des stimuli non verbaux (bruits blancs modulés en amplitude), Lizarazu et collègues (2015) ont montré que les enfants et adultes dyslexiques présentaient une synchronisation neuronale atypique au niveau des rythmes syllabique (thêta, 4 Hz) et phonémique (gamma bas, 30 Hz). Alors que le groupe contrôle montrait une synchronisation latéralisée à droite pour la bande thêta et, bilatérale en gamma, les dyslexiques présentaient le pattern opposé avec une synchronisation bilatérale en thêta, et latéralisée à droite en gamma. Ces résultats suggèrent une altération de la spécialisation hémisphérique droite pour traiter le rythme syllabique chez les dyslexiques, qui serait compensée par un recrutement bilatéral des

régions auditives. Cette interprétation est d'autant plus plausible qu'en espagnol, le rythme syllabique est particulièrement marqué.

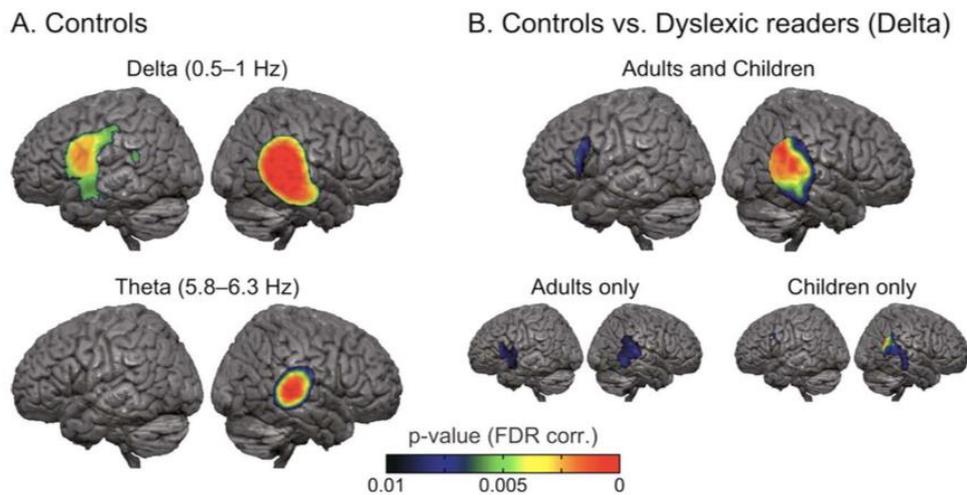


Figure 26 : Analyses de sources révélant une augmentation significative de la cohérence lors de la perception de parole, comparée à la *baseline*, dans les bandes de fréquence delta (0.5-1 Hz) et thêta (5.8-6.3 Hz) (A) pour le groupe contrôle et (B) pour la comparaison contrôle vs. dyslexiques. Figure empruntée à Molinaro et al. (2016).

En conclusion, un nombre croissant d'éléments penche en faveur d'un dysfonctionnement oscillatoire dans la dyslexie, plus particulièrement dans les bandes de fréquence delta et gamma. Un tel dysfonctionnement serait à la base d'un échantillonnage incomplet ou imprécis de l'information rythmique prosodique et phonémique contenue dans le signal de parole, conduisant par conséquent à la formation de représentations phonologiques atypiques. Si des déficits de traitement du rythme (en production et en perception) ont été décrits dans la dysphasie et que des modèles suggérant un échantillonnage oscillatoire inefficace du signal de parole ont été proposés, à notre connaissance, aucune étude à ce jour n'a examiné la dynamique oscillatoire dans les bandes de fréquence caractéristiques des indices rythmiques de l'information linguistique chez des dysphasiques, que ce soit au repos ou lors du traitement de la parole.

VI. Objectifs et problématiques de la thèse

L'objectif principal de ce travail de thèse est de mieux comprendre les processus neurocognitifs qui sous-tendent la perception du rythme de la parole naturelle chez les enfants. Pour ce faire, nous avons étudié les capacités de perception de parole produite naturellement à un débit normal et rapide, et la dynamique corticale oscillatoire sous-jacente en MEG, de deux groupes d'enfants : des enfants présentant un développement typique (dits neurotypiques) et des enfants présentant une dysphasie touchant majoritairement le versant expressif du langage.

Le rythme de la parole contribue de manière essentielle au traitement de l'information linguistique. Les indices rythmiques permettent la segmentation du flux continu de parole en segments de plus petite taille, facilement analysables par le système langagier, les phonèmes, syllabes et mots ou groupes de mots. La syllabe correspond, pour un grand nombre d'auteurs, à l'unité perceptive de base de la parole (Ferrand et al., 1996; Massaro, 1974; Mehler et al., 1981; Ridouane et al., 2009; Segui et al., 1991). Elle joue un rôle majeur dans l'organisation temporelle de la parole en véhiculant l'information rythmique, qui participe de la structure prosodique du signal (Astesano, 2001; Houtgast & Steeneken, 1985). La syllabe porte en effet la plupart des indices rythmiques, telles que les variations d'intensité et de durée des unités, que l'auditeur, dès son plus jeune âge, va exploiter pour segmenter le flux linguistique ; la syllabe caractérise également le débit de parole d'un locuteur (Di Cristo, 2013; Meynadier, 2001). L'information temporelle de la parole est principalement véhiculée au sein de l'enveloppe d'amplitude ou enveloppe temporelle du signal de parole, qui correspond aux modulations lentes d'amplitude du signal englobant l'information rythmique syllabique et prosodique (Drullman et al., 1994a, 1994b; Greenberg et al., 2006; Rosen, 1992). L'intégrité de l'enveloppe temporelle pour l'intelligibilité de la parole a largement été étudiée et démontrée chez l'adulte (Drullman, Festen, & Plomp, 1994a, 1994b; Ghitza, 2012; Steven Greenberg, Arai, & Grant, 2006). Établissant la correspondance étroite entre les rythmes corticaux et le rythme de la parole, les modèles neurobiologiques récents stipulent que la perception de la parole et par là-même sa compréhension seraient permises par la synchronisation des oscillations de basse fréquence delta et thêta sur les modulations lentes de l'enveloppe temporelle de la parole (Ghitza, 2011; Giraud & Poeppel, 2012; Poeppel, 2003). Un tel alignement conduirait à un échantillonnage du signal en unités prosodiques et

syllabiques respectivement ; en parallèle, les oscillations de plus haute fréquence dans la bande gamma permettraient de découper le signal en segments de la taille des phonèmes, qui seraient alors intégrés aux unités syllabiques. Si les recherches autour de cette problématique ont été particulièrement fructueuses ces dernières années, rares sont les études ayant investigué la dynamique oscillatoire sous-tendant le traitement de la parole chez l'enfant, encore moins si l'on considère le cas de la parole naturelle. Le rythme constitue pourtant l'un des fondements de l'acquisition du langage oral et examiner l'alignement entre les rythmes du cerveau et de la parole chez les enfants présentant un développement langagier typique mais aussi atypique est crucial pour améliorer notre connaissance des bases cérébrales du langage.

De plus, tout auditeur sait à quel point le rythme de la parole est particulièrement variable. D'une langue à une autre, d'un locuteur à un autre, mais aussi selon la situation de communication, le signal de parole ne sera en effet pas produit de la même façon et les caractéristiques rythmiques (parmi d'autres) telles que l'accentuation ou le débit varieront (Browman & Goldstein, 1992; Grosjean & Deschamps, 1975; Kohler, 1992). Qui ne s'est pas senti perdu au début d'une conversation avec un interlocuteur ayant un fort accent étranger ou parlant très vite ? Malgré ces difficultés, le système cortical perceptif est capable de traiter cette variabilité du signal de parole (et notamment les variations de débit) afin de décoder le message qui nous est adressé, impliquant notamment un recrutement de la voie dorsale sensori-motrice du modèle à double voie de la perception de la parole (Hickok & Poeppel, 2007; voir Adank & Devlin, 2010; Adank & Janse, 2009 pour des résultats expérimentaux sur de la parole accélérée artificiellement). Là encore, la manière dont les enfants gèrent les variations naturelles de rythme de la parole, en particulier lorsqu'elles concernent le débit, et la dynamique corticale oscillatoire sous-jacente, sont des questions n'ayant, à notre connaissance, jamais fait l'objet de recherches. Les enfants, comme les adultes, sont pourtant confrontés à des interlocuteurs variés, parlant avec des débits différents, et doivent constamment s'adapter. Au vu du déficit de traitement des indices rythmiques de la parole décrit dans certains troubles développementaux du langage (voir Cumming et al., 2015a), ceci pourrait être particulièrement difficile chez les enfants dysphasiques et se traduire par un pattern d'activité corticale oscillatoire atypique.

Dans ce travail de thèse, nous nous sommes ainsi attachés à adresser cette problématique en examinant

1. l'alignement entre les oscillations cérébrales et le signal de parole produit *naturellement* à un débit normal ou rapide chez des enfants francophones neurotypiques ;
2. les capacités d'enfants francophones dysphasiques, comparés à des enfants neurotypiques, pour traiter les variations de débit ;
3. la dynamique des oscillations corticales sous-jacente chez les dysphasiques.

Pour cela, nous avons créé un protocole en MEG afin d'étudier la cohérence entre un signal de parole entrant (phrases produites naturellement en langue française) et le rythme des oscillations corticales chez les enfants. Deux conditions de débit de parole ont été créées : une première condition comportant des phrases produites naturellement à un débit normal (en moyenne 6.6 syllabes/s) et une seconde correspondant à ces mêmes phrases produites naturellement à un débit rapide (en moyenne 9.1 syllabes/s).

Une première étude a été réalisée chez des enfants francophones neurotypiques âgés entre 8 et 13 ans (langage oral acquis) afin de déterminer si le traitement oscillatoire de la parole naturelle chez l'enfant suit le même schéma que celui observé chez l'adulte, à savoir si une synchronisation cortico-acoustique est présente au sein des régions auditives dans la bande de fréquence thêta (4-7 Hz). Nous avons examiné cette synchronisation dans les deux conditions de débit de parole, afin de voir à quel point le rythme cérébral *syllable-like* est capable de s'aligner sur le rythme syllabique variable de la parole. Notre hypothèse était que l'alignement cerveau/parole devrait se produire dans les deux conditions mais que, la parole accélérée naturellement étant plus difficile à décoder, elle pourrait conduire à un recrutement supplémentaire des régions de la voie dorsale sensori-motrice. Outre l'analyse de la bande thêta « classique », nous avons également analysé la cohérence cortico-acoustique dans deux bandes de fréquence centrées sur le propre rythme syllabique de nos signaux de parole (à débit normal et rapide). Ceci nous a permis d'examiner (1) si les oscillations cérébrales suivent précisément le rythme syllabique normal de la parole et (2) si elles sont capables de le suivre lorsqu'il est rapide. Cette première étude constitue le premier chapitre de la partie expérimentale de cette thèse (chapitre VII).

Le second chapitre expérimental (chapitre VIII) porte sur la perception de la parole produite à un débit normal ou rapide, et la dynamique oscillatoire associée, chez des enfants présentant un trouble du langage oral, et notamment une dysphasie touchant le versant expressif du langage. Les enfants (ou adultes) dysphasiques sont connus pour présenter des difficultés à traiter le rythme de stimuli auditifs, langagiers ou non, particulièrement lorsqu'il s'agit d'indices syllabiques contenus dans l'enveloppe temporelle du signal (« *prosodic phrasing hypothesis* » ; Cumming, Wilson, & Goswami, 2015a). Nous avons alors examiné, dans une seconde étude, les capacités de perception de parole à débit normal ou rapide d'enfants francophones âgés de 8 à 13 ans présentant une dysphasie expressive. Ce protocole comportemental consistait en un jugement sémantique de phrases produites dans trois conditions : des phrases produites naturellement à un débit normal, des phrases produites naturellement à un débit rapide, et des phrases compressées temporellement (i.e. accélération artificielle) pour atteindre le débit rapide. Notre hypothèse était que les enfants dysphasiques, comparés à un groupe d'enfants neurotypiques, pourraient présenter une difficulté particulière pour traiter la parole rapide, naturelle et compressée, qui serait marquée par une sensibilité plus faible à l'incohérence sémantique dans les deux conditions de parole rapide.

Enfin, dans une troisième étude, nous avons réalisé le même protocole MEG que celui de l'étude 1 chez les enfants neurotypiques sur un autre groupe d'enfants dysphasiques, afin d'identifier si les troubles de traitement du rythme ainsi que les difficultés supposées pour traiter la parole rapide dans ce trouble du langage oral étaient sous-tendus par une dynamique corticale oscillatoire particulière. Un fonctionnement oscillatoire atypique dans les bandes de fréquence delta et thêta a déjà été observé dans la dyslexie lors de la perception de stimuli verbaux et non verbaux (Hämäläinen et al., 2012; Lehongre et al., 2013; Lizarazu et al., 2015; Molinaro et al., 2016; Power et al., 2013; Soltész et al., 2013), ce qui pourrait être à l'origine du déficit phonologique décrit dans ce trouble (Goswami, 2011). De la même manière, la « *prosodic phrasing hypothesis* » (Cumming et al., 2015a) suggère qu'un dysfonctionnement oscillatoire pourrait sous-tendre les troubles de traitement du rythme dans la dysphasie. En accord avec ce modèle, nous avons émis l'hypothèse que l'alignement oscillatoire dans la bande de fréquence thêta sur le signal de parole, même produit à un débit normal, pourrait être atypique chez notre groupe d'enfants dysphasiques par rapport à des enfants neurotypiques. Nous nous attendions en outre à observer un dysfonctionnement oscillatoire plus important pour la parole produite naturellement à un débit rapide, qui pourrait être reflété par un recrutement atypique de la voie dorsale sensori-motrice.

Ces trois études, s'inscrivant dans le champ de recherches très actif autour des liens entre rythmes cérébraux et parole, fournissent ainsi, pour la première fois à notre connaissance, les premières données sur les processus oscillatoires sous-tendant le traitement du rythme syllabique de la parole *naturelle* chez l'enfant neurotypique et l'enfant présentant une dysphasie expressive entre l'âge de 8 et 13 ans. Nous soulignons également qu'afin d'examiner le couplage entre rythmes cérébraux et langagier au niveau des générateurs corticaux, le choix de la MEG nous a semblé crucial puisque cette technique permet d'obtenir une estimation précise du domaine fréquentiel au niveau des sources. L'utilisation d'un tel outil chez les enfants, qui plus est dans la pathologie, était un défi de taille que nous avons néanmoins choisi de relever dans ce travail de thèse. Dans la suite, les contributions expérimentales seront présentées sous forme d'articles scientifiques rédigés en anglais. Pour chacun, une présentation du travail ainsi qu'une synthèse des résultats en français seront proposées.

CONTRIBUTIONS EXPERIMENTALES

VII. Dynamique oscillatoire lors de la perception de la parole chez les enfants neurotypiques

Ce premier chapitre expérimental se compose d'une étude réalisée en MEG chez les enfants neurotypiques. L'article relatif à cette étude est actuellement soumis dans la revue *NeuroImage*. Les résultats ont par ailleurs fait l'objet de communications affichées dans divers congrès internationaux (dont Biomag 2016, *Cognitive Neuroscience Society meeting* 2016, *Organization for Human Brain Mapping* 2016, *Society for the Neurobiology of Language* 2016).

1. Présentation de l'étude 1

La parole est un signal quasi-périodique et cette rythmicité, caractérisée notamment par les modulations lentes de l'enveloppe temporelle, joue un rôle fondamental dans le décodage de l'information verbale dès le plus jeune âge (Arnal & Giraud, 2016; Drullman, 1995; Goswami, 2011). Par ailleurs, tout auditeur aura constaté que la parole est également un signal très variable, dépendant en particulier du locuteur et de la situation de communication dans laquelle on se trouve. L'une des variations du signal rencontrées quasi-quotidiennement porte sur le débit (défini par le rythme syllabique) auquel la parole est produite. Les modèles neurolinguistiques récents (Ghitza, 2011; Poeppel, 2003; Hickok & Poeppel, 2007) suggèrent l'existence d'un lien entre le rythme des oscillations corticales et l'information rythmique (fluctuations lentes d'amplitude) véhiculée par l'enveloppe temporelle de la parole. Ces deux « structures » seraient caractérisées par des fenêtres d'intégration temporelle comparables, ce qui permettrait aux oscillations corticales de se coupler aux modulations lentes de l'enveloppe et ainsi de faciliter le découpage et le traitement de l'information. Les processus sous-tendant la dynamique des oscillations corticales lors du traitement de la parole produite *naturellement* à un débit rapide n'ont jamais été étudiés chez l'adulte (mais voir Hincapié, Guiraud, Jerbi et Boulenger, en préparation présenté dans la Discussion générale de cette thèse, cf. poster en Annexe 1) et encore moins chez l'enfant. Le rythme de la parole jouant un rôle primordial dans l'acquisition et le développement du langage chez l'enfant (Goswami, 2011, 2015; Jusczyk, 1997), comprendre les processus neurocognitifs sous-tendant son traitement semble pourtant fondamental. Chez l'adulte, les travaux s'étant intéressés à la synchronisation des oscillations corticales sur le signal de parole accéléré ont utilisé une méthode artificielle d'accélération, la compression temporelle, rendant les investigations moins écologiques. De la même manière, la seule étude (à notre connaissance) à ce sujet chez l'enfant a utilisé de la parole compressée temporellement, avec la répétition de la même phrase tout au long de l'expérience (Abrams et al., 2008).

Dans la présente étude, nous avons examiné les processus oscillatoires corticaux mis en jeu chez l'enfant lors de la perception de parole naturelle produite à un débit normal ou rapide. Pour ce faire, nous avons réalisé une étude en MEG (275 capteurs) et enregistré l'activité cérébrale d'enfants francophones de langue maternelle française, âgés entre 8 et 13 ans, droitiers et présentant un développement typique. Les enfants avaient pour tâche d'écouter des phrases produites naturellement par un locuteur francophone dans les deux

conditions de débit de parole (naturel normal : 6.61 syllabes/s, $n = 150$, et naturel rapide : 9.03 syllabes/s, $n = 150$) et de repérer des stimuli « *filler* » non analysés (phrases dans lesquelles un bruit modulé en fréquence était ajouté) en appuyant sur un bouton avec leur index gauche. Cette tâche était réalisée afin de maintenir l'attention des enfants tout au long de l'expérience. L'ensemble des stimuli étaient présentés dans un ordre pseudo-aléatoire. Une expérience pilote menée chez 11 enfants neurotypiques (différents des enfants ayant réalisé l'expérience en MEG) qui devaient décider si les phrases avaient un sens ou non a permis de montrer des performances d'intelligibilité de 84 % pour la parole à débit normal et 74 % pour la parole à débit rapide. Ainsi, la parole accélérée naturellement était globalement plus difficile à comprendre pour les enfants mais restait néanmoins intelligible et donnait lieu à des performances supérieures au hasard.

Dans l'étude en MEG, nous avons réalisé différentes analyses, au niveau des capteurs puis des sources, afin de mesurer les modulations de puissance et la cohérence entre les modulations lentes de l'enveloppe du signal de parole (obtenue à l'aide de la transformée de Hilbert) et les oscillations corticales dans différentes bandes de fréquence. Trois bandes nous intéressaient tout particulièrement : la bande thêta « classique » (4-7 Hz) et deux bandes centrées sur le rythme syllabique moyen de nos stimuli (5.6-7.6 Hz centrée sur le débit de parole normal, et 8-10 Hz centrée sur le débit de parole rapide).

Nos hypothèses étaient que les enfants neurotypiques devraient présenter une dynamique oscillatoire similaire à celle observée dans la littérature chez l'adulte lors de la perception de parole à débit normal, soit une synchronisation des oscillations corticales dans les régions auditives, latéralisées dans l'hémisphère droit voire bilatérales (*e.g.*, Gross et al., 2013; Peelle, Gross, & Davis, 2013), sur l'enveloppe temporelle du signal dans la bande de fréquence « classique » thêta et la bande 5.6-7.6 Hz centrée sur le débit normal moyen des phrases. Lors de l'écoute de parole rapide, une augmentation de cohérence dans la bande thêta pourrait également être observée au vu des études montrant un alignement thêta tant que la parole reste intelligible (*e.g.*, Ahissar et al., 2001). En outre, compte-tenu des résultats de l'expérience pilote et d'une étude menée chez des enfants de CE2 et CM2 montrant une adaptation à la parole rapide (Guiraud et al., 2013), nous nous attendions à observer une synchronisation des oscillations corticales dans les régions auditives sur le rythme syllabique du signal de parole entrant, soit dans la bande de fréquence 8-10 Hz centrée sur le débit rapide. Néanmoins, la parole accélérée naturellement étant moins intelligible que la parole produite à un débit normal, un recrutement des régions motrices et prémotrices de la voie

dorsale, censées assister la perception de la parole dans des conditions difficiles (Hickok & Poeppel, 2007), pourrait être observé en condition de débit rapide. Ceci s'accorderait avec le rôle primordial des régions articulatoires et des interactions sensori-motrices dans le développement du langage.

Auditory and premotor neuronal entrainment to normal and fast speech rhythms: a MEG study in typically developing children.

Hélène Guiraud^{1*}, Ana-Sofia Hincapié^{2,3}, Karim Jerbi^{2a}, Véronique Boulenger^{1a*}

¹ Laboratoire Dynamique Du Langage, UMR 5596 CNRS/Université de Lyon, France

² Département de Psychologie, Université de Montréal, Montréal, Québec, Canada

³ Pontificia Universidad Catolica de Chile, Santiago de Chile, Chile

^a These authors contributed equally to this work.

* Corresponding authors

Correspondence should be addressed to:

Hélène GUIRAUD

Laboratoire Dynamique Du Langage
CNRS/Université de Lyon UMR5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON Cedex 07 - FRANCE
helene.guiraud@univ-lyon2.fr

Véronique BOULENGER

Laboratoire Dynamique Du Langage
CNRS/Université de Lyon UMR5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON Cedex 07 - FRANCE
veronique.boulenger@cnrs.fr

Abstract

Multi-time resolution models suggest that neural oscillations in the auditory cortex synchronize to the slow modulations of speech temporal envelope, which convey rhythmic syllabic and prosodic information. Whereas supporting evidence has been provided in adults, only scarce studies have investigated the dynamics of cortical oscillations during natural speech perception in typically-developing children. The present MEG study aimed at tackling this issue by recording cortical activity in 8-13 year-old French-speaking children while they listened to sentences *naturally* produced either at a normal rate (average 6.6 syllables/s) or at a fast rate (average 9 syllables/s). Source-level coherence analyses revealed that theta-band (4-7 Hz) oscillations in right associative auditory cortex aligned to speech envelope in the normal rate condition, in agreement with adult studies and models. For naturally accelerated speech, increased coherence in this frequency range was seen in left premotor and motor regions as well as in Broca's area, consistent with an involvement of the dorsal stream in more difficult speech perception. Our data furthermore demonstrated neural entrainment to the specific syllabic rate of sentences when produced at normal rate. Significant cortico-acoustic coherence was indeed found in left anterior temporal regions in the 5.6-7.6 Hz frequency band, which corresponds to the mean speech rate of normal rate sentences. In line with previous studies in adults, we interpreted this left-lateralized theta oscillatory activity as reflecting an integrative process where syllabic units would integrate phonemic units in left auditory regions in order to achieve higher-level linguistic processing. Altogether our findings unravel, for the first time to our knowledge, that cortical oscillations in auditory but also in left articulatory regions entrain to natural speech rate modulations in typically-developing children.

Keywords : Speech rhythm; Brain oscillations; Children; Entrainment; Syllabic rate, Speech, Magnetoencephalography (MEG)

Highlights

- Children's theta-band brain oscillations track sentences' temporal envelope
- This occurs in right auditory and fronto-parietal regions for normal rate speech
- Entrainment to natural fast speech is observed in left (pre)motor regions
- Temporal cortex aligns to speech in the normal syllabic rate frequency band

1. Introduction

The last decade or so has witnessed a growing interest in the way the listener's brain tracks the speech signal for comprehension. Emphasizing the striking correspondence between speech and brain intrinsic rhythms, a large body of evidence suggests a functional role of endogenous cortical oscillations in speech segmentation and decoding. Electroencephalography (EEG) and magnetoencephalography (MEG) studies have indeed revealed that brain oscillatory activity synchronizes, or is entrained to speech modulations mainly in the delta-theta (delta: < 3 Hz; theta: 4-7 Hz) and low gamma ranges. This would allow the listener to parse the incoming signal into relevant linguistic units (i.e. phrases, syllables and phonemes respectively), therefore facilitating comprehension (Ding and Simon, 2014; Ghitza, 2011; Giraud and Poeppel, 2012; Peelle and Davis, 2012). Neural entrainment to speech has largely been investigated in the theta frequency band (Cogan and Poeppel, 2011; Ding and Simon, 2013; Doelling et al., 2014; Gross et al., 2013; Howard and Poeppel, 2012; Peelle et al., 2013; Vander Ghinst et al., 2016; Zion Golumbic et al., 2013), wherein brain activity aligns with low-frequency (4-8 Hz) quasi-rhythmic fluctuations in the speech amplitude envelope, which conveys prosodic content and syllabic structure and plays a key role in speech intelligibility (Ghitza and Greenberg, 2009; Peelle and Davis, 2012; Zion Golumbic et al., 2012, for reviews). The necessity of speech comprehension for brain-to-speech alignment is however currently still debated (Ding and Simon, 2014; Peelle and Davis, 2012). In a seminal work, Ahissar and colleagues (2001) demonstrated that auditory oscillatory activity was phase-locked to the temporal envelope only for intelligible sentences that were moderately time-compressed (compression ratios of 0.75 of the original signal duration, with envelope fluctuations at ~ 4 Hz, and of 0.50 corresponding to ~ 6 Hz). When sentences were heavily accelerated and therefore unintelligible (ratios of 0.35 ~ 9 Hz and of 0.20 ~ 15 Hz), synchronization between the two signals failed. This measure furthermore strongly correlated with sentence comprehension: when participants' performance declined with increasing time compression, so did envelope tracking. The authors suggested that the match between speech rate and the modulation capacities of brain auditory regions is a prerequisite for successful comprehension. Other studies later provided evidence that neural entrainment to speech is indeed modulated by intelligibility but can still occur for unintelligible speech (Gross et al., 2013; Howard and Poeppel, 2010; Luo and Poeppel, 2007; Nourski et al., 2009; Park et al., 2015; Peelle et al., 2013; Pefkou et al., 2017; Zoefel and Vanrullen, 2015). In a MEG study using noise-vocoding, which reduces spectral details but

preserves temporal envelope, Peelle and colleagues (2013) showed bilateral theta-band phase-locking to amplitude envelope for unintelligible sentences (1 frequency channel). Entrainment was however enhanced in the left middle temporal gyrus when speech was moderately or fully intelligible (4 and 16 frequency channels respectively). Gross and collaborators (2013) also reported that delta and theta oscillations were more tightly coupled to continuous speech envelope, as compared to reversed (i.e. unintelligible) speech, in bilateral fronto-temporal cortex. Speech low-frequency modulations additionally affected gamma amplitude more strongly for intelligible speech in bilateral temporal, frontal and parietal regions. Altogether, these data therefore provide evidence for cortical entrainment to speech syllabic rhythm that is not only governed by sensory cues but also by linguistic information. Such brain-to-speech coupling is thought to play a functional role in speech comprehension by enabling efficient extraction of key speech components in the speech stream (Peelle and Davis, 2012).

Despite extensive research on neural entrainment to speech, evidence for alignment between brain and speech rhythms in children is surprisingly scarce. Temporal structure information is nevertheless a foundation for language acquisition and development (Goswami, 2011, 2015; Jusczyk, 1997). Sensitivity to linguistic rhythm and prosody, already present in newborns, is thought of as a bootstrapping cue to language discrimination and segmentation as well as to morphosyntactic processing (Mehler et al., 1988; Nazzi & Ramus, 2003). Successful perception of rhythmic patterns in speech temporal envelope is also critical for phonological development (Goswami, 2011, 2015) and growing evidence points toward impaired rhythmic skills in children with language impairments (see Goswami and Leong, 2013 for a review). A very limited number of studies have investigated brain synchronization to slow amplitude modulations in speech in typically developing children. Using EEG, Abrams and colleagues (2008) presented children aged 9-13 years old repeatedly with a sentence produced in three modes: conversational, clear or compressed (to half original duration). Their results revealed phase-locking to speech amplitude envelope over temporal electrodes (reflecting activity from secondary auditory regions) irrespective of the speech condition (i.e. whether envelope modulations were large as in clear speech or more rapid as in time-compressed speech). In addition, the right auditory cortex followed the envelope contours more precisely than the left (see also Abrams and al., 2009). In another EEG study conducted in children (mean age of 13 years old), Power and collaborators (2012) showed neural auditory entrainment to a 2 Hz rhythmic syllabic signal (repetition of /ba/ presented in the auditory and/or visual modality) in the delta (1-3 Hz) as well as in the low-theta frequency range (3-5 Hz, however only in the

auditory modality). The strength of theta coupling was further positively correlated with children's written (and to a lesser extent spoken) language skills.

These findings offer preliminary evidence for cortical entrainment to amplitude envelope in children, yet our knowledge of the brain oscillatory signatures of speech temporal structure is still incomplete. Given that speech rate varies substantially between speakers and contexts in daily life, the ability of neural oscillations to dynamically adjust to rhythmic changes in signal envelope may be crucial for language comprehension throughout development. In particular, the question of whether and how neural oscillations in children track the pace of fast syllabic patterns when speech is naturally uttered more rapidly remains to be tackled. While Abrams and colleagues (2008, 2009) started to examine this question in their study, they used only one stimulus which was repeated many times and, most importantly, artificially accelerated.

In the present MEG study, we sought to unravel how the brain of typically developing children tracks slow modulations in speech temporal envelope when exposed to sentences naturally produced both at normal and fast rates.

2. Materials and Methods

2.1. Participants

Seventeen children aged between 8 to 13 years old (3 girls; mean age: 10.52, standard deviation (SD): 1.60) participated in the experiment. All children were native French speakers, right-handed, with no known history of developmental, language, neurological or psychiatric disorder. All children were screened for audiometric pure-tone thresholds (audiometer 600 M, Electronica Technologies) and no one exceeded 30 dB Hearing Level (HL) at any frequency in the 250-6000 Hz range. Three children had to be excluded from MEG analyses due to head movements in the helmet (see below), thus impoverishing signal quality. Consequently, the final sample included 14 children (5 girls; mean age: 11.12, SD: 1.65). All children underwent a battery of neuropsychological and French language tests to assess their verbal and non-verbal abilities. Linguistic abilities were evaluated with a series of tests from the BALE (Batterie Analytique du Langage Ecrit; Jacquier-Roux et al., 2010 : word, pseudo-word and non-word repetition, phonemic fluency, picture naming, metaphonological tasks) and two tests from the ELO (Evaluation du Langage Oral; Khomsi, 2001 : oral comprehension (i.e. sentence-to-picture matching) and morpho-syntax). Non-verbal intellectual abilities were assessed with the Raven's coloured Progressive Matrices (Raven et al., 1998) and the forward and backward digit span tests. In the different subtests,

all children obtained scores higher than -2 SD of the population mean. Finally, each child's speech rate was measured in a narrative task using the picture book "Frog where are you?" (Mayer, 1969). In this task, children are presented with a 24-page wordless picture book and are asked to tell the story to the experimenter. Children's narratives were recorded with a microphone (audio-technica AT2020USB, 44 kHz). We then computed the speech rate for each child/story by calculating the number of syllables per second after removing silence periods longer than 150 ms (Pellegrino et al., 2011). The mean speech rate of our sample of 14 children was 3.93 syllables/s (SD: 0.50).

The protocol conformed to the Declaration of Helsinki and was approved by the local ethical committee (Comité de Protection des Personnes Lyon Sud-Est II; ID RCB: 2012-A00857-36). All children and their parents signed a consent form before the experiment.

2.2. Materials

The stimuli were composed of 300 sentences (7-9 words) following the same syntactic structure (e.g. "Le public applaudit le joueur pour sa victoire" / The public applauds the player for his victory). Sentences were recorded by a French native male speaker (44.1 kHz, mono, 16 bits) in a sound-attenuated booth using ROCme! Software (Ferragne et al., 2012). Each sentence was recorded twice, at a normal and then at a fast rate. For the latter, the speaker was instructed to produce the sentences as fast as possible while remaining intelligible. The durations of the 2×300 sentences (average duration: 1.68 ms (SD: 0.17) for normal rate sentences and 1.23 ms (SD: 0.14) for fast rate sentences) and the number of produced syllables for each sentence were then calculated with Praat software (Boersma and Weenink, 2012). The average speech rate was 6.61 syllables/s (SD: 0.47) for normal rate sentences and 9.03 syllables/s (SD: 0.56) for fast sentences (Table 1). An 80 Hz high-pass filter was applied to all sentences, the amplitude envelope was smoothed sentence-initially and finally. The intensity of the sound files was finally peak normalized.

The 600 sentences (300 x 2 rate variants) were divided into two experimental lists of 300 items each (150 normal rate sentences and 150 fast rate sentences) using a Latin square design so that each stimulus appeared in each rate condition across all participants but only once per list (to avoid repetition effects). Forty filler sentences with the same syntactic structure (20 normal rate and 20 fast rate; not analyzed), in which sound effects (<http://www.sound-fishing.net/>) with a duration of 539 ms were superimposed at the end of the sentence file, were also added to each experimental list (for a total of 340 trials).

Within each list, the 340 stimuli were pseudo-randomly divided into 10 blocks of 34 stimuli each (15 normal rate sentences, 15 fast rate sentences, 2 normal rate fillers, 2 fast rate fillers, with no more than 3 consecutive stimuli from the same condition).

Table 1 : Mean syllabic rate (syllables/s) of sentence stimuli in the normal and fast rate conditions. The frequency bands defined for coherence and power analyses were centered on these values (+/- 1 Hz).

	<i>Speech Rate Conditions</i>	
	<i>Normal</i>	<i>Fast</i>
<i>Meansyllabic rate (syllables/s)</i>	6.61 (SD 0.47)	9.03 (SD 0.56)
<i>Syllabic-rate-centered frequency bands (Hz)</i>	5.6 - 7.6	8 - 10

In order to measure the coherence between the syllabic rate of the incoming speech signal and brain oscillatory activity, we extracted the amplitude envelope for each sentence in each condition using the method developed by Peelle and collaborators (2013). The speech signal was rectified (full wave rectification) and low-pass filtered at 30 Hz (fifth order Butterworth filter, zero-phase digital filtering; Fig 1). This envelope served as the acoustic signal for all coherence analyses.

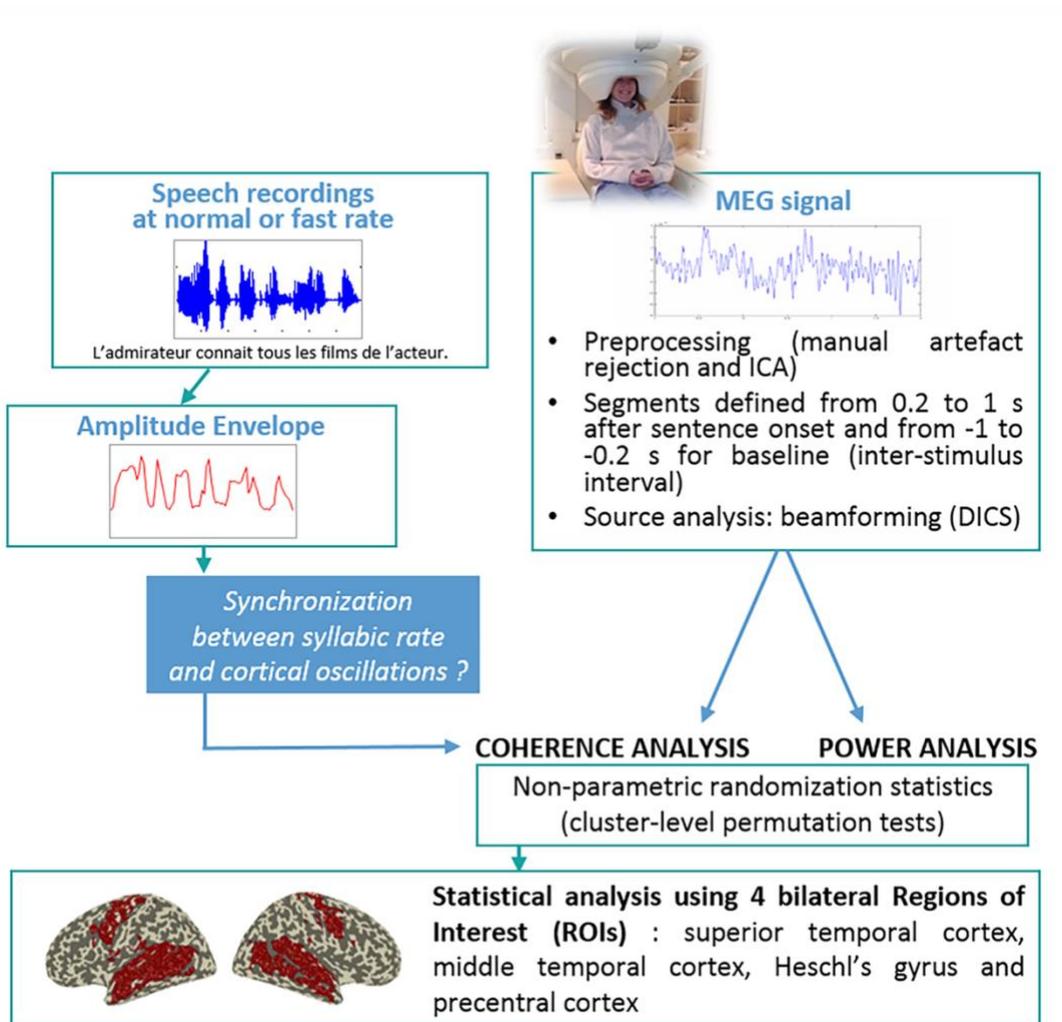


Fig 1 : Overview of MEG data analysis.

2.3. Procedure

Each child came to the lab twice for the purpose of the study: once to undergo the battery of neuropsychological and language tests as well as the audiometry testing, and a second time to participate in the MEG experiment. During the MEG recording, the children were comfortably lying in a sound-attenuated, magnetically-shielded recording room with a screen 40 cm in front of them. To minimize head movements, a piece of foam was placed between the back of their heads and the MEG helmet. Stimuli were presented binaurally through air-conducting tubes with foam insert earphones (Etymotic ER2 and ER3). Prior to the MEG recording, children's auditory detection thresholds were determined for each ear with a 1-minute pure tone of 400 to 3000 Hz; the level was then adjusted so that the stimuli were presented at 50 dB Sensation Level with a central position (stereo) with respect to the children's head.

Once the instructions had been displayed on the screen and explained verbally, a training phase composed of five sentences (not included in the main experiment and not analyzed) was proposed to the children. They were instructed to listen carefully to the sentences while looking at a fixation cross on the screen, and press a button (button response Neuroscan) with their left index finger whenever they heard a filler trial. To maintain children's attention throughout the experiment, they were told that five questions would be asked on the heard sentences at the end of the experiment. Moreover, the experiment was presented as a game: each correct answer to a filler trial uncovered a piece of a treasure map displayed on the screen (each incorrect response was on the other hand associated with a negative feedback reminding them the task). At the end of the experiment, the children could therefore pick up the treasure (a game and a book) thanks to the discovered map. The experiment was run with Presentation software (Neurobehavioral Systems).

One of the two experimental lists was assigned to a child. Each block lasted 230 seconds (less than four minutes), with several short breaks. Each trial started with a white fixation cross presented on the screen for 1500 ms; a sentence stimulus which the child had to attentively listen to was then presented. This was followed by an inter-trial interval (grey screen) of 1250 ms. In the case of filler stimuli, the child was instructed to press the button response as quickly as possible and a feedback (positive or negative) was displayed on the screen for 6250 ms. The next trial was then presented after an interval (grey screen) of 500 ms. The total duration of the experiment was 50 minutes.

2.4. MEG data collection

The different steps of MEG data collection and analysis are summarized in Fig. 1. MEG signals were acquired on a 275-channel whole head sensor system (OMEGA MEG CTF 275, VSM MedTech, Canada) in CERMEP, Lyon, France (www.cermep.fr). The data were continuously recorded at a sampling frequency of 1200 Hz. Three fiducial coils (nasion, left and right pre-auricular points) were placed on each child to determine head position within the MEG helmet. Reference head position was recorded before and after the experiment. Head movements were tracked throughout the experiment using continuous head position identification (HPI). Maximum head movements were calculated and children with mean head displacement exceeding 1 cm from the reference position were not included in the analysis ($n = 3$).

2.5. MEG data analysis

All analyses were performed with Matlab (Mathworks Inc., MA, USA) and the Fieldtrip open-source software (Oostenveld et al., 2011). Data were down-sampled to 300 Hz and artefacts from the mains power supply at 50, 100 and 150 Hz were removed using discrete Fourier transform. Signal slow drifts were also suppressed. The data were then segmented into periods of 2.5 s (1.5 s after sentence onset and 1 s before for the baseline) and were cleaned manually and semi-automatically to remove jumps and muscular artefacts. Finally, independent component analysis (ICA) was applied to correct for heart beat artefacts and eye movements (20 ICA components). Across all subjects, analyses were performed on an average of 139 trials (out of 150) per speech rate condition (SD: 21.51 for the normal rate condition; SD: 19.18 for the fast rate condition).

2.6. Source localization

A template child brain (age 7.5-13.5 years old, Fonov et al., 2011), realigned to the children's fiducial points, was used to create individual children's head model. The homogeneity of head circumference across the children was verified through measurements performed before the recording sessions (average head circumference: 53.14 cm, SD: 1.64).

In order to perform source analysis in a common space and to use an adult atlas for statistics by Regions Of Interest (ROIs), we constructed a source model in a common coordinate system (MNI) from an adult template (common grid with 5780 vertices). Individual children's head model was then transposed on the MNI brain. We chose to use an adult template because the brain of a child at 8-13 years old is anatomically close to that of an adult; it has indeed already reached 90% of its maximum size at the age of 6 (Giedd, 2004; Shapiro and Janzen, 1960).

We measured power and cortico-acoustic coherence at the source level based on the frequency domain beamformer method (DICS; Gross et al., 2001). Coherence corresponds to the linear relationship between the amplitude of two signals in the frequency domain. Its value ranges from 0 to 1, and mathematically defined by:

$$Coh(r_{ef}, r_c, f) = |Cs(r_{ef}, r_c, f)|^2 / (Cs(r_{ef}, r_{ef}, f) Cs(r_c, r_c, f))$$

Where ref is the reference signal, namely the amplitude envelope of the speech signal; rc is the signal at each vertex of the anatomical grid estimated with DICS; f is the frequency bin and Cs is the cross-spectral density matrix. Coherence in a given frequency band was

obtained by averaging the coherence values obtained over all frequency bins within the band of interest.

For each trial of each speech rate condition, we defined an active period (0.2 to 1 s after sentence onset) and a baseline taken from the inter-stimulus interval (-0.2 to -1 s before onset), for which the cross spectral density matrix was computed using fast Fourier transform (FFT) and a multitaper-based spectral smoothing with +1 Hz smoothing for each frequency band (cf. Table 1). We used the standard FieldTrip settings for the DICS algorithm (regularization parameter is the largest singular value of the cross spectral density matrix). Coherence between the MEG signals and the speech signal was calculated across all active and baseline periods (defined above) for each condition and frequency-band of interest. Finally, an average over all children in each condition was computed and the statistics were performed with non-parametric randomization Monte-Carlo tests (1000 repetitions) using Fieldtrip. Results were corrected with cluster-level permutation tests ($\alpha = .05$; Maris and Oostenveld, 2007). The same methods (FFT, multitapering and non-parametric statistics) were also used to compute and assess task-based spectral power modulations.

Two main analyses were carried out. First, we conducted a whole-brain analysis to measure coherence between speech amplitude envelope and MEG signals, as well as power modulations, in the theta band (4-7 Hz). This band corresponds to slow modulations typically found in regular speech signal envelope (Pelle and Davis, 2012; Thompson et al., 2016) and is classically used in the literature to assess brain-to-speech synchronization (e.g., Luo and Poeppel, 2007, 2012; Pelle et al., 2013; see Giraud and Poeppel, 2012 for a model). Second, we focused on the syllabic rate of our stimuli by defining two narrower frequency bands centred on mean computed rates (± 1 Hz): 5.6-7.6 Hz for the normal condition (mean rate = 6.61 syllables/s) and 8-10 Hz for the fast condition (mean rate = 9.03 syllables/s). For each frequency band and condition, we analysed cortico-acoustic coherence and power modulations in four bilateral ROIs. These regions were defined based on models of speech perception (Giraud and Poeppel, 2012; Hickok and Poeppel, 2007) and on the results obtained in the theta band (4-7 Hz) coherence analysis. They included Heschl's gyrus, superior temporal gyrus, middle temporal gyrus and precentral cortex, according to the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002; see Figure 27).

3. Results

3.1. Tracking of speech temporal envelope in the theta band

The results of the whole-brain coherence analysis between speech envelope and oscillatory activity in the theta band (4-7 Hz) are presented in Fig 2A. They revealed tracking of the temporal envelope of normal rate speech in the right hemisphere, more precisely in the associative auditory cortex (anterior middle temporal gyrus – Brodmann area (BA) 21, and anterior superior temporal gyrus – BA22, part of Wernicke’s area) as well as in the superior frontal (BA6/8) and parietal cortices (superior parietal lobule – BA7 and somatosensory cortex – BA1). Small but significant peaks of coherence were also found in the right dorsolateral prefrontal cortex (BA9/10), the right inferior frontal gyrus (BA44) and the right supramarginal gyrus (BA40). As shown by the coherence maps in Fig 2A, the same analysis carried out for the fast rate condition revealed a slightly different pattern of results. Cortico-acoustic coherence indeed mostly increased in left premotor (BA6) and primary motor regions (BA4) as well as in Broca’s area (BA44); oscillatory activity in the right anterior superior temporal (BA22) and right inferior frontal cortex (BA44) also showed enhanced coherence with fast speech signals.

To ensure that the significant coherence observed reflects a genuine synchronization phenomenon and is not driven by power modulations (at the same frequencies), power analyses for both conditions in the theta band were conducted. No significant modulations of power were observed in any of the speech rate conditions as compared to the baseline (Fig 2B).

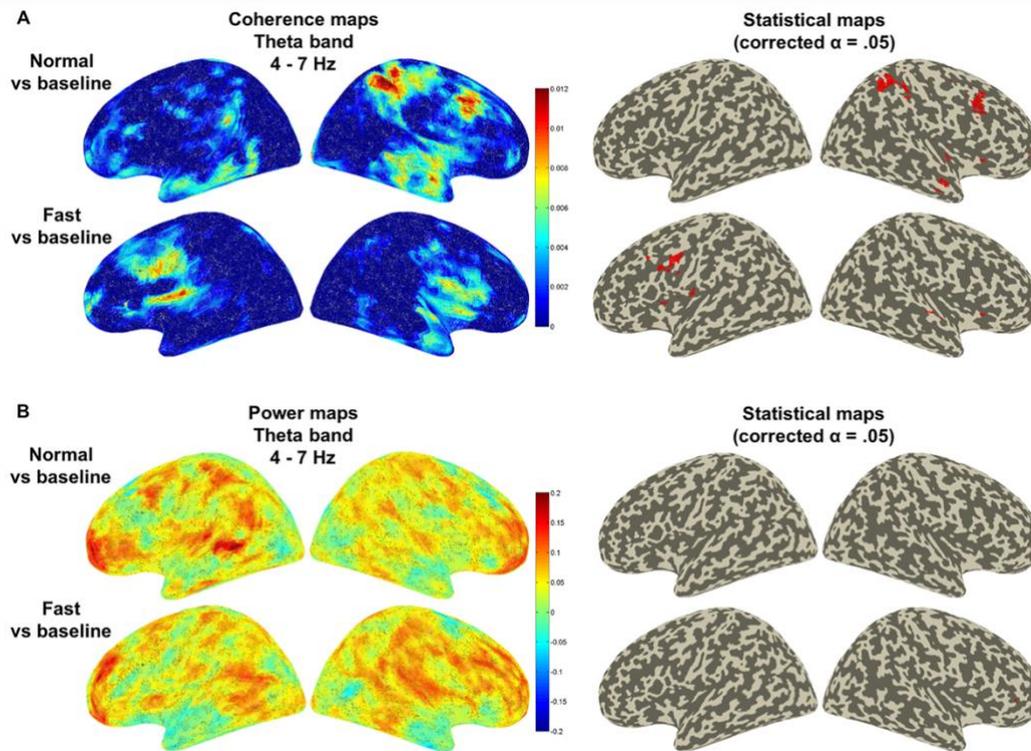


Fig 2 : Cortico-acoustic coherence in theta band.(A) Coherence maps between speech amplitude envelope and cortical oscillations and (B) power modulations maps in the theta frequency band (4-7 Hz) for the normal rate and fast rate conditions, as compared to baseline (relative power). Statistical maps (corrected $\alpha = .05$ and uncorrected $\alpha = .05$) for whole-brain analyses are also presented.

These results in typically-developing children provide evidence for synchronization of theta cortical oscillations to speech envelope when speech is uttered both at a normal rate (~ 6.5 syllable/s) and at a fast rate (~ 9 syllables/s). Yet different neural networks with different patterns of hemispheric lateralization are engaged.

3.2. Tracking of sentence syllabic rhythm by cortical oscillations

Since we were primarily interested in the way the children's brain tracks syllabic rhythm, we examined the coherence between speech amplitude envelope and MEG signals in frequency bands centred on the syllabic rate of our stimuli (see Method). As highlighted in Fig 3A, the ROI-based analysis revealed significant enhanced cortico-acoustic coherence in the 5.6-7.6 Hz frequency band only for the normal rate condition, namely in the frequency band specifically matching the mean syllabic rate of the stimuli. This was mainly observed in the left anterior temporal cortex (BA21/22). No significant coherence between cortical

oscillations and temporal envelope of fast rate sentences was found in this frequency band. As to the analysis within the frequency band centred on fast syllabic rate (8-10 Hz), coherence between cortical oscillations and speech envelope did not reach significance in either condition (Fig 3B).

These data suggest that oscillations mainly in the left anterior temporal cortex of children synchronize to the specific syllabic rhythm of the speech signal at a normal rate; by contrast, such specific tracking does not seem to occur in our sample of children when sentences are produced faster.

Results of the power analyses carried out in the syllabic rate-based frequency bands for each condition, as compared to baseline, are presented in Figure 30. In the normal rate frequency band (5.6-7.6 Hz), no significant power modulations were observed in any of the speech rate conditions (Fig 4A). By contrast, in the fast rate frequency band (8-10 Hz), power decreased in the left superior and middle temporal cortex (BA21/22) for both speech rate conditions (power decrease was observed bilaterally for uncorrected data, Fig 4B). Note that this frequency band also corresponds to the (low) alpha band (8-13 Hz), wherein desynchronization in the auditory cortex is classically observed during auditory stimulus processing (Obleser and Weisz, 2012; Strauß et al., 2014).

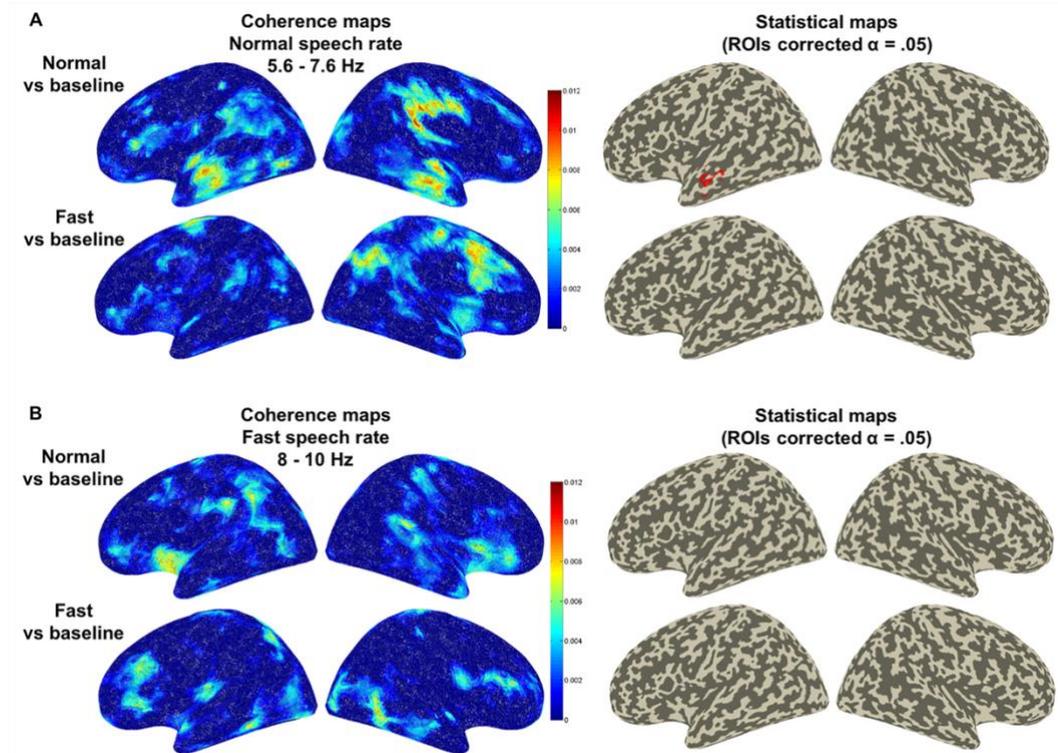


Fig 3 : Cortico-acoustic coherence in frequency bands centered on the mean sentence syllabic rate. Coherence maps between speech amplitude envelope and cortical oscillations, as compared to baseline, together with statistical maps (corrected $\alpha = .05$ and uncorrected $\alpha = .05$) for ROIs analyses, are presented for the normal rate and fast rate conditions in frequency bands centered on the mean sentence syllabic rate, namely (A) 5.6-7.6 Hz and (B) 8-10 Hz.

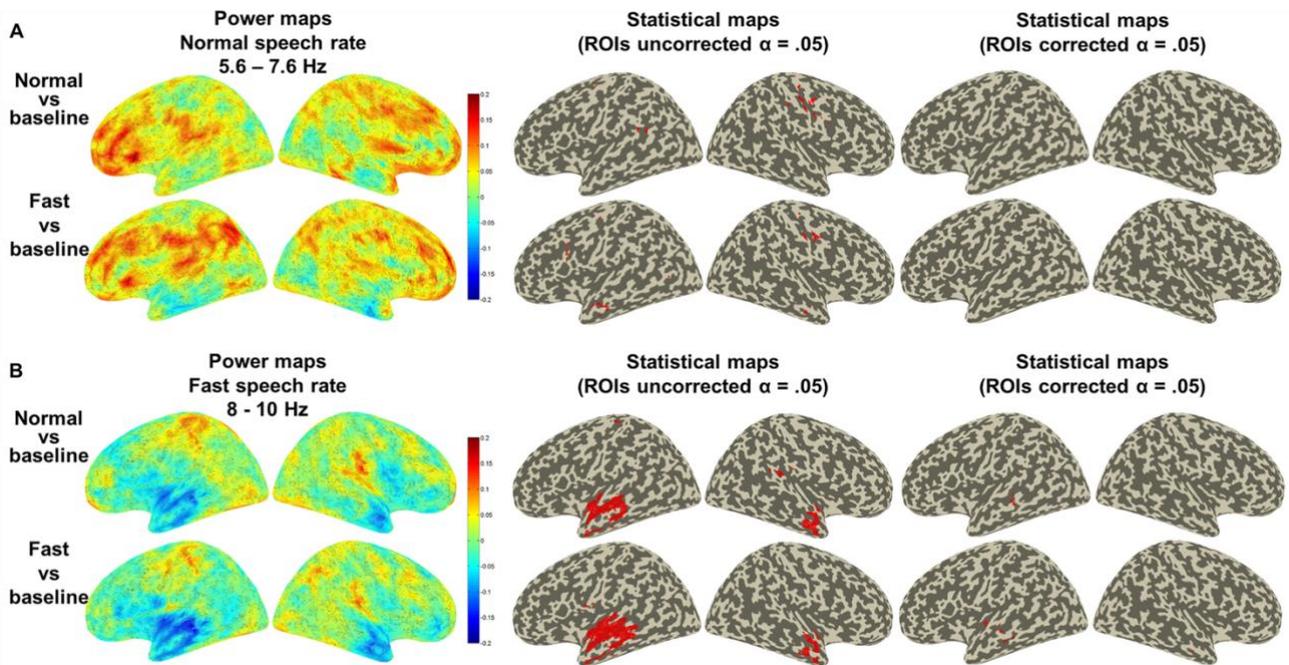


Fig 4 : Power modulations in frequency bands centered on the mean sentence syllabic rate. Power modulations maps in the normal rate and fast rate conditions, as compared to baseline, (A) in the 5.6-7.6 Hz and (B) the 8-10 Hz frequency bands (relative power). The corresponding statistical maps for ROIs analyses are presented (corrected and uncorrected, $\alpha = .05$).

4. Discussion

A growing literature has convincingly demonstrated alignment of theta-band cortical oscillations to slow modulations in speech temporal envelope, which characterize syllabic patterns. However, brain-speech rhythmic coupling in typically developing children is still poorly understood and studies so far have not addressed the issue of naturally varying speaking rates in trial-unique stimuli (previous studies have indeed used either repeated time-compressed sentences or syllables; (Abrams et al., 2008; Alan James Power et al., 2012)). The present study provides first MEG evidence in 8-13 year old children that theta oscillatory entrainment occurs for naturally-produced sentences, both when they are produced at a normal and a fast rate. This however involves slightly different neural networks: whereas theta oscillatory activity in right temporal and fronto-parietal regions tracks amplitude envelope of normal rate sentences, brain alignment to fast rate speech in this frequency range specifically occurs in left premotor and motor regions. Our data furthermore reveal cortico-acoustic coupling in left anterior temporal regions that is specific to the syllabic rate of sentences, at least when they are uttered at a normal rate.

4.1. Theta-band cortical entrainment to speech envelope

Significant cortico-acoustic coherence in the theta range (4-7 Hz) was found in the right auditory association cortex (BA21/22) when children listened to sentences naturally produced at a normal rate (Fig 2A). Right-lateralized brain oscillatory responses to speech amplitude envelope, previously reported in adults (e.g., Gross et al., 2013; Luo and Poeppel, 2007; Vander Ghinst et al., 2016) and to a lesser extent in children for simple verbal and non-verbal stimuli (Abrams et al., 2008; Boemio et al., 2005; Molinaro et al., 2016; Telkemeyer et al., 2009, 2011; Vanvooren et al., 2014), nicely fit with the Asymmetric Sampling in Time (AST) theory (Poeppel, 2003; Poeppel et al., 2008). This multi-time resolution model proposes that slow acoustic fluctuations (conveying syllabic and phrasal structure) would be primarily processed in the right auditory cortex, whereas the left auditory cortex would concurrently extract fast phonemic information. In support of this assumption, resting-state theta oscillations have been shown to correlate more strongly with activity in right auditory regions, while higher-frequency gamma modulations (28-40 Hz) were biased leftward (Giraud et al., 2007; Morillon et al., 2010). To our knowledge, no such resting-state investigation has been conducted in children of the same age as our participants; the only study we are aware of targeted early childhood (3-5 years old) and showed symmetric distribution of endogenous low-frequency (3-7 Hz) oscillations (Thompson et al., 2016). Although this would need to be tested directly, our functional data on sentence processing could indicate that spontaneous theta oscillations may become right-lateralized in later childhood (while the auditory cortex is still maturing; Paus et al., 1999), resulting in stronger coupling to speech envelope in the right hemisphere.

Theta-band brain alignment to normal rate speech was also observed in right frontal and parietal regions, in agreement with previous studies in adults. Gross and colleagues (2013) indeed reported significant correlation between delta/theta neural oscillations and speech amplitude envelope in the parietal cortex (although with a left hemisphere advantage). In the work by Keitel et al. (2017), delta auditory cortical entrainment to speech envelope interacted with oscillatory power in a fronto-parietal network in theta, alpha and beta ranges (see also Park et al., 2015 for frontal modulation on auditory entrainment to speech). Such fronto-parietal activity may reflect the role of these regions in language comprehension and prosodic processing (see Friederici, 2011).

As to fast rate speech, our results in 14 children reveal significant theta entrainment to amplitude envelope in left premotor and primary motor cortex as well as in Broca's region (Fig 2A). Increased coherence was in addition found in right anterior temporal and inferior frontal regions; note that these latter regions were part of the network showing brain-speech rhythmic synchrony in the normal rate condition. Oscillatory entrainment to temporal envelope in (pre)motor regions when speech is uttered more rapidly may indicate the involvement of the left-lateralized dorsal auditory stream, which connects posterior temporal to dorsal premotor regions and Broca's area via the inferior parietal cortex (Saur et al., 2008) and is thought to play a key role in language development and speech sensorimotor integration (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009). Premotor regions have been particularly shown to be involved in speech perception under adverse listening conditions, when bottom-up information alone is not sufficient (Adank and Devlin, 2010; Alho et al., 2012, 2014; D'Ausilio et al., 2009; D'Ausilio et al., 2012; Du et al., 2014; Hervais-Adelman et al., 2012; Osnes et al., 2011; Sato et al., 2009). An fMRI (functional Magnetic Resonance Imaging) study for instance showed activity in the left premotor cortex during adaptation to time-compressed sentences in adults (Adank and Devlin, 2010). Using MEG, Alho and collaborators (2014) also reported increased oscillatory synchrony between left premotor and left auditory and temporo-parietal cortices in the beta frequency band when participants listened to noisy syllables. This synchrony furthermore positively correlated with syllable identification. To account for these results, it has been proposed that the motor system would generate internal forward models to facilitate phoneme categorization by auditory regions when communication conditions are degraded (Alho et al., 2014; Iacoboni, 2008; Rauschecker and Scott, 2009; Wilson and Iacoboni, 2006).

To the best of our knowledge, no study to date has demonstrated brain-to-speech alignment in premotor regions for naturally produced sentences, at least in children. Note that the study by Peelle and collaborators (2013) in adults showed significant phase-locking to the amplitude envelope of unintelligible (1-band noise-vocoded) speech in the bilateral motor cortex. However, no such coupling was described for intelligible (16- and 4-bands noise-vocoded) sentences in this experiment. Other studies examining oscillations in the adult brain have reported auditory-motor coupling or top-down control of auditory cortex from left motor cortex during speech processing. Assaneo and Poeppel (2017) for instance very recently showed enhanced phase-locking between left motor and right auditory regions during processing of syllable trains at a rate of 4.5 Hz. In the work by Park and colleagues (2015), activity in left frontal and motor areas modulated the phase of delta and theta oscillations in

left auditory cortex during continuous speech perception and additionally enhanced brain-to-speech coupling in these regions. In a study using audio-visual speech, the same authors (Park et al., 2016) showed entrainment to speaking lips in left motor cortex which predicted comprehension accuracy. Although these works demonstrate an involvement of the left motor system in oscillatory coupling to speech in auditory cortex, they do not directly report any motor entrainment to speech amplitude envelope. Our findings in typically developing children reveal for the first time that left premotor and primary motor regions are not only activated when speech perception is more challenging (by naturally accelerating the speaking rate), but that theta-band oscillatory activity in these regions entrains to the slow modulations in speech envelope. Such alignment may reflect articulatory mappings of heard sentences, most likely at the syllabic level, which would help speech parsing by auditory regions under more difficult conditions such as in the case of fast rate speech. Interestingly, studies in adults have demonstrated, at rest, alignment between intrinsic auditory temporal and articulatory (pre)motor oscillatory rhythms in the theta (i.e. dominant syllabic rate) and gamma frequency ranges (Giraud et al., 2007; Morillon et al., 2010). This was observed more strongly in the left than in the right hemisphere for the theta band (Morillon et al., 2010). Hence, endogenous theta and gamma neuronal oscillations correlated with metabolic activity in brain regions involved in speech perception and production. This has been taken as evidence that motor areas naturally express oscillatory rhythms that characterize cyclic jaw movements (~4 Hz) underlying syllable production, in line with the Frame/Content theory of speech (MacNeilage and Davis, 2001). Such inherent auditory-motor oscillatory tuning in speech-relevant frequency bands has, to our knowledge, not been reported at rest in children. However, given the strong perception-action linkage during speech development (Kuhl, 2010; Westermann and Miranda, 2004), we can assume that it may be especially important for children during more demanding speech perception. When speech is uttered at a fast rate, auditory tracking of speech rhythmic envelope could therefore be complemented by entrainment in (pre)motor regions that store articulatory representations (see Strauß and Schwartz, 2017 for similar arguments). This is also in line with studies showing a role of the motor system in detecting auditory temporal regularities (Arnal et al., 2015; Arnal and Giraud, 2012; Fujioka et al., 2009; Morillon and Schroeder, 2015; Morillon et al., 2014).

4.2. Frequency-specific cortical entrainment to syllabic rate

Besides coupling in the theta range (4-7 Hz) classically examined in previous works, the present findings also demonstrate that cortical oscillations track normal rate speech envelope in a frequency-specific manner, that is, in a range that matches sentences' syllabic rate. Significant cortico-acoustic coherence for normal rate sentences indeed increased in the 5.6-7.6 Hz frequency band in the left anterior temporal cortex (BA 21/22; Fig 3). No such coupling was seen for fast rate speech in this range nor in the frequency band matching sentences' mean fast rate (8-10 Hz). To the best of our knowledge, this is the first study in typically-developing children showing frequency-specific cortical entrainment at the syllabic rate using naturalistic stimuli.

Left-lateralized cortico-acoustic coherence for normal syllabic rate may seem at odds with the AST model (Poeppel, 2003). However, previous works reported spontaneous oscillatory responses in the theta range also in the left anterior temporal cortex (Giraud et al., 2007; Morillon et al., 2010; Morillon et al., 2012). In a study with intracranial recordings in an epileptic patient, Morillon and colleagues (2012) indeed showed, at rest and during processing of isolated syllables, theta oscillatory activity (peaking at 7.5 Hz) in left primary auditory cortex that persisted in left anterior superior temporal gyrus (BA22), when moving up the processing hierarchy. Co-modulation between the phase of this theta activity and gamma band power was furthermore observed in the same left anterior region. In line with Giraud et al. (2007), the authors suggested that such left-dominant theta activity could play an integrative rather than a parsing role, where gamma-discretized phonemic units in left posterior auditory regions would be integrated into longer syllable-length units parsed at the theta rate more anteriorly (Giraud and Poeppel, 2012; Peelle and Davis, 2012). In this view, left-lateralized theta-gamma coupling would allow the transition from acoustic-feature tracking to more abstract phonological processing (Giraud and Poeppel, 2012). In agreement with this, the results by Gross et al. (2013) showed stronger coupling between theta phase and gamma amplitude in the left temporal cortex. Peelle et al. (2013) also reported that coherence between theta-band cortical oscillations and speech envelope was enhanced in the left anterior middle temporal gyrus for more intelligible speech, suggesting that linguistic content, and not only acoustic features, affects the neural lateralization for speech processing. In our study in children, cortico-acoustic coherence in left BA22 in the frequency band centred on normal syllabic rate (5.6-7.6 Hz) could thus reflect higher order integrative processes at work during

connected speech comprehension. Future investigations are however needed to fully support this interpretation.

Note that methodological considerations such as the use of many trial-unique, attended sentences with rich linguistic content in our study, which was also the case in the work by Peelle et al. (2013) in adults, may also influence the pattern of hemispheric lateralization as compared with other studies using unattended speech or repeated stimuli which reported right brain-to-speech coupling in the theta range (Abrams et al., 2008; Luo and Poeppel, 2007).

When children listened to fast rate sentences, no significant cortico-acoustic coherence was observed in the 8-10 Hz range corresponding to the mean syllabic rate of sentences. This may be related to decreased sentence intelligibility with increasing speech rate. Only a very limited number of studies have investigated the effect of speech rate on brain-to-speech alignment. As mentioned in the introduction, Ahissar and colleagues (2001) reported that coupling between oscillatory activity and speech envelope occurred for time-compressed sentences only when they were intelligible; when the compression ratio increased (leading to modulations around 9 Hz and 15 Hz), synchronization failed. In a recent EEG work using compressed sentences as well, Pefkou et al. (2017) however showed cortical tracking of speech that followed the syllabic rate from 4 up to 14 Hz, namely beyond the upper limit of the classical theta range. Cortico-acoustic coupling was indeed stronger at 3-5 Hz for uncompressed sentences, at 7-9 Hz for intelligible sentences compressed by a factor of 2 and at 10-14 Hz for unintelligible sentences compressed by 3. Similarly, a frequency shift (from 5.5-7.5 Hz to 8-10 Hz) in cortico-acoustic coherence was observed in auditory regions of the adult brain during processing of sentences naturally produced at a normal rate (6.5 syllables/s) or at a fast rate (9 syllables/s ; Hincapié et al., in preparation). In the present study in children, no evidence for brain tracking of speech at the fast syllabic rate (8-10 Hz) was provided. It could therefore be that children's brains were not able to adjust to the increased syllabic rate the way adult brains do. This could either be explained by the incomplete brain maturational processes, or alternatively by the fact that the fast sentences were too fast for the children, leading to lower intelligibility. Note that we did not measure sentence intelligibility here, which may be seen as a limitation of our work. Due to already long MEG recording, it was indeed difficult to assess whether children understood each stimulus during the experiment. However, a behavioral pilot testing in 11 typically-developing children of the same age as our participants showed that fast rate sentences led to scores of 74% in a sentence judgment task (vs 84% for normal rate sentences). We can therefore rule out the possibility that all of our fast rate

sentences were not understandable, although intelligibility certainly decreased as compared to the normal rate condition. Such a decrease in intelligibility may account for, or result from, the lack of cortical entrainment to speech envelope in the syllabic rate frequency range. Greater difficulty to understand naturally accelerated sentences was on the other hand reflected by the previously-discussed alignment to speech in left (pre)motor regions of the dorsal stream. Replicating this study by providing intelligibility scores would nevertheless be valuable to examine in children whether and how intelligibility correlates with cortico-acoustic coherence in the theta band as well in frequency ranges centred on sentences' syllabic rate.

4.3. Alpha-band desynchronization during sentence processing

While we observed no power modulations in the frequency range matching the mean normal syllabic rate (5.6-7.6 Hz), power decreased in the 8-10 Hz band for both speech rate conditions in left middle and superior temporal cortex (Fig 4). This is in agreement with the growing body of evidence showing desynchronization of the auditory alpha rhythm (also called "tau rhythm"; Lehtelä et al., 1997) in the temporal cortex (mainly on the left) during sound and speech processing (Becker et al., 2013; Dimitrijevic et al., 2017; Krause et al., 1994; Krause et al., 1997; Meyer et al., 2013; Obleser and Weisz, 2012). Such alpha suppression in task-relevant regions is thought to reflect enhanced mental operations on the speech signal and thus more active cognitive processing (Obleser and Weisz, 2012; Strauß et al., 2014; Weisz et al., 2011). In other words, through a release of functional inhibition (Jensen and Mazaheri, 2010), alpha oscillatory activity would decrease in brain areas required to actively process information. Evidence for alpha desynchronization during auditory processing in children is still scarce, and most studies so far used non-verbal stimuli or single words (Fujioka and Ross, 2008; Krause and al., 2001; Krause et al., 2007). Our study reveals that lower alpha oscillations (8-10 Hz) in the left temporal cortex of 8-13 year old children desynchronize during processing of sentences naturally produced both at a normal and fast rate, as they do in adults.

5. Conclusion

The present study provides novel contributions to auditory neuroscience, as it is among the first to investigate cortical entrainment to naturally produced speech at various rates in typically developing children. Our findings demonstrate that theta-band cortical oscillations in

the children's brain track speech temporal structure both when sentences are naturally produced at a normal and a fast rate. Whereas brain-to-speech coupling is observed in right auditory and fronto-parietal regions for normal rate speech, it occurs in left premotor and motor regions of the dorsal stream when speech is uttered faster. We furthermore provide evidence that oscillatory activity in the left anterior temporal cortex of children aligns to the specific syllabic rate of the stimuli when it falls within the limits of the theta range. More studies are needed to directly compare the results reported here in children to those obtained in adults with a comparable paradigm. Furthermore, it will also be of high interest to explore putative alterations of the reported oscillatory entrainment phenomena, in children with speech processing or production deficits. Such studies might help identify in children, neural features (or markers) that could be targeted with neuro-cognitive training to reduce language impairments.

Acknowledgments

We would like to thank parents and their children for participating in this study. We also thank D. Bayle for his invaluable help for analyses with the Fieldtrip toolbox, E. Ferragne and N. Bedoin for helping in the construction of the stimuli and D. Gouy for recording the sentences. This study was conducted as part of the ODYSSEE project (PI: V.B; n° 11 JSH2 005 1) funded by the French National Research Agency (ANR). H.G is funded by a doctoral fellowship from the LabEx ASLAN (ANR-10-LABX-0081) of Université de Lyon within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the ANR, and by the Erasmus Mundus, Student Exchange Network in Auditory Cognitive Neuroscience program. K.J acknowledges funding from the Canada Research Chairs program and NSERC Discovery Grant (RGPIN-2015-04854).

References

- Abrams, D., Nicol, T., Zecker, S., & Kraus, N., 2008. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(15), 3958–65.
- Abrams, D., Nicol, T., Zecker, S., & Kraus, N., 2009. Abnormal cortical processing of the syllable rate of speech in poor readers. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(24), 7686–93.
- Adank, P., & Devlin, J. T., 2010. On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage*, 49(1), 1124–1132.

- Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M., 2001. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 13367–13372.
- Alho, J., Lin, F.-H., Sato, M., Tiitinen, H., Sams, M., & Jääskeläinen, I. P., 2014. Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Frontiers in Psychology*, 5.
- Alho, J., Sato, M., Sams, M., Schwartz, J.-L., Tiitinen, H., & Jääskeläinen, I. P., 2012. Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. *NeuroImage*, 60(4), 1937–46.
- Arnal, L. H., Doelling, K. B., & Poeppel, D., 2015. Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.
- Arnal, L. H., & Giraud, A.-L., 2012. Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.
- Assaneo, M. F., & Poeppel, D., 2017. Listening to speech induces coupling between auditory and motor cortices in an unexpectedly rate-restricted manner. *bioRxiv Beta*, preprint.
- Becker, R., Pefkou, M., Michel, C. M., & Hervais-Adelman, A. G., 2013. Left temporal alpha-band activity reflects single word intelligibility. *Frontiers in Systems Neuroscience*, 7, 1–12.
- Boemio, A., Fromm, S., Braun, A., & Poeppel, D., 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8(3), 389–395.
- Boersma, P., & Weenink, D., 2012. Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org/>.
- Cogan, G. B., & Poeppel, D., 2011. A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *J Neurophysiol*, 106, 554–563.
- D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L., 2012. The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, 48, 882–887. 7
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L., 2009. The motor somatotopy of speech perception. *Current Biology*, 19, 381–385.
- Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R., 2017. Cortical Alpha Oscillations Predict Speech Intelligibility. *Frontiers in Human Neuroscience*, 11, 1–10.
- Ding, N., & Simon, J. Z., 2013. Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *Journal of Neuroscience*, 33(13), 5728–5735.
- Ding, N., & Simon, J. Z., 2014. Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience*, 8(May), 311.
- Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D., 2014. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, 85, 761–8.
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 7126–31.
- Ferragne, E., Flavier, S., & Fressard, C., 2012. ROCme! Recording of Oral Corpora Made Easy: Logiciel d'enregistrement de corpus. Lyon, France: DDL.

- Fonov, V., Evans, A. C., Botteron, K., Almli, R. C., McKinstry, R. C., & Collins, L. D., 2011. Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1), 313–327.
- Friederici, A. D., 2011. The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4), 1357–92.
- Fujioka, T., & Ross, B., 2008. Auditory processing indexed by stimulus-induced alpha desynchronization in children. *International Journal of Psychophysiology*, 68(2), 130–140.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B., 2009. Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, 1169, 89–92.
- Ghitza, O., 2011. Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, 2(130).
- 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), 113–26.
- Giedd, J. N., 2004. Structural Magnetic Resonance Imaging of the Adolescent Brain. *New York Academy of Sciences*, 1021, 77–85.
- 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.
- Giraud, A.-L., & Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.
- Goswami, U., 2011. A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10.
- Goswami, U., 2015. Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews. Neuroscience*, 16, 43–54.
- Goswami, U., & Leong, V., 2013. Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, 4(1), 67–92.
- 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).
- Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R., 2001. Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2), 694–699.
- Hervais-Adelman, A. G., Carlyon, R. P., Johnsrude, I. S., & Davis, M. H., 2012. Brain regions recruited for the effortful comprehension of noise-vocoded words. *Language and Cognitive Processes*, 27(7–8), 1145–1166.
- Hickok, G., & Poeppel, D., 2007. The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Howard, M. F., & Poeppel, D., 2010. Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500–11.
- Howard, M. F., & Poeppel, D., 2012. The neuromagnetic response to spoken sentences: co-

- modulation of theta band amplitude and phase. *NeuroImage*, 60(4), 2118–27.
- Iacoboni, M., 2008. The role of premotor cortex in speech perception: evidence from fMRI and rTMS. *Journal of Physiology*, 102, 31–34.
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M., 2010. BALE: Batterie Analytique du Langage Ecrit. Grenoble.
- Jensen, O., & Mazaheri, A., 2010. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience*, 4(186).
- Jusczyk, P. W., 1997. The discovery of spoken language. MIT Press, Cambridge.
- 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.
- Khomsy, A., 2001. ELO: évaluation du langage oral. Paris.
- Krause, C. M., Lang, H. A., Laine, M., Helle, S. I., Kuusisto, M. J., & Pörn, B., 1994. Event-Related Desynchronization evoked by auditory stimuli. *Brain Topography*, 7(2), 107–112.
- Krause, C. M., Pörn, B., Lang, A. H., & Laine, M., 1997. Relative alpha desynchronization and synchronization during speech perception. *Cognitive Brain Research*, 5(4), 295–299.
- Krause, C. M., Sillanmäki, L., Häggqvist, A., & Heino, R., 2001. Test-retest consistency of the event-related desynchronization/event-related synchronization of the 4-6, 6-8, 8-10 and 10-12 Hz frequency bands during a memory task. *Clinical Neurophysiology*, 112(5), 750–757.
- Krause, C. M., Pesonen, M., Björnberg, C. H., Hämäläinen, H., 2007. Effects of pulsed and continuous wave 902 MHz mobile phone exposure on brain oscillatory activity during cognitive processing. *Bioelectromagnetics*, 28(4), 296–308.
- Kuhl, P. K., 2010. Brain Mechanisms in Early Language Acquisition. *Neuron*, 67(5), 713–727.
- Lehtelä, L., Salmelin, R., & Hari, R., 1997. Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neuroscience Letters*, 222, 111–114.
- Luo, H., Husain, F. T., Horwitz, B., & Poeppel, D., 2005. Discrimination and categorization of speech and non-speech sounds in an MEG delayed-match-to-sample study. *NeuroImage*, 28, 59–71.
- Luo, H., & Poeppel, D., 2007. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001–10.
- 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(170).
- MacNeilage, P. F., & Davis, B. L., 2001. Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology*, 11, 696–700.
- Maris, E., & Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164(1), 177–190.
- Mayer, M., 1969. *Frog, where are you?* Dial Press, New York.
- Mehler, J., Jusczyk, P. W., Lamsertz, G., Halsted, N., Bertocini, J., & Amiel-Tison, C.,

1988. A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Meyer, L., Obleser, J., & Friederici, A. D., 2013. Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, 49, 711–721.
- Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M., 2016. Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, 37(8), 2767–83.
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., & Kleinschmidt, A., 2010. Neurophysiological origin of human brain asymmetry for speech and language. *PNAS*, 107(43), 18688–93.
- 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 intra-cortical study. *Frontiers in Psychology*, 3, 248.
- Morillon, B., & Schroeder, C. E., 2015. Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annual New York Academy of Sciences*, 1337(1), 26–31.
- Morillon, B., Schroeder, C. E., & Wyart, V., 2014. Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5, 1–9.
- Nazzi, T., & Ramus, F., 2003. Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233–243.
- Nourski, K. V., Reale, R. a, Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., ... Brugge, J. F., 2009. Temporal envelope of time-compressed speech represented in the human auditory cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(49), 15564–74.
- Obleser, J., & Weisz, N., 2012. Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex*, 22(11), 2466–77.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience - Special Issue on Academic Software Applications for Electromagnetic Brain Mapping Using MEG and EEG*, 2011(1).
- Osnes, B., Hugdahl, K., & Specht, K., 2011. NeuroImage Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *NeuroImage*, NeuroImage(54), 2437–2445.
- Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J., 2015. Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, 25(12), 1649–1653.
- Park, H., Kayser, C., Thut, G., & Gross, J., 2016. Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife*, 5, 1-17.
- Paus, T., Zijdenbos, A., Worsley, K., Collins, L. D., Blumenthal, J., Giedd, J. N., ... Evans, A. C., 1999. Structural Maturation of Neural Pathways in Children and Adolescents : In Vivo Study. *Science*, 283, 1908–1911.
- Peelle, J. E., & Davis, M. H., 2012. Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, 3.
- Peelle, J. E., Gross, J., & Davis, M. H., 2013. Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, 23(6), 1378–1387.

- Pefkou, M., Arnal, L. H., Fontolan, L., & Giraud, A.-L., 2017. Theta- and beta-band neural activity reflect independent syllable tracking and intelligibility of time-compressed speech. *The Journal of Neuroscience*, 37(33), 7930–7938.
- Pellegrino, F., Coupé, C., & Marisco, E., 2011. A Cross-Language Perspective on Speech Information Rate. *Language*, 87(3), 539–558.
- Poeppel, D., 2003. The analysis of speech in different temporal integration windows : cerebral lateralization as asymmetric sampling in time. *Speech Communication*, 41, 245–255.
- 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.
- Power, A. J., Mead, N., Barnes, L., & Goswami, U., 2012. Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Frontiers in Psychology*, 3.
- Rauschecker, J. P., & Scott, S. K., 2009. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–24.
- Raven, J., Raven, J., & Court, J., 1998. *Coloured progressive matrices*. Oxford Psy, Oxford.
- Sato, M., Tremblay, P., & Gracco, V. L., 2009. A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, 111(1), 1–7.
- Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.-S., ... Weiller, C., 2008. Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105(46), 18035–18040.
- Shapiro, R., & Janzen, A. H., 1960. *The normal skull: a Roentgen study*. Hoeber, London.
- Strauß, A., & Schwartz, J.-L., 2017. The syllable in the light of motor skills and neural oscillations. *Language, Cognition and Neuroscience*, 32(5), 562–569.
- Strauß, A., Wöstmann, M., & Obleser, J., 2014. Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, 8(350), 1-7.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., ... Wartenburger, I., 2009. Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. *Journal of Neuroscience*, 29(47), 14726–14733.
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., & Wartenburger, I., 2011. Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near-infrared spectroscopy and EEG study. *Frontiers in Psychology*, 2.
- 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(19737), 1-6.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15, 273–89.
- Vander Ghinst, M., Bourguignon, M., Op de Beeck, M., Wens, V., Marty, B., Hassid, S., ... de Tiège, X., 2016. Left Superior Temporal Gyrus Is Coupled to Attended Speech in a Cocktail-Party Auditory Scene. *The Journal of Neuroscience*, 36(5), 1596–1606.
- Vanvooren, S., Poelmans, H., Hofmann, M., Ghesquière, P., & Wouters, J., 2014. Hemispheric asymmetry in auditory processing of speech envelope modulations in

- prereading children. *The Journal of Neuroscience*, 34(4), 1523–1529.
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J., 2011. Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in Psychology*, 2(73), 1-15.
- Westermann, G., & Miranda, E. R., 2004. A new model of sensorimotor coupling in the development of speech. *Brain and Language*, 89(2), 393–400.
- Wilsch, A., Henry, M. J., Herrmann, B., Maess, B., & Obleser, J., 2015. Alpha Oscillatory Dynamics Index Temporal Expectation Benefits in Working Memory. *Cerebral Cortex*, 25(7), 1938–1946.
- Wilson, S. M., & Iacoboni, M., 2006. Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33, 316–325.
- Zion Golumbic, E. M., Cogan, G. B., Schroeder, C. E., & Poeppel, D., 2013. Visual Input Enhances Selective Speech Envelope Tracking in Auditory Cortex at a “Cocktail Party.” *Journal of Neuroscience*, 33(4), 1417–1426.
- Zion Golumbic, E. M., 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), 151–61.
- Zoefel, B., & Vanrullen, R., 2015. Selective Perceptual Phase Entrainment to Speech Rhythm in the Absence of Spectral Energy Fluctuations. *Journal of Neuroscience*, 35(5), 1954–1964.

2. Synthèse des résultats de l'étude 1

Les résultats obtenus chez des enfants francophones âgés entre 8 et 13 ans présentés dans cet article révèlent une synchronisation (*entrainment*) des oscillations thêta (4-7 Hz) sur le rythme du signal de parole produit naturellement, que le débit syllabique soit normal ou accéléré. Une augmentation de la cohérence dans la bande de fréquence thêta est en effet observée au sein des régions temporales et fronto-pariétales droites lors de la perception de parole à débit normal. Ces résultats sont en accord avec les modèles de résolution multi-temporelle de la parole suggérant un rôle préférentiel de l'hémisphère droit pour le traitement des basses fréquences dans le signal de parole (Ghitza, 2011; Giraud & Poeppel, 2012; Poeppel, 2003). Lors de l'écoute de parole rapide, un alignement oscillatoire thêta sur l'enveloppe d'amplitude du signal de parole est également observé, néanmoins dans des régions corticales différentes. L'augmentation de cohérence est ainsi mise en évidence au sein de la voie dorsale (régions prémotrices et motrices gauches et aire de Broca) de traitement de la parole. D'après la littérature, cette voie est particulièrement recrutée lorsque le signal de parole est plus difficile à traiter, ce qui est le cas ici avec la parole rapide (voir Adank & Devlin, 2010 pour des résultats en IRMf chez l'adulte avec de la parole compressée temporellement). Nos résultats s'accordent également avec les travaux ayant démontré la présence d'un rythme oscillatoire endogène thêta dans les régions motrices (Giraud et al., 2007; Morillon et al., 2012). Ainsi, chez notre groupe d'enfants neurotypiques, l'alignement cerveau/parole ne se produit pas dans les régions auditives lorsque la parole est produite à un débit rapide (comme c'est le cas à débit normal), mais dans les régions de la voie dorsale, corroborant l'hypothèse que ces régions jouent un rôle de renfort dans la perception de la parole.

Lorsqu'on s'intéresse au rythme syllabique précis de nos stimuli de parole, on observe une synchronisation entre les oscillations corticales dans la région temporelle antérieure gauche dans la bande 5.6-7.6 Hz et l'enveloppe temporelle de la parole à débit normal. Si cette latéralisation à gauche semble contradictoire avec le modèle AST de Poeppel (2003), elle rappelle cependant les résultats d'études ayant rapporté une cohérence cortico-acoustique dans une région très proche pour de la parole intelligible (*e.g.*, Peelle et al., 2013). Elle s'accorde également avec les études de repos montrant la présence d'oscillations thêta également dans l'hémisphère gauche, au niveau des régions responsables de traitements linguistiques plus avancés (Giraud et al., 2007; Morillon et al., 2012). L'alignement

thêta/parole dans le cortex temporal antérieur gauche pourrait alors refléter un processus intégratif de haut niveau (Giraud & Poeppel, 2012; Peelle et al., 2013), qui serait réalisé dans la bande de fréquence englobant le rythme syllabique de nos stimuli. Concernant la parole rapide, aucune augmentation significative de cohérence n'est observée dans la bande de fréquence correspondante (8-10 Hz), suggérant une difficulté particulière de traitement de ce type de parole chez les enfants.

Enfin, aucune modulation significative de puissance n'est observée dans les bandes de fréquence thêta et 5.6-7.6 Hz, suggérant que l'augmentation de cohérence mise en évidence précédemment soit liée à un réel entraînement des oscillations corticales dans ces bandes de fréquence et non une activation évoquée suite à la présentation des phrases. Toutefois, dans la bande de fréquence 8-10 Hz, correspondant notamment à une partie de la bande alpha, une désynchronisation émerge au niveau des régions auditives gauches (et bilatérales en non corrigé) lors de l'écoute de la parole à débit normal et rapide. Ce type de désynchronisation au niveau du cortex auditif correspond probablement au rythme tau (Lehtelä et al., 1997), qui reflète un mécanisme de levée d'inhibition fonctionnelle permettant un traitement auditif actif dans les régions dans lesquelles la diminution de puissance alpha est observée (Jensen & Mazaheri, 2010; Obleser & Weisz, 2012). Ce rythme n'a à notre connaissance jamais été décrit chez les enfants pour de la parole naturelle (phrases).

Enfin, bien que ces résultats ne soient pas présentés dans l'article présenté ci-dessus, nous avons également commencé à analyser les modulations de puissance dans les bandes beta (15-30 Hz) et gamma basse (25-45 Hz). Les premières données révèlent une désynchronisation significative au sein des régions temporelles antérieures, bilatérales en beta (cf. Annexe 2), suggérant un rôle de ce rythme oscillatoire dans le traitement de la parole chez l'enfant. Cependant, aucune modulation significative de puissance gamma n'est observée chez ce groupe d'enfants. Des analyses supplémentaires avec un groupe de participants plus important sont nécessaires afin d'analyser les modulations de puissance beta et gamma et ainsi de pouvoir réaliser des analyses de connectivité cortico-corticale dans les différentes bandes de fréquence impliquées dans la perception de la parole (couplage phase-phase, phase-amplitude etc.).

En conclusion, les résultats de cette étude chez l'enfant neurotypique révèlent pour la première fois à notre connaissance un alignement des oscillations thêta dans les régions auditives droites, mais aussi dans les régions (pré)motrices gauches, lors de la perception de parole naturellement produite à un débit normal et rapide respectivement. Ils montrent

également une synchronisation spécifique au rythme syllabique de la parole dans les régions temporales antérieures gauches, au moins lorsque la parole perçue est produite à un débit normal.

VIII. Perception de la parole rapide chez des enfants présentant une dysphasie expressive : études comportementale et en MEG.

Ce deuxième chapitre expérimental se compose de deux études réalisées chez des enfants présentant une dysphasie expressive appariés à des enfants neurotypiques. La première étude est une étude comportementale et l'article relatif à cette étude est en cours de révision dans la revue *PloS One*. Les résultats ont par ailleurs été présentés lors du séminaire international LSCD (*Workshop on Late Stages in Speech and Communication Development*) 2014. La seconde étude est une étude en MEG et l'article présenté ici est en cours de préparation. Les résultats ont été présentés lors du congrès international IASCL (*International Association for the Study of Child Language*) 2017.

1. Présentation de l'étude 2 : Étude comportementale

Le traitement des indices rythmiques de la parole est primordial pour un traitement efficace de l'information verbale. Le système auditif est ainsi capable d'extraire l'information rythmique contenue dans les modulations lentes d'amplitude au sein de l'enveloppe temporelle (qui véhiculent l'information syllabique et donc le débit de parole), afin de segmenter et de décoder le message linguistique. Un certain nombre d'études révèlent des difficultés particulières de traitement des indices rythmiques de la parole chez les enfants présentant une dysphasie (Cumming, Wilson, & Goswami, 2015a; Cumming, Wilson, Leong, et al., 2015b). Cette difficulté serait particulièrement marquée pour les indices rythmiques de courte durée (*rise time* et segments rapides tels que les transitions formantiques), aboutissant probablement à la formation de représentations phonologiques imparfaites qui affecteraient alors les capacités d'expression et/ou de compréhension verbale des enfants. Par ailleurs, comme nous l'avons mentionné précédemment, le rythme de la parole est un indice particulièrement variable dans la vie de tous les jours, les interlocuteurs sont en effet capables de s'exprimer à des débits plus ou moins rapides, modifiant fortement les indices rythmiques du signal acoustique de parole (raccourcissement des durées, phénomènes de réduction, etc). Ces variations inhérentes au signal de parole peuvent d'autant plus entraver le décodage du message linguistique chez les enfants présentant déjà des difficultés à traiter les indices temporels de la parole, comme c'est le cas dans la dysphasie. Aucune étude à notre connaissance ne s'est intéressée aux capacités de perception de la parole produite à un débit rapide chez ces enfants.

Dans cette étude comportementale, nous avons examiné les performances de perception de la parole rapide chez un groupe d'enfants présentant une dysphasie expressive apparié à un groupe contrôle d'enfants neurotypiques, tous deux de langue maternelle française et âgés de 8 à 13 ans. Les enfants avaient pour tâche d'écouter des phrases dont le mot cible final était cohérent ou incohérent avec le contexte de la phrase, et de juger de la cohérence sémantique de ces phrases en appuyant sur deux boutons d'un clavier. Ces phrases étaient par ailleurs présentées dans trois conditions de débit : d'abord de la parole naturellement produite à un débit normal (6.76 syll/s, n = 25), puis de la parole naturellement produite à un débit rapide (9.15 syll/s, n = 25) et enfin de la parole accélérée artificiellement au même débit que le débit naturel rapide par compression temporelle linéaire (n = 25). Nous avons ensuite calculé

l'indice de sensibilité (d' , théorie de détection du signal) à l'incohérence sémantique pour chaque groupe d'enfants.

Notre hypothèse était que, de par leurs difficultés à traiter l'information temporelle en général, les enfants présentant une dysphasie devraient présenter des difficultés spécifiques à « suivre » (*track*) les changements temporels plus rapides dans l'enveloppe d'amplitude de la parole accélérée, conduisant ainsi à un découpage moins efficace de l'information verbale et donc à des performances dans la tâche de jugement sémantique réduites par rapport à leurs pairs. Ceci était attendu quel que soit le type d'accélération (naturelle ou artificielle) puisqu'à la fois la parole naturelle rapide et la parole compressée induisent ces changements rythmiques.

Don't speak too fast! Processing of fast rate speech in children with specific language impairment

Hélène Guiraud^{1*}, Nathalie Bedoin¹, Sonia Krifi-Papoz², Vania Herbillon^{3,4}, Aurélia Caillot-Bascoul⁵, Sibylle Gonzalez-Monge⁶, Véronique Boulenger^{1*}

¹ Laboratoire Dynamique Du Langage, CNRS/Université de Lyon UMR5596, Lyon, France

² Service de Neurologie Pédiatrique, Hôpital Femme Mère Enfant, Bron, France

³ Service Épilepsie, Sommeil et Explorations Fonctionnelles Neuropédiatriques, Hôpital Femme Mère Enfant, Bron, France

⁴ Centre de Recherche en Neurosciences de Lyon, DYCOG, INSERM U1028 / CNRS UMR5292, Bron, France

⁵ Service ORL chirurgie cervico-faciale, Centre Hospitalier Universitaire Gabriel Montpied, Clermont-Ferrand, France

⁶ Centre de Référence Troubles des Apprentissages, Service de Rééducation pédiatrique, Hôpital Femme Mère Enfant, Bron, France

* Corresponding authors

helene.guiraud@univ-lyon2.fr (HG)

veronique.boulenger@cnrs.fr (VB)

Abstract

Background: Perception of speech rhythm requires the auditory system to track temporal envelope fluctuations, which carry syllabic and stress information. Reduced sensitivity to rhythmic acoustic cues has been evidenced in children with Specific Language Impairment (SLI), impeding syllabic parsing and speech decoding. Our study investigated whether these children experience specific difficulties processing fast rate speech as compared with typically developing (TD) children. **Method:** Sixteen French SLI children (8-13 years old) with mainly expressive phonological disorders and with preserved comprehension and 16 age-matched TD children performed a judgment task on sentences produced 1) at normal rate, 2) at fast rate or 3) time-compressed. Sensitivity index (d') to semantically incongruent sentence-final words was measured. **Results:** Overall children with SLI perform significantly worse than TD children. Importantly, as revealed by the significant Group \times Speech Rate interaction, SLI children find it more challenging than TD children to process both naturally or artificially accelerated speech. The two groups do not significantly differ in normal rate speech processing. **Conclusion:** Consistent with rhythm-processing deficits in atypical language development, our results suggest that children with SLI face difficulties adjusting to rapid speech rate. These findings are interpreted in light of temporal sampling and prosodic phrasing frameworks and of oscillatory mechanisms underlying speech perception.

Introduction

Every listener has noticed how speech rate can vary considerably between speakers and contexts and how this can be particularly challenging, at least in the first minutes of a conversation. Speaking at a fast rate inevitably elicits qualitative changes in the speech signal at both temporal and spectral levels. With the increase in speech rate, length of acoustic cues, formant transitions and pauses are shortened [1]. Articulatory gestures are achieved more quickly and less accurately, resulting in reduction phenomena, enhanced coarticulation (i.e., increased gestural overlap) and assimilation, which may even lead to the suppression of whole segments [2]. Moreover, these changes operate nonlinearly, partly because of articulatory restrictions [3]: in English and French for instance, consonants and stressed syllables (in English) are less reduced than vowels and unstressed syllables [4,5].

Such spectro-temporal modifications typically occur in naturally accelerated speech, whereas the spectral and pitch content in time-compressed speech – an artificial reduction of signal duration often used in experimental phonetics – remain intact. Processing of naturally accelerated speech therefore puts high demands on the listener's perceptual system to constantly normalize for a wide range of spectro-temporal parameters, in order to adapt to various speech rates and understand the message. The study by Janse [3] in adults highlighted the greater difficulty to process natural fast speech in comparison to time-compressed speech. Janse also compared two types of time compression: strict linear compression, where all syllables in the signal are compressed to the same degree to match the natural fast rate, and non-linear selective compression, which follows the exact temporal pattern (at the syllable level) of naturally produced fast speech. Results in a phoneme detection task showed a processing advantage for linearly time-compressed speech over both non-linearly time-compressed and natural fast speech. Linear time compression further led speech to be judged more pleasant to listen to. Hence, the increased segmental overlap combined with the changed temporal pattern which characterize natural fast speech make it more difficult to understand than artificially accelerated speech.

Studies focusing on the adaptation process involved in fast speech perception have revealed that young and older adults are able to adapt rather quickly to both types of accelerated speech (natural and artificial) [6-8]. For instance, Dupoux and Green [7] reported that listening to ten sentences compressed at 45% of their original duration was sufficient to improve performance in terms of the number of reported words in a sentence recall task. When sentences were made even faster (compressed at 38% of the original version),

participants were still able to adapt, though after a slightly longer exposure time (15 sentences), therefore reflecting rapid perceptual adjustment even for highly accelerated speech. In a more recent study, Adank and Janse [6] compared adaptation to natural and artificial fast speech in young Dutch adults asked to perform a sentence verification task. Results revealed longer adaptation when listening to natural fast speech (exposure to 18 sentences was necessary) than to time-compressed speech (exposure to six to 12 sentences was sufficient). The same pattern of results, also consistent with the work of Janse [3], was observed in French-speaking typically developing children (aged 8-9 years; [9]), suggesting a qualitative difference between the processing of artificially accelerated speech (i.e., temporal change) and the processing of naturally-produced rapid speech (i.e., spectro-temporal changes). To account for adaptation processes to fast speech, it has been suggested that listeners must learn to associate new acoustic representations with stored phonemic categories. In other words, adaptation would require listeners to recalibrate (i.e., adjust) their phonemic boundaries according to those of the speaker, which could provide them with phonological representations that can be more easily processed by their perceptual system [9,10].

At the cortical level, one way this adjustment to varying speech rates can be achieved may entail oscillatory mechanisms. Studies using electro- and magnetoencephalography (E/MEG) have indeed shown that neuronal oscillations in the theta range (4-7 Hz), mainly in the right auditory cortex, synchronize to the slow modulations (3-5 Hz) dominating in the speech amplitude envelope, which mainly characterize syllabic rate (see [11] for a review). Such “entrainment” was observed for speech that was time-compressed until 50% of the original duration of the signal. When speech was no longer intelligible (at compression rates of 35% and 20%), cortical oscillations could however not align their phase to signal modulations [12]. These findings can be explained in the framework of the Asymmetric Sampling in Time (AST) model [13,14], which states that auditory cortices simultaneously sample the speech signal based on at least two temporal integration windows that correspond to the fundamental units of speech, namely phonemes and syllables. This temporal sampling is furthermore thought to occur asymmetrically. The left auditory cortex, on the basis of its spontaneous oscillatory activity lying around 40 Hz, may preferentially extract information from short integration windows (20-40 ms) and parse the signal into phonemic units (segmental level). In parallel, the right auditory cortex, naturally oscillating at approximately 4 Hz, would extract information from larger windows (150-250 ms) to process slower acoustic fluctuations in the amplitude envelope such as syllabic structure and prosodic cues (suprasegmental level).

Segmental and suprasegmental information would then be integrated to access phonological representations.

The parsing of the continuous speech signal into segmental and suprasegmental features is of key importance for successful comprehension. Specifically, information about speech rhythm and linguistic stress carried by the temporal envelope is critical for intelligibility [11,15]. When speech is accelerated, the listener's auditory system has therefore to deal with faster temporal modulations in the continuous stream: it has to entrain to the new input rhythm to extract shortened units, hence enabling efficient signal decoding. Accurate perception of rhythmic patterns conveyed by temporal changes in the amplitude envelope is also crucial for language acquisition, especially for the development of well-specified phonological representations [16]. In agreement with this, rhythm-processing deficits have been described in children with developmental dyslexia – characterized by persisting major impairment in reading and spelling abilities [16] – and in children with specific language impairment (SLI; note that SLI is now classified under the label “Language Disorder” as one form of neurodevelopmental communication disorder in the Fifth edition of the Diagnostic and Statistical Manual of Mental Disorders - DSM-5 [17], but see Bishop [18] for a discussion).

SLI, comprising 5 to 7% of the kindergarten and primary school population [19,20], is a specific, severe and long-lasting developmental language disorder, affecting the production and/or comprehension of spoken language, despite normal hearing development and non-verbal intelligence, and without any other cognitive or neurological deficits [21-23]. As children with SLI have been shown to present various linguistic symptoms, several classifications have been proposed to account for this heterogeneity. In France, clinicians mostly refer to the classification proposed by Rapin and Allen [24] who proposed three large sub-types of developmental disorders, including six profiles of language deficits that affect phonological, lexical, morpho-syntactic or pragmatic abilities. Mixed expressive-receptive disorders (including phonological-syntactic syndrome and verbal auditory agnosia) are distinguished from expressive disorders (verbal dyspraxia and speech programming deficit disorder) and from higher-order processing disorders (lexical deficit disorder and semantic-pragmatic disorder). Note that in this taxonomy, disorders are categorized according to the affected modality; however they may rather be seen as points on a continuum of language impairment than as discrete entities with clear boundaries between sub-types (see [25]).

Children with SLI have been shown to be impaired in non-verbal auditory discrimination tasks involving two important suprasegmental cues for speech rhythm and stress patterns,

namely rise time of the amplitude envelope (i.e., rate of change in the envelope, corresponding to onsets of successive syllables, [26]) and signal duration [27]. Moreover, their rise time and duration processing abilities accounted for unique variance in several language and phonological tasks (see also [28]). These children however did not fail in non-verbal tasks that do not involve rhythm processing (e.g., intensity discrimination) as compared with chronological age-matched TD children. Note that reduced sensitivity to rise time variations has also been reported in dyslexic children [29], with or without associated oral language disorders [30]. Finally, a recent study in children with SLI [31] confirmed poor skills in discriminating rise time and duration, and these difficulties predicted performance in perceiving lexical stress patterns. In the so-called “temporal sampling framework” (proposed for dyslexia but encompassing SLI), Goswami [16] suggested that poor perception of the rhythmic structure in the speech envelope (i.e. slow modulations in speech) may compromise syllabic and word segmentation skills and consequently the formation of stable phonological representations. Cumming and collaborators [32] extended this notion to SLI with the “prosodic phrasing” hypothesis, emphasizing that perceptual difficulties for rhythmic cues such as rise time may also affect the extraction of prosodic patterns and lead to impaired comprehension as well as to morpho-syntactic impairments during language production.

Interestingly, rhythm-processing deficits in developmental language disorders have been shown to also occur at the motor level [29,33,34]. Children with SLI and dyslexia indeed exhibit difficulties tapping along in time to the slow tempo of a metronome (1.5 and 2 Hz, matching speech stress pattern). Again, children’s motor rhythmic performance can account for their phonological and literacy skills ([33] for SLI, [35] for dyslexia; see also [36] for evidence in pre-schoolers and [37] in adolescents). Recent works similarly demonstrated impaired entrainment to rhythm in nursery rhyme sentences or in music in dyslexic adults and SLI children [34,38]. Timing deficits have also been described in children and adults with verbal dyspraxia both for the production of manual rhythms (e.g., clapped rhythm imitation and tapping) and of spoken rhythms (e.g., vocal rhythm reproduction) [39,40]. Beat synchronization implies fine auditory-motor synchrony in which precise timing information (i.e. sound’s onset) extracted in the auditory system is integrated with motor regions [41]. This also requires monitoring and appropriately adjusting the motor commands to match the auditory input. As previously mentioned, perception and production of rhythmic and temporal patterns are also crucial for language acquisition [16]. Auditory-motor coupling indeed plays a key role in the development of phonological skills as it requires the child to synchronize the timing of auditory perception with the motor networks involved in the representation and

production of speech sounds [41]. Given these similarities, it has been suggested that a neural mechanism underlying the perception and expression of rhythm and timing, and involving auditory and motor brain regions [42,43], may be impaired in SLI and affect both language and motor development [33]. This is in agreement with the strong comorbidity observed between specific language and motor disorders in SLI [44], particularly when speech production is affected [45]. Most interestingly, structural and functional abnormalities have been described in SLI (verbal dyspraxia) in motor cortical and sub-cortical regions such as the cerebellum, the caudate nucleus and the supplementary motor area [46-48], regions that are not only involved in motor planning/sequencing and speech production but also in timing processing, both at the motor [43,49] and the linguistic levels [50-52].

Accumulating evidence therefore speaks in favour of impaired processing of acoustic cues to speech rhythm (i.e., stress and syllable prominence) and impaired rhythmic expression in children with SLI. The issue of how these children deal with natural speech processing when the rhythmic structure of the ongoing signal is altered however remains to be tackled. As previously mentioned, amongst the multiple variations that speech can take in daily life, speech rate changes are ubiquitous both between and within speakers and between communication conditions. Accordingly, given the deficits to process rhythmic information in SLI children, it appears of crucial importance to examine the performance of these children to understand speech at varying syllabic rates. Increasing the rate of speech provides the listener with higher frequency modulations in the amplitude envelope, shorter segment durations and modified rise times due to the increased rate of change in the envelope (i.e. more frequent syllable “beats”). If SLI children show poor perception of rhythmic cues in speech amplitude envelope, accelerating the syllabic rate (and thus affecting speech rhythmic temporal structure) may represent a great challenge for these children and specifically affect their performance to understand speech. The present study set out to investigate the processing of fast rate speech in children with SLI showing difficulties mainly at the expressive level with preserved comprehension skills (phonological-syntactic syndrome and verbal dyspraxia) as compared with chronological age-matched children with typical development. The phonological-syntactic syndrome is the most prevalent form of SLI and is known to mainly affect verbal expression due a phonological programming disorder. The child is often hardly intelligible, shows altered phonology and tends to produce short and ungrammatical sentences [24,53,54]. Verbal dyspraxia affects the programming of articulatory movements despite intact neuromuscular system; impoverished phonological representations and phonological

programming can also be part of the disorder [55]. Primary features of verbal dyspraxia include a restricted phonemic repertoire, inconsistency in articulation errors, difficulties in sequencing speech movements, prosodic abnormalities (affecting rate, rhythm, stress and intonation) and delayed expressive language [24,56]. In our experiment, children with predominant expressive disorders along the SLI continuum were included (i) because rhythmic processing deficits appear both at the acoustic and the motor levels (affecting auditory-motor coupling) and (ii) to ensure that they could perform the task and that their performance could not be merely attributed to poor general comprehension abilities.

We addressed this issue with a semantic judgment task on sentences either 1) naturally produced at a normal rate, 2) naturally produced at a fast rate or 3) artificially time-compressed (at the same rate as in 2). Our hypothesis was that SLI children may experience specific difficulty tracking fast temporal changes in the amplitude envelope. This may impair the parsing of the continuous speech stream into syllables, leading to reduced performance in the sentence judgment task as compared with TD children. Such a rhythm-processing deficit was assumed to be present for natural as well as artificial acceleration, since both lead to changes in rhythmic acoustic cues present in the speech envelope.

Materials and methods

Participants

Sixteen children with SLI mostly affecting the expressive modality (6 girls) aged 8-13 years old (mean = 10.9, SD 1.67) participated in the experiment. Each child was matched by age and sex with a typically developing (TD) child. All SLI children were recruited through neuropsychologists and neurologists from a neuropediatric hospital unit where they had previously been diagnosed using standardized batteries of French verbal tests (e.g., ELO: Evaluation du Langage Oral [57]; N-EEL: Nouvelles Epreuves pour l'Examen du Langage [58]; BILO: Bilans Informatisés de Langage Oral) [59]) and non-verbal tests (WISC IV: Wechsler Intelligence Scale for Children [60]; WPPSI: Wechsler Preschool and Primary Scale of Intelligence) [61]). Verbal tests assessed phonological skills, production, morpho-syntax, vocabulary and oral comprehension. All SLI children underwent their last neuropsychological evaluation (which confirmed the diagnosis) at the hospital within the six months preceding the experiment. Note that the referring neuropsychologists were not asked to provide any scores obtained by children in the different standardized tests. Their main

involvement consisted of referring those children on their caseloads who met the criteria for inclusion in the study, by providing the parents of these children with an information flyer. The inclusion criteria (decided in collaboration with neuropsychologists and neurologists) encompassed age range from 8 to 13 years old, French native language, right-handedness, SLI with predominant expressive impairment, absence of hearing problem and non-verbal IQ > 70 (as recommended in the classification by Rapin and Allen [24]; see also [21,23,62,63] for a discussion on comparable language profiles between SLI children with non-verbal IQ between 70 and 85 and SLI children with IQ above 85). The exclusion criteria were bilingualism, mental retardation, the presence of attention-deficit/hyperactivity disorder (ADHD) and/or autism spectrum disorder. In our sample, thirteen children had been diagnosed with a phonological-syntactic syndrome affecting phonological programming (e.g., repetition and spontaneous language tested with the ELO [57]) and morpho-syntactic expression (e.g., sentence completion), whereas lexical processing and oral comprehension were preserved. Three other children had verbal dyspraxia with a disorder in orofacial motor programming affecting phonological production (one child had additional phonological deficits), combined with deficits in sensory-motor praxis (e.g., fingertip tapping, manual motor sequences and imitation of hand positions as assessed by sub-tests from the NEPSY, a developmental neuropsychological assessment [64]). To ensure the preservation of certain aspects of the diagnosis (expressive disorders with well-preserved comprehension and no mental retardation) at the time of the experiment, children's performance was additionally assessed with a battery of neuropsychological and French language tests. Linguistic abilities were examined with sub-tests from the BALE (Batterie Analytique du Langage Ecrit; [65]). Non-verbal abilities were assessed with the Raven's coloured Progressive Matrices [66] and the forward and backward digit span tests. Scores below 2 SD of the mean of the population were defined as pathological (see Table 1).

The typically developing children were recruited from elementary and secondary schools. All children were right-handed, French monolinguals and they were included in the study provided that they had no history of language and auditory deficit, nor of any other cognitive or neurological disorder (i.e. no deficit had ever been suspected by physicians, teachers or parents and no child had ever received language or cognitive treatment). Accordingly, neither IQ nor verbal abilities had been assessed with standardized batteries by (neuro)psychologists

in these children²³. Because of strong time constraints (the experiment was carried out during school time), non-verbal and verbal tests could not be administered to TD children; however, to guarantee typical language processing, we examined their (out loud) reading skills with the rapid (three minutes) French test “L’Alouette” [67] before the experiment. All children included in our experiment performed in the normal range (i.e. they were able to read the text within the allocated time without any mistakes): their reading age was comprised between 8.1 and 13.1, thus matching their chronological age range. This confirms general intact language abilities in these children. The protocol conformed to the Declaration of Helsinki and was approved by the local ethical committee (Comité de Protection des Personnes Lyon Sud-Est II; ID RCB: 2012-A00857-36). All children and their parents signed a consent form before the experiment.

²³ In France, IQ is tested in children by qualified clinical psychologists only when there is suspicion of a cognitive or intellectual deficit. IQ could not be assessed in our sample of TD children because of time constraints and most importantly because H.G. (who tested the children at school) is not allowed to administer this test as she is not a psychologist.

Table 1 : Verbal and non-verbal abilities of children with SLI as assessed by standardized neuropsychological tests.

	PHONOLOGICAL- SYNTACTIC SLI	VERBAL DYSPRAXIA
VERBAL TESTS		
Word repetition	-5.25 (6.37)	-0.21 (0.64)
Pseudo-word repetition	-3.44 (3.13)	-1.42 (0.31)
Non-word repetition	-3.83 (3.39)	-1.17 (1.18)
Phonemic fluency	-0.64 (0.95)	-0.43 (2.17)
Picture naming	-1.69 (2.97)	0.4 (0.05)
Expressive vocabulary (word definition)	-0.97 (0.95)	-0.03 (0.74)
Oral comprehension (sentence-to-picture matching)	-0.92 (1.28)	-1.67 (0.06)
Receptive vocabulary (word-to-picture matching)	-1.7 (1.43)	-0.22 (0.08)
Phonemic discrimination	-2.5 (3.50)	-0.1 (0.50)
Rhymes	-1.78 (1.79)	-1.81 (0.84)
Syllabic suppression	-1.29 (1.55)	-0.96 (0.38)
Phonemic suppression	-0.93 (1.18)	-0.91 (1.05)
NON-VERBAL TESTS		
Raven's Progressive Matrices	89.61 (9.25) (range 71-103)	99 (11.31) (range 91-107)
Forward digit span	-0.82 (1.31)	-0.15 (1.20)
Backward digit span	-0.49 (0.88)	-0.2 (0)

Mean standard deviations for several sub-tests of the BALE [65], the Raven's coloured Progressive Matrices [66] and the forward and backward digit span tests are reported for SLI children with a phonological-syntactic syndrome and SLI children with verbal dyspraxia. Standard errors are indicated in brackets. Tests for which children deviated from 2 SD from the mean of the population are highlighted in bold.

Materials

Three hundred sentences (7-9 words) were created following the same syntactic structure: Determiner – Noun 1 – Verb – Determiner – Noun 2 – Preposition – Determiner – Noun 3. The semantic content of Noun 3 (disyllabic target word) was chosen to be contextually

congruent with the beginning of the sentence in half of the sentences ($n = 150$) and incongruent in the other half ($n = 150$). Each target word appeared both in a congruent and an incongruent context (e.g., “Sa fille déteste la nourriture de la cantine” / His daughter hates the food at the canteen and “Le public applaudit le joueur pour sa cantine” / The public applauds the player for his canteen; see S2 Supporting Information for examples). The semantic incongruity of the sentences was obvious as revealed by a pilot questionnaire proposed to twelve healthy adults. Sentence-final target words were controlled for lexical frequency, number of phonemes and number of phonological neighbours using the French lexical databases Lexique 3 [68] and Manulex (database on words in children books, [69]).

Sentences were recorded by a French native male speaker (44.1 kHz, mono, 16 bits) in a sound-attenuated booth using ROCme! Software [70]. Each sentence was recorded twice, at a normal and then at a fast rate. The procedure was the following: the sentence was first displayed on a computer screen in front of the speaker who was instructed to silently read it and to subsequently produce it aloud as a declarative statement at a normal rate. Once all sentences had been produced at a normal rate, they had to be produced at a faster rate using the same procedure. The speaker could produce each sentence several times so that the recorded version was as fluent as possible. The durations of the 2×300 sentences and the number of produced syllables for each sentence were then calculated with Praat software [71]. The average speech rate was 6.76 syllables/sec (SD 0.57) for natural normal rate sentences and 9.15 syllables/sec (SD 0.60) for natural fast sentences. Thus, the overall fast-to-normal ratio was 0.74 (i.e., speed-up factor of 1.35). Subsequently, the time-compressed sentences were computed by digitally shortening them with a PSOLA algorithm (Pitch Synchronous Overlap and Add [72]), as implemented in Praat. The compression rate was calculated for each sentence, and every individual time-compressed sentence was precisely matched by rate with its paired natural fast sentence. Compression was achieved by the re-synthesis of the normal rate stimulus, changing only the temporal structure without affecting the pitch. For the 900 sound files (300×3 rate variants), an 80 Hz high-pass filter was applied and the amplitude envelope was smoothed sentence-initially and finally. The intensity of the sound files was finally peak normalized.

The 900 sentences were divided into twelve experimental lists of 75 items each using a Latin square design so that each stimulus appeared in each rate condition across all participants but only once per list (to avoid repetition effects). No congruent/incongruent sentence pair was used within the same list. Sentence-final target words were matched for the above-mentioned psycholinguistic variables between lists. Each list was composed of three

experimental blocks (25 items each, 13 semantically congruent/12 incongruent) corresponding to the three speech rate conditions. The blocks were always presented in the same order, namely normal rate sentences, then natural fast sentences and finally time-compressed sentences to avoid potential transfer of learning (see [10]). Each participant was presented with one of the twelve lists. Across the 12 lists, all target words were presented in the six different conditions (3 speech rates \times 2 semantic incongruity). Within each experimental block, the order of the sentences was randomized across participants.

Procedure

Children were comfortably seated in front of a laptop in a silent room and received oral instructions (the testing began once the experimenter ensured that children understood the instructions well). They were asked to attentively listen to the sentences and to perform a sentence semantic verification task. Each trial began with a white fixation cross presented at the centre of a black screen; after 1 s, the sentence was played diotically via headphones at a comfortable listening level, with the fixation cross remaining on the screen. The children then had to decide as quickly and accurately as possible whether the sentence made sense or not by pressing one of two pre-specified keys on the keyboard with their right index and middle fingers. Once they gave their response, the next trial was automatically played. If no response was given within 7 s, the trial was recorded as “no response” and the next trial was presented. Participants could listen to each stimulus only once. Before the testing phase, they were given five practice items (different from the experimental stimuli and produced by the same speaker at a normal or fast rate). The total duration of the experiment was 15 minutes. Stimulus presentation, response times and error measurements were performed using E-Prime 2 software (Psychology Software Tools, Inc., Pittsburgh, PA).

Data analysis

Response times (RTs: time-interval between the onset of the sentence-final target word and the button press, in milliseconds) and accuracy (% of correct responses) were measured. Trials for which participants made no response or erroneous responses were considered as errors and were not included in RTs analysis. Trials with RTs below or above 2.5 standard deviations from the individual mean of the condition were further excluded from the analysis of RTs. We also computed d' as an index of sensitivity to semantically incongruent sentence-

final words and β as an index of response bias. This calculation is based on the proportion of hits (i.e., correct responses for incongruent sentences, $p[\text{hits}]$) and false alarms (i.e., errors for congruent sentences, $p[\text{FAs}]$). d' is defined as $z(p[\text{hits}]) - z(p[\text{FAs}])$ and β as $z(p[\text{hits}]) / z(p[\text{FAs}])$ (see [73] for more details).

Before statistical analyses, the normality of the data was tested with the Shapiro-Wilk test. Both RTs and d' followed a standard normal distribution ($W = .98, p = .110$ for RTs; $W = .99, p = .407$ for d'). A repeated-measures analysis of variance (ANOVA) was performed on RTs with Speech Rate (normal, fast, time-compressed) as the within-subject factor and Group (SLI, TD) as the between-subjects factor. As this analysis did not reveal any significant main effect nor interaction ($ps > 0.1$), the results on RTs are not presented here. A similar analysis was conducted on d' . In case of significant interaction, Tukey post-hoc comparison tests were performed. To estimate effect sizes, partial η_p^2 were further calculated [74].

To assess the influence of age and non-verbal abilities (as assessed with the Raven's Progressive Matrices in SLI children) on children's performance in the sentence judgment task, we carried out a mixed-effects linear regression with d' as the dependent measure, Age and Raven's scores as fixed factors and Participant as a random factor for each group (SLI and TD) and each speech rate condition (normal, natural fast and time-compressed).

Finally, as our SLI group was composed of 13 children diagnosed with a phonological-syntactic syndrome and three children with verbal dyspraxia, we also performed the d' analysis for the phonological-syntactic subgroup only (in order to check that the effects in the global analysis were not attributable to the subgroup with verbal dyspraxia, as not all classifications consider this disorder to be part of the SLI profile [75] contrary to the classification by Rapin and Allen [24] used in France). d' did not follow a standard normal distribution in this subgroup ($W = .98, p = .311$); accordingly we used non-parametrical tests (Friedman chi-squared test, χ^2) including the factor Speech Rate, with post-hoc Wilcoxon test and Mann-Whitney test.

Results

The analysis on the d' index revealed that children with SLI exhibited significantly reduced sensitivity to semantic incongruity in sentences ($d' = 0.96$, standard error (SE) 0.09) as compared with TD children ($d' = 1.98$, SE 0.05), $F(1,30) = 15.47, p < .001, \eta_p^2 = .52$. The main effect of Speech Rate was also significant, $F(2,60) = 6.85, p < .003, \eta_p^2 = .23$, with

normal rate sentences leading to higher d' values ($d' = 1.73$, SE 0.22) than fast rate ($d' = 1.25$, SE 0.24) and time-compressed sentences ($d' = 1.43$, SE 0.28), although the differences did not turn out to be significant as assessed with post-hoc tests. Most importantly, Speech Rate significantly interacted with Group, $F(2,60) = 4.50$, $p < .020$, $\eta_p^2 = .62$, as illustrated in Fig 1. Tukey post-hoc tests revealed that SLI children performed significantly worse than their TD age peers when listening to natural fast ($p < .030$) and time-compressed sentences ($p < .001$; Fig 1 and Table 1), whereas the difference between the two groups did not reach significance in the normal rate condition ($p = .246$).

The ANOVA on the β index indicated a significant Group effect, $F(1,30) = 26.66$, $p < .001$, $\eta_p^2 = .89$, due to a more conservative decision bias in the TD group ($\beta = 1.41$, SE 0.22) than in the SLI group ($\beta = 1.33$, SE 0.21) (Table 2). There was no significant Speech Rate effect and no Speech Rate \times Group interaction on the β index.

Results of the mixed-effects linear regression did not reveal any significant effect of Age or of non-verbal abilities (scores at the Raven's Matrices), nor any interaction between the two, in any of the three speech rate conditions in SLI children. The same analysis conducted in TD children with Age as a fixed factor also showed that this variable did not significantly affect d' in the normal, natural fast and time-compressed conditions.

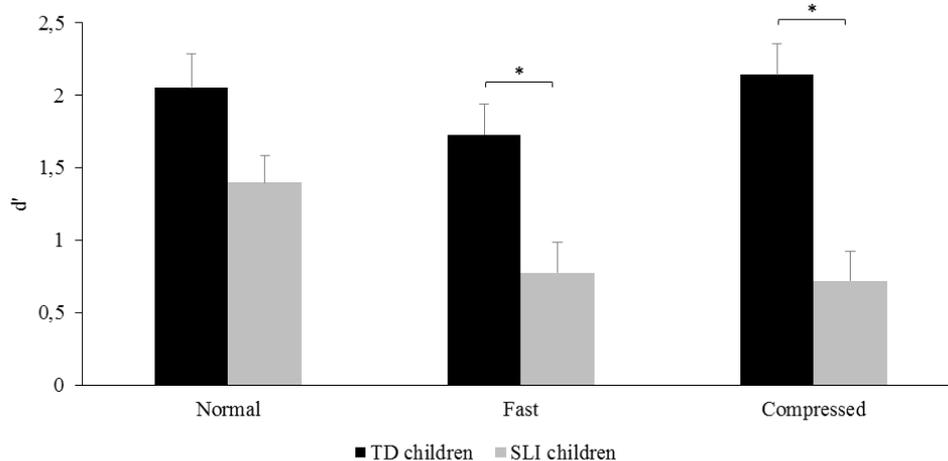


Fig 1 : Children's sensitivity to semantic incongruity during normal and fast rate speech processing. Mean d' values in children with SLI and in chronological age-matched typically-developing(TD) children are presented as a function of Speech Rate (natural normal, natural fast and time-compressed). (*) indicates a significant difference between conditions ($p < .05$). Error bars indicate standard errors.

Table 2 : Mean d' and β as a function of Group and Speech Rate.

Group	Mean d'			Mean β		
	Normal	Fast	Time-compressed	Normal	Fast	Time-compressed
SLI children	1.31 (0.20)	0.71 (0.21)	0.63 (0.21)	1.52 (0.21)	1.07 (0.06)	1.36 (0.26)
TD children	2.12 (0.23)	1.77 (0.20)	2.22 (0.22)	1.43 (0.16)	1.06 (0.17)	1.59 (0.27)

Mean d' values are reported for children with SLI and age-matched TD children in the three speech rate conditions: natural normal, natural fast and time-compressed. Standard errors are indicated in brackets.

To make sure that the results presented above were not merely related to the inclusion of children with verbal dyspraxia, we performed a non-parametric analysis separately for the subgroup of children with a phonological-syntactic syndrome. The results are presented in Fig 2. The Friedman chi-squared analysis on the d' index revealed a significant Speech Rate effect on children's performance, $\chi^2(2) = 7.54, p < .023$. Sensitivity to sentence semantic incongruity was higher in the normal rate condition ($d' = 1.28, SE 0.20$) than in the natural fast ($d' = 0.64, SE 0.21, V = 76.00, p < .040$) and the time-compressed conditions ($d' = 0.73, SE 0.18, V = 82.00, p < .009$) as revealed by a Wilcoxon post-hoc test; the latter two conditions did not significantly differ from each other. The same analysis in the group of 13 paired TD children showed that the effect of Speech Rate only approached significance, $\chi^2(2) = 5.69, p = 0.058$ (Fig2), mainly reflecting a trend towards higher sensitivity in the normal rate condition ($d' = 1.85, SE 0.24$) than in the natural fast rate condition ($d' = 1.49, SE .17; V = 71.00, p = .080$ as shown by the Wilcoxon test). To compare the performance of the two groups (TD children and children with a phonological-syntactic syndrome) in each condition, U Mann-Whitney tests were carried out. Analyses highlighted that children with a phonological-syntactic disorder were less sensitive to semantic incongruity in sentences than their TD peers in the fast ($W = 30.50, p < .007$) and time-compressed conditions ($W = 20.00, p < .002$). In the normal rate condition, no significant difference was found between the two groups ($W = 51.50, p = .100$), thus confirming the results obtained when all SLI children were included in the analysis (Fig1).

Friedman chi-squared tests on the β index in each group of children did not reveal any significant difference between speech rate conditions. However, a significant Group effect emerged when analysing the three conditions separately with U Mann-Whitney tests (W normal = 6.00, $p < .001$; W fast = 20.00, $p < .001$; W time-compressed = 22.00, $p < .002$),

reflecting a more conservative bias decision in the TD group ($\beta = 1.40$, SE 0.20) than in the group with a phonological-syntactic syndrome ($\beta = 1.31$, SE 0.18). Note that this was also observed in the analysis including all children with SLI.

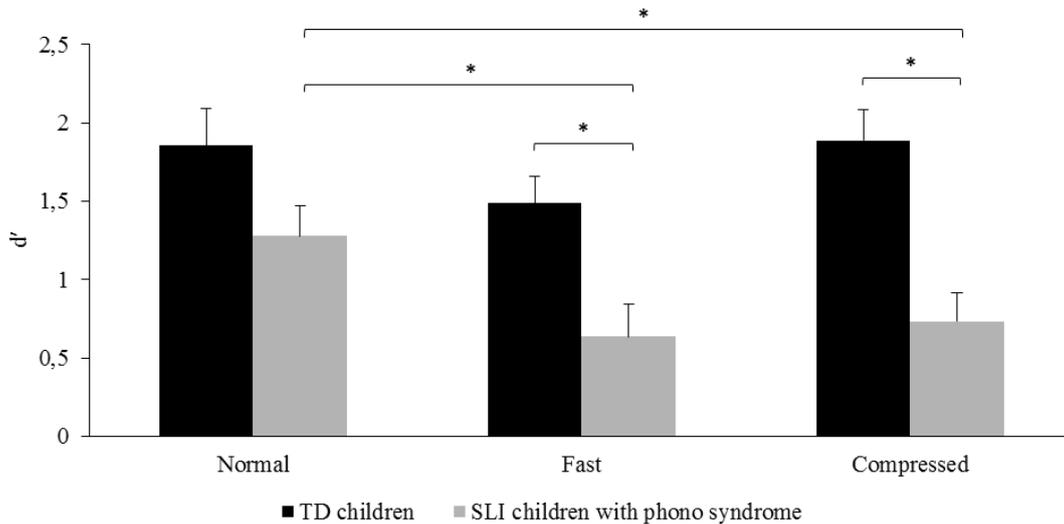


Fig 2 : Sensitivity to semantic incongruity in SLI children diagnosed with a phonological-syntactic syndrome as a function of Speech Rate. Mean d' values are presented for children with a phonological-syntactic syndrome (phono SLI) and chronological age-matched TD children for the three speech rate conditions (natural normal, natural fast and time-compressed). (*) indicates a significant difference between conditions ($p < .05$). Error bars indicate standard errors.

Discussion

The present study investigated the processing of fast rate speech in children who suffer from SLI mainly at the expressive level and in paired children with typical development (TD). Speech processing was assessed through a semantic judgment task performed on aurally presented sentences at either natural normal, natural fast or time-compressed rate. The results revealed that SLI children were less sensitive to semantic incongruity than TD children, and that this sensitivity varied with speech rate. Crucially, the interaction between Group and Speech Rate highlighted that the deficit of the SLI group regarding semantically incongruent word detection was significant only in the case of rapid speech signals. It is therefore more challenging to detect semantic incongruity in sentences for children with SLI than for their TD age-peers when speech is accelerated, either naturally or artificially. No difference emerged between the two groups when sentences were produced at a normal rate.

Children with two profiles of language impairment mainly affecting the expressive modality, namely phonological-syntactic syndrome and verbal dyspraxia [24], participated in the present study. To ascertain that the observed results were not related to the inclusion of children with verbal dyspraxia, we also analyzed the data in the sub-group of children with a phonological-syntactic syndrome only. The results replicated those found when all children with SLI were included: children with a phonological-syntactic syndrome were indeed less sensitive to semantic incongruity in sentences than TD children when speech was naturally or artificially accelerated. This suggests that the general findings cannot only be attributed to the inclusion of children with verbal dyspraxia in our experimental sample, and that they are clearly consistent with the performance of children diagnosed with a phonological-syntactic syndrome. Note that in our experiment, children were aged between 8 and 13 years old; given this rather wide age range, performance in the sentence judgment task may thus be assumed to vary with age. The regression analysis including age as a fixed factor did not reveal any significant influence of this variable on performance in the TD and SLI groups. However, it may be important for future studies including a larger number of participants to systematically compare age ranges (i.e., 8-10 vs. 11-13 years old) to investigate the developmental trajectory of fast rate speech processing in these children. Similarly, in SLI children, no significant influence of non-verbal abilities (as assessed with the Raven's Progressive Matrices) was found on sensitivity to semantic incongruity in this analysis. Although this needs to be more thoroughly assessed with more participants for a well-powered analysis, this suggests that performance of SLI children was not affected by their perceptual intellectual abilities.

In the present study, a cut-off value of 70 for non-verbal IQ was determined as an exclusionary criterion for SLI children based on the classification proposed by Rapin and Allen [24] used by clinicians (see also for instance [76]). This may be considered rather low compared to most other studies using international classifications where only SLI children with a non-verbal IQ higher than 85 (i.e., 1 standard deviation) are included in order to avoid confounds created by intellectual deficits (note however that 13 out of our 16 SLI children had scores above 85 at the Raven's Progressive Matrices at the time of the experiment). The issues of whether a large discrepancy between verbal and non-verbal abilities is needed and whether to extend the exclusion criterion to an IQ of 70 or above have been recently debated in the literature [21,77,78]. As pointed out by Parisse and Maillart [23], the strict integrity of non-verbal skills in SLI seems very demanding and not well justified. First, the same pattern of language deficit can characterize children with low (≤ 85) or high (> 86) non-verbal IQs [20]. Second, studies have shown that SLI children with higher non-verbal IQ do not benefit

more from language interventions than SLI children with lower skills [79-81]. Third, non-verbal IQ can drop or fluctuate considerably over time in SLI with a decline up to 20 points (see [82] for a review), possibly due to persistent language impairment which makes it difficult to obtain non-verbal scores in the normal range. Altogether, this casts doubt on the meaning of non-verbal IQ of 85 in SLI [77] and makes us believe that in the present study, SLI children did not perform poorer for fast rate sentences because of their non-verbal skills.

To our knowledge, our findings provide the first piece of evidence for the impact of speech rate on sentence understanding in children, consistent with previous effects shown in adults [6]. Speeding up the rate of speech, by requiring the listener to cope with temporal changes (as well as with spectral ones in the case of natural acceleration) [3], is thus more demanding and can impact sentence-level integration processes. The processing of fast speech becomes even more problematic for SLI children, even when their developmental language disorders predominantly affect verbal expression with rather preserved comprehension skills, as diagnosed by neuropsychologists and neurologists with standardized tests. This result suggests that neuropsychological measurements of receptive linguistic skills in children should include more ecological tests using fast speech material to assess the actual difficulties experienced by young patients in their daily life.

It has been previously reported that natural fast speech is more difficult to understand than time-compressed speech, mostly due to increased gestural overlap that occurs only when speech is naturally accelerated [6,9]. Although pointing toward this effect, the present pattern of results does not provide clear support for this difference in children. Nevertheless, qualitative inspection of Fig 1 suggests that TD controls tended to perform slightly better when listening to time-compressed speech as compared with natural fast speech, which was not the case for the group of SLI children. Care has to be taken in interpreting non-significant results, but one tentative explanation for this observation is that, unlike children with typical development, SLI children do not take advantage of the preservation of spectral cues in the signal in the case of artificially time-compressed speech. Another potential explanation is that, in the context of this experiment, only TD children benefitted from listening to natural fast sentences to subsequently process time-compressed speech. Such transfer of learning has been reported in adults [6], however it occurred in the reverse order, from the artificial to the natural condition. This was interpreted according to the Reverse Hierarchy Theory [83] which suggests that skills learned in an easier condition (i.e., time-compressed speech) can be subsequently applied to a more complex one (i.e., natural fast speech) but not the reverse. Listening to time-compressed speech would imply learning at high processing levels, then

enabling to focus attention on lower-level, more specific cues and to subsequently adapt to natural fast speech. In the present study, natural fast speech was presented before time-compressed speech to avoid such a bias. Our data suggest slightly better processing of time-compressed than of natural fast speech in TD children and show poorer-than-normal performance to decode both naturally and artificially accelerated speech in children with SLI. Nevertheless, it is not altogether impossible that some kind of learning occurred in TD children. An experiment varying the order of presentation of the two types of fast speech in two sub-groups of participants would be necessary to more precisely assess transfer of learning in TD and SLI children.

One could argue that the low performance of the SLI group to achieve the task in our experiment may result from overall limited speed of information processing. Several studies have indeed shown longer reaction times to sentence-embedded words in SLI, thought to reflect limitations in real-time computing of cognitive operations. Stark and Montgomery [84] for instance reported that SLI children were slower than age-matched controls in detecting target words in spoken sentences, independently of speech rate. In the same word-monitoring paradigm, Montgomery [85] demonstrated lengthened reaction times in these children as compared with TD children for immediate processing of fast rate sentences, whereas the reverse pattern was observed for slower rate stimuli, suggesting that language-impaired children need more time to accurately complete the required cognitive operations. In our study, children with SLI exhibited significant lower sensitivity to semantic incongruity with increasing speech rate as compared with TD children, however no significant difference emerged between the two groups regarding response times (note that qualitative data observation even suggested numerically faster mean response times in SLI children with a phonological-syntactic syndrome than in their TD peers for accelerated speech). This finding is at odds with the results by Montgomery [85] and speaks against processing speed limitation as the main cause of the deficit to process natural fast and time-compressed speech in our SLI children. In addition, the lack of any significant interaction between Group and Speech Rate on the response criterion β suggests that this is not due to decisional strategies. Note also that the task was different from that used in the aforementioned studies [84,85] where children had to remember a target word and provide a timed response immediately upon recognition of this word within a sentence. In our experiment, children had to make a semantic judgment about the entire sentence. Although one could suggest that reduced d' in SLI children stems from their poor understanding of fast linguistic information, we remind the reader that our sentence material was carefully selected in accordance with the children's age (see Method) to ensure

comprehension. The children with SLI were furthermore mainly impaired at the expressive level rather than at the receptive level, and their non-verbal intelligence was also satisfactory, as diagnosed by neuropsychologists and neurologists and assessed with the Raven's Progressive Matrices. The decrease in performance observed in the SLI group – especially for natural fast and time-compressed speech – reported here can therefore not be interpreted as merely reflecting semantic impairment nor altered non-verbal reasoning skills. To fully rule out the semantic explanation, future studies on fast rate speech processing in SLI children should yet make use of other tasks tackling phonological (e.g., sentence-embedded phoneme or syllable identification) and morpho-syntactic processing (e.g., grammaticality judgment). In addition, examining whether performance of SLI children for fast rate speech in these tasks correlates with (and can be predicted by) auditory and motor rhythmic abilities (e.g., rise time and duration perception, finger tapping, beat perception) would allow determining the rhythmic nature of the observed deficit. This would also discard the alternative interpretation that less efficient perceptual processing, as is seen for speech-in-noise perception [86-89], underlies poorer-than-normal fast rate speech processing in SLI.

Our pattern of results is in agreement with the Rapid Auditory Processing (RAP) theory [90,91], suggesting that children with SLI primarily suffer from deficient temporal processing of sequential brief acoustic cues, such as brief complex tones separated by short temporal intervals or phonemes that acoustically differ only by rapid transient formants (e.g., /b/ vs. /d/). Improved speech processing was for instance observed in SLI children when transitions were lengthened [91]. Accordingly, the authors concluded that RAP deficits in SLI may lead to ill-formed phonological representations and thus to atypical language development. In this view, poorer performance of our SLI group for natural fast and time-compressed speech could therefore stem from difficulties dealing with shortened phonemic units in the sentences. However, along with the deficit for very brief cues themselves, children's poor performance could also reflect a deficit to process long ongoing sequences of short segments such as in the fast sentences used in our experiment. Rapid succession of brief acoustic cues particularly impairs SLI children, who have been shown to present with temporal processing deficits [33,92]. As a matter of fact, increasing speech rate does not only reduce the duration of acoustic cues but it additionally affects the dynamics of connected speech, in particular the low-frequency amplitude modulations which characterize signal temporal structure and syllabic rhythm.

Rhythm is a hallmark for speech communication [93,94] and accurate perceptual sensitivity to acoustic rhythmic information in speech is fundamental for language development [36,95]. Processing of rhythm entails, besides the identification of short cues, the detection of temporal regularities in the unfolding signal in order to predict upcoming relevant events such as onsets of syllables, words or phrases. Such anticipation of events' occurrence may imply a temporal shift of attention. According to the Dynamic Attending Theory [52,96], attention is not distributed uniformly over time but it is periodic, with high levels of energy oriented towards salient external events. Self-sustained internal oscillators would synchronize (or entrain) to external sensory rhythms, therefore generating temporal expectancies on the occurrence of future events, whose temporal integration would then be optimized. Crucially, attentional rhythms are assumed to adapt their phase and period to the rate changes in external rhythms so as to maintain synchronicity [52,96]. As far as speech is concerned, temporal characteristics are mainly conveyed by slow amplitude fluctuations which are approximately the duration of syllables and appear at rather regular and thus predictable intervals. The listener's internal oscillators can therefore track this quasi-periodicity and align their phase to that of the amplitude envelope, allowing to anticipate the onset time of forthcoming syllables. This is supposed to facilitate speech segmentation and comprehension ([11] for a review) and to also occur when the rate of speech is accelerated.

As mentioned in the Introduction, deficits in the perception and expression of rhythm have been reported in atypical language development [16,27,31-33]. SLI children may therefore be impaired in detecting temporal regularities in external sensory stimuli, which could explain why language-impaired children in the present study performed worse than their TD peers when listening to accelerated speech. Poor entrainment to rapid speech rhythm may actually impede efficient extraction of syllable onsets to parse the speech stream into key units for decoding. Uttering speech at a fast rate alters signal temporal structure (as well as spectral content for natural acceleration), which is reflected in the amplitude envelope by faster (i.e., higher frequency) and smaller (i.e., less sharp) modulations than in clear, well-articulated speech, as well as by shorter acoustic cues such as rise time [97,98]. When exposed to accelerated speech, the listener's auditory system has therefore to entrain to more frequent syllable "beats" [12] to ensure accurate speech segmentation, a process thought to be fundamental for language acquisition as it promotes prosodic processing [99]. Our results reveal that with the increase in speech rate, children with SLI are less able to track the rhythmic syllable structure than children with typical development, affecting their performance in a semantic judgment task. Our study does not presently allow disentangling

between the rapid auditory processing hypothesis [90,91] and this rhythmic tracking interpretation, which however may not exclude one another. One attempt to do so would be to use the paradigm developed by Ghitza and Greenberg [100] with normal rate sentences, accelerated sentences (i.e., shorter speech segments and faster syllabic rhythm) and accelerated sentences with periodic insertion of silences so as to restore the original sentence rhythm (i.e., shorter segments but normal rhythm). Following our current results, SLI children should perform poorer for fast rate sentences; however, if their deficit stems from difficulties in tracking speech temporal structure (more frequent syllable beats), they should improve when silences are inserted periodically and restore sentence's temporal dynamics. On the contrary, if their deficit is mainly due to impaired processing of shortened speech units, then inserting silences should have no beneficial effect in these children. Additionally, including slowed speech (i.e., longer segments and slower rhythm) and slowed speech with silences (i.e., normal length segments and slower syllabic rhythm) would also be interesting to investigate the processing speed issue as well as to potentially provide clinical recommendations for SLI.

In the context of the rhythmic interpretation, our findings may be explained in the context of multi-time resolution models of speech processing [13,14] suggesting that phase-locking between theta (4-7 Hz) oscillations in auditory cortex and slow fluctuations in the amplitude envelope is critical for intelligibility [11]. This alignment between brain rhythms and speech rhythms, already present in infants [101,102], is assumed to guarantee efficient detection of syllable prominence and therefore reliable speech understanding [103,104]. EEG/MEG studies have revealed abnormal patterns of low-frequency oscillatory activity in auditory regions during processing of rhythmic noise or speech stimuli in adults with developmental dyslexia [105,106] and in children with poor reading skills, especially for time-compressed speech [109]. To the best of our knowledge, no such investigation has been conducted in SLI children so far; however, impaired processing of speech rhythm in these children may be expected to result, at least partly, from atypical functional neural entrainment to slow modulations in speech, hence hindering syllabic parsing and prosodic processing [32]. As suggested by our results, this may be even more evident for fast rate speech perception.

As rhythm-processing deficits arise at the motor level as well, oscillatory abnormalities could furthermore not be circumscribed to auditory cortical regions but could spread to articulatory regions. Perception of acoustic rhythm, in music and also in speech, implies tight reciprocal coupling between temporal auditory and frontal motor planning regions [42,43]. In agreement with the Dynamic Attending Theory [96], such interactions would enable

predicting upcoming beats and therefore enhance rhythmic input processing [41,108]. At the neuroanatomical level, auditory-motor coupling may be mediated by the dorsal stream, which connects posterior temporal to premotor regions via the inferior parietal cortex [109,110]. This pathway is thought to play a key role in language development and to underlie speech sensorimotor integration, particularly under compromised acoustic conditions [111,112]. At the functional level, communication between auditory and motor areas could operate through oscillatory synchrony [111]. Alignment between endogenous oscillations in temporal and premotor cortices has indeed been described, at rest, in theta and gamma ranges, two frequency bands relevant for syllabic and phonemic sampling respectively [113]. Interestingly, Lehongre and coworkers [114] further reported reduced low-gamma (25-35 Hz) entrainment to amplitude-modulated white noise in auditory but also in left articulatory motor and somatosensory regions in dyslexic adults. Accordingly, one may hypothesize that impaired oscillatory dynamics, most likely in the theta syllabic range, within the dorsal sensorimotor network also exists in SLI children, affecting their abilities to process speech rhythm. As the dorsal stream is specifically involved in the perception of degraded speech (e.g., time-compressed or noisy speech, [115,116]), potential abnormal neural communication between auditory and premotor cortices in SLI children may be even more detrimental to the perception of fast rate speech. Future investigations are nevertheless needed to examine brain auditory-motor tracking of speech at various rates together with rhythmic abilities in children with SLI.

To the best of our knowledge, our study is the first to compare the perception of fast rate speech, accelerated naturally or artificially, in French speaking SLI children. Results reveal that these children perform worse than their TD age peers to decode both types of fast speech, suggesting poor entrainment to slow amplitude fluctuations that characterize speech rhythm. These findings, consistent with previous works on rhythm-processing deficits in developmental language disorders [31,32], provide arguments that maintaining a normal rate when talking to children with language disorder, in classrooms but also during neuropsychological assessment and remediation, is essential to help speech processing and communication. Besides, processing of fast speech in SLI children could take advantage of the recent line of research on the impact of musical interventions on language processing. Growing evidence indeed points towards the benefit offered by rhythmic stimulation and music on phonological and morpho-syntactic abilities in typically developing children but also in children with dyslexia or SLI [92,117,118]. If poor decoding of fast rate speech in SLI children results, at least partly, from an inability to accurately adjust to accelerated speech

rhythm, entrainment to musical rhythms may be expected to improve subsequent language performance in these children (see [119]).

Acknowledgments

We would like to thank parents and their children for participating in this study. We also thank Lydie Merle for helping to recruit SLI children, Emmanuel Ferragne for his help in the construction of the stimuli and Damien Gouy for recording the sentences. This study was conducted as part of the ODYSSEE project (PI: V.B; n° 11 JSH2 005 1) funded by the French National Research Agency (ANR). H.G is funded by a doctoral fellowship from the LabEx ASLAN (ANR-10-LABX-0081) of Université de Lyon within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the ANR.

References

1. Duez D, Nishinuma Y. Vitesse d'élocution et durée des syllabes et de leurs constituants en français parlé. *Travaux de l'Institut de Phonétique d'Aix*. 1987;11: 157-180. French.
2. Koreman J. Perceived speech rate: the effects of articulation rate and speaking style in spontaneous speech. *J Acoust Soc Am*. 2006;119(1): 582-96.
3. Janse E. Word perception in fast speech: artificially time-compressed vs. naturally produced fast speech. *Speech Commun*. 2004;42(2): 155-73.
4. Duez D. Consonant sequences in spontaneous French speech. In *Proceedings of the ESCA Sound patterns of Spontaneous Speech (SPoSS) workshop*. 1998; 63-68.
5. Max L, Caruso AJ. Acoustic measures of temporal intervals across speaking rates: Variability of syllable- and phrase-level relative timing. *J Speech Lang Hear Res*. 1997;40(5): 1097-1110.
6. Adank P, Janse E. Perceptual learning of time-compressed and natural fast speech. *J Acoust Soc Am*. 2009;126(5): 2649-59.
7. Dupoux E, Green K. Perceptual adjustment to highly compressed speech: effects of talker and rate changes. *J Exp Psychol Hum Percept Perform*. 1997;23(3): 914-27.
8. Gordon-Salant S, Zion DJ, Espy-Wilson C. Recognition of time-compressed speech does not predict recognition of natural fast-rate speech by older listeners. *J Acoust Soc Am*. 2014;136(4): 268-74.
9. Guiraud H, Ferragne E, Bedoin N, Boulenger V. Adaptation to natural fast speech and time-compressed speech in children. In *Proceedings of Interspeech*. 2013; 2-6.
10. Francis AL, Baldwin K, Nusbaum HC. Effects of training on attention to acoustic cues. *Percept Psychophys*. 2000;62(8): 1668-80.
11. Peelle JE, Davis MH. Neural oscillations carry speech rhythm through to comprehension. *Front Psychol*. 2012;3: 320.

12. Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM. Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci U S A*. 2001;98(23): 13367-72.
13. Poeppel D. The analysis of speech in different temporal integration windows : cerebral lateralization as asymmetric sampling in time. *Speech Commun*. 2003;41: 245-55.
14. Giraud AL, Poeppel D. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*. Nature Publishing Group; 2012;15(4): 511-7.
15. Ghitza O. The theta-syllable: A unit of speech information defined by cortical function. *Front Psychol*. 2013;4: 1-5.
16. Goswami U. A temporal sampling framework for developmental dyslexia. *Trends Cogn Sci*. 2011;15(1): 3-10.
17. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Association. 2013.
18. Bishop DVM. Ten questions about terminology for children with unexplained language problems. *Int J Commun Disord*. 2014;49(4): 381-415.
19. Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for primary speech and language delay: a systematic review of the literature. *International journal of language & communication disorders*. 1998;33(sup1), 21-23.
20. Tomblin JB, Zhang X. Language patterns and etiology in children with specific language impairment. I. In H. Tager-Flusberg editors, *Neurodevelopmental disorders*. Cambridge, MA: MIT Press / Bradford Books. 1999.
21. Leonard LB. *Children with specific language impairment*. Cambridge, MA: The MIT press. 2014.
22. Maillart C, Leclercq AL, Quemart P. La répétition de phrases comme aide au diagnostic des enfants dysphasiques. *Les Entretiens d'Orthophonie*. 2012; 22-30
23. Parisse C, Maillart C. Nouvelles propositions pour la recherche et l'évaluation du langage chez les enfants dysphasiques. In C. G. C. Jacquet-Pfau editor. *Autour du mot: pratiques et compétences*. Limoge, France: Lambert- L. 2010.
24. Rapin I, Allen DA. Developmental language disorders: Nosologic considerations. In U. Kirk, editors. *Neuropsychology of language, reading, and spelling*. New York, NY: Academic. 1983.
25. Tomblin JB, Nippold M.A. Features of language impairment in the school years. In J.B. Tomblin and M.A. Nippold, editors. *Understanding individual differences in language development across the school years*. New York, NY: Psychology Press/Taylor & Francis. 2014.
26. Scott SK. The point of P-centres. *Psychol Res*. 1998;61(1): 4-11.
27. Corriveau K, Pasquini E, Goswami U. Basic auditory processing skills and Specific Language Impairment: A new look at an old hypothesis. *J Speech Lang Hear Res*. 2007;50: 647-666.
28. Fraser J, Goswami U, Conti-Ramsden G. Dyslexia and Specific Language Impairment: The Role of Phonology and Auditory Processing. *Sci Stud Read*. 2010;14(1): 8-29.
29. Thomson JM, Goswami U. Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *J Physiol*. 2008;102: 120-129.

30. Beattie RL, Manis FR. Rise time perception in children with reading and combined reading and language difficulties. *J Learn Disabil.* 2012;46(3): 200-209.
31. Richards S, Goswami U. Auditory Processing in Specific Language Impairment (SLI): Relations With the Perception of Lexical and Phrasal Stress. *J Speech Hear Res.* 2015;58: 1292-305.
32. Cumming R, Wilson A, Goswami U. Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: a new look at a perceptual hypothesis. *Front Psychol.* 2015;6.
33. Corriveau K, Goswami U. Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex.* 2009;45: 119-30.
34. Cumming R, Wilson A, Leong V, Colling LJ, Goswami U. Awareness of rhythm patterns in speech and music in children with specific language impairments. *Front Hum Neurosci.* 2015;9(672).
35. Muneaux M, Ziegler JC, Truc C, Thomson J, Goswami U. Deficits in beat perception and dyslexia: evidence from French. *Neuroreport.* 2004;15(8): 1255-9.
36. Woodruff Carr K, White-Schwoch T, Tierney AT, Strait DL, Kraus N. Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proc Natl Acad Sci U S A.* 2014;111(40): 14559-64.
37. Tierney AT, Kraus N. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and language.* 2013;124(3): 225-231.
38. Leong V, Goswami U. Assessment of rhythmic entrainment at multiple timescales in dyslexia: evidence for disruption to syllable timing. *Hear Res.* 2014;308: 141-61.
39. Alcock KJ, Passingham RE, Watkins K, Vargha-Khadem F. Pitch and timing abilities in inherited speech and language impairment. *Brain and language.* 2000;75(1): 34-46.
40. Peter B, Stoel-Gammon C. Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics.* 2008;22(3): 171-198.
41. Tierney A, Kraus N. Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Front Hum Neurosci.* 2014;8.
42. Chen JL, Penhune VB, Zatorre RJ. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J Cogn Neurosci.* 2008;20(2): 226-39.
43. Grahn JA, Brett M. Rhythm and Beat Perception in Motor Areas of the Brain. *J Cogn Neurosci.* 2007;19(5): 893-906.
44. Hill EL. Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language and Communications Disorders.* 2001;36: 149-171
45. Bishop DV. Motor immaturity and specific speech and language impairment: Evidence for a common genetic basis. *American Journal of Medical Genetics Part A.* 2002;114(1): 56-63.
46. Belton E, Salmond CH, Watkins KE, Vargha-Khadem F, Gadian DG. Bilateral brain abnormalities in an autosomal dominant speech and language disorder. *Human Brain Mapping.* 2003;18: 194-200.
47. Vargha-Khadem F, Watkins K, Price CJ, Ashburner J, Alcock KJ, Connelly A, Frackowiak RS, Friston KJ, Pembrey ME, et al. Neural basis of an inherited speech

- and language disorder. Proceedings of the National Adacemy Sciences USA. 1998;95(21): 12695-12700
48. Watkins KE, Dronkers NF, Vargha-Khadem F. Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*. 2002;125(3): 452-464.
 49. Pastor MA, Day BL, Macaluso E, Friston KJ, and Frackowiak RSJ. The functional neuroanatomy of temporal discrimination. *The Journal of Neuroscience*. 2004;24: 2585–2591.
 50. Geiser E, Zaehle T, Jancke L, Meyer M. The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of Cognitive Neuroscience*. 2008;20: 541-552.
 51. Kotz SA, Schmidt-Kassow M. Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex*. 2015;68: 48-60.
 52. Kotz SA, Schwartze M, Schmidt-Kassow M. Non-motor basal ganglia functions: A review and proposal for a model of sensory predictability in auditory language perception. *Cortex*. 2009;45: 982-990.
 53. de Weck G, Rosat, MC. *Troubles dysphasiques*. Paris: Masson. 2003.
 54. Rapin I. Practitioner review : Developmental language disorders : A Practitioner Review : Developmental Language Disorders : A Clinical Update. *J Child Psychol Psychiatry*. 1996;37(6): 643-55.
 55. Marquardt TP, Sussman HM, Snow T, Jacks A. The integrity of the syllable in developmental apraxia of speech. *Journal of Communication Disorders*. 2002;35(1): 31-49.
 56. Marquardt TP, Sussman HM, Davis BL. Developmental apraxia of speech: Advances in theory and practice. *Treating disordered motor speech control*. 2001;413-473.
 57. Khomsi A. *ELO: évaluation du langage oral*. ECPA. Paris, France: Editions du Centre de Psychologie Appliquée. 2001.
 58. Chevrie-Muller C, Plaza M. *N-EEL. Nouvelles Epreuves pour l'Examen du Langage*. Paris, France: Editions du Centre de Psychologie Appliquée. 2001.
 59. Khomsi A, Khomsi J, Pasquet F, Parbeau-Guéno A. *Bilan Informatisé de Langage Oral au cycle 3 et au Collège (BILO-3C)*. Paris, France: Éditions du Centre de Psychologie Appliquée. 2007.
 60. Wechsler D. *Échelle d'intelligence de Wechsler pour enfants: WISC-IV*. Paris, France: Éditions du Centre de Psychologie Appliquée. 2005.
 61. Wechsler D. *WPPSI-IV. Echelle d'intelligence de Wechsler pour enfants—4ème édition*. Paris, France: Éditions du Centre de Psychologie Appliquée. 2014.
 62. Leonard L. *Children with Specific Language Impairment*. Cambridge, MA: MIT Press. 1998.
 63. Tomblin JB. *Understanding individual differences in language development across the school years*. Hove, England: Psychology Press. 2014.
 64. Korkman M, Kirk U, Kemp S. *NEPSY: Bilan neuropsychologique de l'enfant [Test et manuel]*. Paris, France: Éditions du Centre de Psychologie Appliquée. 1998.
 65. Jacquier-Roux M, Lequette Ch, Pouget G, Valdois S, Zorman M. *BALE: Batterie Analytique du Langage Ecrit*. Grenoble: Laboratoire des Sciences de l'Education,

- groupe Cogni-Sciences, et Laboratoire de Psychologie et NeuroCognition, CNRS/Université Pierre Mendès France. 2010.
66. Raven J, Raven JC, Court JH. Coloured progressive matrices. Oxford, England: Oxford Psychologists Press. 1998.
 67. Lefavrais P. Test de l'Alouette. 1967. French.
 68. New B, Brysbaert M, Veronis J, Pallier C. The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*. 2007;28(04): 661-677.
 69. Lété B, Sprenger-Charolles L, Colé P. MANULEX: A grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments and Computers*. 2004;36(1): 156-166.
 70. Ferragne E, Flavier S, Fressard C. ROCme! Recording of Oral Corpora Made Easy: Logiciel d'enregistrement de corpus. DDL. 2012. Available from: <http://www.ddl.ish-lyon.cnrs.fr/Download/Index.asp>.
 71. Boersma P, Weenink D. Praat: Doing phonetics by computer [Computer program]. Version 5.3.05. 2012. Available from: <http://www.praat.org/>.
 72. Moulines E, Charpentier F. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Commu*. 1990;9: 453-467.
 73. Macmillan NA, Creelman CD. Detection theory: A user's guide. Cambridge, England: Cambridge University Press. 1991.
 74. Cohen J. Statistical power analysis for the behavioral sciences. 2nd Ed. Hillsdale, NJ: Erlbaum. 1988.
 75. Mazeau M. Reconnaître une dyspraxie. *Réadaptation*. 2005;522: 8-13.
 76. Hansson K, Sandgren O, Sahlén B. Changing labels for a concept in change. Commentary on Bishop, DVM, 2014, Ten questions about terminology for children with unexplained language problems. *Int J Lang Comm Disord*. 2014;49: 381-415.
 77. Leonard LB, Replacing one imperfect term with another. *International Journal of Language and Communication Disorders*. 2014;49(4): 436 –437.
 78. Reilly S, Bishop DV, Tomblin B. Terminological debate over language impairment in children: Forward movement and sticking points. *Int J Commun Disord*. 2014;49(4): 452-462.
 79. Bowyer-Crane C, Snowling MJ, Duff F, Hulme C. Response to early intervention of children with specific and general language impairment. *Learning Disabilities: A Contemporary Journal*. 2011;9(2): 107-121.
 80. Fey ME, Long SH, Cleave PL, Watkins RV, Rice ML. Reconsideration of IQ criteria in the definition of specific language impairment. *Specific language impairments in children*. 1994;4: 161-178.
 81. Wake M, Tobin S, Levickis P, Gold L, Ukoumunne OC, Zens N, et al. Randomized trial of a population-based, home-delivered intervention for preschool language delay. *Pediatrics*. 2013;132(4): 895-904.
 82. Botting N. Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry*. 2005;46(3): 317-326.
 83. Ahissar M, Hochstein S. The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci*. 2004;8(10): 457-464.

84. Stark RE, Montgomery JW. Sentence processing in language-impaired children under conditions of filtering and time compression. *Applied Psycholinguistics*. 1995;16(2): 137-154.
85. Montgomery JW. Effects of input rate and age on the real-time language processing of children with specific language impairment. *Int J Commun Disord*. 2005;40(2): 171-188.
86. Leybaert J, Macchi L, Huyse A, Champoux F, Bayard C, Colin C, et al. Atypical audio-visual speech perception and McGurk effects in children with specific language impairment. *Frontiers in psychology*. 2014;5.
87. Robertson EK, Joanisse MF, Desroches AS, Ng S. Categorical speech perception deficits distinguish language and reading impairments in children. *Developmental Science*. 2009;12(5): 753-767.
88. Ziegler JC, Pech-Georgel C, George F, Alario FX, Lorenzi C. Deficits in speech perception predict language learning impairment. *Proceedings of the national academy of sciences of the united states of america*. 2005;102(39): 14110-14115.
89. Ziegler JC, Pech-Georgel C, George F, Lorenzi C. Noise on, voicing off: Speech perception deficits in children with specific language impairment. *Journal of experimental child psychology*. 2011;110(3): 362-372.
90. Tallal P, Piercy M. Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality. *Neuropsychologia*. 1973;11: 389-398.
91. Tallal P, Piercy M. Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*. 1975;13(1): 69-74.
92. Bedoin N, Brisseau L, Molinier P, Roch D, Tillmann B. Temporally regular musical primes facilitate subsequent syntax processing in children with Specific Language Impairment. *Front Neurosci*. 2016;10: 245.
93. MacNeilage PF. The frame/content theory of evolution of speech production. *Behav Brain Sci*. 1998;21(4): 499-546.
94. Port RF. Meter and speech. *J Phon*. 2003;31: 599-611.
95. Mehler J, Nespor M. Linguistic rhythm and the development of language. *Structures and beyond: The cartography of syntactic structures*. 2004;3: 213-222.
96. Large EW, Jones MR. The dynamics of attending: How people track time-varying events. *Psychol Rev*. 1999;106(1): 119.
97. Abrams D, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci*. 2008;28(15): 3958-65.
98. Greenberg S, Arai T, Grant KW. The Role of Temporal Dynamics in Understanding Spoken Language. In: Divenyi P, Vicsi K, Meyer G, editors. *Dynamics of Speech Production and Perception*. Amsterdam, Holland: IOS Press. 2003.
99. Goswami U, Leong V. Speech rhythm and temporal structure: Converging perspectives? *Lab Phonol*. 2013;4(1).
100. Ghitza O, Greenberg S. On the possible role of brain rhythms in speech perception: intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*. 2009; 66: 113-126.
101. Telkemeyer S, Rossi S, Koch SP, Nierhaus T, Steinbrink J, Poeppel D, et al. Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. *J Neurosci*. 2009;29(47): 14726-33.

102. Telkemeyer S, Rossi S, Nierhaus T, Steinbrink J, Obrig H, Wartenburger I. Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near-infrared spectroscopy and EEG study. *Front Psychol.* 2011;2.
103. Doelling KB, Arnal LH, Ghitza O, Poeppel D. Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage.* 2014;85: 761-8.
104. Gross J, Hoogenboom N, Thut G, Schyns P, Panzeri S, Belin P, Garrod S. Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS biology.* 2013;11(12).
105. Hämäläinen J, Rupp A, Soltész F, Szücs D, Goswami U. Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. *Neuroimage.* 2012;59(3): 2952-61.
106. Soltész F, Szücs D, Leong V, White S, Goswami U. Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. *PLoS One.* 2013;8(10).
107. Abrams D a, Nicol T, Zecker S, Kraus N. Abnormal cortical processing of the syllable rate of speech in poor readers. *J Neurosci.* 2009;29(24): 7686-93.
108. Iversen JR, Repp BH, Patel AD. Top-down control of rhythm perception modulates early auditory responses. *Ann N Y Acad Sci.* 2009;1169: 58-73.
109. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci.* 2007;8(5): 393-402.
110. Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat Neurosci.* 2009;12(6): 718-24.
111. Alho J, Lin F-H, Sato M, Tiitinen H, Sams M, Jääskeläinen IP. Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Front Psychol.* 2014;5.
112. Fadiga L, Bufalari I, D'Ausilio A, Salmas P. The role of the motor system in discriminating normal and degraded speech sounds. *Cortex.* 2012; 48(7): 882-7.
113. Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RSJ, Laufs H. Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron.* 2007;56(6): 1127-34.
114. Lehongre K, Ramus F, Villiermet N, Schwartz D, Giraud AL. Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron.* 2011;72(6): 1080-1090.
115. Adank P, Devlin JT. On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *Neuroimage.* 2010;49(1): 1124-32.
116. Peelle JE, McMillan C, Moore P, Grossman M, Wingfield A. Dissociable patterns of brain activity during comprehension of rapid and syntactically complex speech: evidence from fMRI. *Brain Lang.* 2004;91(3): 315-25.
117. Schön D, Tillmann B. Short- and long-term rhythmic interventions: perspectives for language rehabilitation. *Ann N Y Acad Sci.* 2015;1337(1): 32-9.
118. Goswami U. Entraining the Brain : Applications to Language Research and Links to Musical Entrainment. *Empirical Musicology Review.* 2012;7(1): 57-63.
119. Goswami U, Cumming R, Chait M, Huss M, Mead N, Wilson AM, et al. Perception of Filtered Speech by Children with Developmental Dyslexia and Children with Specific Language Impairments. *Frontiers in psychology.* 2016;7.

2. Synthèse des résultats de l'étude 2

Les résultats de cette étude révèlent l'impact du débit syllabique sur les capacités de traitement de la parole naturelle chez des enfants francophones présentant une dysphasie expressive. Ils montrent que notre groupe d'enfants dysphasiques est moins sensible à l'incohérence sémantique dans des phrases que le groupe d'enfants neurotypiques lorsque le débit de parole est accéléré, que ce soit de manière naturelle ou artificielle. Aucune différence significative n'est observée entre les enfants contrôles et les dysphasiques lors de la perception de parole à débit normal. Ces résultats révèlent ainsi que les enfants dysphasiques ne tirent aucun avantage de la préservation des indices spectraux lors de l'écoute de parole compressée par rapport à la parole naturelle rapide. Il semble donc que ces enfants présentent une difficulté de traitement de l'information temporelle, rythmique du signal de parole, lorsque la fréquence des modulations d'amplitude de l'enveloppe augmente avec le débit syllabique.

Ces résultats sont ainsi en accord avec la l'hypothèse « *prosodic phrasing hypothesis* » (Cumming, Wilson, & Goswami, 2015a) suggérant que les troubles morpho-syntaxiques souvent observés chez les enfants dysphasiques résulteraient en partie de leur sensibilité réduite aux indices de *risetime* et de durée. Cette insensibilité affecterait l'extraction des indices rythmiques de basse fréquence contenus dans l'enveloppe temporelle du signal de parole, entravant la segmentation du flux continu en syllabes et mots, et par conséquent le traitement de la structure phonologique à la base des représentations syntaxiques. Les résultats de notre étude démontrent que lorsque le débit de parole augmente, les enfants dysphasiques auraient plus de difficultés à détecter les « battements » syllabiques, diminuant leurs capacités à découper le signal et ainsi à le comprendre. Nous avons proposé, en accord avec Cumming et collaborateurs (2015a) et Goswami (2011, pour la dyslexie), que ce déficit pourrait être sous-tendu par un dysfonctionnement de l'alignement des oscillations corticales sur l'enveloppe d'amplitude du signal de parole. Cette hypothèse fait l'objet de l'étude 3 présentée ci-après.

3. Présentation de l'étude 3 : Étude en MEG

L'étude 2 comportementale a montré que les enfants présentant une dysphasie expressive éprouvaient des difficultés, par rapport à des enfants neurotypiques appariés en âge et sexe, à traiter le signal de parole produit à un débit rapide, que l'accélération soit naturelle ou artificielle, suggérant une difficulté de traitement des indices temporels rythmiques de la parole. Le rythme de la parole est principalement véhiculé au sein de l'enveloppe temporelle et il serait « traqué » par les oscillations corticales dans les régions auditives permettant un échantillonnage en unités syllabiques (et phrastiques) et un décodage plus facile (Ghitza, 2011; Giraud & Poeppel, 2012). Selon les résultats de plusieurs études, la sensibilité réduite des enfants dysphasiques aux indices rythmiques de la parole expliquerait les troubles phonologiques et morpho-syntaxiques fréquemment observés chez ces enfants, et elle pourrait être la conséquence d'un déficit de synchronisation des oscillations corticales sur l'enveloppe temporelle du signal de parole (Cumming, Wilson, & Goswami, 2015a). Bien que cette hypothèse soit particulièrement séduisante, elle n'a jamais fait l'objet d'investigations à l'heure actuelle et à notre connaissance. Un certain nombre de travaux ont néanmoins révélé un fonctionnement oscillatoire atypique dans la dyslexie, caractérisé par un trouble de la synchronisation des oscillations corticales sur le signal de parole dans les basses fréquences (delta et thêta ; Lizarazu et al., 2015; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016) et au niveau de l'amplitude des oscillations gamma (Lehongre et al., 2013, 2011).

Dans la présente étude, nous avons examiné, en MEG, la dynamique des oscillations corticales lors de la perception de parole naturelle produite à un débit normal ou rapide chez un groupe d'enfants présentant une dysphasie expressive apparié à un groupe contrôle d'enfants neurotypiques, âgés de 8 à 13 ans et de langue maternelle française. Le protocole était identique à celui décrit dans l'étude 1 chez les enfants neurotypiques et, le groupe d'enfants contrôles correspond à un sous-groupe des enfants de l'étude 1. Les enfants dysphasiques avaient ainsi pour tâche d'écouter attentivement des phrases produites dans les deux conditions de débit de parole et devaient repérer des stimuli *filler* non analysés. Différentes analyses ont été réalisées au niveau des capteurs et des sources afin de mesurer les modulations de puissance et la cohérence entre les modulations lentes d'amplitude de l'enveloppe du signal de parole (obtenue à l'aide de la transformée de Hilbert) et les oscillations corticales. Comme pour l'étude 1, ces analyses ont été réalisées dans 3 bandes de fréquence différentes : la bande thêta (4-7 Hz), une bande centrée sur le rythme syllabique

moyen des phrases à débit normal (5.6-7.6 Hz) et une bande de fréquence centrée sur le rythme syllabique moyen des phrases à débit rapide (8-10 Hz).

Nous avons formulé l'hypothèse que les enfants présentant une dysphasie expressive devraient montrer un alignement atypique (i.e. réduit ou présentant un pattern de latéralisation hémisphérique différent) des oscillations thêta sur l'enveloppe d'amplitude, ainsi qu'un déficit de couplage cortico-acoustique dans les bandes de fréquence centrées sur le débit syllabique moyen des phrases. Ce déficit devrait être d'autant plus marqué lors de l'écoute de parole rapide, avec notamment un recrutement plus important des régions motrices et prémotrices de la voie dorsale. Ainsi, ces régions étant surtout impliquées en condition d'écoute difficile (Hickok & Poeppel, 2007), elles pourraient être particulièrement mobilisées chez les enfants dysphasiques lors de la perception de parole rapide. Toutefois, les études montrent que ces enfants présentent un déficit de traitement du rythme à la fois en perception et en production, dans des tâches de *tapping* par exemple (Corriveau & Goswami, 2009; Corriveau et al., 2007), qui pourrait être sous-tendu par un couplage auditivo-moteur non fonctionnel. Nous pourrions donc alternativement observer un défaut de recrutement de ces régions (pré)motrices lors de la perception de parole rapide, qui pourrait alors contribuer au déficit observé au niveau comportemental dans l'étude 2 (Guiraud, Bedoin, Krifi-Papoz, Herbillon, Caillot-Bascoul, Gonzalez-Monge et Boulenger, en révision).

Oscillatory brain dynamics and cortico-acoustic interaction deficits in Developmental Language Disorder during normal and fast rate speech perception: A MEG study

Hélène Guiraud^{1*}, Ana-Sofia Hincapié, Karim Jerbi^{2a} and Véronique Boulenger^{1a*}

¹ Laboratoire Dynamique Du Langage, CNRS/Université Lyon 2 UMR 5596, Lyon, France

² Département de Psychologie, Université de Montréal, Montréal, Québec, Canada

³ Pontificia Universidad Catolica de Chile, Santiago de Chile, Chile

^a These authors contributed equally to this work.

* Corresponding authors

Correspondence should be addressed to:

Hélène GUIRAUD
Laboratoire Dynamique Du Langage
CNRS/Université de Lyon UMR5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON Cedex 07 - FRANCE
helene.guiraud@univ-lyon2.fr

Véronique BOULENGER
Laboratoire Dynamique Du Langage
CNRS/Université de Lyon UMR5596
Institut des Sciences de l'Homme
14 avenue Berthelot
69363 LYON Cedex 07 - FRANCE
veronique.boulenger@cnrs.fr

Abstract

The ‘prosodic phrasing hypothesis’ (Cumming, Wilson, & Goswami, 2015) suggests that children with Developmental Language Disorder (DLD) experience specific difficulty extracting the low-frequency rhythmic cues carried by the amplitude envelope of speech. This would hinder the segmentation of the speech stream into syllables and words, leading to imperfect phonological representations. A large body of evidence has revealed that the parsing of the incoming signal into syllables may be underpinned by the alignment of auditory cortical oscillations within the theta frequency range (4-7 Hz) to speech amplitude envelope. The cortical oscillatory dynamics underlying the temporal processing of speech has, to our knowledge, never been investigated in children with DLD. To tackle this crucial issue, we compared the cortical activity recorded with MEG of French-speaking children with DLD (aged from 8 to 13 years) to a group of typically-developing (TD) children matched for age and sex. Children had to carefully listen to sentences naturally produced either at a normal rate (~6.6 syllables/s) or fast rate (~9 syllables/s). Analyses of power modulations and coherence between cortical oscillations and speech amplitude envelope in both rate conditions were performed in three distinct frequency bands: the “classical” theta band (4-7 Hz), and two frequency bands centered on the mean syllabic rate of our stimuli (5.6-7.6 Hz for the normal rate and 8-10 Hz for the fast rate conditions). The results reveal atypical brain-to-speech alignment in children with DLD as compared to TD children during normal rate speech perception. Significantly reduced coherence is indeed observed for these children in right auditory regions in the theta band, and in left auditory regions in the normal-rate frequency band (5.6-7.6 Hz). Neural entrainment during the fast rate condition was also impaired in children with DLD: while dorsal stream oscillations in left (pre)motor regions in TD children aligned to speech envelope, this entrainment was absent in the group of children with DLD. These results are, to the best of our knowledge, the first to unravel atypical brain entrainment to the slow fluctuations in speech in Developmental Language Disorder, suggesting a deficit in processing syllabic information in these children.

Keywords : Developmental Language Disorder; Cortical oscillations; Speech rhythm; Entrainment; Syllabic rate.

Introduction

Developmental Language Disorder (DLD; previously labeled Specific Language Impairment or SLI, see *DSM-5*, 2013; Bishop, 2017; Leonard, 2014; Reilly, Tomblin, et al., 2014; Reilly, Bishop, & Tomblin, 2014) is a severe and long-lasting neurodevelopmental disorder affecting the production and/or comprehension of language and not attributable to hearing or other sensory impairment, motor dysfunction, or another medical or neurological condition. It is not better explained by intellectual disability (intellectual developmental disorder) or global developmental delay (*DSM-5*, 2013; Gérard, 1993; Maillart & Schelstraete, 2012). DLD would affect 3-7% of school-aged children (Tomblin et al., 1997), with a higher prevalence in boys.

Despite the heterogeneity of linguistic symptoms in children with DLD, expressive skills are often more impaired than receptive skills, with deficits affecting predominantly phonological and grammatical (i.e. morpho-syntactic) processing. To account for these deficits, Cumming and colleagues (Cumming, Wilson & Goswami, 2015) have developed the «*prosodic phrasing hypothesis*» according to which children with DLD would show insensitivity to speech rhythmic and prosodic information carried by the amplitude envelope, thus altering language development (see also Goswami, 2011 for a close proposal, the ‘temporal sampling framework’, to account for dyslexia). Speech temporal envelope contains low-frequency amplitude modulations at different temporal rates but is dominated by modulations at 4-7 Hz corresponding to the speaker’s syllable rate (Ghitza & Greenberg, 2009; Rosen, 1992). According to Cumming et al. (2015), children with DLD would be particularly impaired at processing two supra-segmental cues important for speech rhythm and stress pattern, namely amplitude envelope rise time (i.e. rate of change in the envelope, corresponding to onsets of successive syllables, Rosen, 1992) and signal duration. This would affect the extraction of the global prosodic phrasing of utterances, thus hindering syllable and word segmentation and consequently the processing of the phonological and morpho-syntactic structure of speech. An increasing number of studies points to such a rhythmic deficit in children with DLD, using both verbal and non-verbal stimuli in which rise time and sound duration varied (Beattie & Manis, 2012; Corriveau, Pasquini, & Goswami, 2007; Cumming, Wilson, & Goswami, 2015; Goswami et al., 2016; Richards & Goswami, 2015; Tallal & Piercy, 1973a, 1973b, 1975). Corriveau and collaborators (2007) for instance examined auditory rhythmic abilities of English-speaking children with DLD, paired with typically-developing children of the same age (7-11 years), in non-verbal auditory

discrimination tasks. Their results revealed that children with DLD performed significantly poorer than their peers in tasks where rise time and duration were manipulated (*e.g.*, decide which of two stimuli had the sharpest onset), while no significant difference between groups was observed regarding sound intensity. Most interestingly, regression analyses demonstrated that the performance of children with DLD in rise time tasks was a good predictor of children's language skills (vocabulary, phonology and reading). A similar pattern of results was observed by Cumming and colleagues (2015) in English-speaking children (aged on average 9 years and 6 months) with DLD associated or not with phonological and reading impairment. The authors used a set of auditory discrimination tasks, including non-speech (amplitude-modulated noise) and speech stimuli (syllable /ba/), aimed at examining stress perception, sensitivity to rise time, duration and fundamental frequency (*i.e.* dynamic pitch changes). Their results first revealed that children with DLD, irrespective of the presence of phonology and reading deficits, performed significantly poorer than their TD peers in the stress perception tasks (see also Richards & Goswami, 2015). Secondly, both sub-groups of children with DLD were impaired at processing rise time for speech stimuli only (contrary to the results reported by Corriveau, Pasquini, & Goswami, 2007) ; children with associated poor phonology and reading had additional difficulty perceiving the duration of syllables and the pitch changes in sine tones as compared to their age peers. Third, the authors reported that sensitivity to rise time and duration was a significant predictor of children's performance to perceive lexical stress. Sensitivity to lexical stress also predicted children's expressive and receptive language skills. In line with previous developmental studies (Christophe, Gout, Peperkamp, & Morgan, 2003; Jusczyk, 1995, 1997; Mehler et al., 1988; Mehler, Sebastian-Galles, & Nespore, 2004; Nazzi & Ramus, 2003) these findings therefore suggest that accurate processing of the temporal structure of speech (syllable onsets and prominence), which appears to be altered in children with DLD, is crucial for prosodic representations and thus for phonological and morpho-syntactic skills to develop efficiently.

Current neurocognitive models of speech perception postulate that encoding of speech rhythmic cues may be underpinned by the entrainment (*i.e.* synchronization) of cortical oscillations to the speech signal in frequency bands crucial for speech segmentation and comprehension (Ghitza, 2011; Giraud & Poeppel, 2012; Poeppel, 2003). The close correspondence between the temporal scales of speech and cortical endogenous oscillations would allow the listener's brain to parse the continuous stream into relevant linguistic units for decoding. While oscillations in the delta (1-3 Hz) and theta (4-7 Hz) frequency bands

would sample the signal into phrasal (~500-2000 ms) and syllable (~100-300 ms) chunks respectively, low-gamma range oscillatory activity (25-45 Hz) would discretize phonemic segments (~20-50 ms). Moreover, as emphasized in the Asymmetric Sampling in Time (AST) model (Poeppel, 2003), syllabic and phonemic parsing would occur in parallel but asymmetrically, with theta-syllable oscillations dominating in the right auditory cortex whereas gamma-phonemic oscillations would be more left-lateralized (Poeppel, 2003; see Giraud et al., 2007 and Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012 for resting-state studies in adults). Convincing empirical evidence from electro- and magnetoencephalography (E/MEG) studies in healthy adults generally support these assumptions (see Peelle & Davis, 2012; Zion Golumbic, Poeppel, & Schroeder, 2012 for reviews). Theta-band oscillations in the auditory cortex have indeed been shown to align to the slow quasi-rhythmic modulations of speech temporal envelope characterizing syllable rate (*e.g.*, Ahissar et al., 2001; Ding & Simon, 2013; Doelling, Arnal, Ghitza, & Poeppel, 2014; Joachim Gross et al., 2013; Hyafil, Fontolan, Kabdebon, Gutkin, & Giraud, 2015; Luo & Poeppel, 2007; Mai, Minett, & Wang, 2016; Peelle, Gross, & Davis, 2013). Whereas some studies reported this brain-to-speech synchrony only for intelligible speech (Ahissar et al., 2001; Luo & Poeppel, 2007) others demonstrated that it occurred when speech was also unintelligible, yet stronger effects were observed for intelligible speech (Gross et al., 2013; Howard & Poeppel, 2010; Luo, Liu, & Poeppel, 2010; Park, Ince, Schyns, Thut, & Gross, 2015; Peelle et al., 2013; Pefkou, Arnal, Fontolan, & Giraud, 2017; Zoefel & Vanrullen, 2015). Although most work has been carried out on the theta band related to the syllable level, evidence for entrainment of gamma oscillations to speech has also been provided. Using continuous speech, Gross and colleagues (2013) demonstrated that low-frequency fluctuations in the speech envelope (3-7 Hz) modulated the amplitude of gamma activity (35-45 Hz) in auditory cortices, with a left hemisphere advantage. In agreement with previous studies (*e.g.*, Peelle, Gross, & Davis, 2013) this effect was stronger when speech was intelligible as compared to reversed (unintelligible) speech. The authors additionally described coupling between theta phase and gamma amplitude in bilateral auditory regions and left fronto-parietal cortex, which was again enhanced for intelligible speech (see also Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012). Such theta-gamma nesting may reflect a hierarchy of cortical oscillations in different frequency bands during speech decoding, allowing syllabic and phonemic parsing to be carried out simultaneously while remaining hierarchically bound (Ghitza, 2011; Giraud & Poeppel, 2012).

Overall, these findings have been taken to suggest that neural tracking of speech temporal (quasi-)regularities at different temporal scales on syllabic rhythm would favor the segmentation of the auditory input into "packets" of information, therefore facilitating linguistic processing (Giraud & Poeppel, 2012; Jones & Boltz, 1989), but also playing a key role in the acquisition and development of language (Benasich, Gou, Choudhury, & Harris, 2008; Goswami, 2011; Thompson et al., 2016). As previously mentioned, successful extraction of syllabic and prosodic rhythms is one important foundation for speech segmentation and phonological as well as morpho-syntactic processing (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Goswami, 2011, 2015; Goswami & Leong, 2013; Nazzi & Ramus, 2003). Strikingly, only few studies have investigated cortical entrainment to speech in children so far. In an EEG study conducted in 9-13 year-old children, Abrams and colleagues (2008) reported phase-locking between temporal cortical activity and amplitude envelope of the same repeated sentence in three modes of production (conversational, clear speech and artificially accelerated speech). In agreement with the AST model, right temporal cortex followed the contours of the envelope more precisely than the left. Power and collaborators (2012) also showed, in children aged 13 years old on average, entrainment of delta (1-3 Hz) and theta (3-5 Hz) oscillations to a rhythmic syllabic sequence presented in the auditory or audio-visual modality. Finally, a recent study (Thompson et al., 2016) in children aged 3-5 years old at rest demonstrated left-lateralized gamma oscillatory activity whereas theta endogenous activity was symmetrically distributed, thus only partially confirming the predictions of the AST model (Poeppel, 2003). The authors however suggested that later development of the right auditory cortex (Paus et al., 1999) may account for the lack of lateralization of theta band activity.

Recent models have emphasized that temporal sampling of speech by cortical oscillations may be defective in disorders of language learning such as dyslexia and DLD (Cumming, Wilson, & Goswami, 2015; Giraud & Poeppel, 2012; Goswami, 2011; Goswami & Leong, 2013). Atypical neural entrainment at one or more of the 'privileged' temporal rates for speech segmentation could indeed lead to a deficit in extracting syllabic and phonemic information, therefore resulting in inaccurate phonological representations. In the 'temporal sampling framework' (TSF), Goswami (2011) proposed that impaired sensitivity to amplitude envelope rise time and prosody reported in dyslexic children (Goswami et al., 2011) could be related to inefficient phase-locking between theta-band (and possibly delta-band) oscillations in the (right) auditory cortex and the slow amplitude modulations in speech temporal

envelope. Through theta-gamma coupling, altered neural entrainment at these low frequencies could then impact brain synchronization to faster phonemic modulations in the speech signal. Giraud and Poeppel (2012) extended this hypothesis to the phonemic level, which is known to be also impaired in dyslexia (*e.g.*, Serniclaes, Sprenger-charolles, Carré, & Demonet, 2001; Serniclaes, Van Heghe, Mousty, Carré, & Sprenger-Charolles, 2004), suggesting that gamma-band oscillatory activity in left temporal regions may also show atypical response to speech. These two models have received support from studies mainly conducted in dyslexic adults and using both speech and non-speech materials, with however conflicting results (Cutini, Szűcs, Mead, Huss, & Goswami, 2016; Hämäläinen, Rupp, Soltész, Szűcs, & Goswami, 2012; Lehongre, Morillon, Giraud, & Ramus, 2013; Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011; Lizarazu et al., 2015; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; Poelmans et al., 2012; Power, Colling, Mead, Barnes, & Goswami, 2016; Vanvooren, Poelmans, Hofmann, Ghesquière, & Wouters, 2014; Vos, Vanvooren, Vanderauwera, Ghesquière, & Wouters, 2017). In a study combining EEG and functional magnetic resonance imaging (fMRI) using audiovisual speech perception, Lehongre and coworkers (2013) demonstrated that whereas French control participants showed right-lateralized oscillatory activity only for the theta frequency range (4-7 Hz), French dyslexic adults showed the same pattern of right-hemisphere advantage for the delta (1-3 Hz), theta and gamma (25-35 Hz) bands. Additional meticulous analyses revealed that this neural sampling impairment was mainly related to a poorer sensitivity of the left auditory cortex to amplitude modulations matching the phonemic rate of speech, in agreement with the predictions by Giraud and Poeppel (2012) (see also Lehongre et al., 2011 and Poelmans et al., 2012 for similar results). Other studies in dyslexic adults and children reported a lack of brain-to-speech alignment in the delta frequency band (Hämäläinen et al., 2012; Lizarazu et al., 2015; Molinaro et al., 2016; Power, Mead, Barnes, & Goswami, 2013; Soltész, Szűcs, Leong, White, & Goswami, 2013). Using MEG, Molinaro and colleagues (2016) examined the oscillatory cortical dynamics of Spanish children (from 8 to 14) and adults with dyslexia as compared to controls when listening to naturally produced sentences. In both control and dyslexic groups, increased coherence between brain oscillations and speech was found in the delta (0.5-1 Hz) and theta bands (5.8-6.3 Hz); this effect was stronger in the right than in the left hemisphere. Dyslexic children and adults however showed reduced delta coherence, as compared to controls, in the right auditory cortex and left inferior frontal gyrus, whereas no difference between groups was seen for the theta band (contrary to the TSF predictions, Goswami, 2011).

These studies therefore suggest atypical oscillatory cortical dynamics in dyslexia, mainly in the delta and gamma frequency bands, which may result in poor prosodic and phonemic processing at the basis of the phonological and reading deficits characteristic of dyslexia. To the best of our knowledge, no study to date has investigated how brain oscillations synchronize with speech signal properties in children with DLD in children with DLD. As Cumming and collaborators (2015) pointed out in their ‘prosodic phrasing hypothesis’, atypical cortico-acoustic synchronization, similar to the one observed in dyslexia, could also account for the phonological and morpho-syntactic deficits observed in this developmental disorder. In particular, inefficient neural entrainment to the prosodic (~2 Hz) and syllabic (~4-5 Hz) rhythms of speech in children with DLD could be associated with insensitivity to syllable stress and rise time, affecting oral language development. In the present MEG study, we sought to shed light on the oscillatory cortical dynamics underlying natural speech perception in children with DLD as compared with typically developing (TD) children matched for age. We were particularly interested in how the developing brain tracks the amplitude envelope of speech when it is uttered at a normal or fast rate, given that adjusting to speech rate variations in daily life appears critical for language processing and development. In agreement with the assumption of Cumming et al. (2015), we hypothesized that children with DLD would show atypical brain-to-speech synchrony (in the form of reduced coupling and/or different cortical networks involved) as compared to TD children in the frequency band encompassing the sentences’ syllabic rate.

Materials and Methods

Participants

Twelve French-speaking children with DLD aged between 8 to 13 years old (2 girls; mean age: 10.29, SD: 1.54) participated in the experiment. These children were recruited from a neuropsychiatric hospital unit where there had been diagnosed using standardized French verbal and non-verbal tests. The inclusion and exclusion criteria were determined in collaboration with their neuropsychologists and paediatric neurologists. The inclusion criteria encompassed age range from 8 to 13 years old, French native language, right-handedness, lack of hearing problem, non-verbal IQ > 70 (as recommended in the classification by Rapin and Allen (1983) mostly used by French clinicians), and predominant impairment at the expressive level with rather preserved comprehension skills (i.e. phonological-syntactic

syndrome²⁴, Rapin & Allen, 1983). The exclusion criteria were bilingualism, the presence of attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder and/or intellectual disability. Prior to inclusion in the study, children's cognitive performance was re-evaluated with a battery of neuropsychological and French language tests in order to ensure the preservation of certain aspects of the diagnosis at the time of the experiment as well as to attest that comprehension was well-preserved (Table 1). Linguistic abilities were assessed with a series of tests from the BALE (Batterie Analytique du Langage Ecrit ; Jacquier-Roux, Lequette, Pouget, Valdois, & Zorman, 2010), including word, pseudo-word and non-word repetition, phonemic fluency, picture naming and metaphonological tasks (phoneme suppression and phoneme fusion). Oral comprehension and morpho-syntactic production were assessed with subtests from the ELO (Evaluation du Langage Oral; Khomsi, 2001). Finally, non-verbal abilities were examined with the Raven's coloured Progressive Matrices (Raven, Raven, & Court, 1998) and the forward and backward digit span tests from the BALE. Each child diagnosed with DLD was matched for chronological age and sex with a typically developing (TD) child (mean age: 10.83, SD: 1.65). TD children met the same inclusion and exclusion criteria as children with DLD except that they did not show any history of developmental language disorder. Verbal and non-verbal abilities of TD children were evaluated with the same battery of neuropsychological and French language tests (Table 1). All children (with DLD and TD) were furthermore screened for audiometric pure-tone thresholds (audiometer 600 M, Electronica Technologies) and none of them exceeded 30 dB Hearing Level (HL) at any frequency in the 250-6000 Hz range. Finally, each child's speech rate was measured in a narrative task using the picture book "Frog where are you?" (Mayer, 1969). Children's narratives were recorded with a microphone (audio-technica AT2020USB, 44 kHz). We then computed two rate measures for each child/story by calculating the number of syllables per second (1) including silences and pauses (corresponding to speech rate; Lacheret-Dujour & Beaugendre, 1999) and (2) after removing silence periods longer than 150 ms (corresponding to articulatory rate Grosjean & Deschamps, 1975; Pellegrino, Coupé, & Marisco, 2011). Whereas the two groups of children showed similar mean articulatory rates (3.84 syllables/s, SD: 0.53 for children with DLD vs 3.85 syllables/s, SD: 0.53 for TD children; Wilcoxon test, $p > .05$), children with DLD had a lower speech rate (1.81 syllables/s, SD: 0.43) than their TD peers (2.45 syllables/s, SD: 0.70; Wilcoxon test, $p = .02$). The

²⁴The phonological-syntactic syndrome is the most prevalent form of DLD (SLI) and is known to mainly affect verbal expression due a phonological programming disorder. The child is often hardly intelligible, shows altered phonology and tends to produce short and ungrammatical sentences (de Weck & Rosat, 2003; Rapin, 1996).

protocol was approved by the French ethical committee (Comité de Protection des Personnes Lyon Sud-Est II) and all children and their parents signed a consent form before the experiment.

Table 1: Mean standard deviations for several subtests of the BALE and the ELO, as well as for the Raven’s coloured Progressive Matrices and the forward and backward digit span tests are reported for children with DLD and their TD peers. Mean speech and articulatory rates (syllables/s) are also presented. Standard errors are indicated in brackets. Tests for which children deviated from 2 SD from the mean of the population are highlighted in bold.

	DLD	TD
VERBAL TESTS		
Word repetition	- 10.97(14.77)	- 0.24 (1.05)
Pseudo-word repetition	- 4.87 (3.99)	0.07 (0.54)
Non-word repetition	- 3.51 (2.42)	0.06 (0.67)
Phonemic fluency	- 0.64 (1.03)	-0.04 (1.04)
Picture naming	- 0.32 (1.41)	0.35 (0.76)
Oral comprehension		
sentence-to-picture matching	0.02 (1.01)	1.34 (0.78)
Morpho-syntactic production		
sentence completion	- 1.50 (1.23)	1.32 (0.14)
Metaphonological tasks		
phoneme suppression	- 1.05 (1.39)	0.56 (0.52)
phonemic fusion	- 1.46 (1.20)	0.32 (0.71)
NON-VERBAL TESTS		
Raven’s Progressive Matrices	95.75 (12.74) (range 71-114)	99.08 (9.19) (range 82-111)
Forward digit span	- 0.81 (0.97)	1.61 (1.37)
Backward digit span	- 0.38 (0.77)	- 0.11 (0.62)
Speech rate (syll/s)	1.78 (0.42)	2.49 (0.69)
Articulatory rate (syll/s)	3.89 (0.85)	3.87 (0.51)

Materials

The stimuli were identical as the ones used in Guiraud et al. (Guiraud, Hincapié, Jerbi & Boulenger, submitted). Three-hundred sentences (7-9 words), with the same syntactic structure (e.g., “Le public applaudit le joueur pour sa victoire” / The public applauds the player for his victory), were recorded by a French native male speaker (44.1 kHz, mono, 16 bits) in a sound-attenuated booth using ROCme! Software (Ferragne, Flavie, & Fressard, 2012). The speaker was asked to record each sentence first at a normal rate and then as fast as possible while

remaining intelligible. The duration of the whole set of sentences (2×300 ; mean duration: 1.68 ms, SD: 0.17 for normal rate sentences and 1.23 ms, SD: 0.14 for fast rate sentences) and the number of articulated syllables for each stimulus were calculated with Praat software (Boersma & Weenink, 2012), allowing to compute the speech rate in each condition. The mean normal speech rate was 6.61 syllables/s (SD: 0.47) and the mean fast rate was 9.03 syllables/s (SD: 0.56; Table 2). An 80 Hz high-pass filter was applied to all stimuli, the amplitude envelope was smoothed sentence-initially and finally, and the intensity of the sound files was peak normalized.

Two experimental lists of 300 items each (150 normal rate sentences and 150 fast rate sentences) were created from the 600 experimental sentences (300 x 2 rate variants). Each stimulus appeared in each rate condition across all participants but only once per list (i.e. no repetition effects). We also created 40 filler sentences (20 normal rate and 20 fast rate; not analysed) following the same syntactic structure but in which 539 ms-long sound effects (<http://www.sound-fishing.net/>) were superimposed at the end of the sentence. These 40 filler stimuli were added to each experimental list, leading to a total of 340 trials. Within each list, the 340 stimuli were pseudo-randomly divided into 10 blocks of 34 stimuli each (15 normal rate sentences, 15 fast rate sentences, 2 normal rate fillers, 2 fast rate fillers, with no more than 3 consecutive stimuli from the same condition) to allow for short breaks during the experiment.

In order to measure the coupling between the syllabic rate of the incoming speech signal and brain oscillatory activity, we extracted the amplitude envelope of each sentence in each condition using the method described in Peelle and coworkers (2013). The speech signal was rectified (full wave rectification) and low-pass filtered at 30 Hz (fifth order Butterworth filter, zero-phase digital filtering). This temporal envelope was used as the acoustic signal for coherence analyses.

Procedure

The experiment followed the same procedure as in Guiraud et al. (submitted). After a first session where each child underwent the battery of neuropsychological and language tests as well as the audiometry testing, the MEG experiment was carried out in a second session (another day). During the MEG recording, the children were comfortably lying in a sound-attenuated, magnetically-shielded recording room with a screen 40 cm in front of them. A

piece of foam placed between the back of their heads and the MEG helmet minimized head movements. Children were presented with binaural stimuli thanks to air-conducting tubes with foam insert earphones (Etymotic ER2 and ER3). Prior to the MEG recording, auditory detection thresholds were determined for each ear with a 1-minute pure tone (400 to 3000 Hz); the level was then adjusted so that the stimuli were presented at 50 dB Sensation Level with a central position (stereo) with respect to the children's head.

Children were instructed to listen carefully to the sentences while looking at a fixation cross on the screen, and press a button (button response Neuroscan) with their left index finger for each filler trial. To maintain their attention all over the experiment, they were first told that they would have to answer five questions pertaining to the heard sentences at the end of the experiment. Second, the experiment was presented as a game: whenever they correctly answered a filler trial, a piece of a treasure map displayed on the screen was uncovered (a negative feedback was on the other hand given to each incorrect response to remind the child of the task). At the end of the experiment, the children could pick up their gift thanks to the discovered map. A training phase (five sentences not included in the main experiment and not analyzed) was first proposed to children so that they could get used to the speaker's voice and experimental procedure. The experiment was run with Presentation software (Neurobehavioral Systems).

Each child was assigned with one of the two experimental lists. Each trial was organized according to the following sequence: after a white fixation cross presented on the screen for 1500 ms, a spoken sentence stimulus which the child had to carefully listen to was delivered. This was followed by an inter-trial interval (grey screen) of 1250 ms. If the child heard a filler sentence, she/he had to press the button response as quickly as possible; a positive or negative feedback was then displayed on the screen for 6250 ms. After a delay (grey screen) of 500 ms, the next trial could then begin. The whole experiment lasted 50 minutes (each of the 10 blocks (34 stimuli) lasted 230 seconds (< 4 minutes) with several short breaks).

MEG data collection

MEG signals were acquired on a helmet-shaped 275-channel whole head sensor system (OMEGA MEG CTF 275, VSM MedTech, Canada) in CERMEP, Lyon (France, www.cermep.fr). The data were continuously recorded at a sampling frequency of 1200 Hz.

Three fiducial coils (nasion, left and right pre-auricular points) were placed on each participant to determine head position within the helmet. Reference head position was recorded before and after the experiment. Head movements were checked throughout the experiment using continuous head position identification (HPI). Maximum head movements were computed and children with movements larger than 1 cm from the head reference position were not included in the analysis ($n = 1$). The results are thus presented for 11 children with DLD and 11 TD children. Figure 1 provides an overview of the different steps of data collection and analysis.

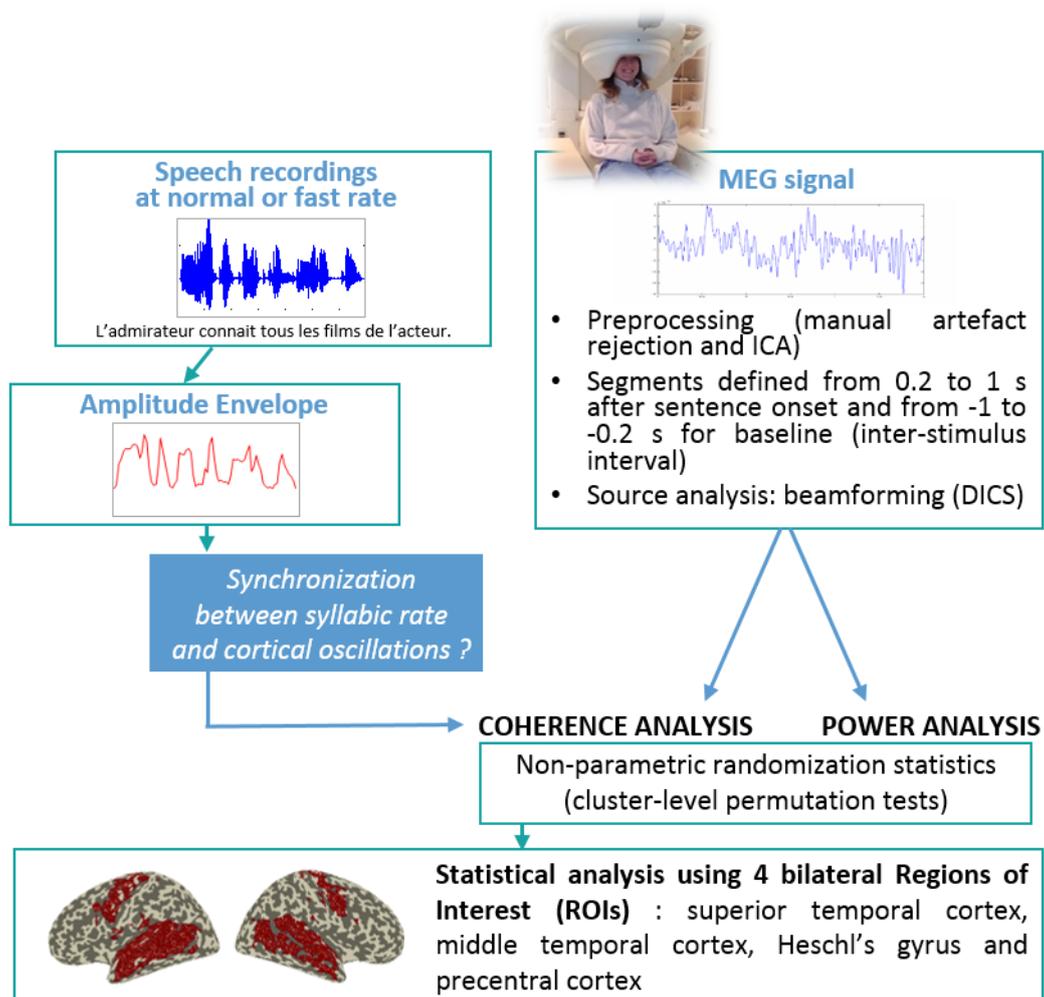


Figure 1 : Overview of MEG data analysis.

MEG data analysis

All analyses were performed with Matlab (Mathworks Inc., MA, USA) and the Fieldtrip open-source software (Oostenveld, Fries, Maris, & Schoffelen, 2011). Data were down-sampled to 300 Hz and artefacts from the mains power supply at 50, 100 and 150 Hz were

removed using discrete Fourier transform. Signal slow drifts were also suppressed. We then segmented the data into periods of 2.5 s (1.5 s after sentence onset and 1 s before for the baseline) and cleaned them first manually and then semi-automatically to remove jumps and muscular artefacts. Finally, we applied independent component analysis (ICA) to correct for heart beat artefacts and eye movements (20 ICA components).

Source localization

A template child brain (age 7.5-13.5 years old, Fonov et al., 2010; Fonov, Evans, McKinstry, Almlie, & Collins, 2009), realigned to the participants' fiducial points, was used to create individual head models. The homogeneity of head circumference across children was verified through measurements performed before each recording session (average head circumference: 52.75 cm, SD: 1.38).

Source analysis was carried out in a common space and we used an adult atlas for statistics by Regions Of Interest (ROIs). To this aim, we constructed a source model in a common coordinate system (MNI) from an adult template (common grid with 5780 vertices). Individual children's head model was then transposed to the MNI brain. We chose to use an adult template because at 8-13 years old, the brain of a child is anatomically close to that of an adult, having reached 90% of its maximum size by the age of 6 (Giedd, 2004; Shapiro & Janzen, 1960).

We measured power and coherence between MEG signals and speech temporal envelope (i.e. cortico-acoustic coherence) at the source level from the frequency domain beamformer method (DICS; Gross et al., 2001). Coherence refers to the linear relationship between the amplitude of two signals in the frequency domain, with values ranging from 0 to 1. It is mathematically defined by:

$$Coh(r_{ef}, r_c, f) = |Cs(r_{ef}, r_c, f)|^2 / (Cs(r_{ef}, r_{ef}, f) Cs(r_c, r_c, f))$$

Where r_{ef} denotes the reference signal (i.e. amplitude envelope of the speech signal); r_c refers to the signal at each vertex of the anatomical grid estimated with DICS; f is the frequency bin and Cs is the cross-spectral density matrix. Coherence in each frequency band was computed by averaging the coherence values obtained over all frequency bins within the band of interest.

For each trial of each speech rate condition, an active period (0.2 to 1 s after sentence onset) and a baseline period (inter-stimulus interval; -0.2 to -1 s before onset) were defined. The cross spectral density matrix was computed using fast Fourier transform (FFT) and a multitaper-based spectral smoothing with ± 1 Hz was used to smooth for each frequency band (cf. Table 2). We used the standard FieldTrip settings for the DICS algorithm (regularization parameter is the largest singular value of the cross spectral density matrix). Coherence between the MEG signals and speech envelope was calculated across all active and baseline periods for each rate condition and frequency band of interest. Finally, averages over each group (DLD and TD) in each of the speech rate conditions were computed. Between-group comparison statistics were performed with non-parametric randomization Monte-Carlo tests (1000 repetitions) using Fieldtrip. Results were corrected with cluster-level permutation tests ($\alpha = .05$). The same methods (FFT, multitapering and non-parametric statistics) were also used to compute and assess task-based spectral power modulations.

Cortico-acoustic coherence and power modulations during normal and fast rate sentence perception were examined in two analyses carried out in different frequency bands. In the first analysis, the theta band (4-7 Hz), which matches the slow fluctuations typically found in regular speech envelope (Pelle and Davis, 2012; Thompson et al., 2016) and has been used in previous studies (*e.g.*, Luo & Poeppel, 2007, 2012; Pelle et al., 2013; see Giraud & Poeppel, 2012 for a model), was chosen as the band of interest. Second, as we were interested in the neural alignment to sentences' syllabic rate, we defined two narrower frequency bands centred on the mean calculated rates (± 1 Hz) in each condition: 5.6-7.6 Hz for the normal condition (mean rate = 6.61 syllables/s) and 8-10 Hz for the fast condition (mean rate = 9.03 syllables/s). For each frequency band and condition, we performed the coherence and power analyses in four bilateral ROIs, defined according to models of speech perception (Giraud & Poeppel, 2012; Hickok & Poeppel, 2007) and consistent with results we observed in the first theta band (4-7 Hz) coherence analyses. These ROIs included Heschl's gyrus, superior temporal gyrus, middle temporal gyrus and precentral cortex, following the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002, see Figure 1). Finally, we conducted Spearman's rank correlations and regression analyses between ROI-based coherence values for which differences between children with DLD and TD children were found, and (i) children's performance in the verbal standardized tests and (ii) children's own speech rate.

Table 2 : Mean syllabic rate (syllables/s) of sentence stimuli used in the normal and fast rate conditions. The frequency bands defined for MEG coherence and power analyses were centered on these values (± 1 Hz).

	<i>Speech Rate Conditions</i>	
	<i>Normal</i>	<i>Fast</i>
<i>Mean syllabic rate (syllables/s)</i>	6.61 (SD 0.47)	9.03 (SD 0.56)
<i>Syllabic-rate-centered frequency bands (Hz)</i>	5.6 - 7.6	8 - 10

Results

Theta-band neural entrainment to speech amplitude envelope

Our first analysis compared cortico-acoustic coherence in the theta frequency range (4-7 Hz) between children with DLD and TD children. As shown by the left panel of Figure 2C (see also Figures 2A and B), the between-group comparison for the normal rate condition revealed significantly stronger brain-to-speech coupling in TD children than in children with DLD in the right anterior middle temporal gyrus corresponding to the associative auditory cortex (Figure 2E). Note that in the uncorrected statistical analyses (Figure 2D), these drops in speech-brain coherence were observed bilaterally as well as in the right frontal cortex. Interestingly, the uncorrected results also depict an abnormal increase in coherence in the right posterior middle temporal cortex in the DLD group compared to controls. When listening to fast rate speech, no significant difference between the two groups was found in this frequency band. Uncorrected results for this condition however suggested enhanced coherence in the left premotor and primary motor cortex as well as in Broca's area in TD children, whereas the group with DLD tended to show stronger coherence in the right posterior superior and middle temporal gyri, in the vicinity of the temporo-parietal junction (Figure 2D). Of course, uncorrected statistics represent a tendency in the data and need to be interpreted with caution.

Regarding power modulations analyses (Figure 2, right panel), between-group comparisons revealed significantly increased power (i.e. more synchronization) at 4-7 Hz in the right anterior middle temporal cortex in children with DLD as compared to TD children in the normal rate condition (Figures 2C and E; stronger power modulations (i.e. desynchronization) were also seen in the posterior part of the middle temporal cortex for

uncorrected results in this condition, Figure 2D; see also Figure 2A and B, right panel). No significant difference was found between the two groups in the fast rate condition; however, when looking at uncorrected results, reduced theta-band power synchronization seemed to emerge in bilateral posterior middle temporal cortex in children with DLD (Figure 2D). Interestingly, the brain regions that exhibit modulations in cortico-acoustic coherence differ from those that show power modulations. This indicates that the speech-brain interactions observed are predominantly phase-based effects that are not merely driven by changes in signal power.

The results of this first analysis in the canonical theta-band suggest atypical entrainment of theta-band cortical oscillations in right (possibly bilateral) auditory regions to speech amplitude envelope in children with DLD when speech is uttered at a normal rate. When speech rate is naturally accelerated (although the results were only observed at $p < 0.05$ uncorrected), different neural networks with different patterns of hemispheric lateralization are involved: left (pre)motor regions in TD children synchronize to the fast rate speech signal whereas this is observed in right posterior temporal regions in children with DLD.

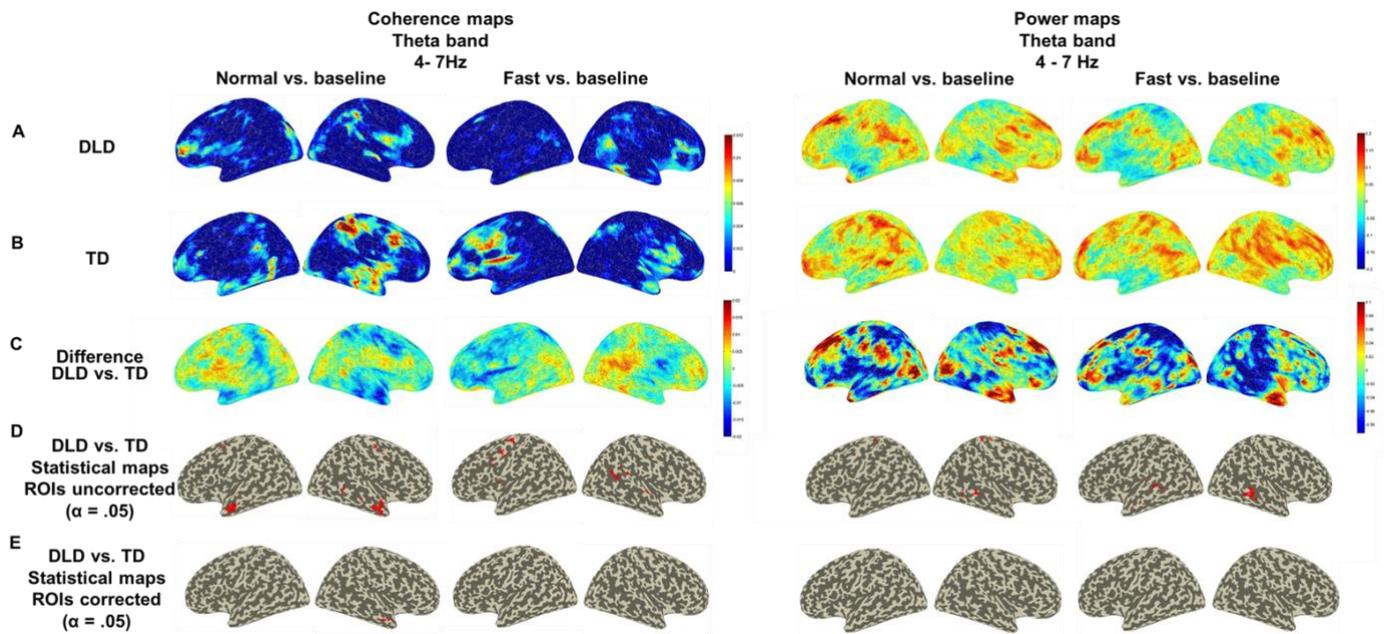


Figure 2 : Theta band analysis (4-7 Hz). (Left) Coherence maps between speech amplitude envelope and cortical oscillations in the two speech rate conditions (normal and fast) against baseline and (Right) Power modulations maps in the same conditions as compared to baseline (relative power) for (A) children with DLD, (B) TD children, (C) the contrast between the two groups. Uncorrected (D) and corrected (E) statistical maps ($\alpha = .05$) for ROI-based analyses are also presented. Coherence maps values for DLD (A left) and TD (B left) go from 0 (blue) to 0.012 (red) and from 0 (blue) to 0.02 (red) for the difference coherence map (C left). Power maps values for DLD (A right) and TD (B right) go from -0.2 (blue) to 0.2 (red) and from -0.1 (blue) to 0.1 (red) for the contrast maps (C right).

Neural entrainment to specific syllabic rate

To investigate how the brain of children with typical and atypical language development tracks syllabic information in natural speech, we compared, between the two groups, the coherence between speech envelope and MEG signals in frequency bands centred on the syllabic rate of our stimuli (see Method and Figure 3). The left panel of Figure 3 illustrates the results of the ROI-based analysis in the 5.6-7.6 Hz frequency band corresponding to the mean normal syllabic rate. In the normal rate condition, uncorrected statistics revealed stronger brain-to-speech coupling in the left anterior superior temporal cortex in TD children than in children with DLD (Figure 3D). No significant increase of coherence in the 5.6-7.6 Hz range was seen in the fast rate condition in any group. Note that none of these results survive the statistical correction.

The coherence analysis in the frequency band centred on the fast syllabic rate (8-10 Hz, Figure 3, right panel) suggested tiny differences between groups only when looking at uncorrected statistics. In the fast speech condition, cortico-acoustic coherence was slightly higher in the right dorsal premotor cortex and right superior temporal cortex in children with DLD as compared with TD children, and in the right posterior middle temporal gyrus in TD children (Figure 3D). During normal rate speech perception, coherence seemed stronger in the right anterior temporal cortex and right (more ventral) premotor cortex in children with DLD than in TD children.

Results of the power analyses carried out for each condition, as compared to baseline, are presented in Figure 4 (left panel for normal rate speech and right panel for fast rate speech). In the two frequency bands, differences in power modulations were observed between the two groups only for uncorrected results (Figure 4D). Children with DLD tended to show stronger oscillatory desynchronization in the right anterior superior temporal gyrus at 5.6-7.6 Hz for normal rate sentences and at 8-10 Hz for normal and fast rate sentences (see also Figures 4A and B). Higher (uncorrected) desynchronization for this group of children was also observed for fast rate speech in the left superior temporal cortex in the 5.6-7.6 Hz frequency range.

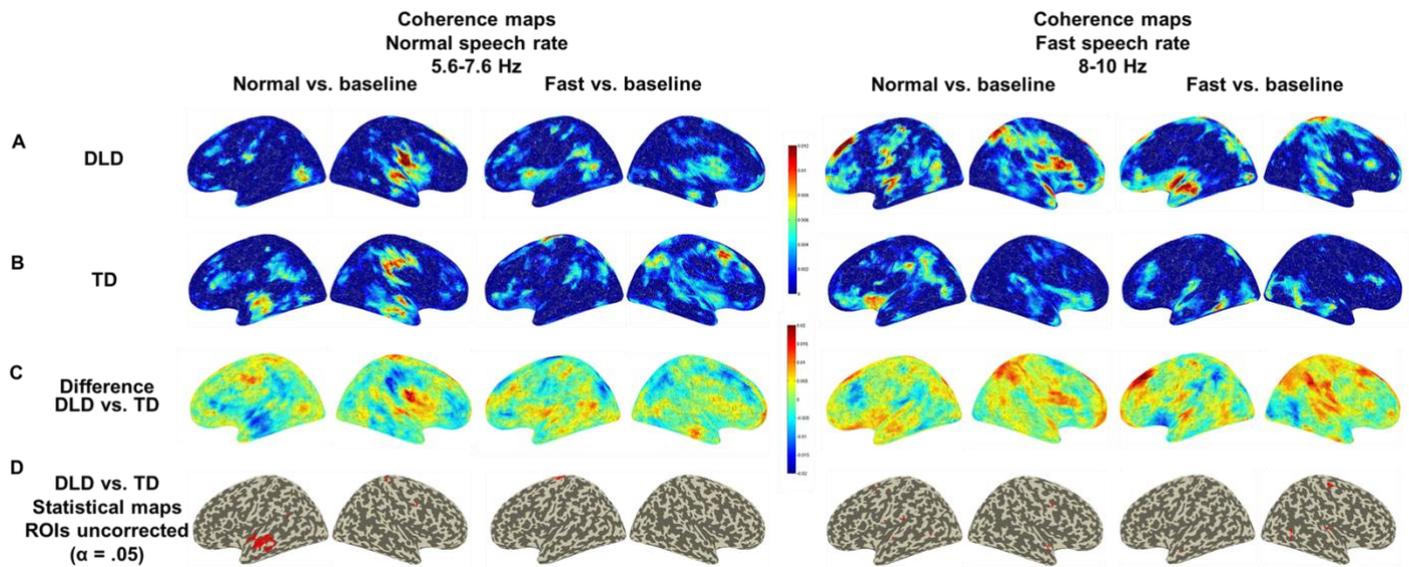


Figure 3 : Coherence maps between speech amplitude envelope and cortical oscillations in the two speech rate conditions (normal and fast) against baseline, in the frequency band centered on the mean syllable rate of normal speech (5.6-7.6 Hz; left panel) and fast speech (8-10 Hz; right panel) for (A) children with DLD, (B) TD children, (C) the difference between the two groups. Figures in D show the uncorrected statistical maps for the ROI-based analysis ($\alpha = .05$; note that corrected maps are not presented here as they did not reveal any significant group difference). Coherence maps values for DLD (A) and TD (B) go from 0 (blue) to 0.012 (red) and from 0 (blue) to 0.02 (red) for the DLD vs. TD difference coherence maps (C).

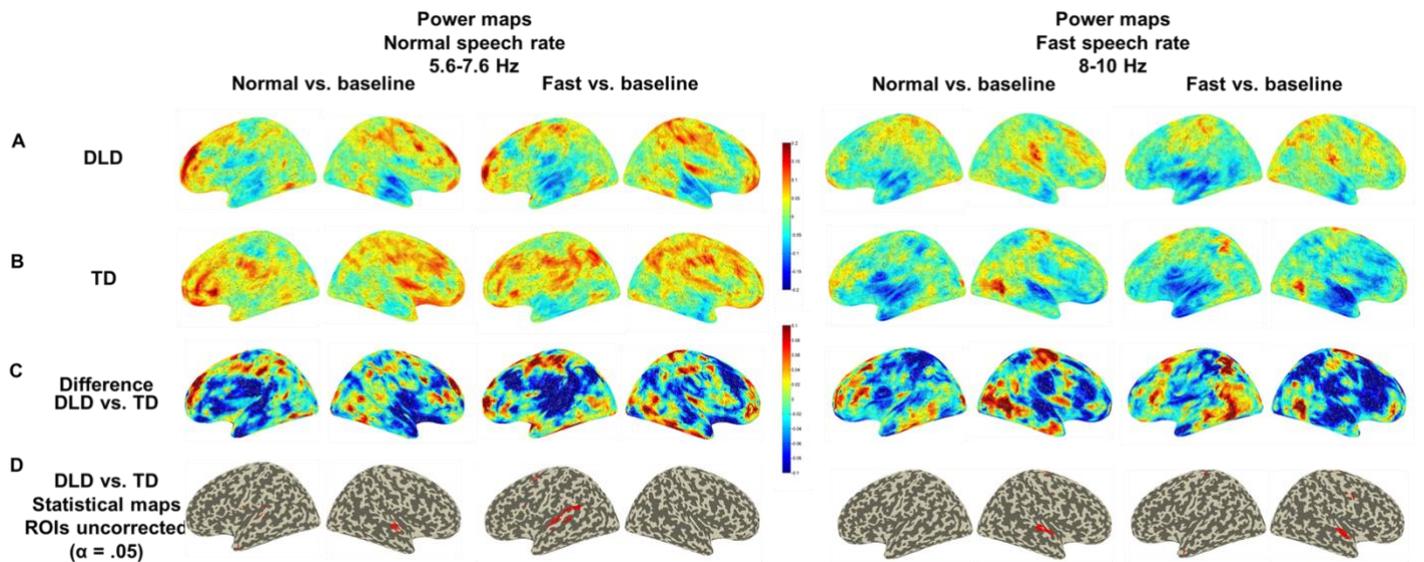


Figure 4 : Power modulation maps for the two speech rate conditions (normal and fast) compared to baseline (relative power), in the frequency bands centered on the mean syllable rate of normal speech (5.6-7.6 Hz; left panel) and fast speech (8-10 Hz; right panel) for (A) children with DLD, (B) TD children, (C) the contrast between the two groups of children. (D) Uncorrected statistical maps ($\alpha = .05$) are also presented (note that corrected maps are not presented here as they did not reveal any significant group difference). Power maps values for DLD (A) and TD (B) go from -0.2 (blue) to 0.2 (red) and from -0.1 (blue) to 0.1 (red) for the DLD vs. TD maps (C).

Correlations between speech rate, phonology and neural oscillations during speech perception

To explore the link between linguistic abilities and the extent of speech-brain coupling, we computed Spearman's rank correlations between cortico-acoustic coherence and speech abilities across individuals. More specifically, we computed correlations between cortico-acoustic coherence in regions where between-group differences were found, for the two speech rate conditions and the three frequency bands of interest, and (1) children's own speech rate in a narrative task and (2) phonological abilities as assessed by different standardized tests (word, pseudo-word and logatome repetition, phoneme suppression and phoneme fusion; see Methods) for each participant. These computations were done separately for children with DLD and TD children.

Whereas no significant correlation was observed for any of the variables tested in the control group, several significant correlations were found in the group of children with DLD.

As shown in Figure 5, cortico-acoustic coherence in the theta band (4-7 Hz) in the right posterior superior and middle temporal gyri for fast rate speech correlated with, respectively, the speech rate of children with DLD ($r = 0.91$, $p < .001$, Figure 5A) and their phonological performance ($r = -0.75$, $p < .01$; Figure 5B)²⁵. Hence, the more children with DLD spoke faster in the narrative task, the more the right posterior temporal cortex aligned its theta oscillatory activity to amplitude envelope of fast rate sentences. In addition, the more impaired they were at phonological tasks, the more coherence in this region increased. Interestingly, right posterior temporal cortex is the region where children with DLD exhibited increased neural theta (4-7 Hz) entrainment to fast rate speech as compared to TD children (see Figure 2E).

Finally, a significant negative correlation between children's own speech rate and coherence was also found in the 5.6-7.6 Hz range (corresponding to the mean normal syllabic rate of the stimuli) for the normal rate condition in the right superior temporal cortex ($r = -0.65$, $p < .05$, Figure 5C) namely in the region where children with DLD showed less oscillatory power in the same frequency band (as well as at 8-10 Hz) and condition as compared to their peers (see Figure 3C right panel).

²⁵ Note that the correlation was still significant ($r = -0.88$; $p = 0.0016$) even when excluding the two children with DLD with very low phonological scores.

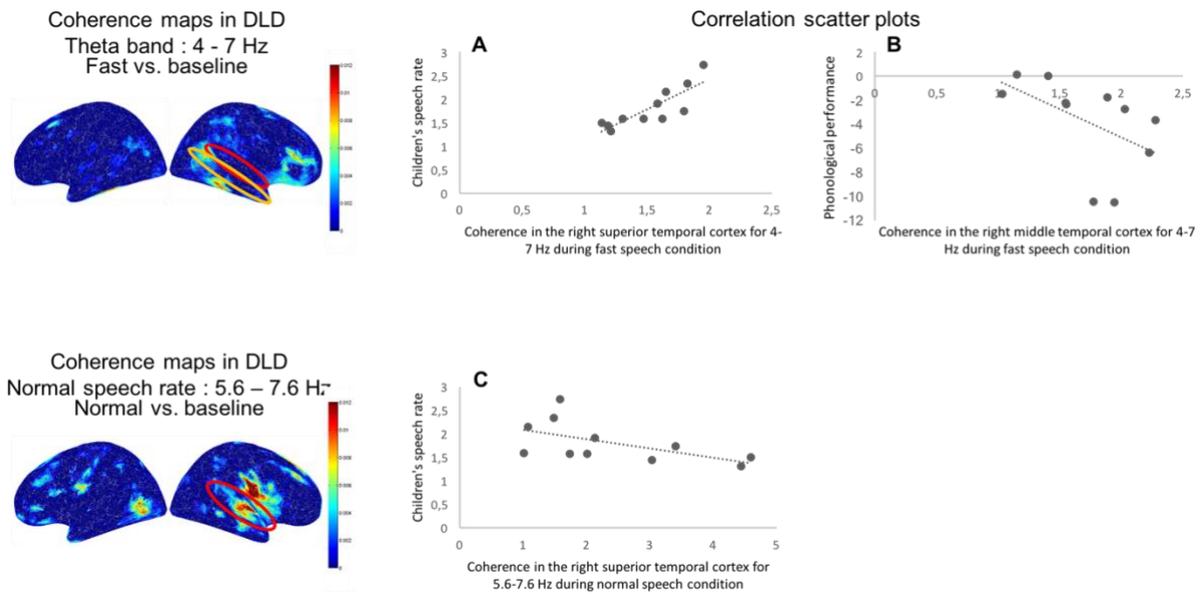


Figure 5 : Spearman's rank correlations (ρ) between theta-band (4-7 Hz) coherence values (A) in the right posterior superior temporal cortex (red circle) for fast rate speech and the own speech rate of children with DLD in a narrative task and (B) in the right posterior middle temporal cortex (orange circle) for fast rate speech and children's phonological performance. (C) Correlation between cortico-acoustic coherence in the right superior temporal cortex in the 5.6-7.6 Hz frequency range for normal rate speech and children's speech rate. No significant results were found with the same analysis in the TD group, we therefore only present the results in the DLD group here. All values correspond to z-scores. On coherence maps, values go from 0 (blue) to 0.012 (red).

To furthermore explore the observed correlations, a series of fixed-order multiple regressions was carried out in children with DLD. A set of regressions including (in a fixed order) age, non-verbal IQ (scores at the Raven's progressive matrices) and cortico-acoustic coherence values in the right superior and middle temporal cortex as dependent variables and, children's own speech rate and phonological performance as independent variables were calculated. The resulting parameter estimates are displayed in Table 3 and Figure 6 along with the unique variance accounted for by each variable (ΔR^2) for each stage of the model. Only one model was significant: theta (4-7 Hz) entrainment in the right posterior superior temporal cortex to fast rate speech was a significant predictor of children's speech rate ($p = .02$), explaining 60% of unique variance.

Table 3 : Hierarchical regression analysis in children with DLD exploring the unique variance in coherence in the right posterior superior temporal cortex accounted for by the children's speech rate, with age and non-verbal skills fixed.

	β	ΔR^2
Models : Coherence at 4-7 Hz for fast rate speech in the right posterior superior temporal cortex		
1. Age	-0.0453	-0.03
2. Raven's matrices scores	0.1640	0.14
3. Children's speech rate	0.6198	0.60

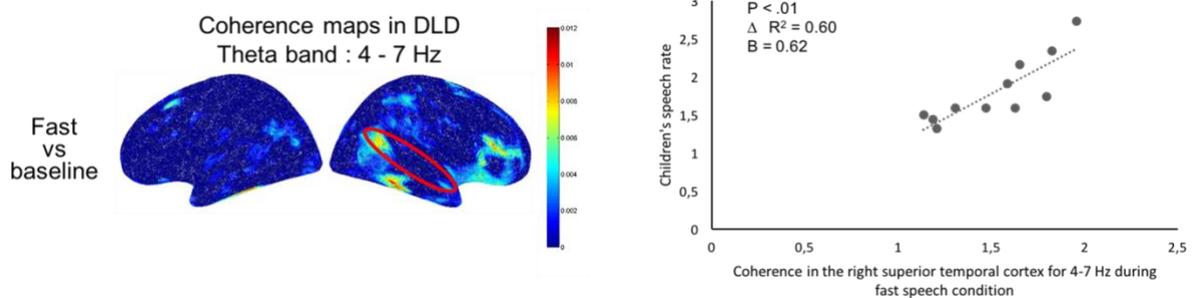


Figure 6 : Hierarchical multiple regression exploring the unique variance in maximum coherence measures (circled in red in coherence maps) in the theta band (4-7 Hz) in the right posterior superior temporal cortex when listening to fast rate sentences explained by the own speech rate of children with DLD in a narrative task. All values correspond to z-scores. On coherence maps, values go from 0 (blue) to 0.012 (red).

Discussion

Increasing evidence points to a perceptual sensitivity deficit to acoustic rhythmic information in speech signal in children with DLD, which may account for the phonological and morpho-syntactic deficits characteristic of this developmental disorder (Corriveau et al., 2007; Cumming, Wilson, & Goswami, 2015). In parallel, it has been convincingly demonstrated that early-step speech processing is governed by entrainment of oscillatory activity in auditory cortex to the slow modulations of speech amplitude envelope, which convey speech rhythm (syllabic and prosodic information; (Ghitza, 2011; Giraud & Poeppel, 2012). Such brain-to-speech alignment in the theta frequency band would enable the

segmentation of the incoming speech signal into syllable-sized segments, thus facilitating speech decoding (Peelle et al., 2013). In the present MEG study, we provide first preliminary evidence to our knowledge in children with DLD for atypical synchronization of theta-range neural oscillations in auditory regions to naturally produced speech. In addition, we demonstrate that whereas left (pre)motor regions entrain their theta oscillatory activity to speech when produced at a fast rate in typically-developing children, this is not the case in children with DLD.

Atypical theta-band oscillatory entrainment to speech envelope in children with DLD

In the normal rate speech condition, stronger cortico-acoustic coherence in the theta band (4-7 Hz) was found in TD children as compared to children with DLD in the right associative auditory cortex (BA22). Results in TD children agree with multi-time resolution models of speech perception (Giraud & Poeppel, 2012; Poeppel, 2003; Poeppel, Idsardi, & van Wassenhove, 2008) as well as with previous studies in adults showing right-lateralized neural entrainment in the theta frequency range to slow acoustic modulations (Hertrich, Dietrich, Trouvain, Moos, & Ackermann, 2012; Howard & Poeppel, 2010; Peelle et al., 2013). Studies in children and infants mostly using simple verbal and non-verbal stimuli have also reported such a pattern of hemispheric specialization (Abrams et al., 2008; Boemio, Fromm, Braun, & Poeppel, 2005; Molinaro et al., 2016; Telkemeyer et al., 2009, 2011; Vanvooren et al., 2014). In their EEG study in 9-13 year-old children, Abrams and colleagues (2008) demonstrated that right auditory cortex phase-locked more precisely to the amplitude envelope of the same repeated sentence than its left counterpart. Vanvooren and coworkers (2014) similarly described stronger cortical responses to amplitude-modulated noises with a modulation frequency of 4 Hz (~ syllabic rate) in the right hemisphere of children aged 5 years old. Our results obtained with naturally produced sentences show that theta neural entrainment in right auditory cortex to envelope modulations is reduced in a group of children with DLD as compared to their peers. As oscillatory activity in this region is assumed to parse the incoming speech signal into syllable-length units (Ghitza, 2011; Poeppel, 2003), the present data suggest that right auditory regions in children with DLD may be less efficient in tracking the low frequency fluctuations in the incoming signal. In other words, theta oscillations would not as capable as in TD children to align to speech temporal (quasi-)regularities at the syllable level, resulting in poorer extraction of rhythmic information. This result fits nicely with the “prosodic phrasing hypothesis” (Cumming, Wilson, & Goswami,

2015) and the previously reported deficits in children with oral language disorders to process rise time of the amplitude envelope (Beattie & Manis, 2012; Corriveau et al., 2007; Richards & Goswami, 2015). To our knowledge, no previous work has examined brain-to-speech alignment in children with DLD, therefore not allowing any direct comparison. In a study comparing poor readers and typically developing readers (aged 9-15 years old), Abrams and collaborators (2009) reported reduced phase-locking between right temporal evoked responses and speech amplitude envelope in poor readers. This was nevertheless only observed when speech was artificially accelerated (by time compression) and thus more difficult to decode. Moreover, poor readers showed later cortical responses in the right hemisphere together with earlier responses in the left as compared to good readers. It is worth mentioning that in the study by Abrams et al., the difference between poor and good readers only emerged for time-compressed speech and not for normal speech, which seems at odds with our result of lack of right-lateralized theta entrainment to speech envelope already in the normal rate condition. Methodological issues may account for this discrepancy. First, we included children with diagnosed oral language disorders whereas children with poor reading abilities (based on word and non-word reading scores) but without dyslexia participated in the experiment by Abrams et al. Second, children in the present experiment listened to 150 different sentences in each rate condition, each sentence being heard only once. On the contrary, Abrams et al. (2008) presented the exact same sentence (“The young boy left home”) either in its clear, conversational or time-compressed version (1000 repetitions each). It is thus not altogether impossible that right temporal regions of poor readers aligned to speech envelope when speech was presented at a normal rate in their work because repetition of the same sentence made processing easier, whereas in our study the right auditory cortex of children with DLD had to align each time to a different amplitude envelope. As mentioned in the Introduction, previous studies examining brain-to-speech synchrony in developmental language disorders have mostly included adults (and children) with dyslexia. Evidence for atypical (i.e. reduced and/or differently-lateralized) synchronization of cortical oscillations to speech envelope has been provided in the delta but not in the theta range so far (e.g., Hämäläinen, Rupp, Soltész, Szücs, & Goswami, 2012a; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016). In the present study, given that our shortest sentences (in the fast rate condition) were 1 second long, we could not examine delta cortico-acoustic coherence properly. Future investigations using continuous speech (e.g., Gross et al., 2013) would be necessary to investigate whether delta-band oscillations also align atypically to speech in children with DLD.

Cortico-acoustic coherence analyses in the fast rate condition only showed differences between groups for uncorrected statistics. Although no clear conclusion can be drawn from the current data, they however reveal an interesting pattern which deserves to be discussed. Whereas TD children showed increased neural entrainment to amplitude envelope of fast rate sentences in left premotor/motor regions and Broca's area, no such alignment in articulatory regions was observed in children with DLD. Previous EEG/MEG or fMRI studies in adults have demonstrated recruitment (pre)motor regions or enhanced connectivity between these regions and auditory cortex during challenging speech processing, such as noise-degraded or time-compressed speech (Adank & Devlin, 2010; Alho et al., 2014; Du, Buchsbaum, Grady, & Alain, 2014; Osnes, Hugdahl, & Specht, 2011). Peelle and colleagues (2013) also reported increased coherence between oscillatory activity in bilateral motor cortex of adults and amplitude envelope of noise-vocoded unintelligible sentences. In agreement with this, our results in TD children suggest that cortical oscillations in left (pre)motor regions align to speech envelope when speech is uttered at a fast rate and thus more difficult to decode. Such a pattern concurs with dual-stream models of speech perception (Hickok & Poeppel, 2007; Rauschecker, 2011; Rauschecker & Scott, 2009; Rauschecker & Tian, 2000) which suggest that a dorsal stream, encompassing temporo-parietal as well as ventral and dorsal premotor regions, underlies speech perception particularly under degraded acoustic conditions (see also Iacoboni, 2008). Our finding in TD children furthermore agrees with studies in adults showing alignment between endogenous theta oscillations in the temporal and (pre)motor cortices (Giraud et al., 2007; Morillon et al., 2012) as well as revealing a role of motor areas in predictive timing of auditory rhythmic information (Arnal, Doelling, & Poeppel, 2015; Morillon, Schroeder, & Wyart, 2014; see Morillon & Schroeder, 2015 for a review). Hence, in TD children, tracking of speech low frequency (syllable-rate) fluctuations could be achieved in motor regions involved in syllable production when speech perception becomes more difficult such as in the case of fast rate speech.

Our group of children with DLD did not exhibit such theta entrainment in left motor regions to fast rate speech. Again, although we need to remain cautious, this might indicate atypical function of the dorsal stream in this population. In a behavioural study (Giraud et al., under revision), we reported poorer performance of children with DLD to understand sentences naturally produced at a fast rate as compared to normal rate sentences and TD children. The present MEG data suggest that this difficulty may be underpinned by inefficient synchronization of theta-band motor cortical oscillations to syllable-rate information in

speech amplitude envelope. Studies investigating rhythmic processing in children with DLD have reported impairment both for the perception of auditory rhythmic information and the production of rhythm, in tapping tasks for instance (Corriveau & Goswami, 2009; Cumming, Wilson, Leong, Colling, & Goswami, 2015; Goswami, 2012; Hill, 2001; Thomson, Fryer, & Maltby, 2006; Thomson & Goswami, 2008). As synchronizing to a beat involves precise temporal auditory-motor coupling (Grahn & Brett, 2007; Tierney & Kraus, 2014), it has been hypothesized that such coupling may be impaired in children with DLD (Corriveau & Goswami, 2009). This may also account for the phonological deficits observed in these children given that rhythmic perception and production are fundamental for language acquisition (see Goswami, 2011; Tierney & Kraus, 2014). To develop efficient phonological representations, the child has indeed to align the timing of perceived auditory information with motor regions implied in speech production. In this view, our data of a lack of (pre)motor cortical alignment to fast rate speech modulations may be indicative of atypical coupling between auditory temporal regions and motor frontal regions in developmental language disorders. Further evidence including a larger sample of participants as well as analyses of auditory-motor connectivity during speech perception in children with DLD is however needed to corroborate this interpretation.

It is also worth mentioning that uncorrected results unexpectedly showed increased cortico-acoustic coherence in the right posterior superior and middle temporal cortex, close to the temporo-parietal junction, in children with DLD in the fast rate condition as compared to TD children (note that this was also observed, albeit to a lesser extent, in the normal rate condition). Bilateral (with sometimes a left-hemisphere bias) posterior superior temporal cortex has been consistently found to be involved in phonological processing both during speech perception and production as well as during active maintenance of phonemic information (Buchsbaum, Hickok, & Humphries, 2001; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Hickok & Poeppel, 2000, 2004; Papathanassiou et al., 2000). The temporo-parietal junction is furthermore thought to underlie speech sensorimotor integration, during which phonological information would be compared with corresponding articulatory-based representations *via* a mechanism of forward mapping, which may be especially important in challenging acoustic conditions (Hickok & Poeppel, 2004, 2007; Rauschecker & Scott, 2009). In line with this, Alho and colleagues (2014) demonstrated using MEG enhanced neural synchrony between the left temporo-parietal junction and left ventral premotor cortex during listening to noisy syllables as compared to intact stimuli. Interestingly, several

neuroanatomical studies in children with DLD have revealed atypical rightward asymmetry in perisylvian areas as compared to children with typical development (e.g., Jernigan, Hessenlink, Sowell, & Tallal, 1991; Plante, Swisher, & Vance, 1989; Plante, Swisher, Vance, & Rapcsak, 1991). In other words, (at least some) children with DLD exhibit a right perisylvian area that is equivalent in size or larger than its left counterpart. Symmetrical functional responses, namely reduced or absent left hemispheric lateralization with sometimes a bias toward the right hemisphere, have also been described in children with DLD in various language tasks (Badcock, Bishop, Hardiman, Barry, & Watkins, 2012; de Guibert et al., 2011; see Whitehouse & Bishop, 2008 in adults with DLD). Hypoactivation of the left temporo-parietal region (among others) has for instance been shown during phoneme discrimination (Tzourio, Heim, Zilbovicius, Gerard, & Mazoyer, 1994) or in a word definition task (de Guibert et al., 2011). Most remarkably, Cutini and colleagues (Cutini, Szucs, Mead, Huss, & Goswami, 2016) recently revealed atypical entrainment to a 2 Hz amplitude-modulated noise in dyslexic children with respect to controls (mean age 12 years 11 months) in the form of stronger hemodynamic response in the right supramarginal gyrus ; increased activation was also reported in the angular gyrus, although bilaterally. The authors interpreted this effect as reflecting atypical function of the dorsal stream in dyslexia, in agreement with studies showing increased connectivity between right supramarginal gyrus and left posterior temporal cortex (Wernicke's area) in teenagers born prematurely and thus known to be at higher risk for developing language disorders (Myers et al., 2010). Accordingly, although it remains speculative, the finding that our group of children with DLD showed increased cortico-acoustic coherence in the *right* posterior superior and middle temporal cortex, extending to the temporo-parietal region (see Figure 2A), for fast rate speech might therefore indicate that auditory-motor integration occurs to some extent in children with DLD but in an atypical (i.e. right-lateralized) way, thus suggesting atypical functional organization of the dorsal stream during speech perception (in line with our results of a lack of coherence in left premotor regions in the same condition). Such an interpretation agrees with the correlation analyses showing that cortico-acoustic coherence to fast rate speech in this same right posterior temporal region correlated positively with children's own speech rate and negatively with their phonological performance. If activity in posterior superior and middle temporal (and parieto-temporal) cortex during fast speech perception can be thought of as reflecting auditory-motor mapping, the stronger theta-band brain-to-speech synchrony in this region for children with DLD who spoke faster in the narrative task suggests that auditory-motor integration occurred especially for the children having better production skills and thus

possibly better articulatory representations. On the other hand, the second correlation showed that the poorer children with DLD performed in phonological tasks, the more cortico-acoustic coupling increased in the right superior temporal cortex. Although no causal relationship can be inferred from this analysis, it however suggests that this pattern of brain-to-speech alignment is atypical.

Entrainment to specific syllable rate also seems atypical in children with DLD

The second analysis we carried out aimed at testing whether neural oscillations track the specific syllabic rate of speech stimuli. In the normal rate condition, uncorrected results revealed that coherence between speech envelope and oscillatory activity in the corresponding frequency band (5.6-7.6 Hz) was stronger in the left anterior superior temporal cortex in TD children as compared to children with DLD who did not show this pattern. Left-lateralized neural entrainment to slow modulations in speech in TD children, although contradictory to the predictions of the AST model (Poeppel, 2003), agrees with previous adult studies showing stronger theta entrainment or theta-gamma coupling in the left temporal cortex for intelligible speech (Gross et al., 2013; Peelle et al., 2013). Resting studies in adults have also revealed that endogenous oscillatory activity also expresses in the theta range in the left anterior temporal cortex, namely in a cortical region that is higher in the processing hierarchy of speech (Giraud et al., 2007; Morillon et al., 2012). This has been taken to suggest that left-lateralized temporal theta oscillations would be involved in the integration of phonemic units, sampled at the gamma rate in left posterior temporal regions, into syllable-size units segmented more anteriorly (Giraud & Poeppel, 2012; Morillon et al., 2012; Peelle & Davis, 2012). Accordingly, our results in TD children for the normal rate condition suggest that the left anterior temporal cortex is particularly sensitive to the specific syllabic rate of the incoming signal. It should yet be noted that coherence maps for these children also showed some cortico-acoustic coupling in the right anterior temporal cortex (Figure 3A). In children with DLD, no significant neural entrainment in the left temporal cortex was observed in the 5.6-7.6 Hz range. If activity in this anterior temporal region is involved in the coupling between theta and gamma oscillations for hierarchical processing of syllables and phonemes (Giraud & Poeppel, 2012; Morillon et al., 2012), one may speculate that it may show atypical entrainment to syllable-rate speech modulations together with atypical left gamma phonemic parsing. Our current results do not allow to corroborate such an interpretation and further research examining the precise phase-locking to syllable onsets in the temporal envelope (see for instance Gross et al., 2013; Hertrich et al., 2012) as well as theta-gamma coupling in both

groups of children is required to more deeply investigate this issue. In addition, similarly as in TD children, cortico-acoustic coherence was observed to some extent in the right temporal cortex of children with DLD (Figure 3A), suggesting that neural alignment might still occur in this frequency band for normal rate speech. Including a larger number of participants in each group is necessary to determine whether the specific syllabic rhythm entrains oscillatory activity in the left anterior temporal cortex in TD children and not in children with DLD or whether alignment occurs in the right region as well in both groups.

Power modulations analyses revealed that children with DLD showed stronger desynchronization (i.e. less power) in the right anterior superior temporal cortex in the 5.6-7.6 Hz range for normal rate speech as compared to their TD peers. In the same region, cortico-acoustic coherence was furthermore negatively correlated with children's own speech rate, namely the children with DLD who had the fastest speech rate in the narrative task exhibited less brain-to-speech alignment in right temporal cortex when normal rate sentences were presented. This result was rather unexpected and is difficult to interpret.

Finally, we did not find evidence for cortical entrainment to fast rate speech in the 8-10 Hz frequency band, corresponding to the mean syllabic rate of fast sentences. This was the case in both groups of children. In children with DLD however, small (uncorrected) peaks of coherence were observed in the right dorsal premotor and temporal cortices. As this effect was relatively small, we prefer not to discuss it further; again, increasing the sample size should allow determining whether this reflects any kind of involvement of the (atypically right-lateralized) dorsal stream for fast rate speech processing.

Conclusion

The present study provides the first piece of evidence to our knowledge for atypical cortical entrainment to naturally produced speech in children with developing language disorders (DLD) compared to typically developing children. Despite the fact that our results would need to be replicated on a larger group of children, we believe that our findings provide the first evidence of impaired brain-to-speech synchronization in the theta-band for normal rate speech. As such our data support the 'prosodic phrasing hypothesis' (Cumming, Wilson, & Goswami, 2015) and suggest a neurobiological account of the deficit that DLD have in processing syllable-rate modulations that are embedded in the speech signal. They also suggest atypical function of the dorsal stream in children with DLD, who exhibit theta-band

entrainment to fast rate speech in right superior temporal cortex in the vicinity of the parieto-temporal junction, and not in left (pre)motor regions as observed in TD children. More analyses are required to examine, in this population, the oscillatory processing of the delta and gamma frequency bands that appear to be deficient in other developmental disorders (dyslexia), as well as the resting oscillatory dynamics of children with dysphasia.

Acknowledgments

We would like to thank parents and their children for participating in this study. We also thank D. Bayle for his invaluable help for analyses with the Fieldtrip toolbox. This study was conducted as part of the ODYSSEE project (PI: V.B; n° 11 JSH2 005 1) funded by the French National Research Agency (ANR). H.G was funded by a doctoral fellowship from the LabEx ASLAN (ANR-10-LABX-0081) of Université de Lyon within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) of the French government operated by the ANR, and by the Erasmus Mundus, Student Exchange Network in Auditory Cognitive Neuroscience program. K.J acknowledges funding from the Canada Research Chairs program and NSERC Discovery Grant (RGPIN-2015-04854).

References

- Abrams, D., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(15), 3958–65. <http://doi.org/10.1523/JNEUROSCI.0187-08.2008>
- Abrams, D., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(24), 7686–93. <http://doi.org/10.1523/JNEUROSCI.5242-08.2009>
- Adank, P., & Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage*, 49(1), 1124–1132. <http://doi.org/10.1016/j.neuroimage.2009.07.032>
- Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 13367–13372.
- Alho, J., Lin, F.-H., Sato, M., Tiitinen, H., Sams, M., & Jääskeläinen, I. P. (2014). Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Frontiers in Psychology*, 5(394). <http://doi.org/10.3389/fpsyg.2014.00394>

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders. 5th ed.* Arlington, VA: American Psychiatric Association.
- Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085. <http://doi.org/10.1093/cercor/bhu103>
- Badcock, N. A., Bishop, D. V. M., Hardiman, M. J., Barry, J. G., & Watkins, K. E. (2012). Co-localisation of abnormal brain structure and function in specific language impairment. *Brain and Language*, 120(3), 310–320. <http://doi.org/10.1016/j.bandl.2011.10.006>
- Beattie, R. L., & Manis, F. R. (2012). Rise time perception in children with reading and combined reading and language difficulties. *Journal of Learning Disabilities*, 46(3), 200–209. <http://doi.org/10.1177/0022219412449421>
- Benasich, A. A., Gou, Z., Choudhury, N., & Harris, K. D. (2008). Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. *Behavioural Brain Research*, 195(2), 215–22. <http://doi.org/10.1016/j.bbr.2008.08.049>
- Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders*. <http://doi.org/10.1111/1460-6984.12335>
- Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, 8(3), 389–395. <http://doi.org/10.1038/nn1409>
- Boersma, P., & Weenink, D. (2012). Praat: Doing phonetics by computer. Retrieved from <http://www.praat.org/>.
- Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, 25(5), 663–678. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0364021301000489>
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of Acoustic Society of America*, 95(3), 1570–1580.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics*, 31, 585–598. [http://doi.org/10.1016/S0095-4470\(03\)00040-8](http://doi.org/10.1016/S0095-4470(03)00040-8)
- Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex*, 45, 119–130. <http://doi.org/10.1016/j.cortex.2007.09.008>
- Corriveau, K. H., Pasquini, E. S., & 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(3), 647–666. [http://doi.org/10.1044/1092-4388\(2007/046\)](http://doi.org/10.1044/1092-4388(2007/046))
- 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). <http://doi.org/10.3389/fpsyg.2015.00972>
- Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with specific language impairments.

- Frontiers in Human Neuroscience*, 9(672). <http://doi.org/10.3389/fnhum.2015.00672>
- Cutini, S., Szucs, 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. <http://doi.org/10.1016/j.neuroimage.2016.08.012>
- de Guibert, C., Maumet, C., Jannin, P., Ferré, J.-C., Tréguier, C., Barillot, C., ... Biraben, A. (2011). Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia). *Brain: A Journal of Neurology*, 134(10), 3044–58. <http://doi.org/10.1093/brain/awr141>
- de Vos, A., Vanvooren, S., Vanderauwera, J., Ghesquière, P., & Wouters, J. (2017). Atypical neural synchronization to speech envelope modulations in dyslexia. *Brain and Language*, 164, 106–117. <http://doi.org/10.1016/j.bandl.2016.10.002>
- de Weck, G., & Rosat, M. C. (2003). *Troubles dysphasiques*. Paris, France: Masson.
- Ding, N., & Simon, J. Z. (2013). Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *Journal of Neuroscience*, 33(13), 5728–5735. <http://doi.org/10.1523/JNEUROSCI.5297-12.2013>
- Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, 85, 761–8. <http://doi.org/10.1016/j.neuroimage.2013.06.035>
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 7126–31. <http://doi.org/10.1073/pnas.1318738111>
- Ferragne, E., Flavier, S., & Fressard, C. (2012). ROCme! Recording of Oral Corpora Made Easy: Logiciel d'enregistrement de corpus. Lyon, France: DDL.
- Fonov, V., Evans, A. C., Botteron, K., Almlic, R., McKinstryef, R. C., Collins, L., & Group, the B. D. C. (2010). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54(1), 313–327.
- Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R., & Collins, D. L. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, 47(S102).
- Gérard, C. (1993). *L'enfant dysphasique*. Bruxelles, Belgique: De Boeck Université.
- Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, 2(130). <http://doi.org/10.3389/fpsyg.2011.00130>
- 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), 113–26. <http://doi.org/10.1159/000208934>
- Giedd, J. N. (2004). Structural Magnetic Resonance Imaging of the Adolescent Brain. *New York Academy of Sciences*, 1021, 77–85.
- 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. <http://doi.org/10.1016/j.neuron.2007.09.038>
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517. <http://doi.org/10.1038/nn.3063>

- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10. <http://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.
- Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews. Neuroscience*, 16, 43–54. <http://doi.org/10.1038/nrn3836>
- Goswami, U., Cumming, R., Chait, M., Huss, M., Mead, N., Wilson, A. M., ... Fosker, T. (2016). Perception of filtered speech by children with developmental dyslexia and children with specific language impairments. *Frontiers in Psychology*, 7(791). <http://doi.org/10.3389/fpsyg.2016.00791>
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives. *Laboratory Phonology*, 4(1), 67–92. <http://doi.org/10.1515/lp-2013-0004>
- Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal Sensory Deficits in Developmental Dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*, 23(2), 325–337. <http://doi.org/10.1162/jocn.2010.21453>
- Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience*, 19(5), 893–906.
- Grosjean, F., & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica*, 31, 144–184.
- 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). <http://doi.org/10.1371/journal.pbio.1001752>
- Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences*, 98(2), 694–699. <http://doi.org/10.1073/pnas.98.2.694>
- Guiraud, H., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Caillot-Bascoul, A., Gonzalez-Monge, S., & Boulenger, V. (n.d.). Don't speak too fast! Processing of fast rate speech in children with specific language impairment. (in revision). *Plos One*.
- Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., & Goswami, U. (2012a). Reduced phase locking to slow amplitude modulation in adults with dyslexia: An MEG study. *NeuroImage*, 59(3), 2952–2961. <http://doi.org/10.1016/j.neuroimage.2011.09.075>
- Hämäläinen, J. A., Rupp, A., Soltész, F., Szücs, D., & Goswami, U. (2012b). Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. *NeuroImage*, 59(3), 2952–61. <http://doi.org/10.1016/j.neuroimage.2011.09.075>
- Hertrich, I., Dietrich, S., Trouvain, J., Moos, A., & Ackermann, H. (2012). Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal. *Psychophysiology*, 49(3), 322–34. <http://doi.org/10.1111/j.1469-8986.2011.01314.x>
- Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–Motor Interaction Revealed by fMRI: Speech, Music, and Working Memory in Area Spt. *Journal of Cognitive Neuroscience*, 15(5), 673–682. <http://doi.org/10.1162/089892903322307393>
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception.

- Trends in Cognitive Sciences*, 4(4), 131–138. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1364661300014637>
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99. <http://doi.org/10.1016/j.cognition.2003.10.011>
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Hill, E. L. (2001). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, 36(2), 149–171. <http://doi.org/10.1080/1368282001001987>
- Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500–11. <http://doi.org/10.1152/jn.00251.2010>
- Hyafil, A., Fontolan, L., Kabdebon, C., Gutkin, B., & Giraud, A.-L. (2015). Speech encoding by coupled cortical theta and gamma oscillations. *eLife*, 4, 1–23. <http://doi.org/10.7554/eLife.06213>
- Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from fMRI and rTMS. *Journal of Physiology*, 102, 31–34. <http://doi.org/10.1016/j.jphysparis.2008.03.003>
- Jacquier-Roux, M., Lequette, C., Pouget, G., Valdois, S., & Zorman, M. (2010). *BALE: Batterie Analytique du Langage Ecrit.*
- Jernigan, T., Hessenlink, J., Sowell, E., & Tallal, P. (1991). Cerebral structure on magnetic resonance imaging in language-and-learning-impaired children. *Archives of Clinical Neuropsychology*, 48, 539–545.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2756068>
- Jusczyk, P. W. (1995). Infant's detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1–23.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.
- Khomsi, A. (2001). *ELO: évaluation du langage oral*. Paris, France: ECPA.
- Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. Paris, France: CNRS Edition.
- 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(454). <http://doi.org/10.3389/fnhum.2013.00454>
- Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered Low-Gamma Sampling in Auditory Cortex Accounts for the Three Main Facets of Dyslexia. *Neuron*, 72(6), 1080–1090. <http://doi.org/10.1016/j.neuron.2011.11.002>
- Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: (The MIT press).
- 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. <http://doi.org/10.1002/hbm.22986>

- Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *Plos Biology*, 8(8). <http://doi.org/10.1371/journal.pbio.1000445>
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001–10. <http://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(170). <http://doi.org/10.3389/fpsyg.2012.00170>
- Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, 133, 516–528. <http://doi.org/10.1016/j.neuroimage.2016.02.064>
- Maillart, C., & Schelstraete, M. A. (2012). *Les dysphasies: de l'évaluation à la rééducation*. Issy-les-Moulineaux, France: Elsevier-Masson.
- Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.
- Mehler, J., Jusczyk, P. W., Lamsertz, G., Halsted, N., Bertoincini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Mehler, J., Sebastian-Galles, N., & Nespors, M. (2004). Biological foundations of language: language acquisition, cues for parameter setting and the bilingual infant. In M. Gazzaniga (Ed.), *The New Cognitive Neuroscience* (p. 825–36.). Cambridge, MA: MIT Press.
- Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping*, 37(8), 2767–83. <http://doi.org/10.1002/hbm.23206>
- 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 intra-cortical study. *Frontiers in Psychology*, 3(248). <http://doi.org/10.3389/fpsyg.2012.00248>
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annual New York Academy of Sciences*, 1337(1), 26–31. <http://doi.org/10.1530/ERC-14-0411>. Persistent
- Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5, 1–9. <http://doi.org/10.1038/ncomms6255>
- Myers, E. H., Hampson, M., Vohr, B., Lacadie, C., Frost, S. J., Pugh, K. R., ... Ment, L. R. (2010). Functional connectivity to a right hemisphere language center in prematurely born adolescents. *NeuroImage*, 51(4), 1445–1452. <http://doi.org/10.1016/j.neuroimage.2010.03.049>
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233–243. [http://doi.org/10.1016/S0167-6393\(02\)00106-1](http://doi.org/10.1016/S0167-6393(02)00106-1)
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience - Special Issue on Academic Software Applications for Electromagnetic Brain Mapping Using MEG and EEG*, 2011(1).
- Osnes, B., Hugdahl, K., & Specht, K. (2011). NeuroImage Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *NeuroImage*, 54(5), 2437–2445. <http://doi.org/10.1016/j.neuroimage.2010.09.078>

- Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A Common Language Network for Comprehension and Production: A Contribution to the Definition of Language Epicenters with PET. *NeuroImage*, *11*(4), 347–357. <http://doi.org/10.1006/nimg.2000.0546>
- Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, *25*(12), 1649–1653. <http://doi.org/10.1016/j.cub.2015.04.049>
- Paus, T., Zijdenbos, A., Worsley, K., Collins, L. D., Blumenthal, J., Giedd, J. N., ... Evans, A. C. (1999). Structural Maturation of Neural Pathways in Children and Adolescents : In Vivo Study. *Science*, *283*, 1908–1911. <http://doi.org/10.1126/science.283.5409.1908>
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*. <http://doi.org/10.3389/fpsyg.2012.00320>
- Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. <http://doi.org/10.1093/cercor/bhs118>
- Pefkou, M., Arnal, L. H., Fontolan, L., & Giraud, A.-L. (2017). Theta- and beta-band neural activity reflect independent syllable tracking and intelligibility of time-compressed speech. *The Journal of Neuroscience*, *37*(33), 7930–7938. <http://doi.org/10.1523/JNEUROSCI.2882-16.2017>
- Pellegrino, F., Coupé, C., & Marisco, E. (2011). A Cross-Language Perspective on Speech Information Rate. *Language*, *87*(3), 539–558. <http://doi.org/10.1353/lan.2011.0057>
- Plante, E., Swisher, L., & Vance, R. (1989). Anatomical correlates of normal and impaired language in a set of dizygotic twins. *Brain and Language*, *37*(4), 643–655. [http://doi.org/10.1016/0093-934X\(89\)90117-X](http://doi.org/10.1016/0093-934X(89)90117-X)
- Plante, E., Swisher, L., Vance, R., & Rapcsak, S. (1991). MRI findings in boys with specific language impairment. *Brain and Language*, *41*(1), 52–66. [http://doi.org/10.1016/0093-934X\(91\)90110-M](http://doi.org/10.1016/0093-934X(91)90110-M)
- Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquière, P., & Wouters, J. (2012). Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. *Ear and Hearing*, *33*(1), 134–143.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows : cerebral lateralization as asymmetric sampling in time. *Speech Communication*, *41*, 245–255. [http://doi.org/10.1016/S0167-6393\(02\)00107-3](http://doi.org/10.1016/S0167-6393(02)00107-3)
- 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. <http://doi.org/10.1098/rstb.2007.2160>
- Power, A. J., Colling, L. J., Mead, N., Barnes, L., & Goswami, U. (2016). Neural encoding of the speech envelope by children with developmental dyslexia. *Brain and Language*, *160*, 1–10. <http://doi.org/10.1016/j.bandl.2016.06.006>
- Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Frontiers in Psychology*, *3*. <http://doi.org/10.3389/fpsyg.2012.00213>
- Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Frontiers in Human Neuroscience*, *7*(777). <http://doi.org/10.3389/fnhum.2013.00777>

- Rapin, I. (1996). Practitioner review : Developmental language disorders : A Practitioner Review : Developmental Language Disorders : A Clinical Update. *Journal of Child Psychology and Psychiatry*, 37(6), 643–655.
- Rapin, I., & Allen, D. A. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling* (p. 155–184). San Diego, CA: Academic Press/
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271(1–2), 16–25. <http://doi.org/10.1016/j.heares.2010.09.001>
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–24. <http://doi.org/10.1038/nn.2331>
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “ what ” and “ where ” in auditory cortex, 97(22).
- Raven, J., Raven, J., & Court, J. (1998). *Coloured progressive matrices*. Oxford, England: Oxford Psy.
- Reilly, S., Bishop, D. V. M., & Tomblin, B. (2014). Terminological debate over language impairment in children: Forward movement and sticking points. *International Journal of Language and Communication Disorders*, 49(4), 452–462. <http://doi.org/10.1111/1460-6984.12111>
- Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., ... Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders*, 49(4), 416–451. <http://doi.org/10.1111/1460-6984.12102>
- 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. <http://doi.org/10.1044/2015>
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences*, 336(1278), 367–373.
- Serniclaes, W., Sprenger-charolles, L., Carré, R., & Demonet, J.-F. (2001). Perceptual Discrimination of. *Journal of Speech Language and Hearing Research*, 44, 384–399.
- 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, 336–361. <http://doi.org/10.1016/j.jecp.2004.02.001>
- Shapiro, R., & Janzen, A. H. (1960). The normal skull: a Roentgen study. *Hoerber*.
- Soltész, F., Szűcs, D., Leong, V., White, S., & Goswami, U. (2013). Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. *Plos One*, 8(10), 1–11. <http://doi.org/10.1371/journal.pone.0076608>
- Tallal, P., & Piercy, M. (1973a). Defects of Non-Verbal Auditory Perception in Children with Developmental Aphasia. *Nature*, 241(5390), 468–469.
- Tallal, P., & Piercy, M. (1973b). Developmental aphasia: impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, 11, 189–398.
- Tallal, P., & Piercy, M. (1975). Developmental aphasia: the perception of brief vowels and extended stop consonants. *Neuropsychologia*, 13(1), 69–74.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., ...

- Wartenburger, I. (2009). Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. *Journal of Neuroscience*, 29(47), 14726–14733. <http://doi.org/10.1523/JNEUROSCI.1246-09.2009>
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., & Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near-infrared spectroscopy and EEG study. *Frontiers in Psychology*, 2. <http://doi.org/10.3389/fpsyg.2011.00062>
- 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(19737). <http://doi.org/10.1038/srep19737>
- Thomson, J. M., Fryer, B., & Maltby, J. (2006). Auditory and motor rhythm awareness in adults with dyslexia, 29(3), 334–348. <http://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), 120–9. <http://doi.org/10.1016/j.jphysparis.2008.03.007>
- Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience*, 8. <http://doi.org/10.3389/fnhum.2014.00949>
- Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research*, 40, 1245–1260.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15, 273–89. <http://doi.org/10.1006/nimg.2001.0978>
- Tzourio, N., Heim, A., Zilbovicius, M., Gerard, C., & Mazoyer, B. M. (1994). Abnormal regional CBF response in left hemisphere of dysphasic children during a language task. *Pediatric Neurology*, 10(1), 20–26. [http://doi.org/10.1016/0887-8994\(94\)90062-0](http://doi.org/10.1016/0887-8994(94)90062-0)
- Vanvooren, S., Poelmans, H., Hofmann, M., Ghesquière, P., & Wouters, J. (2014). Hemispheric asymmetry in auditory processing of speech envelope modulations in prereading children. *The Journal of Neuroscience*, 34(4), 1523–1529. <http://doi.org/10.1523/JNEUROSCI.3209-13.2014>
- Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Cerebral dominance for language function in adults with specific language impairment or autism. *Brain*, 131(12), 3193–3200. <http://doi.org/10.1093/brain/awn266>
- Zion Golumbic, E. M., 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), 151–61. <http://doi.org/10.1016/j.bandl.2011.12.010>
- Zoefel, B., & Vanrullen, R. (2015). Selective Perceptual Phase Entrainment to Speech Rhythm in the Absence of Spectral Energy Fluctuations. *Journal of Neuroscience*, 35(5), 1954–1964. <http://doi.org/10.1523/JNEUROSCI.3484-14.2015>

4. Synthèse des résultats de l'étude 3

Les résultats de cette dernière étude suggèrent pour la première fois à notre connaissance un fonctionnement oscillatoire atypique chez les enfants présentant une dysphasie (8-13 ans) lors de l'écoute de parole naturelle à débit normal ou rapide, par rapport à un groupe d'enfants neurotypiques du même âge.

Dans la bande de thêta (4-7 Hz), nous avons mis en évidence une réduction de la synchronisation cortico-acoustique dans le cortex auditif associatif droit chez les enfants dysphasiques par rapport aux enfants contrôles lors de la perception de parole à débit normal. Ces premiers résultats s'accordent avec l'hypothèse d'un défaut d'extraction des modulations de basse fréquence dans l'enveloppe temporelle de la parole, affectant la segmentation de l'information syllabique chez ces enfants (d'après les modèles de Ghitza, 2011 et Poeppel, 2003; voir aussi la « *prosodic phrasing hypothesis* » de Cumming, Wilson, & Goswami, 2015a). Par ailleurs, nos résultats révèlent que contrairement aux enfants contrôles, les enfants dysphasiques ne présentent aucune cohérence dans les régions (pré)motrices gauches et l'aire de Broca lors de la perception de parole rapide. Ces régions articulatoires, impliquées dans la voie dorsale des modèles de perception de la parole (Hickok & Poeppel, 2007; Rauschecker, 2011; Rauschecker & Scott, 2009; Rauschecker & Tian, 2000), sont fortement engagées lors de conditions d'écoute difficiles (Adank & Devlin, 2010; Alho et al., 2014; Du et al., 2014; Osnes et al., 2011), comme c'est le cas ici lors de l'écoute de parole accélérée naturellement, et semblent donc, selon nos résultats, sous-recrutées dans la dysphasie. De manière intéressante, dans cette même condition de parole rapide, une augmentation de cohérence cortico-acoustique est observée dans le cortex temporal postérieur droit chez les dysphasiques uniquement. Cette région, proche de la jonction temporo-pariétale impliquée dans l'intégration sensorimotrice de la parole, est connue pour participer au traitement phonologique en perception et production. En outre, des études anatomiques ont révélé une latéralisation hémisphérique différente des régions périsylviennes (incluant le cortex temporal postérieur) chez les dysphasiques, la région droite étant souvent de taille équivalente ou supérieure à la région gauche (Plante, Swisher, Vance, & Rapcsak, 1991). Des anomalies fonctionnelles ont également été décrites dans ces régions lors du traitement de phonèmes (Badcock et al., 2012). Les données en MEG de notre étude suggèrent donc que l'intégration sensorimotrice se produise malgré tout chez les enfants dysphasiques, bien que dans une région atypique latéralisée à droite (voir Cutini, Szucs, Mead, Huss, & Goswami, 2016 pour

des résultats dans la bande delta chez des dyslexiques). Ainsi, nos résultats, qui méritent d'être confirmés avec un échantillon plus important, penchent plutôt en faveur d'un dysfonctionnement de la voie dorsale, et potentiellement du couplage cortical auditivo-moteur (qui serait néanmoins à tester directement), chez les enfants dysphasiques. Ces résultats pourraient par ailleurs expliquer les faibles performances de ces enfants dans la tâche comportementale de perception de parole rapide de l'étude 2 de cette thèse (Guiraud et al., en révision).

Enfin, dans la bande de fréquence centrée sur le débit syllabique moyen de la parole normale (5.6-7.6 Hz), nous observons un alignement plus fort des oscillations corticales du cortex temporal antérieur gauche sur l'enveloppe temporelle de la parole à débit normal chez les enfants neurotypiques par rapport aux enfants dysphasiques. Chez les enfants contrôles, nous avons suggéré que cette synchronisation à gauche sur le rythme syllabique moyen des phrases soit liée à un processus de plus haut niveau permettant d'intégrer l'information phonémique échantillonnée dans la bande gamma au sein des unités syllabiques (cf. Étude 1, Guiraud, Hincapié, Jerbi & Boulenger, soumis) en accord avec les études et modèles précédents (Giraud & Poeppel, 2012; Morillon, Liégeois-Chauvel, Arnal, Bénar, & Giraud, 2012; Peelle & Davis, 2012). Un tel processus pourrait alors être défaillant chez les enfants dysphasiques. Toutefois un échantillon plus grand de participants dans les deux groupes serait nécessaire afin de valider cette hypothèse.

L'ensemble de ces résultats, même s'ils restent préliminaires, suggèrent donc un couplage atypique des oscillations corticales sur le rythme du signal de parole chez les enfants dysphasiques, suggérant un trouble du traitement de l'information syllabique qui pourrait être à la base des troubles phonologiques et morpho-syntaxiques observés chez ces enfants.

DISCUSSION GÉNÉRALE

L'objectif principal de ce travail de thèse était de mieux comprendre les processus neurocognitifs sous-tendant la perception de la parole naturelle chez les enfants présentant un développement langagier typique et les enfants présentant une dysphasie majoritairement sur le versant expressif du langage. Nous avons axé plus précisément notre travail sur la dynamique des oscillations corticales mise en jeu en réponse à des variations du rythme de la parole et notamment du débit syllabique, à l'aide de stimuli phrastiques produits naturellement à un débit normal (environ 6.61 syll/s) ou rapide (parole accélérée naturellement, environ 9.03 syll/s). Pour ce faire, malgré le défi méthodologique que cela constituait, nous avons utilisé une technique d'imagerie cérébrale totalement non-invasive et offrant une excellente résolution temporelle couplée à une résolution spatiale permettant une bonne localisation des sources corticales : la MEG. Un vaste courant de recherches menées principalement chez l'adulte montre que les oscillations corticales, dans les bandes delta, thêta et gamma principalement, sont capables de s'aligner sur les différentes échelles temporelles du signal de parole, favorisant la segmentation et ainsi le décodage de l'information linguistique. En dépit du rôle primordial des informations rythmiques dans le développement du langage oral, les études chez l'enfant sont particulièrement rares, surtout si l'on considère le cas de la parole naturelle. Notre premier objectif était donc de déterminer si le couplage cerveau/parole mis en évidence dans la bande thêta chez l'adulte se produit de la même manière chez l'enfant neurotypique. En outre, des troubles du traitement du rythme de stimuli auditifs, langagiers ou non, ayant été décrits chez les enfants dysphasiques, notre deuxième objectif était de tester l'hypothèse selon laquelle ces troubles rythmiques seraient sous-tendus par un déficit de l'alignement des rythmes cérébraux sur le rythme syllabique de la parole. Nous avons donc examiné les capacités de perception de la parole produite à un débit rapide ainsi que la dynamique oscillatoire sous-jacente. A notre connaissance, les résultats présentés dans ce travail constituent les premières preuves (1) d'une synchronisation des oscillations thêta au sein des régions auditives mais aussi prémotrices sur le rythme syllabique de la parole naturelle chez l'enfant neurotypique et (2) d'un déficit de perception de la parole naturelle rapide associé à un couplage atypique entre l'activité oscillatoire thêta et le signal de parole chez l'enfant dysphasique.

IX. Rythmes corticaux et perception de la parole chez les enfants neurotypiques

Nos résultats, présentés dans le chapitre VII, révèlent pour la première fois à notre connaissance la dynamique des oscillations corticales lors de la perception de parole naturelle produite à débit normal et rapide chez des enfants neurotypiques francophones âgés de 8 à 13 ans. Une augmentation de la cohérence entre les modulations lentes de l'enveloppe temporelle de la parole et les oscillations corticales dans la bande thêta est observée pour les deux types de débit de parole. Cependant, un réseau cortical différent, incluant les régions temporales ou (pré)motrices, semble être impliqué selon le rythme syllabique du signal de parole perçu. De plus, nos données mettent en évidence un couplage cortico-acoustique dans des régions temporales sensiblement différentes lorsqu'on s'intéresse à l'ensemble de la bande thêta (4-7 Hz) ou lorsqu'on restreint nos analyses à une bande de fréquence centrée sur le débit syllabique des stimuli. Dans la suite, nous détaillerons ces différents résultats en commençant par le traitement de la parole à débit normal, puis de la parole à débit rapide.

1. Dynamique oscillatoire de la perception de la parole à débit normal

Nous avons étudié la dynamique oscillatoire sous-tendant le traitement de la parole dans la bande thêta (4-7 Hz), correspondant à la bande de fréquence couvrant les modulations lentes de l'enveloppe temporelle de la parole et dans laquelle la synchronisation cortico-acoustique se produit d'après la théorie de l'échantillonnage asymétrique de la parole (AST) de Poeppel (Poeppel, 2003; Poeppel, Idsardi, & van Wassenhove, 2008). Les résultats obtenus chez les enfants neurotypiques lors de l'écoute de phrases naturellement produites à débit normal confirment les prédictions du modèle AST et s'accordent avec les études précédemment réalisées chez les adultes (Boemio, Fromm, Braun, & Poeppel, 2005; Gross et al., 2013; Luo & Poeppel, 2007; Vander Ghinst et al., 2016) et les quelques résultats rapportés chez les enfants à partir de stimuli verbaux ou non-verbaux (stimuli verbaux répétés : Abrams, Nicol, Zecker, & Kraus, 2008; Molinaro, Lizarazu, Lallier, Bourguignon, & Carreiras, 2016; stimuli non-verbaux : Telkemeyer et al., 2009; Telkemeyer et al., 2011; Vanvooren, Poelmans, Hofmann, Ghesquiere, & Wouters, 2014), révélant une préférence

générale de traitement des basses fréquences de l'enveloppe d'amplitude du signal de parole dans les régions auditives associatives de l'hémisphère droit. Nos résultats mettent également en évidence l'implication d'un réseau oscillatoire pariéto-frontal droit se couplant à l'enveloppe de la parole dans la bande thêta. Un tel réseau a déjà été observé dans différentes études lors de la perception de parole naturelle chez l'adulte (Gross et al., 2013; Keitel et al., 2017). Gross et collègues (2013) ont notamment rapporté une synchronisation des oscillations delta/thêta sur l'enveloppe d'amplitude de la parole au niveau des régions pariétales (droite et gauche) lors de la perception de parole intelligible. Une modulation de la puissance des oscillations thêta au sein de la région pariéto-frontale droite (cuneus (bilatéral), précuneus, gyrus pariétal supérieur et une partie du lobule frontal inférieur) par les oscillations delta (1-4 Hz) du gyrus temporal supérieur postérieur droit a par ailleurs été observée par Keitel et al. (2017) lors de la perception de parole continue chez les adultes. Ces différentes régions sont connues pour être impliquées dans la mémoire de travail phonologique et sémantique mais aussi dans le traitement de l'information prosodique (D'Ausilio et al., 2009; Friederici, 2011; Meyer, Alter, Friederici, Lohmann, & Von Cramon, 2002; Bastiaansen & Hagoort, 2006 pour une revue). La charge de mémoire de travail serait en outre associée à une augmentation de l'activité thêta pariétale (Moran et al., 2010; Sauseng et al., 2010). Enfin, des études développementales ont montré que l'acquisition du langage est régie par les liens entre les systèmes de perception et de production de la parole, impliquant très probablement les régions auditives et articulatoires de la voie dorsale de la perception de la parole (Doupe & Kuhl, 1999; Hickok & Poeppel, 2007; Pulvermüller & Fadiga, 2010). Les premières productions de l'enfant correspondent à des structures syllabiques qui émergent grâce au mouvement oscillatoire de la mandibule. C'est la comparaison entre le signal produit par ce mouvement articulatoire et le signal perçu, nécessitant de fortes interactions sensori-motrices, qui permet à l'enfant de créer son propre répertoire phonologique (Bertoncini & Mehler, 1981). Dans notre étude, l'alignement de l'activité oscillatoire des régions fronto-pariétales droites sur l'enveloppe d'amplitude lors de l'écoute de parole à débit normal pourrait ainsi refléter l'implication d'un réseau langagier sous-tendant à la fois le traitement de l'information syllabique et prosodique (véhiculée dans l'enveloppe d'amplitude), mais aussi potentiellement un appariement auditivo-moteur qui pourrait être particulièrement mobilisé dans le cerveau en développement afin de traiter la parole naturelle. Il faut noter que dans leur modèle, Hickok et Poeppel (2007) suggèrent que la voie dorsale d'intégration sensori-motrice soit fortement latéralisée dans l'hémisphère gauche. Nos résultats montrent une implication du réseau perception/action de l'hémisphère droit également, en accord avec l'étude récente de Sammler

et collaborateurs suggérant l'existence d'une voie dorsale droite pour le traitement de la prosodie (Sammler, Grosbras, Anwander, Bestelmeyer, & Belin, 2015; voir aussi les interprétations de Cutini, Szucs, Mead, Huss, & Goswami, 2016 et Keitel, Ince, Gross, & Kayser, 2017), ainsi qu'avec l'étude de Giraud et collègues (2007) montrant un couplage entre l'activité oscillatoire dans les bandes thêta et gamma et l'activité hémodynamique du cortex (pré)moteur bilatéral.

Nos analyses montrent par ailleurs que l'information linguistique contenue dans les modulations lentes de l'enveloppe temporelle de la parole entraînerait également les oscillations corticales dans le cortex auditif gauche dans une bande de fréquence plus restreinte et centrée sur le débit moyen des stimuli de parole. Nos résultats chez les enfants neurotypiques francophones lors de la perception de parole à débit normal révèlent en effet une augmentation de la cohérence cortico-acoustique au niveau du cortex temporal antérieur gauche (gyri temporaux supérieur et médian) dans la bande 5.6-7.6 Hz, correspondant au débit moyen normal des phrases. Ces résultats contredisent le modèle AST (Poeppel, 2003) et la seule étude à notre connaissance dans une langue syllabique ayant utilisé des stimuli phrastiques naturels (Molinaro et al., 2016). Dans leur étude chez des enfants et adultes hispanophones, Molinaro et collègues (2016) ont en effet observé une augmentation de la cohérence entre l'enveloppe du signal de parole et les oscillations thêta au niveau des cortex auditif primaire et temporal supérieur droits. Aucune cohérence significative n'a été mise en évidence au sein de l'hémisphère gauche dans cette étude. Cependant, contrairement à notre étude, les auteurs n'ont pas analysé la cohérence dans une bande de fréquence centrée sur le rythme syllabique de leurs phrases mais sur le pic de cohérence maximale. Ainsi leurs résultats sont plus proches de notre analyse sur l'ensemble de la bande thêta (4-7 Hz) que de celle portant sur le rythme syllabique précis des phrases.

Nos résultats semblent ainsi être particulièrement novateurs puisque d'après nos connaissances, ils sont les premiers à révéler, chez des enfants, une synchronisation des oscillations corticales sur le rythme syllabique spécifique d'un signal de parole naturel. Une première interprétation est que ce couplage entre les modulations lentes de l'enveloppe temporelle du signal de parole et l'activité oscillatoire du cortex temporal antérieur gauche au rythme de l'information syllabique pourrait refléter l'alignement spécifique des rythmes cérébraux sur l'attaque (*onset*) ou le noyau des syllabes (Gross et al., 2013). Gross et collaborateurs (2013) ont récemment décrit une telle synchronisation, en montrant que les

crêtes (*edges*) au sein de l'enveloppe d'amplitude de la parole (correspondant aux attaques syllabiques) recalait la phase de l'activité corticale thêta. Ce couplage était observé au sein des régions auditives bilatérales et dans une fenêtre temporelle précoce (100-300 ms) mais aussi plus tardive (400-600 ms), suggérant qu'il ne correspond pas simplement à une réponse évoquée transitoire pour chaque crête mais bien à un *entraînement* oscillatoire (nous reviendrons sur ce point un peu plus tard). Dans notre étude, une analyse similaire examinant le couplage précis sur chaque syllabe dans le signal verbal serait intéressante à réaliser chez les enfants afin de déterminer précisément l'origine de l'alignement cortico-acoustique dans la bande de fréquence centrée sur le débit syllabique des stimuli. Il faut par ailleurs noter qu'alors que Gross et collègues (2003) ont rapporté un couplage de phase entre la réponse corticale et les attaques des syllabes dans l'enveloppe au niveau bilatéral, la cohérence cortico-acoustique sur le débit syllabique a augmenté dans le cortex temporal gauche dans notre étude. Bien que la comparaison entre les résultats soit délicate en raison de paradigmes, de populations et d'analyses différentes, le fait que leur étude ait été réalisée en anglais et la nôtre en français doit être un facteur à considérer (les caractéristiques de la langue devraient d'ailleurs être prises en compte pour toute comparaison entre études). L'anglais est en effet une langue accentuelle alors que le français est une langue syllabique, c'est-à-dire que sa structure rythmique est régie par les syllabes qui présentent une durée très régulière (Laver, 2003; Meynadier, 2001; Ridouane et al., 2009) ; ceci pourrait laisser supposer un rôle encore plus fort de la syllabe et en particulier du débit syllabique (nombre de syllabes par seconde) à proprement parler dans la segmentation et le traitement du message linguistique. Le ratio entre débit syllabique et débit d'information (taux d'information porté par la syllabe, en *bits/syllabe*) constitue également probablement une donnée à prendre en compte. Pellegrino et collègues (Pellegrino, Coupé, & Marisco, 2011) ont montré, à partir d'un corpus multilingue, que plus le débit syllabique d'une langue est rapide, moins les syllabes sont porteuses d'information linguistique. Le français présente ainsi un débit syllabique un peu plus rapide que l'anglais mais un débit d'information un peu plus faible. A l'inverse, comparé à l'espagnol (langue syllabique) et au japonais (langue moraïque) par exemple, le débit syllabique moyen du français est plus lent mais la syllabe véhicule un taux d'information plus important. Ces caractéristiques du français, entre autres, pourraient alors potentiellement expliquer la latéralisation gauche du couplage cerveau/parole observé dans la bande de fréquence centrée sur le débit syllabique de nos phrases. Il serait pertinent de comparer les résultats issus du même paradigme dans des langues syllabiques (*e.g.*, français, espagnol, italien) et accentuelles (*e.g.*, anglais, allemand) afin de vérifier l'importance de l'information

rythmique syllabique dans la dynamique oscillatoire sous-tendant la perception de la parole. Il serait par ailleurs intéressant d'étudier la synchronisation cortico-acoustique lors du traitement d'une langue seconde, notamment dans le cas où les langues maternelle et seconde n'appartiennent pas à la même classe rythmique. Par exemple, l'étude du français en tant que langue seconde d'individus anglophones permettrait d'examiner la dynamique corticale oscillatoire sous-tendant le traitement de langues caractérisées par des structures prosodiques différentes dans lesquelles la syllabe peut jouer un rôle plus ou moins important dans la segmentation du signal continu. En effet, si le rythme oscillatoire du système cortical auditif dépend de (ou est corrélé à) la structure rythmique du système de production de la parole, le traitement d'une langue seconde de structure rythmique différente de la langue maternelle pourrait entraîner une dynamique oscillatoire différente. Peu d'études à ce jour se sont intéressées à cette question. Une étude de Pérez et collaborateurs (Pérez, Carreiras, Gillon Dowens, & Duñabeitia, 2015) chez des adultes hispanophones ayant l'anglais pour langue seconde, a néanmoins montré une augmentation de la synchronisation cortico-acoustique dans la bande thêta lors de l'écoute d'un discours dans leur langue maternelle, alors que le couplage était plus faible lorsque le discours était produit dans leur langue seconde (voir aussi Jin, Díaz, Colomer, & Sebastián-Gallés, 2014; Peña & Melloni, 2012).

Une deuxième interprétation de l'alignement des oscillations corticales gauches centré sur le débit syllabique des phrases est qu'il pourrait refléter l'activité thêta liée à un mécanisme d'intégration de plus haut niveau, comme suggéré dans différentes études (Giraud et al., 2007; Giraud & Poeppel, 2012; Morillon et al., 2012), plutôt qu'un simple processus de découpage de l'information verbale. A l'aide d'enregistrements EEG intracrâniens chez une patiente adulte épileptique francophone, Morillon et collaborateurs (2012) ont montré une activité thêta (avec un pic à 7.5 Hz) au sein du cortex auditif primaire gauche mais aussi du gyrus temporal supérieur antérieur gauche (BA 22), lorsque la patiente était au repos ou écoutait des syllabes isolées. De plus, les auteurs ont observé, au sein de cette région associative antérieure gauche, une co-modulation entre l'activité thêta et la puissance des oscillations gamma, qui pourrait refléter le couplage (*nesting*) thêta-gamma postulé dans les différentes modèles (Ghitza, 2011; Giraud et al., 2007; Giraud & Poeppel, 2012; Peelle & Davis, 2012). D'autres études utilisant des stimuli verbaux intelligibles (phrases) ont également rapporté une synchronisation cortico-acoustique dans la bande thêta au sein des régions temporales gauches et la présence d'un couplage de ces oscillations avec la puissance des oscillations gamma (Gross et al., 2013; Peelle et al., 2013). Peelle et collaborateurs (2013)

ont par ailleurs montré que le couplage cortico-acoustique thêta était d'autant plus fort et latéralisé à gauche que la parole était intelligible (parole normale vs vocodée), suggérant une sensibilité particulière des régions temporales antérieures gauches au contenu linguistique de la parole plus qu'à son contenu acoustique (voir aussi Ahissar et al., 2001; Hertrich, Dietrich, Trouvain, Moos, & Ackermann, 2012; Luo & Poeppel, 2012; Mai, Minett, & Wang, 2016; Nourski et al., 2009; Park, Ince, Schyns, Thut, & Gross, 2015; Pefkou, Arnal, Fontolan, & Giraud, 2017). A la lumière de ces données et de l'idée d'un couplage hiérarchique thêta-gamma (Giraud et al., 2007; Giraud & Poeppel, 2012; Peelle & Davis, 2012), nos résultats chez les enfants suggèrent ainsi que la cohérence cortico-acoustique gauche centrée sur le débit syllabique puisse correspondre à un processus d'intégration des segments phonémiques échantillonnés dans la bande gamma dans le cortex temporal postérieur gauche au sein de segments plus longs de taille syllabique échantillonnés dans une région plus antérieure (voir Giraud et al., 2007 et Morillon et al., 2012). Cette intégration de l'information syllabique dans les régions temporales antérieures gauches permettrait ainsi un traitement linguistique plus efficace, évitant une communication inter-hémisphérique pour traiter les informations syllabiques et phonémiques. En effet, alors que le découpage de l'information syllabique à droite et de l'information phonémique à gauche est défini, d'après le modèle AST (Poeppel, 2003), comme se réalisant au sein des régions proches du cortex auditif primaire, le mécanisme thêta d'intégration de second ordre centré sur le rythme syllabique pourrait se produire, comme suggéré par Giraud et collègues (2007), dans les régions plus antérieures (i.e. plus élevées dans la hiérarchie de traitement), dédiées au traitement de la parole intelligible et contenant les cartes topographiques des syllabes et des voyelles (Obleser et al., 2006; Rauschecker & Tian, 2000; Scott, 2000).

L'ensemble de nos résultats concernant la dynamique oscillatoire sous-tendant le traitement de la parole à débit normal chez l'enfant neurotypique francophone suggère donc deux types d'*entraînement* dans la bande thêta. Un premier processus de synchronisation entre l'enveloppe d'amplitude du signal de parole et les oscillations thêta au sein des régions auditives droites, semblable à celui proposé par le modèle AST (Poeppel, 2003), est observé. En accord avec la littérature (Giraud et al., 2007; Giraud & Poeppel, 2012; Gross et al., 2013; Morillon et al., 2012; Peelle & Davis, 2012; Peelle et al., 2013), nous proposons que ce processus soit un prérequis nécessaire au traitement initial du message linguistique, permettant de découper le signal de parole à partir de l'information prosodique, mais

insuffisant pour permettre l'accès à la compréhension (Ding & Simon, 2014; Peelle & Davis, 2012; Zoefel & Vanrullen, 2015). Un processus de plus haut niveau entrerait ainsi en jeu au sein de l'hémisphère gauche, permettant d'échantillonner l'information syllabique dans la bande thêta au sein des régions temporales antérieures afin d'y intégrer l'information phonémique extraite par la bande gamma dans les régions plus postérieures et de réaliser les traitements linguistiques permettant notamment l'accès aux représentations phonologiques plus abstraites (Giraud et al., 2007; Giraud & Poeppel, 2012; Morillon et al., 2012; Obleser et al., 2006; Peelle et al., 2013; Scott, 2000). Un tel processus permettrait un traitement plus efficace et moins coûteux au niveau cognitif, par rapport à un traitement où chaque type de segment de parole serait traité par un seul hémisphère, comme proposé dans le modèle AST (Poeppel, 2003). A notre connaissance, aucune étude jusqu'à présent n'a montré un couplage inter-hémisphérique entre l'activité oscillatoire thêta à droite et gamma à gauche lors du traitement du signal de parole, ce qui s'accorderait avec l'hypothèse que le *nesting* thêta-gamma, processus de plus haut niveau, soit réalisé en intra-hémisphérique, au sein de l'hémisphère gauche. Notre étude fournit les premiers éléments à notre connaissance en faveur d'une activité thêta centrée sur le débit syllabique et latéralisée à gauche, qui pourrait participer à ce couplage thêta/gamma, chez l'enfant neurotypique âgé de 8 à 13 ans.

Des analyses supplémentaires sont toutefois nécessaires afin de valider notre hypothèse : il serait ainsi crucial d'examiner les couplages inter-fréquences thêta et gamma (i.e. couplage phase-phase, phase-amplitude, amplitude-phase) dans et entre les deux hémisphères lors du traitement de la parole à débit normal, afin de corroborer le rôle des oscillations dans les régions temporales antérieures gauches centrées sur le rythme syllabique de la parole dans le *nesting* thêta/gamma. Etudier l'activité oscillatoire au repos chez les enfants permettrait en outre de vérifier la prédisposition des régions auditives postérieures droites et antérieures gauches à osciller de manière endogène au rythme thêta. Comme mentionné dans le chapitre IV.2.1 de l'Introduction, peu de travaux à notre connaissance ont étudié la dynamique des oscillations corticales chez l'enfant au repos (voir Uhlhaas, Roux, Rodriguez, Rotarska-Jagiela, & Singer, 2010 pour une revue) et une seule s'est particulièrement intéressée aux bandes de fréquence pertinentes pour la parole, révélant une latéralisation préférentielle des oscillations gamma à gauche alors que les oscillations thêta semblent être réparties de façon bilatérale chez l'enfant entre 3 et 5 ans (Thompson et al., 2016). Dans notre étude en MEG, nous avons enregistré l'activité oscillatoire des enfants au repos, 3 minutes avant et après la tâche de perception de parole. Ces données n'ont à ce jour

pas encore été analysées ; néanmoins nous avons pour prochain objectif d'analyser cette activité de repos afin de déterminer si, comme chez l'adulte (Giraud et al., 2007; Morillon et al., 2012), les cortex auditifs droit (postérieur) et gauche (antérieur) présentent une activité oscillatoire intrinsèque dans la bande thêta chez les enfants neurotypiques âgés de 8 à 13 ans. Nous pourrions ainsi comparer ces mesures avec nos analyses dans la tâche de perception de parole et vérifier la prédisposition des régions auditives pour le rythme syllabique.

Enfin, nous tenons à préciser que tous les auteurs ne partagent pas la même hypothèse concernant les processus à la base de l'alignement cortical sur les modulations lentes d'amplitude de l'enveloppe temporelle de la parole. Dans ce travail de thèse, nous partageons l'hypothèse proposée principalement par Ghitza (2011, 2013) et Giraud et Poeppel (2012) selon laquelle le couplage cortico-acoustique dans la bande thêta reflèterait un *entrainement* de l'activité corticale oscillatoire endogène. Ainsi, chaque cycle oscillatoire cortical thêta s'alignerait (par un recalage de phase) sur un segment syllabique (plus précisément sur la portion du signal entre deux voyelles) et les schémas acoustiques contenus dans un cycle thêta seraient ensuite utilisés pour décoder l'information phonétique du signal. Dans cette optique, l'*entrainement* oscillatoire thêta ne consisterait pas à « suivre » (*track*) passivement les caractéristiques acoustiques du signal verbal, mais jouerait un rôle fonctionnel de découpage linguistique du signal en paquets syllabiques. Toutefois, pour d'autres auteurs, ce couplage cerveau/parole dans la bande thêta serait plutôt lié à une activité oscillatoire supplémentaire, à savoir une modification de l'activité corticale en réponse à la stimulation auditive et notamment à la superposition des réponses auditives évoquées par l'attaque (*onset*) des syllabes (Ding & Simon, 2012, 2014). Une telle hypothèse suggère ainsi un rôle plus passif du système cortical oscillatoire dans le traitement du signal de parole. A ce jour, aucun consensus ne semble encore avoir été atteint quant à ces deux hypothèses. Cependant, un certain nombre de travaux semblent aller dans le sens de la première hypothèse, suggérant l'existence d'un processus actif de synchronisation des oscillations corticales sur la parole. Tout d'abord, l'activité thêta mesurée au repos chez l'adulte au sein des régions auditives suggère bien une prédisposition de ces régions à osciller à un rythme proche du rythme syllabique de la parole (Giraud et al., 2007; Morillon, Lehongre, Frackowiak, Ducorps, & Kleinschmidt, 2010). Lorsque le signal de parole est présenté à un rythme trop élevé par rapport au rythme oscillatoire endogène thêta des régions auditives, le traitement est plus difficile et la parole devient incompréhensible (Ahissar et al., 2001). Par ailleurs, comme mentionné

précédemment, Gross et collègues (2013) ont montré que la réponse évoquée après chaque crête (*edge*) dans l'enveloppe d'amplitude (indépendamment du signal verbal) suivait une dynamique similaire à la synchronisation cortico-acoustique sur le signal de parole. Cependant, dans une fenêtre temporelle plus tardive, alors que la synchronisation oscillatoire sur la parole perdurait, la réponse évoquée s'atténuait. Ces résultats suggèrent donc que les oscillations corticales sont entraînées par la parole au-delà d'un simple recalage de phase transitoire évoqué par les crêtes dans l'enveloppe. « *Speech continuously entrains brain rhythms beyond a stereotypical short-lived phase reset evoked by edges* » (Gross et al., 2013). Les résultats de notre étude montrant une augmentation de cohérence cortico-acoustique dans la bande thêta (4-7 Hz) et la bande de fréquence centrée sur le débit syllabique normal (5.6-7.6 Hz), non accompagnée d'une augmentation de puissance dans les régions corticales concernées, vont également dans le sens d'un processus de synchronisation oscillatoire. Enfin, les études révélant l'effet facilitateur de séquences rythmiques (auditives ou motrices) en amont du traitement d'un stimulus langagier suggèrent également un *entrainment* oscillatoire plutôt qu'une superposition de réponses évoquées à chaque syllabe ou événement saillant (Bedoin, Brisseau, Molinier, Roch, & Tillmann, 2016; Cason & Schön, 2012; Falk & Dalla Bella, 2016; Przybylski et al., 2013). Nous reviendrons sur ces études dans la partie 3 du chapitre X de cette Discussion générale.

2. Dynamique oscillatoire lors de la perception de la parole à débit rapide

Jusqu'à présent, très peu d'études se sont intéressées à la synchronisation des oscillations corticales sur le rythme de la parole naturelle chez les enfants. Notre étude est ainsi la première à examiner les conséquences d'une accélération naturelle du débit syllabique sur la dynamique oscillatoire sous-tendant la perception de la parole. Nous avons mesuré la cohérence entre l'enveloppe temporelle de la parole produite à débit rapide et les oscillations corticales dans deux bandes de fréquence, la bande thêta (4-7 Hz) et une bande centrée sur le débit syllabique rapide moyen de nos phrases (8-10 Hz).

Dans la bande thêta, nos résultats chez les enfants neurotypiques francophones révèlent un couplage entre l'enveloppe d'amplitude du signal de parole rapide et l'activité oscillatoire des régions motrices, prémotrices et de Broca latéralisées à gauche. Une synchronisation à droite est également observée au sein des mêmes régions que celles mises en évidence dans la

condition de parole à débit normal, à savoir dans les cortex temporal supérieur antérieur et frontal inférieur. La parole naturelle rapide correspond à un signal acoustique que l'on peut considérer comme dégradé. Elle est caractérisée par des modifications à la fois temporelles et spectrales : la longueur des indices acoustiques est réduite et les gestes articulatoires sont effectués plus rapidement aboutissant à des phénomènes de réduction (Duez, 1995, 1998, 1999; Kohler, 1992; Max & Caruso, 1997; Van Son & Pols, 1996, 1999). Afin de traiter ces variations du signal de parole, des régions cérébrales supplémentaires appartenant à la voie dorsale sensori-motrice seraient recrutées (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009). Cette voie, supposée être prédominante à gauche (Hickok & Poeppel, 2007), connecte les régions temporales postérieures aux régions prémotrices dorsales et à l'aire de Broca, *via* le cortex pariétal inférieur (aire Spt au niveau de la jonction temporo-pariétale, cf. chapitre I.4 et Figure 7 p.65). Elle permet ainsi d'apparier les représentations phonologiques aux schémas articulatoires de l'auditeur afin de faciliter la catégorisation des sons de parole et d'accéder à la compréhension du message linguistique. Le recrutement de cette voie lors de l'écoute de parole dégradée a largement été démontré chez l'adulte que ce soit lors de l'écoute de parole vocodée ou inversée (Hervais-Adelman et al., 2012; Osnes et al., 2011), de parole dans du bruit (Alho et al., 2012, 2014; D'Ausilio et al., 2012; Du et al., 2014) ou de parole compressée temporellement de manière artificielle (Adank & Devlin, 2010; cf. chapitre I.3).

D'après le modèle proposé par Strauß et Schwartz (2017) chez l'adulte, nous faisons l'hypothèse que chez les enfants, lorsque la parole entendue est produite à un débit rapide, les modulations lentes de l'enveloppe d'amplitude, qui véhiculent l'information syllabique et sont donc fortement liées aux mouvements articulatoires sous-tendant la production des syllabes (MacNeilage & Davis, 2001), entraîneraient les oscillations corticales thêta au sein des régions (pré)motrices gauches de la voie dorsale afin de pallier la difficulté de traitement engendrée par l'accélération naturelle du débit. Dans cette optique, l'alignement des régions motrices sur le signal de parole, en mobilisant les propres représentations articulatoires de l'auditeur, permettrait de faciliter l'extraction de l'information syllabique, qui serait plus difficile à réaliser pour les régions temporales auditives (droites selon le modèle AST de Poeppel, 2003). Comme nous l'avons mentionné précédemment, les études précédentes en IRMf ou en EEG/MEG ont rapporté une activation des régions (pré)motrices ou une communication renforcée entre ces régions et les régions auditives lors du traitement de parole dégradée. Cependant, aucune à notre connaissance n'a montré un couplage direct entre le rythme thêta dans le cortex moteur et le rythme de la parole ; nos résultats chez l'enfant

neurotypique sont ainsi les premiers à mettre en lumière un tel couplage. Nos données s'accordent par ailleurs avec les études chez l'adulte examinant l'activité oscillatoire au repos, révélant la présence d'un rythme thêta intrinsèque au sein des régions motrices (gauches) impliquées dans les mouvements articulatoires de la parole (Giraud et al., 2007; Morillon et al., 2010, 2012). De plus, des analyses de corrélation ont montré que l'activité oscillatoire endogène au sein de ces régions gauches était couplée à l'activité mesurée au niveau de la région temporo-pariétale gauche (BA 40, proche de l'aire Spt), suggérant l'existence d'un couplage auditivo-moteur intrinsèque (Morillon et al., 2010). Nos résultats seraient là encore à approfondir en examinant la dynamique oscillatoire de repos chez les enfants de 8 à 13 ans afin de déterminer si les régions (pré)motrices gauches expriment le rythme thêta comme chez l'adulte. Compte-tenu du fait que les enfants sont encore en plein développement cognitif et langagier, et du rôle joué par les régions de production de la parole dans son acquisition et son développement (Doupe & Kuhl, 1999; Goswami, 2011; Hickok & Poeppel, 2007; Pulvermüller & Fadiga, 2010; Tierney & Kraus, 2014), nous pouvons en outre supposer un couplage thêta entre les régions auditives et (pré)motrices d'autant plus fort chez les enfants que chez les adultes au repos.

Si les oscillations corticales des régions auditives sont capables de suivre le rythme syllabique, nous nous attendions à une augmentation de la cohérence entre l'enveloppe temporelle du signal de parole rapide et l'activité oscillatoire dans la bande de fréquence correspondant au débit syllabique moyen de la parole rapide, à savoir 8-10 Hz. Cependant, nos résultats ne montrent aucune augmentation significative de la synchronisation des oscillations corticales sur le rythme syllabique de la parole rapide chez notre groupe d'enfants neurotypiques. Très peu d'études ont étudié l'*entrainement* cortico-acoustique sur de la parole rapide et aucune à notre connaissance ne s'est intéressée à l'accélération naturelle du débit. Dans une étude en EEG chez des adultes francophones utilisant de la parole compressée temporellement, Pefkou et collaborateurs (2017) ont récemment rapporté que les oscillations corticales (sur 5 électrodes « auditives ») étaient capables de suivre le rythme syllabique de phrases non compressées, compressées à un facteur 2 et à un facteur 3. Ainsi, la densité cross-spectrale était plus forte à 3-5 Hz pour la parole non compressée, à 7-9 Hz pour la parole compressée à un facteur 2 et à 10-14 Hz pour la parole compressée à un facteur 3. Ces résultats suggèrent que le rythme syllabique peut entraîner les oscillations corticales même lorsque le signal de parole devient inintelligible. Les auteurs ont néanmoins rapporté que

l'alignement entre la réponse corticale et le signal de parole était meilleur pour un rythme syllabique lent que rapide. Ces données suggèrent donc que l'activité neuronale s'adapte au rythme syllabique de la parole au-delà de la limite supérieure de la bande thêta (telle qu'elle est définie classiquement), à savoir jusqu'à 14 Hz, contredisant ainsi les prédictions de Ghitza (2014) pour qui l'alignement cerveau/parole se produirait jusqu'à 9 Hz (« *thêta-syllable* »), c'est-à-dire lorsque la parole reste intelligible. Les résultats de Pefkou et al. sont également en contradiction avec ceux rapportés par Ahissar et collègues (Ahissar et al., 2001) chez des adultes anglophones, montrant un couplage entre les réponses du cortex temporal et l'enveloppe temporelle de la parole compressée uniquement lorsque celle-ci restait intelligible (rythme des modulations lentes de l'enveloppe < 9 Hz). Dans une étude réalisée dans le cadre du projet ODYSSEE (ANR JC ; PI : V. Boulenger) chez des adultes francophones et utilisant les mêmes phrases produites naturellement que celles présentées aux enfants dans la présente étude, nous avons observé une augmentation de cohérence entre les oscillations des régions auditives et l'enveloppe de la parole dans la bande 8-10 Hz dans la condition de parole à débit rapide (Hincapié, Guiraud, Jerbi et Boulenger, en préparation, Annexe 1). Contrairement à ce que nous avons observé chez les enfants neurotypiques, un *shift* (changement) de bande de fréquence se produit donc chez les adultes lorsque les oscillations corticales doivent s'aligner sur un débit syllabique normal ou rapide. Nos résultats chez les enfants pourraient être expliqués de deux manières. Ils peuvent tout d'abord évoquer une immaturité du système cortical auditif des enfants à s'ajuster à l'augmentation naturelle du débit syllabique de la parole. Ceci s'accorderait avec l'alignement oscillatoire thêta observé dans les régions (pré)motrices gauches de la voie dorsale pour la parole rapide. Les études développementales ont montré la présence d'une période charnière au niveau de l'organisation corticale oscillatoire. Les oscillations delta et thêta seraient prédominantes par rapport aux oscillations alpha et beta chez les enfants avant l'âge de 10 ans ; au-delà de cet âge, la proportion s'inverserait (Campbell & Feinberg, 2009; Feinberg & Campbell, 2010; Whitford et al., 2007). Notre groupe d'enfants ayant un âge compris entre 8 et 13 ans et se situant donc autour de cette période charnière, l'hypothèse d'un système oscillatoire encore immature est tout à fait plausible. Une comparaison de la cohérence cortico-acoustique observée chez les enfants âgés de 8 à 10 ans et les enfants âgés de 11 à 13 ans pourrait alors être intéressante, nécessitant toutefois d'augmenter la taille de notre échantillon pour des analyses suffisamment robustes. Nos résultats peuvent également suggérer une difficulté de compréhension des phrases produites à un débit rapide chez les enfants testés. Dans notre étude, en raison d'une durée d'expérience déjà longue, nous n'avons pas pu vérifier la

compréhension des phrases entendues. D'après les résultats d'Ahissar et collègues (2001) et de différentes études ayant montré un lien entre intelligibilité et synchronisation cortico-acoustique (Ahissar et al., 2001; Luo & Poeppel, 2007), le couplage cerveau/parole ne se produirait pas, ou serait très faible, lorsque les stimuli sont inintelligibles. Il est donc possible que dans notre étude, aucune augmentation de cohérence n'ait été observée dans la bande 8-10 Hz (centrée sur le débit rapide) pour les phrases accélérées naturellement parce-que les enfants n'ont pas réussi à les décoder. Toutefois, une expérience comportementale pilote chez des enfants neurotypiques francophones du même âge avait permis de montrer des scores d'intelligibilité de 74% pour les phrases à débit rapide contre 84% pour les phrases à débit normal. Les performances dans la condition de débit rapide ne semblent donc pas être liées au hasard ou particulièrement faibles. Nous pouvons donc supposer que dans l'étude en MEG, les phrases produites à débit rapide étaient, au moins pour une grande majorité, compréhensibles. L'inclusion d'un nombre supplémentaire d'enfants serait nécessaire afin d'augmenter la puissance statistique de nos analyses et ainsi de vérifier si une augmentation de cohérence dans la bande 8-10 Hz lors de la perception de parole rapide, identique à celle observée chez les adultes dans le même paradigme (Hincapié et al., en préparation), peut être mise en évidence ou non chez les enfants.

Pour terminer, les analyses de modulations de puissance dans les bandes de fréquence d'intérêt révèlent une désynchronisation alpha (dans la partie basse, 8-10 Hz) au sein des régions temporales supérieures et médianes gauches (bilatérales en non-corrigé) lors de l'écoute de parole à débit à la fois normal et rapide. Ce résultat est en accord avec la désynchronisation du rythme alpha auditif, aussi appelé rythme tau (Lehtelä, Salmelin, & Hari, 1997), décrite lors du traitement de stimuli auditifs verbaux ou non (Dimitrijevic, Smith, Kadis, & Moore, 2017; Obleser & Weisz, 2012; Scharinger, Bendixen, Trujillo-Barreto, & Obleser, 2012; Strauß, Wöstmann, & Obleser, 2014; Weisz, Hartmann, Müller, Lorenz, & Obleser, 2011). Lors du traitement d'un signal auditif, ce rythme est supprimé dans les régions temporales supérieures afin d'orienter l'attention vers le stimulus et donc permettre un traitement cognitif plus actif (Dimitrijevic et al., 2017; Obleser & Weisz, 2012b; Strauß, Wöstmann, & Obleser, 2014; Weisz, Hartmann, Müller, Lorenz, & Obleser, 2011). Cette désynchronisation alpha a largement été observée chez l'adulte (Becker, Pefkou, Michel, & Hervais-Adelman, 2013; Dimitrijevic et al., 2017; Krause, Pesonen, Björnberg, & Hämäläinen, 2007; Krause, Sillanmäki, Häggqvist, & Heino, 2001; Meyer, Obleser, &

Friederici, 2013; Obleser & Weisz, 2012), mais très peu d'études l'ont décrite chez l'enfant (Fujioka & Ross, 2008). Par ailleurs, la plupart ont utilisé des stimuli non verbaux ou des mots isolés et la seule étude à notre connaissance chez l'enfant a révélé des résultats similaires aux nôtres chez des enfants anglophones de 6 ans lors de l'écoute de sons de violon ou de bruits blancs (Fujioka & Ross, 2008). Nos résultats sur de la parole naturelle fournissent ainsi des preuves supplémentaires que le rythme alpha chez l'enfant de 8 à 13 ans se désynchronise dans le cortex auditif lors de l'écoute de phrases, quel que soit le débit auquel elles sont produites.

3. Perspectives chez les enfants neurotypiques

Les résultats de ce travail chez les enfants neurotypiques apportent de premiers éléments expérimentaux quant au couplage entre oscillations corticales dans la bande thêta et rythme syllabique de la parole naturelle. Plusieurs questions restent toutefois en suspens et mériteraient des analyses et études complémentaires.

Nous avons mentionné l'importance des interactions auditivo-motrices dans l'acquisition et le développement du langage chez l'enfant et leur rôle dans les processus de perception de la parole, *via* la voie dorsale temporo-pariéto-prémotrice (Doupe & Kuhl, 1999; Goswami, 2011; Hickok & Poeppel, 2007; Pulvermüller & Fadiga, 2010; Tierney & Kraus, 2014). Nos résultats en MEG chez les enfants francophones de 8 à 13 ans suggèrent que la voie dorsale soit recrutée lors de la perception de parole à débit normal, avec une latéralisation à droite (cortex pariétal et frontal supérieurs), ainsi que lors de la perception de parole rapide, avec une implication des régions prémotrices gauches. Un certain nombre d'études chez les adultes suggèrent que le cortex moteur soit impliqué dans un codage prédictif temporel de l'information verbale : il « surveillerait » le rythme du signal auditif de parole afin de générer des prédictions sur le *timing* des unités linguistiques pertinentes, optimisant ainsi le traitement perceptif (Arnal & Giraud, 2012; Fujioka, Trainor, Large, & Ross, 2009; Morillon & Baillet, 2017; Morillon & Schroeder, 2015; Morillon et al., 2014). Ainsi, le cortex moteur aurait une influence *top-down* sur les régions auditives dans la bande de fréquence thêta, dans le but de prédire l'information articulatoire correspondant au signal entrant et ainsi aider au décodage du message linguistique au niveau des régions auditives (Abrams et al., 2009; Ahissar et al., 2001; Park et al., 2015; Pefkou et al., 2017). Une

poursuite logique de notre travail serait donc de réaliser des analyses de connectivité cortico-corticale au sein des régions auditives et (pré)motrices dans la bande thêta afin de préciser le rôle des régions de la voie dorsale dans la perception et le traitement de la parole, à la fois à débit normal et rapide, chez l'enfant. Par ailleurs, des études combinant EEG et IRMf au repos chez l'adulte ont montré un couplage entre l'activité oscillatoire spontanée thêta enregistrée au niveau des régions auditives et celle enregistrée au niveau du cortex moteur ventral impliqué dans les mouvements de la bouche et dans la production des syllabes (Giraud et al., 2007; Morillon, Lehongre, Frackowiak, Ducorps, & Kleinschmidt, 2010). Comme souligné précédemment, la dynamique oscillatoire au repos n'a que très peu été étudiée chez l'enfant et il serait particulièrement intéressant d'examiner le rythme oscillatoire endogène des enfants de notre étude au sein des régions temporales auditives, motrices et prémotrices afin de vérifier si comme chez l'adulte, ces régions oscillent naturellement dans la bande de fréquence thêta.

De récentes études suggèrent que les influences *top-down* des régions (pré)motrices de la voie dorsale ne soient pas limitées à la bande thêta. Alho et collaborateurs (2014) ont ainsi mis en évidence une connectivité renforcée dans la bande beta (~20 Hz) au sein de l'hémisphère gauche, entre le cortex prémoteur dorsal et la jonction temporo-pariétale, ainsi qu'entre le cortex auditif et le cortex prémoteur ventral, dans une tâche où des adultes devaient identifier des syllabes bruitées. Morillon et Baillet (2017) ont également montré un effet *top-down* des oscillations beta (18-24 Hz) du cortex sensorimoteur gauche, supposées représenter les prédictions temporelles, sur le cortex auditif associatif droit lors de l'écoute passive de stimuli rythmiques (i.e. séquences de tons). Keitel et collègues (2017) ont quant à eux révélé que le couplage des oscillations delta du cortex auditif sur l'enveloppe d'amplitude de la parole était modulé par l'activité oscillatoire de différentes régions corticales dans des bandes de fréquence différentes. Les oscillations beta dans les régions orbito-frontales (bilatérales avec une prédominance gauche) modulaient ainsi le couplage cortico-acoustique dans le gyrus temporal supérieur gauche, les oscillations alpha des régions centrales bilatérales modulaient négativement le couplage au niveau du gyrus de Heschl et du gyrus temporal supérieur antérieur gauche et enfin, les oscillations thêta au niveau pariétal modulaient négativement le couplage dans le gyrus temporal postérieur droit. Aucune étude à notre connaissance n'a investigué le couplage oscillatoire auditivo-moteur dans ces différentes bandes de fréquence chez l'enfant. Nos analyses préliminaires de modulations de puissance dans la bande beta (15-30 Hz) suggèrent une désynchronisation dans le cortex

temporal antérieur bilatéral (significative dans les analyses statistiques corrigées) mais aussi possiblement dans le cortex frontal inférieur, ainsi qu'une augmentation de puissance dans le cortex (pré)moteur gauche, en particulier pour la parole à débit rapide, chez notre groupe d'enfants (cf. Annexe 2). Il serait important d'approfondir ces analyses, notamment en augmentant la taille de notre échantillon, ainsi que d'étudier la connectivité entre les régions (pré)motrices et temporales dans la bande beta.

Un dernier type d'analyse cortico-corticale (phase-amplitude), déjà mentionnée dans la partie 1 de ce chapitre, serait également à réaliser afin de corroborer notre interprétation d'un rôle de l'activité oscillatoire gauche thêta dans le couplage avec les oscillations gamma. Ce couplage a été défini comme le *theta-gamma nesting* (Giraud & Poeppel, 2007) et est supposé refléter le rôle intégratif de haut niveau des oscillations thêta dans le cortex temporal antérieur gauche dans le découpage phonémique par les oscillations gamma dans les régions plus postérieures et dans l'accès aux représentations phonologiques (Giraud et al., 2007; Morillon et al., 2012; Peelle & Davis, 2012). Nos résultats révélant une augmentation de la cohérence cortico-acoustique au sein de la région temporale antérieure gauche lors de la perception de parole à débit normal dans la bande de fréquence centrée sur le débit syllabique moyen des phrases correspondant par ailleurs à une partie de la bande thêta, pourraient s'accorder avec les études précédentes. Des analyses complémentaires mesurant les modulations de puissance dans la bande gamma sont toutefois à réaliser afin d'analyser le couplage thêta-gamma au sein des régions temporales gauches lors de la perception de parole intelligible chez les enfants.

Dans cette étude en MEG, nous nous sommes intéressés au traitement d'un signal de parole produit par un adulte chez l'enfant. Mais le débit moyen de parole des enfants n'est pas le même que celui des adultes : chez le jeune enfant, il est en effet de 2-3 syllabes/s puis augmente jusqu'à l'âge adulte, atteignant une moyenne de 5-7 syll/s (Logan, Byrd, Mazzocchi, & Gillam, 2011; Pellegrino et al., 2011; Pindzola, Jenkins, & Lokken, 1989; Sturm & Seery, 2007). Ainsi, le rythme oscillatoire des cortex auditif et (pré)moteur de l'enfant pourrait ne pas être adapté parfaitement au traitement du rythme syllabique d'un adulte, ce qui pourrait expliquer la cohérence cortico-acoustique observée dans les régions de la voie dorsale lors de l'écoute des phrases à débit normal et surtout rapide. Dans notre étude, le débit syllabique moyen des enfants dans la tâche narrative de « L'histoire de la Grenouille » (Mayer, 1969) était de 3.85 syllabes/s, ce qui est 1.6 fois inférieur au débit de

parole moyen des stimuli de la condition normale (6.61 syllabes/s) et près de 2.5 fois inférieur au débit rapide (9.03 syllabes/s). Il serait alors intéressant de réaliser une étude identique à celle-ci mais avec des stimuli phrastiques produits par des enfants du même âge que nos participants, afin d'observer si la synchronisation cortico-acoustique est plus forte lorsque les enfants écoutent un enfant parler, comparé à un locuteur adulte. De tels résultats pourraient appuyer d'autant plus le possible rôle prédictif des régions motrices sur l'alignement oscillatoire des régions auditives sur l'enveloppe temporelle de la parole afin de traiter efficacement le signal entrant.

Enfin, nous aimerions approfondir un dernier type d'analyse que nous avons initiée dans le but de comparer la dynamique oscillatoire sous-tendant le traitement de la parole naturelle chez l'adulte et l'enfant francophones. Dans notre étude utilisant le même paradigme chez des adultes (Hincapié et al., en préparation, cf. Annexe 1), nous avons observé un *entrainement* des oscillations gamma hautes centrées sur la valeur moyenne de la fréquence fondamentale (F0) de nos stimuli à débit normal (76.8-80.8 Hz) et rapide (86.1-90.1 Hz) au niveau du cortex auditif droit (condition normale) ou bilatéral (condition rapide). La F0 est caractérisée par la fréquence de vibration de base des cordes vocales, définissant la hauteur (ou *pitch*) au niveau acoustique d'un son. Différentes études ont montré que l'information temporelle rythmique proche de la F0 d'un signal acoustique de parole ou d'un signal non verbal se reflétait au niveau de la fréquence des oscillations corticales au sein des régions auditives chez l'adulte (Griffiths et al., 2010; Hertrich, Dietrich, & Ackermann, 2013; Hertrich et al., 2012; Hertrich, Mathiak, Lutzenberger, & Ackermann, 2004; Kerlin et al., 2010). Dans nos analyses chez les enfants neurotypiques, nous avons observé une augmentation de la cohérence entre le signal acoustique de parole et les oscillations au niveau des gyri temporaux supérieur et médian, dans les bandes de fréquence correspondant à la F0 moyenne des deux conditions de débit de parole (cf. Annexe 3). Ainsi, une augmentation significative de la cohérence cortico-acoustique au sein de ces régions bilatérales est observée lors de la perception de parole à débit normal (et pas à débit rapide) dans la bande de fréquence englobant la F0 moyenne des phrases de cette condition (76.8-80.8 Hz). De la même manière une augmentation significative de cohérence est observée dans ces mêmes régions, même si principalement latéralisées à droite, dans la bande de fréquence englobant la F0 moyenne des phrases de la condition rapide (86.1-90.1 Hz), lors de la perception de parole rapide. Notons toutefois qu'un pic significatif de cohérence est également observé au sein du cortex temporal

supérieur et médian gauche pour les phrases à débit normal dans cette bande de fréquence centrée sur la F0 de la parole rapide. Ces résultats suggèrent que l'alignement des oscillations cérébrales sur le signal de parole dans les régions auditives ne se limite pas aux modulations de basses fréquences de l'enveloppe d'amplitude de la parole, mais qu'il se produit également dans des fréquences plus hautes englobant notamment la fréquence fondamentale du locuteur. De plus, cet *entraînement* semble « suivre » précisément la F0 du signal entrant puisque nous observons un *shift* dans la bande de fréquence dans laquelle la synchronisation corticale est observée, qui correspond à l'augmentation de la F0 avec l'augmentation du débit de parole. Des résultats comparables ont été rapportés chez l'adulte, notamment par Hertrich et collaborateurs (2012, 2013) qui ont montré une synchronisation des oscillations corticales sur la F0 d'un signal de parole continue (~ 100 Hz) et cela indépendamment du rythme syllabique du signal de parole (la synchronisation était présente que le débit soit normal ou très rapide par accélération artificielle). Nos premiers résultats chez l'enfant semblent donc intéressants et méritent d'être approfondis afin de révéler le rôle précis joué par la F0 dans la dynamique oscillatoire sous-tendant le traitement de la parole chez l'enfant.

Notre principal intérêt était d'étudier les processus sous-tendant la perception de la parole des enfants ayant acquis le langage oral et écrit (âgés de plus de 8 ans) et avant leur entrée dans l'adolescence (13 ans). Notre étude 1 est la première à mettre en évidence, chez des enfants de cet âge, la dynamique oscillatoire lors du traitement de stimuli phrastiques naturels produits à différents débits de parole. Les études antérieures chez l'enfant ont pour la plupart utilisé des stimuli non-verbaux (musicaux ou bruits blancs modulés en amplitude ; Telkemeyer et al., 2009; Telkemeyer et al., 2011; Vanvooren, Poelmans, Hofmann, Ghesquière, & Wouters, 2014) ou verbaux simples (répétition de syllabes ; Power, Mead, Barnes, & Goswami, 2012) produit à un rythme normal, et ont observé une synchronisation du cortex auditif bilatéral dans les basses fréquences oscillatoires (delta/thêta) sur le signal auditif, que ce soit chez les nouveau-nés, les enfants de 5 ans ou de 13 ans. Une seule étude, déjà présentée dans la partie 1 de ce chapitre, a montré une préférence hémisphérique droite chez des enfants hispanophones âgés en moyenne de 11.3 ans lors du couplage des oscillations corticales thêta sur l'enveloppe temporelle de phrases produites à un débit normal (Molinario et al., 2016). Cette étude a été réalisée dans la langue espagnole qui est une langue syllabique comme le français, alors que les autres travaux ont été menés sur les langues anglaise ou allemande, qui sont des langues accentuelles. Il serait alors intéressant de tester,

dans un même paradigme, différents groupes d'enfants parlant des langues syllabiques différentes (français, espagnol, italien, etc.), afin de déterminer si la préférence hémisphérique à droite dans la bande « classique » thêta (4-7 Hz), mais également à gauche dans la bande de fréquence centrée sur le rythme syllabique moyen des stimuli est observée dans toutes les langues syllabiques, et si cette dynamique diffère de celle observée pour des langues non syllabiques (i.e. accentuelles).

Par ailleurs, une perspective à plus long terme pourrait être d'étudier la trajectoire développementale de la dynamique oscillatoire du traitement du rythme de la parole, et notamment comment le système cortical oscillatoire de l'enfant francophone à des âges différents (dès 4-5 ans par exemple) réagit face à un signal de parole produit naturellement à un débit normal ou rapide. Comprendre les processus sous-tendant le traitement du rythme de la parole chez l'enfant est primordial puisque comme les adultes, les enfants sont en permanence confrontés à des interlocuteurs différents parlant à des vitesses différentes, mais contrairement aux adultes, leur système langagier est encore en plein développement et il est tout à fait possible que les variations du rythme de la parole aient un impact plus fort sur l'efficacité du traitement de l'information linguistique. Déterminer à quel point les rythmes cérébraux sont capables d'être entraînés par le signal de parole à différents débits chez des enfants d'âges différents pourrait alors permettre de donner des recommandations à l'école notamment, pour sensibiliser les enseignants à l'importance du rythme de leurs productions afin de favoriser une compréhension et un apprentissage efficace chez les enfants.

Dans l'ensemble, nos résultats chez enfants neurotypiques fournissent donc les premières preuves expérimentales d'un alignement des oscillations corticales sur l'enveloppe de la parole naturelle. Des analyses supplémentaires sont toutefois envisagées afin de mieux comprendre le fonctionnement de cet *entraînement* et ses implications dans les processus de traitement de la parole.

X. Rythme et perception de la parole dans la dysphasie

L'un des intérêts de ce travail de thèse était d'étudier les processus de perception du rythme de la parole dans les troubles développementaux du langage oral et notamment dans la dysphasie. Pour cela, nous avons réalisé deux études, l'une comportementale et l'autre en MEG, présentées dans le chapitre VIII de la partie expérimentale, chez des enfants francophones présentant une dysphasie touchant majoritairement le versant expressif de la parole.

La dysphasie est un trouble neurodéveloppemental sévère et durable de l'élaboration du langage oral. Elle peut toucher l'ensemble des composantes langagières de l'expression et/ou de la compréhension du langage de manière plus ou moins sévère, lui conférant un certain nombre de formes différentes, répertoriées dans différentes classifications. La plus utilisée par les cliniciens en France est celle de Rapin et Allen (1983). Dans ce travail de thèse, au vu des liens étroits entre systèmes de perception et de production de la parole, nous avons choisi d'étudier les capacités de traitement du rythme de la parole des enfants présentant une dysphasie touchant majoritairement le versant expressif de la parole, correspondant au syndrome phonologico-syntaxique (le plus répandu, caractérisé par un trouble de la programmation phonologique de la parole, la compréhension étant relativement épargnée).

1. Capacités de perception de la parole rapide des enfants dysphasiques : étude comportementale

La perception efficace des indices rythmiques de la parole est primordiale pour le développement du langage et notamment pour la mise en place des représentations phonologiques. Un certain nombre d'études ont révélé un trouble de la perception des indices rythmiques de la parole dans la dysphasie, que ce soit au niveau de la perception des segments de durée brève (déficit de traitement temporel ; Tallal et al., 1996; Tallal & Piercy, 1973a, 1973b, 1975) ou des indices suprasegmentaux (*rise time* de l'enveloppe d'amplitude; Corriveau, Pasquini, & Goswami, 2007; Cumming, Wilson, & Goswami, 2015a; Richards & Goswami, 2015). À l'aide d'un paradigme expérimental original, nous avons montré que les enfants présentant une dysphasie expressive avaient des difficultés à traiter un signal de parole entrant lorsque celui-ci était produit à un débit rapide. Ces difficultés étaient par

ailleurs observées que l'accélération soit naturelle ou artificielle (par compression temporelle). Dans cette étude, les enfants avaient pour tâche de juger de la cohérence sémantique de phrases produites à un débit naturel normal, naturel rapide ou compressées linéairement. Nos résultats ont ainsi montré que les enfants dysphasiques sont moins sensibles à l'incohérence sémantique des phrases lorsqu'elles sont produites à un débit rapide (naturel ou compressé) comparé à des enfants neurotypiques du même âge. Aucune différence entre les deux groupes n'a été observée dans la condition de parole naturelle normale. Ces résultats sont ainsi les premiers à notre connaissance à montrer l'effet négatif d'une augmentation de débit de parole sur les capacités de compréhension des enfants dysphasiques francophones. Le traitement de la parole accélérée, obligeant l'auditeur à faire face et à s'adapter à des modifications temporelles (et spectrales dans le cas d'une accélération naturelle) au sein du signal de parole, est un processus exigeant qui peut influencer sur les processus d'intégration de haut niveau tels que l'intégration sémantique des mots dans leur contexte. Ce processus semble ainsi particulièrement difficile chez les enfants présentant une dysphasie dont les capacités générales de compréhension sont par ailleurs préservées. Les évaluations neuropsychologiques actuelles examinant la compréhension du langage utilisent des mots ou des phrases généralement produits par l'évaluateur (neuropsychologue ou orthophoniste) à un rythme normal voire lent. Nos résultats suggèrent qu'il serait pertinent d'inclure des tests plus écologiques portant sur de la parole naturelle produite à des débits différents, plus rapides notamment, afin de mettre en évidence les difficultés potentielles rencontrées au quotidien par les enfants dysphasiques, même si leurs troubles s'expriment *a priori* principalement sur le plan expressif. Entraîner les enfants à la reconnaissance de parole à un rythme de plus en plus rapide pourrait également permettre de les aider à faire face plus facilement aux variations de débit auxquelles ils sont confrontés lors de leurs interactions. En outre, sensibiliser l'ensemble des professionnels de santé intervenant dans le suivi des enfants dysphasiques mais aussi les enseignants sur l'importance du débit de parole pourrait être important afin de faciliter la compréhension et ainsi la communication chez ces enfants.

Notre étude avait également pour objectif de comparer deux types d'accélération de la parole, naturelle et artificielle, ces deux accélérations différant par le type de modification structurelle qu'elles entraînent. La parole compressée (accélération artificielle) correspond à un simple raccourcissement de la durée des segments du signal de parole, soit à une modification exclusive de la structure temporelle. La parole naturellement produite à un débit

rapide quant à elle implique des changements temporels mais aussi spectraux du signal de parole. Lorsque le locuteur accélère son discours, il va raccourcir la durée de certains segments et ses gestes articulatoires seront moins précis (phénomènes de réduction ; Kohler, 1992), compliquant le traitement et le décodage du message linguistique (Janse, 2004; Janse et al., 2003). Nos résultats comparant les capacités de perception des différentes conditions de débit de parole entre les deux groupes d'enfants révèlent que les enfants dysphasiques ne bénéficient pas de la préservation de l'information spectrale dans le signal lors de la perception de parole compressée, contrairement aux enfants contrôles qui ont des performances sensiblement supérieures en parole compressée qu'en parole naturelle rapide. Cet avantage du traitement de la parole accélérée artificiellement par rapport à la parole accélérée naturellement a par ailleurs également été observé dans la littérature chez l'adulte sain (e.g., Janse, 2004). D'après les résultats de notre étude, nous pouvons supposer une difficulté générale de traitement de l'information temporelle de la parole lorsque celle-ci est présentée à un rythme rapide chez les enfants dysphasiques. Autrement dit, les enfants dysphasiques pourraient présenter des difficultés particulières pour détecter les régularités temporelles (i.e. battements syllabiques ici), véhiculées principalement par les modulations lentes de l'enveloppe d'amplitude, dans le signal de parole accéléré. Toutefois, afin de confirmer cette hypothèse, une étude complémentaire utilisant par exemple le même paradigme que celui développé par Ghitza & Greenberg (2009) pourrait être réalisée. Ces auteurs ont démontré, chez des adultes, une amélioration des performances d'intelligibilité de phrases compressées temporellement lorsque des silences étaient insérés périodiquement entre les segments de parole compressés de sorte à restaurer le rythme original des stimuli. Si nos résultats comportementaux sont le reflet d'un trouble d'entraînement sur le rythme syllabique plus rapide de la parole dans la dysphasie, les performances des enfants dysphasiques devraient également être améliorées dans cette condition. En effet, malgré la compression, chaque syllabe est présentée au moment attendu (en accord avec la théorie de l'attention dynamique, Large & Jones, 1999), ce qui facilite l'alignement et ainsi le découpage du flux continu de parole. En revanche, une absence d'amélioration des performances des enfants dans cette condition suggèrerait au contraire un trouble du traitement des unités raccourcies de la parole, indépendamment du rythme de présentation de l'information syllabique. Ceci s'accorderait alors plutôt avec la théorie du traitement rapide de l'information ("*Rapid Auditory Processing (RAP) theory*"; Tallal & Piercy, 1973a, 1975) selon laquelle le trouble développemental du langage oral (dysphasie) serait lié à un traitement déficitaire des indices acoustiques brefs (i.e. sons complexes brefs séparés par de courts intervalles ou phonèmes

différant de par leurs transitions formantiques rapides), conduisant à des représentations phonologiques imparfaites. Dans notre étude, les faibles performances des enfants dysphasiques lors de l'écoute de parole rapide, accélérée naturellement et artificiellement, pourraient donc être liées à un déficit de traitement des indices acoustiques rapides et non à un déficit de « suivi » (*tracking*) de la structure temporelle de la parole. Il faut noter que ces deux interprétations ne sont toutefois pas exclusives : les enfants dysphasiques pourraient en effet être perturbés à la fois par le raccourcissement des indices de la parole et par l'enchaînement des unités syllabiques à un rythme plus rapide.

L'accélération du débit de parole affecte ainsi non seulement la durée des segments mais aussi la dynamique temporelle de la parole en ce qu'elle accélère le rythme des modulations lentes d'amplitude de l'enveloppe temporelle. Ces modulations de basse fréquence sont caractérisées par des pics d'énergie, saillants dans le signal de parole, qui correspondent au noyau vocalique des syllabes : le rythme de ces modulations représente donc le rythme syllabique et ainsi le débit de parole. Comme nous l'avons déjà souligné, le rythme syllabique constitue un fondement de l'acquisition et du développement du langage oral : sur le versant expressif, il est présent très tôt dans les productions du jeune enfant (stade du babillage), et sur le plan perceptif, il permet de segmenter le signal continu de parole en segments de plus petites tailles correspondant aux syllabes (Goswami & Leong, 2013). Une accélération du débit syllabique se répercutera donc directement sur la structure temporelle de l'enveloppe d'amplitude, et le système auditif de l'auditeur devra s'adapter rapidement en s'alignant sur les battements syllabiques plus fréquents. Cet alignement permettrait de prédire l'occurrence des prochaines syllabes dans le flux de parole et ainsi de segmenter efficacement le signal verbal (Abrams et al., 2008; Greenberg et al., 2006). Les résultats de notre étude suggèrent que les enfants présentant une dysphasie expressive auraient plus de difficultés à extraire les informations pertinentes dans la structure temporelle de la parole lorsqu'elle est accélérée. Un certain nombre d'études ont révélé un déficit de perception du *rise time* (i.e. temps de montée précédant le pic d'énergie de la syllabe dans l'enveloppe d'amplitude ; Rosen, 1992) chez les enfants présentant des troubles développementaux du langage tels que la dysphasie et la dyslexie (Beattie & Manis, 2012; Cumming, Wilson, & Goswami, 2015a; Goswami, 2011; Pasquini, Corriveau, & Goswami, 2007). Cumming et collaborateurs (2015a) ont ainsi montré que les enfants dysphasiques présentaient des performances significativement plus faibles que des enfants contrôles dans des tâches où les indices rythmiques (*rise time* et durée) de stimuli verbaux étaient manipulés. Ces performances

étaient en outre fortement corrélées aux capacités de conscience phonologique des enfants. Afin de corroborer notre interprétation d'un trouble du traitement de la structure temporelle de la parole rapide dans la dysphasie sous-tendu par un mauvais alignement sur le rythme syllabique, il serait intéressant (1) d'examiner les capacités des enfants dysphasiques pour discriminer le *rise time* et la durée de stimuli verbaux simples et (2) de déterminer si elles sont corrélées à leurs performances de perception de phrases produites à un débit rapide comme dans la présente étude. A cet égard, l'une des critiques qui pourrait être faite sur notre étude est que les enfants devaient juger la cohérence sémantique des phrases et qu'un déficit sémantique, plutôt qu'un déficit de traitement rythmique, pourrait sous-tendre leurs difficultés à décoder la parole rapide. Cette interprétation n'est toutefois que peu plausible puisque les enfants dysphasiques testés présentaient des performances de compréhension verbale dans la moyenne (comme révélé par les tests neuropsychologiques) et qu'ils ne différaient pas des enfants contrôles pour comprendre les phrases produites à débit normal. L'une des manières de contourner ce possible facteur confondant serait d'inclure des stimuli phrastiques ne possédant pas de contexte sémantique très marqué (i.e. ne permettant pas aux enfants de deviner les mots ; e.g., « Le verre vrai ouvre le coin », voir Benoît, Grice, & Hazan, 1996), à l'instar de l'étude de Ghitza et Greenberg (2009) par exemple. Pour revenir aux troubles du rythme, des études ont rapporté qu'ils pouvaient se manifester non seulement au niveau perceptif mais aussi au niveau expressif. Les dysphasiques ainsi que les dyslexiques, sans trouble de la dextérité motrice, présentent en effet des difficultés à taper du doigt en rythme avec un signal auditif lent et ces performances sont fortement corrélées aux capacités de conscience phonologique des enfants (Corriveau & Goswami, 2009; Cumming, Wilson, Leong, Colling, & Goswami, 2015b; Goswami, 2012; Hill, 2001; Thomson, Fryer, & Maltby, 2006; Thomson & Goswami, 2008). Vérifier les capacités de synchronisation motrice des enfants sur un rythme normal ou rapide permettrait d'étudier l'envergure du trouble du rythme dans les différentes modalités et de préciser les liens entre systèmes moteur et perceptif lors du traitement d'un signal auditif rythmique.

Les hypothèses actuelles expliquant les troubles de la perception du rythme de la parole dans la dysphasie (et la dyslexie avec la théorie de l'échantillonnage temporel proposée par Goswami, 2011) suggèrent que la sensibilité réduite aux indices de *rise time* et de durée observée chez ces enfants affecterait l'extraction des indices rythmiques de basse fréquence contenus dans l'enveloppe temporelle du signal de parole, entravant la segmentation des

syllabes et des mots, et par conséquent le traitement de la structure phonologique à la base des représentations syntaxiques (Cumming, Wilson, & Goswami, 2015a). Une étude comportementale de Leong et Goswami (2014) a révélé un défaut de synchronisation sur le rythme syllabique de stimuli phrastiques chez des enfants dyslexiques par rapport à des contrôles, caractérisé par une synchronisation systématique sur la phase oscillatoire précédente. Ces résultats supportent l'idée d'un trouble de synchronisation sur les rythmes lents de la parole (rythmes syllabique et prosodique) dans la dyslexie, contribuant probablement au développement de représentations phonologiques atypiques, caractéristiques de ce trouble du langage écrit. D'après ces hypothèses et au vu de nos résultats, nous pouvons supposer l'existence d'un déficit d'entraînement des oscillations corticales de basse fréquence sur les modulations lentes de l'enveloppe temporelle de la parole dans la dysphasie, conduisant à une extraction inefficace de l'information syllabique et donc à un décodage altéré de l'information linguistique. Cet entraînement oscillatoire serait d'autant plus déficitaire lors de la perception de parole produite à un débit rapide, les indices rythmiques de la parole étant encore plus difficiles à extraire par les enfants dysphasiques. Notre seconde étude en MEG chez un autre groupe d'enfants dysphasiques a permis de tester cette hypothèse, en examinant la dynamique oscillatoire sous-tendant le traitement de la parole naturelle à débit normal et rapide.

2. Dynamique oscillatoire sous-tendant le traitement de la parole chez l'enfant dysphasique

La troisième et dernière étude de cette thèse avait pour objectif d'observer la synchronisation cortico-acoustique des enfants présentant une dysphasie expressive âgés de 8 à 13 ans comparés à des enfants neurotypiques du même âge. Pour cela, dans le même paradigme que celui de l'étude 1, nous avons mesuré la cohérence entre le signal MEG et l'enveloppe temporelle de phrases naturellement produites dans deux conditions de débit de parole (normal avec une moyenne de 6.61 syllabes/s, et rapide avec une moyenne de 9.03 syllabes/s), dans trois bandes de fréquences différentes : thêta (4-7 Hz), 5.6-7.6 Hz (centrée sur le débit moyen normal) et 8-10 Hz (centrée sur le débit moyen rapide). Nos résultats ont révélé un profil oscillatoire atypique chez les enfants dysphasiques lors de la perception de la parole dans les deux conditions de débit. Avant de discuter plus avant de ces résultats, nous tenons à préciser qu'il s'agit là de données relativement préliminaires, obtenues chez un

groupe de 11 enfants dysphasiques comparés à 11 enfants neurotypiques (sous-groupe des 14 enfants ayant pris part à l'étude 1). Afin de valider nos hypothèses et nos conclusions, l'inclusion de quelques enfants supplémentaires dans chaque groupe sera sans aucun doute nécessaire. Nous précisons également qu'afin d'éviter toute redondance, les résultats obtenus chez les enfants neurotypiques ayant été discutés dans le chapitre IX de cette Discussion générale, ils ne seront globalement pas repris ici.

2.1. Le cas de la parole à débit normal

Lors du traitement de la parole à débit normal, nos résultats ont mis en évidence que contrairement aux enfants contrôles, les enfants dysphasiques ne présentaient pas d'alignement de l'activité oscillatoire du cortex auditif associatif droit dans la bande thêta sur l'enveloppe d'amplitude. Le système oscillatoire des régions auditives droites chez les enfants dysphasiques semblerait donc moins sensible à l'information rythmique syllabique véhiculée par l'enveloppe temporelle du signal de parole entrant. Les modèles actuels de perception de la parole (*multi-time resolution models* ; Giraud & Poeppel, 2012; Poeppel, Idsardi, & van Wassenhove, 2008; Poeppel, 2003) suggèrent un couplage préférentiel de l'activité oscillatoire du cortex auditif droit sur les modulations lentes de l'enveloppe temporelle de la parole, en accord avec les résultats observés chez les enfants neurotypiques. Ce couplage serait permis par une prédisposition du système cortical oscillatoire interne pour détecter les régularités temporelles d'un stimulus acoustique afin de prédire les événements à venir tels que les syllabes, les mots et les phrases dans le cas de la parole (théorie de l'attention dynamique; Jones, 1976; Jones, Moynihan, MacKenzie, & Puente, 2002; Kotz, Schwartz, & Schmidt-Kassow, 2009; Large & Jones, 1999). Ce système oscillatoire interne serait notamment dépendant de processus attentionnels, variant eux-mêmes de façon périodique dans le temps et dont les pics d'énergie seraient orientés vers les événements extérieurs saillants. Ceci permettrait de générer des attentes temporelles sur l'occurrence d'événements futurs, dont l'intégration temporelle serait alors optimisée (cf. chapitre III.3 de l'Introduction). Les rythmes attentionnels sont supposés adapter leur phase et leur période aux variations des rythmes externes afin de maintenir la synchronie (Kotz et al., 2009; Lakatos, Karmos, Mehta, Ulbert, & Schroeder, 2008; Large & Jones, 1999; Morillon & Schroeder, 2015). Concernant la parole, le couplage cortico-acoustique serait réalisé sur l'attaque des modulations successives de l'enveloppe, englobant l'information relative au *rise time* (Goswami, 2012), ce qui permettrait de faciliter la segmentation du flux continu et par conséquent la compréhension

(Giraud & Poeppel, 2012; Poeppel, 2003; Peelle & Davis, 2012 pour une revue). Nous avons vu précédemment que les enfants dysphasiques présentaient des difficultés à extraire des indices rythmiques comme le *rise time* (Beattie & Manis, 2012; Corriveau et al., 2007; Cumming, Wilson, & Goswami, 2015a; Richards & Goswami, 2015) ainsi qu'à traiter l'information rythmique de la parole lorsqu'elle est produite à un débit rapide (étude 2, Guiraud et al., en révision). Ainsi, si cette information rythmique n'est pas correctement perçue, le système oscillatoire interne ne pourrait se synchroniser dessus (ou se synchroniserait mal), empêchant la mise en place de prédictions temporelles fines sur les futures unités syllabiques et prosodiques, et conduisant à un mauvais découpage et décodage de l'information verbale. Nos résultats obtenus en MEG chez les enfants dysphasiques semblent les premiers à démontrer un tel dysfonctionnement oscillatoire chez les enfants dysphasiques, en accord avec la « *prosodic phrasing hypothesis* » proposée par Cumming et collègues (2015a), elle-même basée sur le modèle d'échantillonnage temporel (*temporal sampling framework*, TSF) de Goswami pour la dyslexie. Il faut cependant noter que l'une des limites de notre travail est que, pour des raisons d'ordre pratique notamment, nous n'avons pas mesuré les capacités de discrimination du *rise time* chez notre groupe d'enfants, ni les performances d'intelligibilité des phrases présentées lors de l'acquisition en MEG. Ces mesures seraient à inclure dans une prochaine étude afin de disposer d'informations plus complètes et complémentaires sur le traitement des indices rythmiques de la parole chez les enfants dysphasiques francophones. Par ailleurs, l'absence de synchronisation cortico-acoustique dans les régions auditives droites pour la parole à débit normal pourrait sembler contradictoire avec les résultats que nous avons obtenus dans l'étude comportementale (étude 2). En effet, dans cette étude, les enfants dysphasiques présentaient des performances comparables à celles des enfants contrôles pour juger les phrases à débit normal. Une différence méthodologique majeure pourrait néanmoins rendre compte de ce pattern. Alors que les phrases étaient présentées par bloc de débit dans l'étude comportementale (d'abord parole normale puis parole rapide puis parole compressée), la présentation des stimuli phrastiques dans l'étude 3 en MEG était pseudo-randomisée (i.e. pas plus de 3 stimuli de la même condition consécutifs) de sorte qu'aucune adaptation ne se mette en place. Autrement dit, l'alignement des oscillations corticales sur l'enveloppe d'amplitude des phrases devait être « réactualisé » pour chaque stimulus. Ceci a probablement constitué une difficulté supplémentaire chez les enfants dysphasiques, pouvant expliquer qu'aucune cohérence cortico-acoustique n'ait été observée dans le cortex auditif droit même pour la parole à débit normal.

Dans ce travail de thèse, outre l'alignement dans la bande thêta classiquement utilisée dans les études précédentes, nous nous sommes également intéressés au traitement particulier du débit syllabique précis du signal de parole entrant, en mesurant la cohérence cortico-acoustique dans deux bandes de fréquences centrées sur le débit moyen de nos stimuli de parole (5.6-7.6 Hz et 8-10 Hz). Alors que chez l'enfant neurotypique nous avons observé un couplage spécifique dans la bande 5.6-7.6 Hz entre les oscillations du cortex temporal antérieur gauche et l'enveloppe temporelle des phrases produites à débit normal, ceci n'a pas été mis en évidence chez notre groupe d'enfants dysphasiques. Si cette activité oscillatoire à gauche reflète un processus d'intégration de plus haut niveau permettant de hiérarchiser les traitements syllabique et phonémique (*thêta-gamma nesting*) au sein du même hémisphère (mais dans des régions distribuées sur l'axe antéro-postérieur) et ainsi d'accéder aux représentations phonologiques (Giraud et al., 2007; Giraud & Poeppel, 2012; Morillon et al., 2012; Peelle et al., 2013; Scott, 2000), nos résultats pourraient donc laisser penser que ce couplage thêta-gamma soit moins fonctionnel chez les enfants dysphasiques, rendant potentiellement compte de leurs déficits phonologiques. Toutefois, cette interprétation reste spéculative et des analyses de couplage inter-fréquences devraient être réalisées avant de pouvoir tirer une telle conclusion.

Les travaux s'intéressant à la dynamique des oscillations cérébrales dans les troubles développementaux du langage, et notamment la dyslexie, ont révélé la présence d'un déficit de synchronisation cortico-acoustique dans la bande gamma basse, censée refléter l'échantillonnage phonémique du signal de parole (Giraud & Poeppel, 2012; Lehongre, Morillon, Giraud, & Ramus, 2013; Lehongre, Ramus, Villiermet, Schwartz, & Giraud, 2011). Lehongre et collaborateurs (2011, 2013) ont notamment observé une activité gamma (25-35 Hz) réduite au sein des régions auditives de l'hémisphère gauche ainsi que des régions articulaires motrices et somatosensorielles gauches chez des adultes dyslexiques francophones comparés à des sujets contrôles. Aucune étude à notre connaissance n'a été menée dans le cas de la dysphasie ; l'hypothèse selon laquelle un tel déficit d'extraction de l'information phonémique serait également présent dans ce trouble du langage oral reste donc à tester directement. Dans le cas où ce déficit serait avéré, nous pouvons alors supposer qu'il soit corrélé au mauvais échantillonnage thêta dans les régions temporales gauches. Dans ce travail de thèse, nous avons commencé à analyser les modulations de puissance dans différentes bandes de fréquence, incluant la bande gamma (Annexe 4), mais faute d'un nombre suffisant de participants, aucun résultat significatif n'a pour l'instant émergé. Le

recrutement de quelques enfants supplémentaires dans les deux groupes nous permettrait à encore d'examiner de façon plus robuste les modulations de puissance dans la bande gamma notamment, ainsi que d'analyser le couplage cortico-cortical de la phase des oscillations thêta sur l'amplitude des oscillations gamma afin de vérifier notre hypothèse d'un dysfonctionnement de ce couplage chez les enfants dysphasiques.

2.2. Le cas de la parole à débit rapide

Lorsque les enfants dysphasiques écoutaient de la parole rapide, nous avons observé une augmentation de la synchronisation entre les oscillations corticales dans la bande thêta au niveau du cortex temporal supérieur postérieur droit et les modulations lentes de l'enveloppe temporelle de la parole. Ceci n'était pas retrouvé chez les enfants neurotypiques chez qui le couplage cerveau/parole était plus fort au sein du cortex prémoteur gauche dans la bande thêta. Ce profil atypique chez les enfants dysphasiques a également été mis en évidence, quoique de façon moindre si l'on se réfère aux cartes statistiques, dans la bande englobant le débit syllabique moyen de la parole rapide (8-10 Hz) : la cohérence cortico-acoustique avait tendance à être augmentée chez les dysphasiques au niveau du cortex prémoteur dorsal droit et chez les contrôles au niveau du gyrus temporal médian postérieur droit.

Ces résultats suggèrent tout d'abord un recrutement plus important des régions motrices gauches lors de la perception de parole rapide chez les enfants neurotypiques par rapport aux enfants dysphasiques, ce qui suggère un dysfonctionnement de la voie dorsale chez ces derniers (Hickok & Poeppel, 2007; Rauschecker, 2011; Rauschecker & Scott, 2009; Rauschecker & Tian, 2000). Nous l'avons vu dans la partie IX.2 de cette Discussion générale, cette voie est impliquée dans le traitement sensori-moteur de la parole, intégrant l'information articulatoire (hébergée dans les régions prémotrices) aux représentations phonologiques (analysées dans les régions temporales postérieures) *via* la jonction temporo-pariétale. Elle est censée être particulièrement recrutée chez l'enfant dont le langage oral est en plein essor (Goswami, 2011; Hickok & Poeppel, 2007; Pulvermüller & Fadiga, 2010) et lors de conditions d'écoute difficile (Adank & Devlin, 2010; Du et al., 2014; Osnes et al., 2011; cf. chapitre II.4 de l'Introduction). Chez les enfants neurotypiques, nous avons donc interprété l'alignement de l'activité oscillatoire thêta des régions prémotrices gauches sur les modulations de basse fréquence de la parole comme reflétant le rôle de renfort de ces régions au cortex auditif pour décoder la parole rapide « dégradée » (cf. partie IX.2 de cette Discussion générale). L'absence d'un tel couplage chez les enfants dysphasiques suggère un

sous-recrutement de la voie dorsale gauche lors de la perception de parole rapide, qui pourrait expliquer les difficultés de perception de parole rapide observées chez un autre groupe d'enfants dysphasiques du même âge dans l'étude 2 de cette thèse (Guiraud et al., en révision). Par ailleurs, nous avons déjà mentionné que le trouble du rythme observé dans la dysphasie ne se limitait pas à la modalité auditive mais était aussi observé au niveau moteur (Corriveau & Goswami, 2009; Cumming, Wilson, Leong, Colling, & Goswami, 2015b; Goswami, 2012; Hill, 2001; Thomson, Fryer, & Maltby, 2006; Thomson & Goswami, 2008). Pour rappel, Corriveau et Goswami (2009) ont par exemple montré que des enfants dysphasiques anglophones âgés entre 7 et 11 ans n'étaient pas capables de battre du doigt correctement en rythme avec un métronome, en particulier lorsque le rythme était lent (1.5 et 2 Hz, correspondant à la « fréquence prosodique ») et que ces performances expliquaient en grande partie les compétences en phonologie et en lecture de ces enfants. Au vu de ces résultats et compte-tenu du fait que la synchronisation d'un mouvement sur un signal auditif rythmique nécessite un couplage cortical auditivo-moteur étroit, il est tout à fait possible d'envisager que le recrutement des régions sensori-motrices et le couplage entre les régions auditives et (pré)motrices soient défaillants chez les dysphasiques. Selon cette hypothèse, ce dysfonctionnement pourrait alors altérer le processus de traitement de la parole rapide chez ces enfants. A notre connaissance, aucune étude ne s'est encore intéressée à la connectivité entre les régions auditives et articulatoires chez les enfants dysphasiques. Examiner le couplage entre l'activité oscillatoire de ces régions lors de la perception de parole, notamment produite à un débit rapide, permettrait d'apporter des arguments expérimentaux en faveur de cette interprétation.

Contrebalançant en quelque sorte l'absence de cohérence dans les régions gauches de la voie dorsale, une synchronisation cerveau/parole a été mise en évidence dans les régions temporales supérieures postérieures droites, à proximité de la jonction temporo-pariétale, lors de la perception de parole rapide chez les enfants dysphasiques. Ce résultat était inattendu et pourrait s'expliquer par l'implication de ces régions, latéralisées à gauche normalement, dans les traitements phonologiques ainsi que dans l'intégration sensori-motrice de la parole notamment (Alho et al., 2014; Buchsbaum et al., 2001; Hickok & Poeppel, 2004; Papathanassiou et al., 2000). Comme nous l'avons décrit dans le chapitre V.3.1 de l'Introduction, des différences anatomiques et fonctionnelles ont été rapportées dans ces régions chez les dysphasiques par rapport à des contrôles (Jernigan et al., 1991; Plante, Swisher, & Vance, 1989; Plante et al., 1991), avec en particulier une absence d'asymétrie en

faveur de l'hémisphère gauche, voire une latéralisation à droite (Cutini et al., 2016; de Guibert et al., 2011; Tzourio-Mazoyer et al., 1994). Nos résultats, s'ils sont confirmés, pourraient alors refléter une « tentative » d'intégration sensori-motrice dans les régions droites pour s'aligner sur l'information rythmique de la parole rapide. Dans cette optique, la voie dorsale pourrait être partiellement fonctionnelle mais atypique chez les enfants dysphasiques : l'alignement oscillatoire sur les modulations de basse fréquence de la parole se produirait dans une région temporelle postérieure droite et non dans les régions prémotrices gauches, mais pourrait ne pas être suffisant pour garantir un décodage efficace de l'information verbale produite à un débit rapide, comme suggéré par les résultats de notre étude 2 comportementale.

De manière intéressante, lorsqu'on s'intéresse à la synchronisation cortico-acoustique dans la bande de fréquence englobant le débit rapide moyen de nos stimuli, les résultats tendent à montrer une cohérence plus forte dans les régions prémotrices droites (dans une partie plus dorsale que ce qui est observé dans la bande thêta chez les enfants contrôles) chez les enfants dysphasiques par rapport à leurs pairs qui présentent quant à eux une synchronisation plus marquée au sein des régions temporelles médianes postérieures droites. Ainsi, contrairement aux enfants contrôles, le système oscillatoire au sein des régions prémotrices chez les enfants dysphasiques semblerait capable de se synchroniser sur un rythme syllabique accéléré dans une bande de fréquence correspondant à ce rythme. Tout comme nous l'avons proposé pour les enfants neurotypiques (cf. partie IX.1. de la Discussion générale), il serait alors intéressant, et c'est l'un de nos objectifs, d'examiner la dynamique oscillatoire de repos chez les enfants dysphasiques afin d'identifier les rythmes exprimés de manière intrinsèque dans les régions auditives et (pré)motrices (Giraud et al., 2007; Morillon, Lehongre, Frackowiak, Ducorps, & Kleinschmidt, 2010; voir partie III.4.2 de l'Introduction). Si ces régions oscillent naturellement à un rythme atypique (notamment différent du rythme thêta), ceci pourrait rendre compte, au moins en partie, du défaut de synchronisation entre le rythme cortical thêta et le rythme syllabique observé chez notre groupe d'enfants dysphasiques.

3. Perspectives dans la dysphasie

L'ensemble de nos résultats chez les enfants présentant une dysphasie touchant majoritairement le versant expressif du langage révèlent pour la première fois à notre connaissance un déficit pour traiter la parole naturellement produite à un débit rapide (étude 2), sous-tendu par une dynamique corticale oscillatoire atypique qui est également observée lorsque la parole est produite à un débit normal (étude 3). Ce travail de recherche offre un certain nombre de perspectives, à court, moyen ou plus long terme, que nous détaillerons dans la suite.

3.1. Analyses et études complémentaires

Comme nous l'avons déjà indiqué, nos données actuelles mériteraient d'être complétées par des inclusions supplémentaires afin d'une part, de confirmer notre pattern de résultats et d'autre part, de mener des analyses supplémentaires. Certaines de ces analyses ont déjà été discutées dans les parties précédentes, il s'agit principalement d'examiner la dynamique oscillatoire de repos dans les bandes de fréquence d'intérêt de la parole (pour rappel, nous disposons d'enregistrements de repos de 3 minutes au début et à la fin de l'expérience de parole chez tous les enfants), ou encore du couplage cortico-cortical entre les régions temporales, temporo-pariétales et prémotrices de la voie dorsale mais aussi entre les bandes thêta et gamma (couplage phase-amplitude notamment) lors de la perception de parole produite dans les deux débits. L'une des principales limites de notre étude 3 en MEG est que nous ne disposons pas de mesures comportementales (intelligibilité des phrases et/ou performances de discrimination d'indices rythmiques tels que le *rise time* et la durée) que nous pourrions corrélérer avec les mesures de cohérence cortico-acoustique dans les différentes conditions de débit. Des données comportementales de capacités de perception de parole rapide ont été recueillies dans une autre étude (étude 2 comportementale, Guiraud et al., en révision) menée chez un autre groupe d'enfants dysphasiques, rendant les interprétations relativement délicates. Nos résultats suggèrent toutefois une corrélation, à confirmer là encore avec un nombre plus important de participants, entre l'alignement cerveau/parole dans le cortex temporal et les performances phonologiques des enfants dysphasiques évaluées par un test de langage standardisé d'une part, et leur propre débit de parole d'autre part (Guiraud et al., en préparation). Approfondir cette ligne de recherche dans un paradigme plus complet incluant des acquisitions en EEG ou MEG, des évaluations neuropsychologiques et

langagières, des mesures des capacités rythmiques en perception (*e.g.*, *rise time* et durée des stimuli verbaux et non verbaux) et en production (*e.g.*, tâches de *tapping* sur le rythme d'un métronome ou d'une comptine), ainsi que des capacités de perception de parole rapide constitue sans aucun doute l'un des défis majeurs à relever pour mieux comprendre les liens entre rythmes cérébraux et linguistiques dans la dysphasie.

Dans notre étude en MEG, nous nous sommes intéressés à la manière dont les oscillations corticales dans la bande thêta s'alignent sur les modulations lentes de l'enveloppe d'amplitude de la parole avec un intérêt tout particulier pour le débit syllabique. Ces modulations véhiculent l'information syllabique (au rythme de 4-7 Hz) mais également l'information suprasegmentale prosodique, correspondant à un rythme plus lent situé autour de 2-3 Hz. Les modèles de perception de la parole suggèrent que le flux continu de parole soit échantillonné par l'activité cérébrale oscillatoire dans différentes bandes de fréquence correspondant aux fenêtres d'intégration temporelle pertinentes pour le traitement linguistique, à savoir les bandes gamma (unités phonémiques), thêta (unités syllabiques) et delta (unités phrastiques ; Ghitza, 2011, 2017; Giraud & Poeppel, 2012; Poeppel, 2003). Des études chez les enfants dysphasiques ont par ailleurs révélé un déficit d'expression du rythme (tâches de *tapping*) à une fréquence de 1.5 et 2 Hz, c'est-à-dire dans la bande delta (Corriveau & Goswami, 2009; voir aussi Thomson, Fryer, & Maltby, 2006 pour des résultats concordants chez des adultes dyslexiques). Ceci s'accorde avec l'hypothèse selon laquelle les dysphasiques présenteraient des difficultés pour traiter les indices rythmiques prosodiques, conduisant à une extraction inefficace des syllabes et des mots et groupes de mots et ainsi aux déficits grammaticaux, morpho-syntaxiques fréquemment décrits dans ce trouble (Corriveau et al., 2007; Cumming, Wilson, & Goswami, 2015a). Il serait alors tout à fait intéressant d'examiner le couplage cortico-acoustique sur de la parole naturelle dans la bande delta chez les enfants dysphasiques afin de déterminer s'il présente un pattern atypique par rapport à une population d'enfants contrôles. Dans notre expérience, la tâche des enfants était d'écouter des phrases isolées dont la durée minimale était d'une seconde (condition de débit rapide). Comme nous l'avons précédemment indiqué, il n'est donc pas possible avec les données actuelles d'examiner précisément le couplage cerveau/parole dans la bande delta. Une étude incluant des extraits de parole de quelques minutes, produits à débit normal ou rapide, serait plus à même d'élucider cette question. Il faut d'ailleurs noter qu'avec l'accélération naturelle du débit, non seulement les « battements » syllabiques étaient plus fréquents mais la fréquence des syllabes accentuées (allongement de durée et augmentation d'intensité,

caractéristiques du français) marquant les groupes rythmiques était également plus élevée. En effet, des mesures préliminaires sur le signal acoustique (i.e. nombre de syllabes allongées par phrase rapporté à la durée totale de la phrase) ont mis en évidence une « fréquence prosodique » moyenne de 2 Hz pour les phrases à débit normal et de 3 Hz pour les phrases à débit rapide. Ceci était également visible dans le spectre de puissance montrant un pic à 2 Hz pour la parole normale et à 3 Hz pour la parole rapide. Examiner la dynamique corticale oscillatoire dans la bande delta permettrait de déterminer si un *shift* de bande de fréquence se produit lors de l’alignement cortico-acoustique dans ces différentes conditions de parole chez les enfants dysphasiques et neurotypiques.

Nous aimerions également souligner que les enfants inclus dans nos deux études (2 et 3) étaient âgés de 8 à 13 ans. Ces enfants n’ont pas tous été diagnostiqués dysphasiques au même âge et n’ont donc pas bénéficié d’un suivi orthophonique équivalent en termes de durée. Il est ainsi possible que ce facteur ait influencé nos résultats : on peut supposer qu’un enfant suivant une rééducation orthophonique depuis 4-5 ans, dont l’expression langagière et les habiletés phonologiques et morpho-syntaxiques se sont améliorées (bien qu’encore en-deçà de la moyenne d’une population contrôle), soit en effet plus apte à traiter la parole rapide et à voir ses oscillations corticales s’aligner un peu mieux sur le rythme syllabique de l’information verbale qu’un enfant venant à peine de commencer sa rééducation. Avec un groupe plus important d’enfants dysphasiques, il pourrait donc être pertinent d’examiner la dynamique oscillatoire sous-tendant la perception de la parole en fonction de la durée du suivi orthophonique, en créant des sous-groupes d’enfants ou en examinant les corrélations potentielles entre les différentes variables. De la même manière que ce que nous avons proposé pour les enfants neurotypiques (partie IX.3. de cette Discussion générale), il serait par ailleurs intéressant d’étudier la trajectoire développementale de la dynamique oscillatoire en comparant les résultats obtenus chez des enfants à des âges différents, à partir de 5-6 ans puisque c’est vers cet âge que le diagnostic est généralement posé (Gérard, 1993 ; chapitre V.1. de l’Introduction). Inclure un groupe d’enfants présentant un risque de développer une dysphasie (i.e. dont un parent et/ou frère/sœur est dysphasique) permettrait en outre de déterminer si l’alignement cerveau/parole est atypique chez ces enfants et peut potentiellement constituer un marqueur neuronal qui pourrait être ciblé par des entraînements neurocognitifs visant à améliorer les performances langagières, en particulier concernant le rythme de la parole.

3.2. Rôle du rythme dans la musique et le mouvement

Lors de la perception de la parole rapide, le décodage et le traitement du signal de parole ainsi que l'*entraînement* des oscillations corticales sur les modulations lentes de l'enveloppe temporelle caractérisant le débit de parole (normal également) semblent altérés, ou atypiques, chez les enfants dysphasiques. Ces résultats ouvrent la voie à des perspectives quant à la remédiation des troubles temporels et morpho-syntaxiques présents dans la dysphasie.

Ces dernières années, la recherche sur le langage et ses troubles a fait l'objet d'un engouement certain autour de la musique et de ses qualités en tant que système de remédiation possible. En effet, l'une des caractéristiques intrinsèques du signal musical, partagée avec le signal de parole, est sa rythmicité. La théorie de l'attention dynamique proposée par Jones et collaborateurs (Jones, 1976; voir aussi Jones & Boltz, 1989; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999) suppose que la perception de l'organisation temporelle d'une séquence auditive repose sur la variation cyclique de l'énergie attentionnelle: les oscillateurs internes se synchroniseraient sur les événements auditifs les plus saillants, permettant de générer des prédictions temporelles sur les événements futurs. Ainsi, la présentation d'un stimulus hautement rythmique va permettre d'entraîner les oscillations cérébrales et d'optimiser le traitement de ce stimulus. D'autre part, un nombre important de travaux suggèrent que le traitement de la musique et celui du langage soient régis par des structures cérébrales communes (Goswami, 2011; Koelsch, 2011; LaCroix, Diaz, & Rogalsky, 2015; Patel, 2007; Patel, Peretz, & Tramo, 1998; Peretz & Coltheart, 2003; Peretz, Vuvan, Lagrois, & Armony, 2015; Schön et al., 2010; Tillmann, 2012; voir Besson, Chobert, & Marie, 2011 pour une brève perspective historique). Certains auteurs proposent d'ailleurs que lors de l'acquisition de la parole, le langage serait traité comme un cas particulier de musique soutenant d'autant plus l'idée de liens anatomiques et fonctionnels étroits entre ces deux systèmes (Koelsch, 2011; Sallat & Jentschke, 2015). De manière intéressante, Halwani et collaborateurs (2011) ont par exemple observé un recrutement du faisceau arqué lors de l'apprentissage d'un instrument de musique ou du chant. Cette structure est connue pour être un élément crucial dans le traitement de la parole, puisqu'elle permet notamment de relier les régions temporales supérieures auditives aux régions frontales motrices au sein de l'hémisphère gauche, correspondant à la voie dorsale d'intégration sensori-motrice (Saur et al., 2008). Cette voie constitue un support clé pour l'acquisition et le développement du langage oral (Hickok & Poeppel, 2007), et son dysfonctionnement peut

potentiellement conduire à des troubles développementaux du langage. Dans cette optique, entraîner les interactions sensori-motrices et le fonctionnement de la voie dorsale à travers la musique semble constituer une option prometteuse pour améliorer les habiletés langagières. « *If there are common underlying processes between music and language, especially between music perception and speech perception, one might assume that improving some of the processes involved in the perception of music can also improve speech perception and reading skills* » (Habib et al., 2016).

Un grand nombre d'études ont révélé les effets positifs d'un entraînement musical sur les capacités langagières des individus, que ce soit au niveau de la segmentation du signal de parole (e.g., François, Chobert, Besson, & Schön, 2013), du traitement phonologique (Cason & Schön, 2012; Degé & Schwarzer, 2011; Moreno et al., 2009; Moritz, Yampolsky, Papadelis, Thomson, & Wolf, 2013) ou encore syntaxique (e.g., Bedoin, Brisseau, Molinier, Roch, & Tillmann, 2016; Jentschke & Koelsch, 2009; Przybylski et al., 2013). Ainsi, l'écoute ou la pratique préalable de musique permettrait de renforcer les capacités de traitement linguistique des individus, qu'ils présentent un trouble du développement du langage écrit ou oral (dyslexie ou dysphasie respectivement) ou non. « *This effect of music training suggests that, akin to physical exercise and its impact on body fitness, music is a resource that tones the brain for auditory fitness* » (Kraus & Chandrasekaran, 2010).

Tout d'abord, l'apprentissage et la pratique de la musique, demandant une coordination précise entre les systèmes sensoriels (visuel, auditif et somatosensoriel) et moteur, sont connus pour modeler les structures corticales et sous-corticales impliquées dans les processus musicaux, augmentant ainsi la plasticité du cerveau et conduisant à des changements fonctionnels (Kraus & Chandrasekaran, 2010; Münte, Altenmüller, & Jäncke, 2002; Pantev et al., 2003; Schlaug, Norton, Overy, & Winner, 2005; voir Besson, Chobert, & Marie, 2011 pour une revue). L'entraînement musical permettrait notamment de renforcer les mécanismes d'ajustement des oscillations corticales sur le signal auditif rythmique dans le but d'augmenter les prédictions temporelles sur un stimulus. Ces modifications structurelles et fonctionnelles de nature sensori-motrice pourraient alors se généraliser à d'autres domaines d'apprentissage tels que le langage (Bangert & Altenmüller, 2003; Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Habib & Besson, 2009; Moreno et al., 2009; Zatorre et al., 2002). Des études ont ainsi montré que l'expertise musicale (après un entraînement intensif sur plusieurs semaines/mois ou chez les musiciens professionnels) améliorerait le traitement auditif

et certaines habiletés cognitives, tels que l’empan de chiffres (Fujioka et al., 2006), le traitement de la hauteur dans la musique et la parole, la lecture de mots complexes ou encore le traitement de la structure métrique de la parole (*e.g.*, Besson, Schön, Moreno, Santos, & Magne, 2007; Magne, Schön, & Besson, 2006; Marie, Magne, & Besson, 2011; Moreno et al., 2009; Moreno & Besson, 2006; voir aussi Lee & Hung, 2008; Marques, Moreno, Luís Castro, & Besson, 2007; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011 pour des résultats en langue étrangère). Ainsi, la maîtrise de la musique permettrait de modifier et d’améliorer les processus impliqués dans le traitement de la parole (Besson, Chobert, & Marie, 2011; Moreno, 2009 pour des revues). Besson et collaborateurs (2011) ont notamment proposé que la sensibilité accrue des musiciens (professionnels ou ayant suivi un entraînement) aux indices acoustiques communs à la musique et à la parole, qui repose sur des processus généraux (non spécifiques), leur permettrait d’élaborer des représentations plus sophistiquées des signaux de parole comparés à des non-musiciens. Ceci optimiserait alors les traitements spécifiques à la parole (*i.e.* non partagés avec la musique), suggérant que l’effet bénéfique de la musique sur les habiletés langagières ne résulte pas simplement du recrutement de substrats neuronaux communs mais bien d’un transfert d’apprentissage.

La musique, comme le langage, étant un processus hautement sensori-moteur, il n’est pas étonnant d’observer des troubles de la perception et de la production de rythmes (musicaux) chez les enfants présentant des troubles développementaux du langage (dyslexie ou dysphasie ; Clément, Planchou, Béland, Motte, & Samson, 2015; Cumming, Wilson, Leong, Colling, & Goswami, 2015b; Huss, Verney, Fosker, Mead, & Goswami, 2011; Overy, Nicolson, Fawcett, & Clarke, 2003; voir aussi Jentschke, Koelsch, Sallat, & Friederici, 2008 pour un déficit du traitement syntaxique dans la musique). Huss et collaborateurs (2011) ont par exemple rapporté que les enfants dyslexiques présentaient de moins bonnes performances que des enfants normo-lecteurs pour percevoir la structure métrique de séquences musicales, et ces performances étaient corrélées à leurs capacités de discrimination du *rise time* et de conscience phonologique ainsi qu’à leurs compétences en lecture. Une corrélation entre leurs performances musicales et leurs scores en conscience phonologique a en outre été observée. Clément et collègues (2015) ont quant à eux démontré que des enfants dysphasiques francophones présentaient plus de difficultés que des enfants neurotypiques à reproduire des notes de musique isolées ou des mélodies, que celles-ci leur soient familières (*e.g.*, « Frère Jacques ») ou non. Une corrélation positive entre leurs capacités de reproduction mélodique et

leurs performances musicales générales, testées avec la MBEMA (*Montreal Battery of Evaluation of Musical Abilities*, Peretz et al., 2013), a par ailleurs été mise en évidence (voir aussi les résultats de Planchou et al., en révision et Planchou, 2014). Dans l'ensemble, ces résultats suggèrent donc un dysfonctionnement possible de la voie dorsale auditivo-motrice chez les enfants dysphasiques (en accord avec nos données en MEG dans l'étude 3), qui pourrait rendre compte à la fois de leurs difficultés pour traiter à la fois la musique et la parole.

En dépit des déficits de traitement du rythme dans les troubles développementaux du langage, des études révèlent un bénéfice de la pratique musicale sur les capacités de traitement linguistique chez des enfants dyslexiques ou dysphasiques. Ainsi, chez les enfants ou adultes dyslexiques, un entraînement musical intensif pendant 3 jours ou sur plusieurs mois (à raison de quelques heures par semaine) améliore la perception de la durée des syllabes incluses dans des mots, la perception catégorielle des phonèmes, la conscience phonologique et le niveau de lecture des enfants (Flaugnacco et al., 2015; Habib et al., 2013, 2016; voir aussi Slater, Tierney, & Kraus, 2013 chez des enfants à risque de développer une dyslexie). Weiss et collègues (2014) ont par ailleurs montré que des adultes dyslexiques musiciens présentaient de meilleures performances que leurs pairs non musiciens dans des tâches de discrimination de hauteur (*pitch*), de *tapping* ou encore de perception de parole bruitée. De manière intéressante, leurs performances dans ces tâches ne différaient pas de celles de musiciens non dyslexiques. Bishop-Liebler et collaborateurs (2014) ont enfin rapporté de meilleures performances de traitement temporel chez des adultes dyslexiques musiciens que non-musiciens, en particulier pour détecter le *rise time* dans l'enveloppe d'amplitude.

D'autres travaux montrent que la perception d'une séquence musicale amorcée améliore les capacités de traitement morpho-syntaxique des enfants présentant une dysphasie ou une dyslexie (voir aussi Cason & Schön, 2012 pour des résultats chez des adultes sains). L'hypothèse sous-jacente se base sur la théorie de l'attention dynamique (Large & Jones, 1999) et suppose que la présentation préalable d'un rythme musical modulerait l'attention temporelle de l'auditeur (i.e. entraînerait les oscillations corticales) de manière prolongée de sorte à augmenter l'efficacité des traitements rythmiques (verbaux) suivants. Przybylski et collègues (2013) ont ainsi rapporté que des enfants dyslexiques ou dysphasiques francophones avaient de meilleures performances de détection d'une anomalie syntaxique dans des phrases lorsque celles-ci étaient précédées d'amorces musicales simples

(percussions) possédant une métrique régulière (i.e. facile à extraire) par rapport à des séquences dont la métrique était irrégulière (i.e. plus difficile à extraire; voir aussi l'étude *follow-up* de Bedoin, Brisseau, Molinier, Roch, & Tillmann, 2016, et Cason, Hidalgo, Isoard, Roman, & Schön, 2015 pour des résultats en production chez des enfants sourds appareillés). Dans la même lignée, Falk et Dalla Bella (2016) ont montré, chez des adultes sains, que taper du doigt en rythme avec les syllabes accentuées d'une phrase améliorerait la détection d'un changement lexical au sein de cette phrase ; ceci n'était pas observé lorsque le mouvement n'était pas aligné avec les syllabes ou lorsque les participants ne réalisaient aucun mouvement (voir aussi Falk et al., 2017). La réalisation d'un mouvement rythmique produirait ainsi une modulation *top-down* en provenance des régions motrices sur le traitement de l'information auditive verbale, favorisant l'alignement sur les événements saillants dans le signal (voir chapitre III.3 de l'Introduction; Morillon & Baillet, 2017; Morillon & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014). Nous n'avons pas connaissance d'études ayant examiné l'influence de la réalisation préalable de mouvements manuels en rythme sur un signal auditif (e.g., métronome ou percussions) sur les capacités de traitement de la parole dans les troubles développementaux du langage.

L'ensemble des résultats de ces différentes études suggère donc que la structure temporelle rythmique de la musique, et le fait qu'elle requière une intégration sensori-motrice pour être traitée, exercent une influence positive sur le traitement du langage, dans la sphère orale notamment, en renforçant l'efficacité des réseaux fonctionnels communs aux traitements musical et verbal, mais aussi probablement en amplifiant l'alignement cortical oscillatoire sur les (quasi-)régularités temporelles de la parole. Les données de la littérature soulignent ainsi le caractère prometteur de l'amorçage rythmique musical, et d'entraînements musicaux favorisant plus particulièrement le traitement du rythme, dans l'amélioration des traitements linguistiques chez les enfants présentant des troubles langagiers. Nous pouvons alors supposer qu'une amorce rythmique dont le tempo et la métrique régulière seraient similaires au rythme syllabique de stimuli phrastiques naturels permettrait de renforcer la synchronisation du système oscillatoire interne sur le rythme d'intérêt et donc de faciliter le traitement de la parole, notamment produite à débit rapide, chez les enfants dysphasiques mais aussi neurotypiques. Dans ce cadre, il serait intéressant de comparer l'influence d'amorces auditives musicales et de mouvements rythmiques des doigts/mains (tâche de *tapping*). L'hypothèse est que l'entraînement des oscillations cérébrales au sein des régions auditives et

(pré)motrices de la voie dorsale sur la métrique régulière des séquences auditives ou motrices orienterait l'attention sur les régularités temporelles du signal et permettrait ainsi de mieux segmenter le signal de parole présenté ultérieurement (par renforcement de la synchronisation cortico-acoustique) et d'en améliorer le traitement et l'intelligibilité. Comme souligné par Cumming et collaborateurs (2015b), la nature des troubles observés chez les enfants dysphasiques devrait néanmoins être prise en compte. Alors que des entraînements basés sur la synchronisation motrice sur un rythme simple pourraient être efficaces chez les enfants dysphasiques présentant des troubles phonologiques uniquement, mettre en place des entraînements utilisant des séquences musicales ayant la même structure prosodique que la parole pourrait s'avérer plus judicieux chez les enfants présentant des troubles morpho-syntaxiques liés à un traitement prosodique moins efficace.

XI. Conclusion

Ce travail de thèse avait pour objectifs principaux d'explorer les processus neurocognitifs de traitement du rythme syllabique de la parole naturelle chez les enfants présentant un développement langagier typique et atypique âgés entre 8 et 13 ans. Trois études, l'une comportementale et les deux autres en MEG, nous ont permis de mettre en lumière, pour la première fois à notre connaissance, (1) la dynamique des oscillations corticales lors du traitement de la parole naturelle produite à débit normal ou rapide chez les enfants neurotypiques, (2) un alignement atypique des rythmes cérébraux sur le rythme syllabique de parole chez des enfants présentant une dysphasie expressive et (3) un déficit de perception de la parole rapide chez ces enfants dysphasiques.

La première étude en MEG chez les enfants neurotypiques a révélé qu'alors que les oscillations au sein du cortex auditif droit se synchronisaient sur les modulations lentes de l'enveloppe d'amplitude de la parole à débit normal dans la bande thêta classique (4-7 Hz), un entrainement spécifique sur le débit syllabique moyen de nos stimuli phrastiques (5.6-7.6 Hz) a été observé au sein des régions temporales antérieures gauches. Ces résultats sont globalement en accord avec les modèles et études précédentes menées chez l'adulte, suggérant que le cortex auditif droit échantillonnerait le flux continu de parole en unités syllabiques. Le couplage cortico-acoustique dans l'hémisphère gauche pourrait quant à lui refléter un processus d'intégration de plus haut niveau, notamment dans le couplage hiérarchique entre les traitements phonémique et syllabique. Par ailleurs, nos résultats dans la condition de parole à débit rapide mettent en évidence un alignement de l'activité oscillatoire des régions prémotrices et motrices gauches sur le signal de parole, révélant le rôle majeur de la voie dorsale sensori-motrice dans le cas où la parole est plus difficile à décoder ainsi que chez l'enfant dont le système langagier est encore en développement.

Les deux études suivantes réalisées au cours de ce travail de thèse ont permis d'examiner les capacités de traitement de la parole à débit normal et rapide chez des enfants dysphasiques ainsi que la dynamique corticale oscillatoire sous-jacente. L'étude 2, comportementale, fournit des arguments supplémentaires en faveur des théories suggérant un trouble du traitement de l'information temporelle, en particulier des indices rythmiques, de la parole chez les enfants dysphasiques, puisqu'un déficit de perception de la parole rapide, accélérée naturellement ou artificiellement a été mis en évidence chez ces enfants en regard d'enfants neurotypiques. Ces résultats d'un trouble de traitement de l'information syllabique

ont été corroborés par la troisième étude, réalisée en MEG, qui a révélé un fonctionnement oscillatoire atypique chez les enfants dysphasiques, comparés à leurs pairs neurotypiques, caractérisé par une faible synchronisation thêta au sein des régions auditives droites lors de l'écoute de parole à débit normal et un recrutement de la voie dorsale gauche absent, voire latéralisé à droite, lors du traitement de la parole rapide.

En conclusion, notre travail, qui est le premier à notre connaissance à examiner le cas de la parole naturelle à débits variables chez les enfants, apporte donc des arguments expérimentaux convaincants quant à un alignement des oscillations cérébrales sur le rythme syllabique de la parole chez les enfants neurotypiques, globalement comparable à celui observé chez l'adulte. Il fournit également les premières preuves d'un déficit de traitement du rythme syllabique chez les enfants présentant une dysphasie, qui serait sous-tendu par un couplage cortico-acoustique atypique. De tels résultats, s'ils sont confirmés par des études complémentaires, pourraient permettre d'identifier des marqueurs neuronaux potentiels de la dysphasie et d'ouvrir la voie à des entraînements neurocognitifs basés sur le rythme visant à améliorer les capacités d'extraction des indices rythmiques de la parole.

BIBLIOGRAPHIE

- Abercombie, D. (1967). *Elements of general phonetics*. Aldine Pub.
- Abrams, D., Nicol, T., Zecker, S., & Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(15), 3958–65.
- Abrams, D., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(24), 7686–93.
- Ackermann, H., & Riecker, A. (2010). The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Structure and Function*, 214, 419–433.
- Adank, P., & Devlin, J. T. (2010). On-line plasticity in spoken sentence comprehension: Adapting to time-compressed speech. *NeuroImage*, 49(1), 1124–1132.
- Adank, P., & Janse, E. (2009). Perceptual learning of time-compressed and natural fast speech. *The Journal of the Acoustical Society of America*, 126(5), 2649–2659.
- Ahissar, E., Nagarajan, S. S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 13367–13372.
- Ajuriaguerra, J., Borel-Maisonny, S., Diatkine, R., Narlian, S., & Stambak, M. (1958). Le groupe des audimutités. *La Psychiatrie de l'Enfant*, 1, 7–51.
- Ajuriaguerra, J., Guignard, F., Jaeggi, A., Kocher, F., Maquard, M., Paunier, A., ... Siotis, E. (1962). Organisation psychologique et troubles du développement du langage. In *Problemes de psycho-linguistique* (p. 109–142). Paris, France: PUF.
- Alcock, K. J., Passingham, R. E., Watkins, K. E., & Vargha-Khadem, F. (2000). Pitch and Timing Abilities in Inherited Speech and Language Impairment. *Brain and Language*, 75, 34–46.
- Alho, J., Lin, F.-H., Sato, M., Tiitinen, H., Sams, M., & Jääskeläinen, I. P. (2014). Enhanced neural synchrony between left auditory and premotor cortex is associated with successful phonetic categorization. *Frontiers in Psychology*, 5(394).
- Alho, J., Sato, M., Sams, M., Schwartz, J.-L., Tiitinen, H., & Jääskeläinen, I. P. (2012). Enhanced early-latency electromagnetic activity in the left premotor cortex is associated with successful phonetic categorization. *NeuroImage*, 60(4), 1937–46.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. 5th ed. Arlington, VA: American Psychiatric Association.
- Anderson, J. M., Gilmore, R., Roper, S., Crosson, B., Bauer, R. M., Nadeau, S., ... Heilman, K. M. (1999). Conduction Aphasia and the Arcuate Fasciculus: A Reexamination of the Wernicke–Geschwind Model. *Brain and Language*, 70(1), 1–12.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*, 28, 105–167.
- Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.

- Arnal, L. H., & Giraud, A.-L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, 16(7), 390–398.
- Arnal, L. H., & Giraud, A.-L. (2016). Neurophysiologie de la perception de la parole et multisensorialité. In S. Pinto and M. Sato(Eds.), *Traité de neurolinguistique: Du cerveau au langage*(p. 97–107). Louvain-la-Neuve, Belgique: De Boeck Supérieur.
- Arsenault, J. S., & Buchsbaum, B. R. (2016). No evidence of somatotopic place of articulation feature mapping in motor cortex during passive speech perception. *Psychonomic Bulletin & Review*, 23(4), 1231–1240.
- Astesano, C. (2001). *Rythme et Accentuation en Français : Invariance et Variabilité Stylistique* Paris, France: L'Harmatta.
- Badcock, N. A., Bishop, D. V. M., Hardiman, M., Barry, J. G., & Watkins, K. E. (2012). Co-localisation of abnormal brain structure and function in specific language impairment. *Brain and Language*, 120(3), 310–320.
- Bailly, G. (1997). Learning to speak. Sensori-motor control of speech movements. *Speech Communication*, 22, 251–267.
- Baker, L., & Cantwell, D. P. (1991). *Psychiatric and developmental disorders in children with communication disorder*. Washington, DC : American Psychiatric Association.
- Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience*, 26(4), 1–14.
- Bastiaansen, M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research*, 159(6), 179–96.
- Beattie, R. L., & Manis, F. R. (2012). Rise time perception in children with reading and combined reading and language difficulties. *Journal of Learning Disabilities*, 46(3), 200–209.
- Becker, R., Pefkou, M., Michel, C. M., & Hervais-Adelman, A. G. (2013). Left temporal alpha-band activity reflects single word intelligibility. *Frontiers in Systems Neuroscience*, 7, 1–12.
- Bedoin, N., Brisseau, L., Molinier, P., Roch, D., & Tillmann, B. (2016). Temporally regular musical primes facilitate subsequent syntax processing in children with Specific Language Impairment. *Frontiers in Neuroscience*, 10(245).
- Belton, E., Salmond, C. H., Watkins, K. E., Vargha-Khadem, F., & Gadian, D. G. (2003). Bilateral brain abnormalities associated with dominantly inherited verbal and orofacial dyspraxia. *Human Brain Mapping*, 18(3), 194–200.
- Ben-shachar, M., Hendler, T., Kahn, I., Ben-bashat, D., & Grodzinsky, Y. (2013). Research Article THE NEURAL REALITY OF SYNTACTIC TRANSFORMATIONS : Evidence From Functional Magnetic Resonance Imaging, 14(5), 433–440.
- Benasich, A. A., Gou, Z., Choudhury, N., & Harris, K. D. (2008). Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. *Behavioural Brain Research*, 195(2), 215–22.
- Benoît, C., Grice, M., & Hazan, V. (1996). The SUS test 1: A method for the assessment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences. *Speech Communication*, 18(4), 381–392.
- Benton, A. (1964). Developmental aphasia and brain damage. *Cortex*, 1, 40–52.
- Berger, H. (1929). Uber das Elektrenkephalogramm des Menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1), 527–570.

- Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L. J., & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, *117*(1), 21–33.
- Bertoncini, J., & Mehler, J. (1981). Syllables as Units in Infant Speech Perception. *Infant Behavior and Development*, *4*, 247–260.
- Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*(94), 1–12.
- 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.
- Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, *29*(4), 711–721.
- Billières, M. (2015). La perception de la parole. Retrieved from <https://www.verbotonale-phonetique.com/perception-parole/>
- Bishop-Liebler, P., Welch, G., Huss, M., Thomson, J. M., & Goswami, U. (2014). Auditory temporal processing skills in musicians with dyslexia. *Dyslexia*, *20*(3), 261–279.
- Bishop, D. V. M. (1992). The underlying nature of specific language impairment. *Journal of Child Psychology and Psychiatry*, *33*, 3–66.
- Bishop, D. V. M. (2002). The role of genes in the etiology of specific language impairment. *Journal of Communication Disorders*, *35*, 311–328.
- Bishop, D. V. M. (2006). What Causes Specific Language Impairment in Children? *Current Directions in Psychological Science*, *15*(5), 217–221.
- Bishop, D. V. M. (2013). Cerebral asymmetry and language development: cause, correlate, or consequence? *Science*, *340*.
- Bishop, D. V. M. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language and Communication Disorders*, *49*(4), 381–415.
- Bishop, D. V. M. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders*.
- Bishop, D. V. M., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: evidences from a twin study. *Developmental Medicine and Child Neurology*, *37*, 56–71.
- Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience*, *8*(3), 389–395.
- Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage*, *32*, 821–841.
- Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical Rhythm Implicitly Orients Attention in Time as Indexed by Improved Target Detection and Left Inferior Parietal Activation. *Journal of cognitive neuroscience*, *26*(3), 593–605.
- Bonfond, M., Kastner, S., & Jensen, O. (2017). Communication between Brain Areas Based on Nested Oscillations. *Eneuro*, *4*(2), 1–14.
- Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, *65*(1), 33–69.
- Botting, N., & Conti-Ramsden, G. (2004). Characteristics of children with specific language

- impairment. In L. Erhoeven & H. van Balkom (Eds.), *Classification of development language disorders: Theoretical issues and clinical implications* (p. 23–38). Nahwah, NJ: Lawrence Edition.
- Bowyer, S. M. (2016). Coherence a measure of the brain networks: past and present. *Neuropsychiatric Electrophysiology*, 2(1), 1.
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde*. Leipzig, Germany: Johann Amb.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49, 155–180.
- Brown, S., Ingham, R. J., Ingham, J. C., Laird, A. R., & Fox, P. T. (2005). Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, 25, 105–117.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *The European Journal of Neuroscience*, 13(2), 400–404.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., ... Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: an FMRI study. *Journal of Cognitive Neuroscience*, 16(1), 114–126.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N. F., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language*, 119(3), 119–128.
- Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, 25(5), 663–678.
- Buzsáki, G. (2006). *Rhythms of the Brain*. *Rhythms of the Brain*. New York, NY: Oxford University.
- Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, 304, 1926–1929.
- Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron*, 80, 751–764.
- Callan, D., Jones, J. A., Callan, A., & Akahane-Yamada, R. (2004). Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage*, 22, 1182–1194.
- Campbell, I. G., & Feinberg, I. (2009). Longitudinal trajectories of non-rapid eye movement delta and theta EEG as indicators of adolescent brain maturation. *Proceedings of the National Academy of Sciences*, 106(13), 5177–5180.
- Campione, E., & Véronis, J. (2004). Pauses et hésitations en français spontané. In *XXVe Journées d'Etudes sur la Parole*. Fès, Maroc.
- Canault, M. (2007). *L'émergence du contrôle articulatoire au stade du babillage. Une étude acoustique et cinématique*.
- Canault, M., Laboissière, R., Perrier, P., & Sock, R. (2008). Development of lingual displacement independence at babbling stage. In *8th International Seminar on Speech Production*. Strasbourg, France.

- Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, *14*(11), 506–515.
- Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, *29*, 102–107.
- Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652–2658.
- Chandrasekaran, C., Trubanova, A., Stillitano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*(7).
- Chemin, B., Mouraux, A., & Nozaradan, S. (2014). Body Movement Selectively Shapes the Neural Representation of Musical Rhythms. *Psychological Science*, *25*(12), 2147–2159.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239.
- Chevillet, M. A., Jiang, X., Rauschecker, J. P., & Riesenhuber, M. (2013). Automatic phoneme category selectivity in the dorsal auditory stream. *The Journal of Neuroscience*, *33*(12), 5208–5215.
- Chiron, C., Pinton, F., Masure, M. C., Duvelleroy-Hommet, C., Leon, F., & Billard, C. (1999). Hemispheric specialization using SPECT and stimulation tasks in children with dysphasia and dystrophia. *Developmental Medicine and Child Neurology*, *41*(8), 512–520.
- Christiansen, T. U., Dau, T., & Greenberg, S. (2007). Spectro-temporal processing of speech – An information-theoretic framework. *Hearing-From Sensory Processing to Perception*, 517–523.
- Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of Acoustic Society of America*, *95*(3), 1570–1580.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. (2003). Discovering words in the continuous speech stream: The role of prosody. *Journal of Phonetics*, *31*, 585–598.
- Christophe, A., Pallier, C., Bertoncini, J., & Mehler, J. (1991). A la recherche d'une unité: segmentation et traitement de la parole. *L'année Psychologique*, *91*, 59–86.
- Clark, A., & Carter, G. (2014). What should we call children with unexplained language difficulties? A practical perspective. Commentary on Bishop, D. V. M., Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders*, *49*, 381–415.
- Clément, S., Planchou, C., Béland, R., Motte, J., & Samson, S. (2015). Singing abilities in children with Specific Language Impairment (SLI). *Frontiers in Psychology*, *6*(April), 1–11.
- Cogan, G. B., & Poeppel, D. (2011). A mutual information analysis of neural coding of speech by low-frequency MEG phase information. *J Neurophysiol*, *106*, 554–563.
- Cohen, M., Campbell, R., & Yaghai, F. (1989). Neuropathological abnormalities in developmental dysphasia. *Annals of Neurology*, *25*, 567–570.
- Conti-ramsdén, G. (2014). What should we call children who struggle to talk? Taking a developmental, global perspective of diagnostic labels reflections on Bishop (2014). Commentary on Bishop, D. V. M., Ten questions about terminology for children with unexplained language problems. *International Journal of Language &*

- Communication Disorders*, 49, 381–415.
- Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex*, 45, 119–130.
- Corriveau, K. H., Pasquini, E. S., & 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(3), 647–666.
- Crosse, M. J., Butler, J. S., & Lalor, E. C. (2015). Congruent Visual Speech Enhances Cortical Entrainment to Continuous Auditory Speech in Noise-Free Conditions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 35(42), 14195–204.
- Cumming, R., Wilson, A., & Goswami, U. (2015a). 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).
- Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015b). Awareness of rhythm patterns in speech and music in children with specific language impairments. *Frontiers in Human Neuroscience*, 9(672).
- Cutini, S., Szücs, 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.
- Cutler, A. (1992). Rhythmic Cues to Speech Segmentation: Evidence from Juncture Misperception. *Journal of Memory and Language*, 31, 218–236.
- Cutler, A., Butterfield, S., & Williams, J. N. (1987). The perceptual integrity of syllabic onsets. *Journal of Memory and Language*, 26(4), 406–418.
- Cutler, A., Norris, D., & Segui, J. (1986). The Syllable 's Differing Role in the Segmentation of French and English, 400, 385–400.
- D'Ausilio, A., Bufalari, I., Salmas, P., & Fadiga, L. (2012). The role of the motor system in discriminating normal and degraded speech sounds. *Cortex*, 48, 882–887.
- D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology*, 19, 381–385.
- D'Imperio, M., Dittinger, E., & Besson, M. (2016). Prosodie et intonation : notions de base et données neuro-psycholinguistiques. In S. Pinto & M. Sato (Eds.), *Traité de neurolinguistique* (p. 133–145). Louvain-la-Neuve, Belgique: De Boeck Supérieur.
- Dalla Bella, S., Białuńska, A., & Sowiński, J. (2013). Why Movement Is Captured by Music, but Less by Speech: Role of Temporal Regularity. *Plos One*, 8(8), 1–16.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51–62.
- Davis, B. L., & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, 38(6), 1199–1211.
- Davis, M. H., Ford, M. A., Kherif, F., & Johnsrude, I. S. (2011). Does semantic context benefit speech understanding through “top-down” processes? Evidence from time-resolved sparse fMRI. *Journal of Cognitive Neuroscience*, 23(12), 3914–32.
- Davis, M. H., & Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 23(8), 3423–3431.
- de Guibert, C., Maumet, C., Jannin, P., Ferré, J.-C., Tréguier, C., Barillot, C., ... Biraben, A.

- (2011). Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia). *Brain*, *134*(10), 3044–58.
- de Weck, G., & Marro, P. (2010). *Les troubles du langage chez l'enfant: Description et évaluation*. Issy-les-Moulineaux, France: Elsevier Masson.
- Degé, F., & Schwarzer, G. (2011). The effect of a music program on phonological awareness in preschoolers. *Frontiers in Psychology*, *2*(124), 1–7.
- Dehaene, S., & Petit, C. (2009). *Parole et musique: Aux origines du dialogue humain. Travaux du collège de France*. Paris, France: Odile Jacob.
- Dellwo, V., & Wagner, P. (2003). Relations between language rhythm and speech rate. In *Proceedings of the International Congress of Phonetics Science* (pp. 471–474). Barcelona, Spain.
- Demonet, J.-F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: Functional neuroimaging. *Physiological Reviews*, *85*, 49–95.
- Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders*, *43*(4), 361–389.
- Dewey, D. (1993). Error analysis of limb and orofacial praxis in children with developmental motor deficits. *Brain and Cognition*, *23*, 203–221.
- Di Cristo, A. (2013). *La prosodie de la parole*. Louvain-la-Neuve, Belgique: de Boeck supérieur.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Brain Research*, *91*, 176–180.
- Dibbets, P., Bakker, K., & Jolles, J. (2006). Functional MRI of Task Switching in Children with Specific Language Impairment (SLI). *Neurocase*, *12*(1), 71–79.
- Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology*, *55*, 149–79.
- Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, *59*(3), 294–311.
- Dimitrijevic, A., Smith, M. L., Kadis, D. S., & Moore, D. R. (2017). Cortical Alpha Oscillations Predict Speech Intelligibility. *Frontiers in Human Neuroscience*, *11*(88).
- Ding, N., & Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proceedings of the National Academy of Sciences*, *109*(29), 11854–11859.
- Ding, N., & Simon, J. Z. (2013). Adaptive Temporal Encoding Leads to a Background-Insensitive Cortical Representation of Speech. *Journal of Neuroscience*, *33*(13), 5728–5735.
- Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*, 311.
- Dodd, B. (1979). Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, *11*(4), 478–484.
- Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–8.
- Dolata, J. K., Davis, B. L., & MacNeilage, P. F. (2008). Characteristics of the rhythmic organization of vocal babbling: Implications for an amodal linguistic rhythm. *Infant*

- Behavior and Development*, 31, 422–431.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22, 567–631.
- Drake, C., Jones, M. R., & Baruch, C. (2000). *The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending*. *Cognition*, 77, 251–288.
- Dronkers, N. F. (2000). The Pursuit of Brain–Language Relationships. *Brain and Language*, 71(1), 59–61.
- Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *The Journal of the Acoustical Society of America*, 97, 585–592.
- Drullman, R., Festen, J. M., & Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. *The Journal of the Acoustical Society of America*, 95(5), 2670–2680.
- Drullman, R., Festen, J. M., & Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, 95(2), 1053.
- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19), 7126–31.
- Duez, D. (1995). On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced stops. *Journal of Phonetics*, 23, 407–427.
- Duez, D. (1998). Consonant sequences in spontaneous french speech. In *Sound patterns of Spontaneous Speech, ESCA Workshop* (p. 63–68). L beaume-les-Aix, France.
- Duez, D. (1999). Effects of Articulation rate on duration in read French speech. In *Eurospeech* (p. 715–718). Budapest, Hungary.
- Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America*, 121(4), 131–136.
- Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: effects of talker and rate changes. *Journal of Experimental Psychology. Human Perception and Performance*, 23(3), 914–927.
- Eimas, P. D. (1999). Segmental and syllabic representations in the perception of speech by infants. *Journal of Acoustic Society of America*, 105, 1901–1911.
- Eimas, P. D., Siqueland, E. P., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Ellis-Weismer, S., Evans, J., & Hesketh, L. J. (1999). An Examination of Verbal Working Memory Capacity in Children With. *Journal of Speech, Language, and Hearing Research*, 42(5), 1249–1260.
- Ellis-Weismer, S., Plante, E., Jones, M., & Tomblin, B. (2005). A Functional Magnetic Resonance Imaging Investigation of Verbal Working Memory in Adolescents With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, 48(2), 405.
- Engel, A. K., Fries, P., & Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top–down processing. *Nature Reviews Neuroscience*, 2, 704–716.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *The European Journal of Neuroscience*, 15(2), 399–402.

- Falk, S., & Dalla Bella, S. (2016). It is better when expected: aligning speech and motor rhythms enhances verbal processing. *Language, Cognition and Neuroscience*, 3798, 1–10.
- Falk, S., Volpi-Moncorger, C., & Dalla Bella, S. (2017). Auditory-motor rhythms and speech processing in French and German listeners. *Frontiers in Psychology*, 8(395).
- Feinberg, I., & Campbell, I. G. (2010). Sleep EEG changes during adolescence: An index of a fundamental brain reorganization. *Brain and Cognition*, 72, 56–65.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8(2), 181–195.
- Ferrand, L., Segui, J., & Grainger, J. (1996). Masked priming of word and picture naming: The role of syllabic units. *Journal of Memory and Language*, 35, 708–723.
- Fisher, S. E., Lai, C. S. L., & Monaco, A. P. (2003). Deciphering the Genetic Basis of Speech and Language Disorders. *Annual Review of Neuroscience*, 26(1), 57–80.
- 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), 1–17.
- Fogassi, L., Ferrari, P., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal Lobe: From Action Organization to Intention Understanding. *Science*, 308, 662–667.
- Fougeron, R., & Jun, S. (1998). Rate effects on French intonation : prosodic organization and phonetic realization. *Journal of Phonetics*, 26, 45–69.
- Fraisse, P. (1956). Les structures rythmiques. In *Studia Psychologica*. Paris, France: Presse Universitaire de France.
- Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception & Psychophysics*, 62(8), 1668–1680.
- François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, 23(9), 2038–2043.
- Friederici, A. D. (2006). The neural basis of language development and its impairment. *Neuron*, 52(6), 941–52.
- Friederici, A. D. (2011a). The Brain Basis of Language Processing: From Structure to Function. *Physiological Reviews*, 91, 1357–1392.
- Friederici, A. D. (2011b). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4), 1357–92.
- Friederici, A. D. (2017). Evolution of the neural language network. *Psychonomic Bulletin & Review*, 24(1), 41–47.
- Friederici, A. D., Bahlmann, J. R., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences*, 103(7), 2458–2463.
- Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474–480.
- Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88(1), 220–235.
- Fujioka, T., & Ross, B. (2008). Auditory processing indexed by stimulus-induced alpha desynchronization in children. *International Journal of Psychophysiology*, 68(2), 130–140.

- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain*, *129*(10), 2593–2608.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92.
- Galantucci, B., Fowler, C., & Turvey, M. T. (2006). “The motor theory of speech perception reviewed”: Erratum. *Psychonomic Bulletin & Review*, *13*(4), 742.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex.pdf. *Brain*, *119*, 593–609.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (2002). Action representation and the inferior parietal lobule. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action: Attention and Performance, Vol. XIX*. (p. 247–266). Oxford, UK: Oxford University Press.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11.
- Gasser, T., Verleger, R., Bacher, P., & Sroka, L. (1988). Development of the EEG of school-age children and adolescents. I. Analysis of band power. *Electroencephalogr. Clin. Neurophysiol.* *69*, 91–99. *Electroencephalogr Clin Neurophysiol*, *69*, 91–99.
- Gathercole, S. E., & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*, *29*, 336–360.
- Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, UK: Lawrence Erlbaum.
- Gauger, L. M., Lombardino, L. J., & Leonard, C. M. (1997). Brain morphology in children with specific language impairment. *J Speech Lang Hear Res* 1997; *40*: 1272–84. *Journal of Speech Language and Hearing Research*, *40*, 1272–1284.
- Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience and Biobehavioral Reviews*, *30*(7), 949–960.
- Gérard, C.-L. (1993). *L'enfant dysphasique*. Bruxelles, Belgique: De Boeck Université.
- Ghitza, O. (2011). Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology*, *2*(130).
- Ghitza, O. (2012). On the role of theta-driven syllabic parsing in decoding speech: intelligibility of speech with a manipulated modulation spectrum. *Frontiers in Psychology*, *3*(238).
- Ghitza, O. (2017). Acoustic-driven delta rhythms as prosodic markers. *Language, Cognition and Neuroscience*, *32*(5), 545–561.
- Ghitza, O., Giraud, A.-L., & Poeppel, D. (2012). Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Frontiers in Human Neuroscience*, *6*(340).
- 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), 113–26.
- 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.
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.
- Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *The Journal of the Acoustical Society of America*, 121(3), 1701–1708.
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10.
- Goswami, U. (2012). Entraining the Brain : Applications to Language Research and Links to Musical Entrainment. *Empirical Musicology Review*, 7(1–2), 57–63.
- Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews. Neuroscience*, 16, 43–54.
- Goswami, U., Cumming, R., Chait, M., Mead, N., Wilson, A., & Fosker, T. (2016). Perception of Filtered Speech by Children with Developmental Dyslexia and Children with Specific Language Impairment. *Frontiers in Psychology*, 7(791).
- Goswami, U., & Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Laboratory Phonology*, 4(1), 67–92.
- Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of cognitive neuroscience*, 19(5), 893–906.
- Green, J. R., Moore, C. A., Higashikawa, M., & Steeve, R. W. (2000). The physiologic development of speech motor control: lip and jaw coordination. *Journal of Speech, Language, and Hearing Research : JSLHR*, 43(1), 239–55.
- Greenberg, S. (1999). Speaking in shorthand - a syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, 29, 159–176.
- Greenberg, S., & Ainsworth, W. (2006). *Listening to speech: An auditory perspective*. Mahwah, NJ: Erlbaum.
- Greenberg, S., & Arai, T. (2004). What are the essential cues for understanding spoken language? *IEICE TRANSACTIONS on Information and Systems*, 87(5), 1059–1070.
- Greenberg, S., Arai, T., & Grant, K. W. (2006). The Role of Temporal Dynamics in Understanding Spoken Language. In P. Divenyi, K. Vicsi, & G. Meyer (Eds.), *Dynamics of Speech Production and Perception (Vol 374)*. Amsterdam, Netherlands: IOS Press.
- Greenberg, S., Hollenback, J., & Ellis, D. (1996). Insights into spoken language gleaned from phonetic transcription of the switchboard corpus. *International Conference on Spoken Language Processing*, 24–27.
- Griffiths, T. D., Kumar, S., Sedley, W., Nourski, K. V, Kawasaki, H., Oya, H., ... Howard, M. A. (2010). Direct recordings of pitch responses from human auditory cortex. *Current Biology*, 20(12), 1128–1132.
- Grodzinsky, Y. (2000). The neurology of syntax: language use without Broca's area. *The Behavioral and Brain Sciences*, 23(1), 1-21-71.
- Groenen, P., Maassen, B., Crul, T., & Thoonen, G. (1996). The specific relation between perception and production errors for place of articulation in developmental apraxia of speech. *Journal of Speech and Hearing Research*, 39, 468–482.
- Grosjean, F., & Deschamps, A. (1975). Analyse contrastive des variables temporelles de l'anglais et du français: Vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonetica*, 31, 144–184.

- Gross, J., Hoogenboom, N., Thut, G., Schyns, P. G., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, *11*(12).
- Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics*, *31*(3–4), 423–445.
- Guiraud, H., Ferragne, E., Bedoin, N., & Boulenger, V. (2013). Adaptation to natural fast speech and time-compressed speech in children. In *Interspeech* (p. 1–5). Lyon, France.
- Habib, M., & Besson, M. (2009). Langage, musique et plasticité cérébrale : perspectives pour la rééducation. *Revue de Neuropsychologie*, *18*(1–2), 103–126.
- Habib, M., Lardy, C., Desiles, T., Commeiras, C., Chobert, J., & Besson, M. (2013). Musique et dyslexie : vers une rééducation cognitivo-musicale intermodale des « troubles dys ». *Développements*, *16–17*(3), 36–60.
- 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*(26).
- Hadar, U., Steiner, T. J., Grant, E. C., & Rose, F. C. (1983). Kinematics of head movements accompanying speech during conversation. *Human Movement Science*, *2*, 35–46.
- Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, *2*(156), 1–9.
- 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.
- Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.
- Hertrich, I., Dietrich, S., & Ackermann, H. (2013). Tracking the speech signal--time-locked MEG signals during perception of ultra-fast and moderately fast speech in blind and in sighted listeners. *Brain and Language*, *124*(1), 9–21.
- Hertrich, I., Dietrich, S., Trouvain, J., Moos, A., & Ackermann, H. (2012). Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal. *Psychophysiology*, *49*(3), 322–34.
- Hertrich, I., Mathiak, K., Lutzenberger, W., & Ackermann, H. (2004). Transient and phase-locked evoked magnetic fields in response to periodic acoustic signals. *Neuroreport*, *15*(10), 1687–90.
- Hervais-Adelman, A. G., Carlyon, R. P., Johnsrude, I. S., & Davis, M. H. (2012). Brain regions recruited for the effortful comprehension of noise-vocoded words. *Language and Cognitive Processes*, *27*(7–8), 1145–1166.
- Hickok, G. (2009). Eight Problems for the Mirror Neuron Theory of Action Understanding in Monkeys and Humans. *Journal of Cognitive Neuroscience*, *21*(7), 1229–1243.
- Hickok, G. (2014). *The Myth of Mirror Neurons: The Real Neuroscience of Communication and Cognition*. New York, NY: W. W. Norton & Company, Inc.
- Hickok, G., Houde, J. F., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*(3), 407–422.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, *4*(4), 131–138.

- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92, 67–99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393–402.
- Hill, E. L. (2001a). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, 36(2), 149–171.
- Hill, E. L. (2001b). Non-specific nature of specific language impairment: a review of the literature with regard to concomitant motor impairments. *International Journal of Language & Communication Disorders*, 2, 149–171.
- Hillis, A. E. (2007). Aphasia: progress in the last quarter of a century. *Neurology*, 69, 200–213.
- Homae, F., Watanabe, H., Nakano, T., Asakawa, K., & Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research*, 54(4), 276–280.
- 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. *The Journal of the Acoustical Society of America*, 77(3), 1069–1077.
- Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology*, 104(5), 2500–11.
- Howard, M. F., & Poeppel, D. (2012). The neuromagnetic response to spoken sentences: co-modulation of theta band amplitude and phase. *NeuroImage*, 60(4), 2118–27.
- Hugdahl, K., Gundersen, H., Brekke, C., Thomsen, T., Rimol, L. M., Ersland, L., & Niemi, J. (2004). fMRI brain activation in a Finnish family with specific language impairment compared with a normal control group. *Journal of Speech Language and Hearing Research*, 47, 162–172.
- Hüneke, A., & Lascelles, L. (2014). Terminology mayhem: why it matters – the ramifications for parents and families. *International Journal of Language & Communication Disorders*, 49(4), 410–412.
- Hurst, J. A., Baraitser, M., Auger, E., Graham, F., & Norell, S. (1990). An extended family with a dominantly inherited speech disorder. *Developmental Medicine and Child Neurology*, 32(4), 352–355.
- Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 47(6), 674–89.
- Hutsler, J., & Galuske, R. A. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends in Neurosciences*, 26, 429–435.
- Hwang, J. W., Lee, J.-B., Kim, B.-N., Lee, H.-Y., Lee, D.-S., Shin, M.-S., & Cho, S.-C. (2006). Regional cerebral perfusion abnormalities in developmental language disorder Statistical parametric mapping analysis. *Eur Arch Psychiatry Clin Neurosci*, 256, 131–137.
- Hyde, K. L., Peretz, I., & Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, 46(2), 632–639.
- Iacoboni, M. (2008). The role of premotor cortex in speech perception: evidence from fMRI and rTMS. *Journal of Physiology*, 102, 31–34.

- International Federation of Societies for Electroencephalography and Clinical Neurophysiology. (1974). *Electroencephalogr Clin Neurophysiol*, 37(5), 521.
- Iverson, J. M. (2010). Developing language in a developing body: the relationship between motor development and language development. *Journal of Child Language*, 37(2), 229–261.
- Iverson, J. M., & Thelen, E. (1999). Hand, mouth and brain: The dynamic emergence of speech and gesture. *Journal of Consciousness Studies*, 6(11–12), 19–40.
- Jakobson, R. (1969). *Langage enfantin et aphasique*. Paris, France: Editions de Minuit.
- Jäncke, L., Siegenthaler, T., Preis, S., & Steinmetz, H. (2007). Decreased white-matter density in a left-sided fronto-temporal network in children with developmental language disorder: evidence for anatomical anomalies in a motor-language network. *Brain and Language*, 102, 91–98.
- Janse, E. (2004). Word perception in fast speech: artificially time-compressed vs. naturally produced fast speech. *Speech Communication*, 42(2), 155–173.
- Janse, E., Nooteboom, S., & Quené, H. (2003). Word-level intelligibility of time-compressed speech: prosodic and segmental factors. *Speech Communication*, 41(2–3), 287–301.
- Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7), 267–269.
- Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *NeuroImage*, 47(2), 735–744.
- Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *Journal of Cognitive Neuroscience*, 20(11), 1940–1951.
- Jerbi, K., Lachaux, J.-P., N'Diaye, K., Pantazis, D., Leahy, R. M., Garnero, L., & Baillet, S. (2007). Coherent neural representation of hand speed in humans revealed by MEG imaging. *Proceedings of the National Academy of Sciences*, 104(18), 7676–7681.
- Jerbi, K., Ossandón, T., Hamamé, C. M., Senova, S., Dalal, S. S., Jung, J., ... Lachaux, J.-P. (2009). Task-related gamma-band dynamics from an intracerebral perspective: Review and implications for surface EEG and MEG. *Human Brain Mapping*, 30(6), 1758–1771.
- Jernigan, T. L., Hessenlink, J. R., Sowell, E., & Tallal, P. (1991). Cerebral structure on magnetic resonance imaging in language-and-learning-impaired children. *Archives of Clinical Neuropsychology*, 48, 539–545.
- Jin, Y., Díaz, B., Colomer, M., & Sebastián-Gallés, N. (2014). Oscillation encoding of individual differences in speech perception. *PLoS ONE*, 9(7).
- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–355.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–91.
- Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal Aspects of Stimulus-Driven Attending in Dynamic Arrays. *Psychological Science*, 13(4), 313–319.
- Jusczyk, P. W. (1995). Infant's detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1–23.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA : MIT Press.
- Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' Preference for the Predominant Stress Patterns of English Words. *Child Development*, 64(3), 675–687.

- Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, M. A. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32(3), 402–420.
- Kail, R. (1994). A Method for Studying the Generalized Slowing Hypothesis in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research*, 37, 418–421.
- 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.
- Keller, S. S., Crow, T., Foundas, A., Amunts, K., & Roberts, N. (2009). Broca's area: nomenclature, anatomy, typology and asymmetry. *Brain and Language*, 109, 29–48.
- Kent, R. D. (2000). Research on speech motor control and its disorders: a review and prospective. *Journal of Communication Disorders*, 33, 391–428.
- Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional gain control of ongoing cortical speech representations in a “cocktail party”. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(2), 620–628.
- Kern, S. (2001). Le langage en émergence. *Approche Neuropsychologique Des Apprentissages Chez l'Enfant*, 13, 8–12.
- Khomsî, A. (1984). Langue maternelle et langage adressé à l'enfant. *Langue Française*, 54, 93–107.
- Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, 2, 1–20.
- Kohler, E., Keysers, C., Umiltà, A. M., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, 297, 846–848.
- Kohler, K. J. (1992). Gestural Reorganization in Connected Speech: A Functional Viewpoint on “Articulatory Phonology.” *Phonetica*, 49(3–4), 205–211.
- Kotz, S. A., & Schmidt-Kassow, M. (2015). Basal ganglia contribution to rule expectancy and temporal predictability in speech. *Cortex*, 68, 48–60.
- Kotz, S. a, Schwartze, M., & Schmidt-Kassow, M. (2009). Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 45(8), 982–90.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11, 599–605.
- Krause, C. M., Pesonen, M., Björnberg, C. H., & Hämäläinen, H. (2007). Effects of pulsed and continuous wave 902 MHz mobile phone exposure on brain oscillatory activity during cognitive processing. *Bioelectromagnetics*, 28(4), 296–308.
- Krause, C. M., Sillanmäki, L., Häggqvist, A., & Heino, R. (2001). Test-retest consistency of the event-related desynchronization/event-related synchronization of the 4-6, 6-8, 8-10 and 10-12 Hz frequency bands during a memory task. *Clinical Neurophysiology*, 112(5), 750–757.
- Krishnan, S., Watkins, K. E., & Bishop, D. V. M. (2016). Neurobiological Basis of Language Learning Difficulties. *Trends in Cognitive Sciences*, 20(9), 701–714.
- Kuhl, P. K., & Miller, J. D. (1975). Speech perception by the chinchilla: voiced-voiceless

- distinction in alveolar plosive consonants. *Science*, 190(4209), 69–72.
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, 63(3), 905–917.
- La Guitton, V., de Graaf, J. B., Cazals, Y., Marquis, P., Chauvel, P., & Liégeois-Chauvel, C. (1997). Traitement d'un indice acoustique de la parole naturelle: le délai de voisement. In J. Lambert & J. L. Despopoulos (Eds.), *Perception auditive et compréhension du langage oral* (p. 99–113). Marseille, France: Solal.
- Labov, W. (1986). “Sources of inherent variation in the speech process”, Invariance and Variability. In J. Perkell & D. Klatt (Eds.), *Speech Processes* (p. 402–423.). Hillsdale, N.J: LEA.
- Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. Paris, France: CNRS Edition.
- LaCroix, A. N., Diaz, A. F., & Rogalsky, C. (2015). The relationship between the neural computations for speech and music perception is context-dependent: an activation likelihood estimate study. *Frontiers in Psychology*, 6, 1–19.
- Ladefoged, P. (1971). *A Course in Phonetics*. Fort Worth, TX: Harcourt B.
- Ladefoged, P. (1975). *A course in phonetics*. New York, NY: Harcourt B.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology VII* (p. 637–675). Berlin, Germany: De Gruyter Mouton.
- Lakatos, P., Chi-ming, C., O’Connell, M. N. O., Mills, A., & Schroeder, C. E. (2007). Neuronal Oscillations and Multisensory Interaction in Primary Auditory Cortex. *Neuron*, 53, 279–292.
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, 320(5872), 110–113.
- 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 An Oscillatory Hierarchy Controlling Neuronal Excitability and Stimulus Processing in the Auditory Cortex. *Journal of Neurophysiology*, 94, 1904–1911.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159.
- Laver, J. (2003). Three semiotic layers of spoken communication. *Journal of Phonetics*, 31(3–4), 413–415.
- Lee, C.-Y., & Hung, T.-H. (2008). Identification of Mandarin tones by English-speaking musicians and nonmusicians. *The Journal of the Acoustical Society of America*, 124(5), 3235–3248.
- 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, 454.
- Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered Low-Gamma Sampling in Auditory Cortex Accounts for the Three Main Facets of Dyslexia. *Neuron*, 72(6), 1080–1090.

- Lehtelä, L., Salmelin, R., & Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neuroscience Letters*, 222, 111–114.
- Leonard, C. M., Eckert, M., Given, B., Virginia, B., & Eden, G. (2006). Individual differences in anatomy predict reading and oral language impairments in children. *Brain*, 129(12), 3329–3342.
- Leonard, C. M., Lombardino, L. J., Walsh, K., Eckert, M., Mockler, J. L., Rowe, L. A., ... DeBose, C. B. (2002). Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. *Journal of Communication Disorders*, 35, 501–531.
- Leonard, L. B. (1988). Lexical development and processing in specific language impairment. In R. L. Scheifelbusch & L. L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation and intervention*. (p. 69–87). Austin, TX: Pro-Ed.
- Leonard, L. B. (2014). *Children with specific language impairment*. Cambridge, MA: MIT press.
- Leong, V., & Goswami, U. (2014a). Assessment of rhythmic entrainment at multiple timescales in dyslexia: evidence for disruption to syllable timing. *Hearing Research*, 308, 141–61.
- Leong, V., & Goswami, U. (2014b). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience*, 8, 96.
- Lieberman, A. M., Cooper, F. E., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461.
- Lieberman, A. M., Harris, K. D., Hoffman, H. S., & Griffit, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.
- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Lieberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187–196.
- 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.
- Liégeois, F. J., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., & Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nature Neuroscience*, 6(11), 1230–1237.
- Lindblom, B. (1990). Explaining Phonetic Variation: A Sketch of the H&H Theory. *Speech Production and Speech Modelling*, 55, 403–439.
- 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.
- Logan, K. J., Byrd, C. T., Mazzocchi, E. M., & Gillam, R. B. (2011). Speaking rate characteristics of elementary-school-aged children who do and do not stutter. *Journal of Communication Disorders*, 44(1), 130–47.
- Lotto, A. J., Hickok, G., & Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends in Cognitive Sciences*, 13(3), 110–114.

- Lou, H. C., Henriksen, L., & Bruhn, P. (1984). Focal cerebral hypoperfusion in children with dysphasia and/or attention deficit disorder. *Archives of Clinical Neuropsychology*, *41*, 825–829.
- Lucci, V. (1983). Phonétique du français contemporain à travers la variation situationnelle. *Langue Française*, *60(1)*, 31–66.
- Lüchinger, R., Michels, L., Martin, E., & Brandeis, D. (2011). EEG-BOLD correlations during (post-)adolescent brain maturation. *NeuroImage*, *56(3)*, 1493–1505.
- Lüchinger, R., Michels, L., Martin, E., & Brandeis, D. (2012). Brain state regulation during normal development: Intrinsic activity fluctuations in simultaneous EEG-fMRI. *NeuroImage*, *60(2)*, 1426–1439.
- Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology*, *8(8)*.
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, *54(6)*, 1001–10.
- 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(170)*.
- MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *The Behavioral and Brain Sciences*, *21(4)*, 499–546.
- MacNeilage, P. F., & Davis, B. L. (2001). Motor mechanisms in speech ontogeny: Phylogenetic, neurobiological and linguistic implications. *Current Opinion in Neurobiology*, *11*, 696–700.
- 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.
- Mai, G., Minett, J. W., & Wang, W. S.-Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage*, *133*, 516–528.
- Maillart, C., & Schelstraete, M.-A. (2012). *Les dysphasies: de l'évaluation à la rééducation*. Issy-les-Moulineaux, France: Elsevier-Masson.
- Malécot, A., Johnston, R., & Kizziar, P. A. (1972). Syllabic rate and utterance length in french. *Phonetica*, *26*, 235–251.
- Malmberg, B. (1967). *Structural Linguistics and Human Communication*. New York, NY: Springer.
- Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53(2)*, 155–180.
- Marcus, G. F., & Fisher, S. E. (2003). FOXP2 in focus: What can genes tell us about speech and language? *Trends in Cognitive Sciences*, *7(6)*, 257–262.
- Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of Musical Expertise on Segmental and Tonal Processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23(10)*, 2701–2715.
- Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23(2)*, 294–305.
- Marques, C., Moreno, S., Luís Castro, S., & Besson, M. (2007). Musicians Detect Pitch Violation in a Foreign Language Better Than Nonmusicians: Behavioral and Electrophysiological Evidence. *Journal of Cognitive Neuroscience*, *19(9)*, 1453–1463.

- Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access - words, phonemes, and features. *Psychological Review*, *101*, 653–675.
- Massaro, D. W. (1974). Perceptual units in speech recognition. *Journal of Experimental Psychology*, *102*(2), 199–208.
- Max, K., & Caruso, A. J. (1997). Acoustic measures of temporal intervals across speaking rates: variability of syllable- and phrase-level relative timing. *Journal of Speech Language and Hearing Research*, *40*(5), 1097–1110.
- Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.
- Mazeau, M. (2005). *Neuropsychologie et troubles des apprentissages : Du symptôme à la rééducation*. Paris, France: Masson.
- Mazeau, M., & Pouhet, A. (2014). *Neuropsychologie et troubles des apprentissages chez l'enfant: du développement typique aux "dys". 2nd Edition*. Paris : Elsevier Masson.
- McArthur, G. M., & Bishop, D. V. M. (2004). Frequency discrimination deficits in people with specific language impairment: reliability, validity, and linguistic correlates. *Journal of Speech Language and Hearing Research*, *47*(3), 527–541.
- McGettigan, C., Faulkner, A., Altarelli, I., Obleser, J., Baverstock, H., & Scott, S. K. (2012). Speech comprehension aided by multiple modalities: behavioural and neural interactions. *Neuropsychologia*, *50*(5), 762–76.
- McGettigan, C., & Scott, S. K. (2012). Cortical asymmetries in speech perception: What's wrong, what's right and what's left? *Trends in Cognitive Sciences*, *16*(5), 269–276.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U., & Segui, J. (1981). The Syllable's Role in Speech Segmentation. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 298–305.
- Mehler, J., Dupoux, E., Nazzi, T., & Dehaene-lambertz, G. (1996). Coping with linguistic diversity: the infant's point of view. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax..* Mahwah, NJ: Erlbaum.
- Mehler, J., Jusczyk, P. W., Lamsertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, *29*, 143–178.
- Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The Essential Role of Premotor Cortex in Speech Perception. *Current Biology*, *17*, 1692–1696.
- Meunier, C. (2001). Le traitement de la variabilité dans la parole. Aspects théoriques et méthodologiques. *Travaux Interdisciplinaires Du Laboratoire Parole et Langage*, *20*, 69–90.
- Meunier, C. (2005). Invariants et Variabilité en Phonétique. In N. Nguyen (Ed.), *Phonologie et phonétique: Forme et substance*. (p. 349–374). *Traité de Cognition et Traitement de l'Information*. Cachan, France: Hermès.
- Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*, 711–721.
- Meyer, M., Alter, K., Friederici, A. D., Lohmann, G., & Von Cramon, D. Y. (2002). fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Human Brain Mapping*, *17*(2), 73–88.
- Meynadier, Y. (2001). La syllable phonétique et phonologique: une introduction. *Travaux Interdisciplinaires Du Laboratoire Parole et Langage d'Aix-En-Provence*, *20*, 91–148.
- Michels, L., Muthuraman, M., Lüchinger, R., Martin, E., Anwar, A. R., Raethjen, J., ... Siniatchkin, M. (2013). Developmental changes of functional and directed resting-state connectivities associated with neuronal oscillations in EEG. *NeuroImage*, *81*, 231–242.

- Miller, C. A., Leonard, L. B., & Tomblin, B. (2001). Speed of Processing in Children With Specific Language Impairment. *Journal of Speech Language and Hearing Research, 44*, 416–433.
- Mishkin, M., Ungerleider, L. G., & Macko, C. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences, 6*, 414–417.
- Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: auditory processing or phonological coding? *Journal of Experimental Child Psychology, 64*, 199–231.
- Molinaro, N., Lizarazu, M., Lallier, M., Bourguignon, M., & Carreiras, M. (2016). Out-of-synchrony speech entrainment in developmental dyslexia. *Human Brain Mapping, 37*(8), 2767–83.
- Montgomery, J. W. (2002). Examining the Nature of Lexical Processing in Children with Specific Language Impairment: Temporal Processing or Processing Capacity Deficit? *Applied Psycholinguistics, 23*(3), 447–470.
- Montgomery, J. W. (2003). Working memory and comprehension in children with specific language impairment : what we know so far, *36*, 221–231.
- Montgomery, J. W. (2005). Effects of input rate and age on the real-time language processing of children with specific language impairment. *International Journal of Language & Communication Disorders, 40*(2), 171–188.
- Montgomery, J. W., & Windsor, J. (2014). Examining the Language Performances of Children With and Without Specific Language Impairment : Contributions and Speed of Processing, *50*(June 2007), 778–797.
- Moraïs, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45–64.
- Moran, R. J., Campo, P., Maestu, F., Reilly, R. B., & Dolan, R. J. (2010). Peak frequency in the theta and alpha bands correlates with human working memory capacity. *Frontiers in Human Neuroscience, 4*, 1–12.
- Moreno, S. (2009). Can Music Influence Language and Cognition? *Contemporary Music Review, 28*(3), 329–345.
- Moreno, S., & Besson, M. (2006). Musical training and language-related brain electrical activity in children. *Psychophysiology, 43*(3), 287–291.
- Moreno, S., Marques, C., Santos, A., Santos, M., Luís Castro, S., & Besson, M. (2009). Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex, 19*(3), 712–723.
- Morgan, J., & Demuth, K. (1996). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahawah, NJ: Erlbaum.
- Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences, 114*(42), E8913-E8921.
- Morillon, B., Lehongre, K., Frackowiak, R. S. J., Ducorps, A., & Kleinschmidt, A. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *PNAS, 107*(43), 18688–93.
- 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 intra-cortical study. *Frontiers in Psychology, 248*(3).
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of

- auditory temporal prediction. *Annual New York Academy of Sciences*, 1337(1), 26–31.
- Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5, 1–9.
- Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal Prediction in lieu of Periodic Stimulation. *Journal of Neuroscience*, 36(8), 2342–2347.
- Moritz, C., Yampolsky, S., Papadelis, G., Thomson, J. M., & Wolf, M. (2013). Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing*, 26, 739–769.
- Mottonen, R., & Watkins, K. E. (2009). Motor Representations of Articulators Contribute to Categorical Perception of Speech Sounds. *Journal of Neuroscience*, 29(31), 9819–9825.
- Möttönen, R., & Watkins, K. E. (2012). Using TMS to study the role of the articulatory motor system in speech perception. *Aphasiology*, 26(9), 1103–1118.
- Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews. Neuroscience*, 3, 473–478.
- Murakami, T., Kell, C. A., Restle, J., Ugawa, Y., & Ziemann, U. (2015). Left dorsal speech stream components and their contribution to phonological processing. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 35(4), 1411–22.
- Murakami, T., Restle, J., & Ziemann, U. (2011). Observation-execution matching and action inhibition in human primary motor cortex during viewing of speech-related lip movements or listening to speech. *Neuropsychologia*, 49, 2045–2054.
- Murakami, T., Ugawa, Y., & Ziemann, U. (2013). Utility of TMS to understand the neurobiology of speech. *Frontiers in Psychology*, 4, 1–7.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance*, 24(3), 756–766.
- Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language Discrimination by English-Learning 5-Month-Olds: Effects of Rhythm and Familiarity. *Journal of Memory and Language*, 43(1), 1–19.
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41, 233–243.
- Nguyen, N. (2005). Perception de la parole. In N. Nguyen, S. Wauquier-Gravelines, & J. Durand (Eds.), *Phonétique et phonologie: forme et substance*. (p. 425–447). Paris : Hermès & Lavoisier.
- Nijland, L. (2009). Speech perception in children with speech output disorders, 23(March), 222–239.
- Noordenbos, M. W., Segers, E., Serniclaes, W., Mitterer, H., & Verhoeven, L. (2012). Neural evidence of allophonic perception in children at risk for dyslexia. *Neuropsychologia*, 50(8), 2010–2017.
- Nourski, K. V., Reale, R. A., Oya, H., Kawasaki, H., Kovach, C. K., Chen, H., ... Brugge, J. F. (2009). Temporal envelope of time-compressed speech represented in the human auditory cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 29(49), 15564–74.
- Obleser, J., Boecker, H., Drzezga, A., Haslinger, B., Hennenlotter, A., Roetlinger, M., ... Rauschecker, J. P. (2006). Vowel sound extraction in anterior superior temporal cortex. *Human Brain Mapping*, 27(7), 562–571.

- Obleser, J., & Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cerebral Cortex*, 22(11), 2466–77.
- Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. *Journal of Acoustic Society of America*, 99(3), 1718–1725.
- Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I. H., Saberi, K., ... Hickok, G. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex*, 20(10), 2486–2495.
- Ors, M., Ryding, E., Lindgren, M., Gustafsson, P., Blennow, G., & Rosén, I. (2005). Spect Findings in Children With Specific Language Impairment. *Cortex*, 41(3), 316–326.
- Osnes, B., Hugdahl, K., & Specht, K. (2011). NeuroImage Effective connectivity analysis demonstrates involvement of premotor cortex during speech perception. *NeuroImage*, 54(5), 2437–2445.
- Otake, T., & Cutler, A. (1996). *Phonological Structure and Language Processing: Cross-Linguistic Studies*. Berlin, Germany: de Gruyter Mouton.
- Overy, K., Nicolson, R. I., Fawcett, A. J., & Clarke, E. F. (2003). Dyslexia and music: measuring musical timing skills. *Dyslexia*, 9(1), 18–36.
- Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., & Schulz, M. (2003). Music and Learning-Induced Cortical Plasticity. *Annals of the New York Academy of Sciences*, 999, 438–450.
- Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A Common Language Network for Comprehension and Production: A Contribution to the Definition of Language Epicenters with PET. *NeuroImage*, 11(4), 347–357.
- Parisse, C., & Maillart, C. (2010). Nouvelles propositions pour la recherche et l'évaluation du langage chez les enfants dysphasiques. In C. Gruaz & C. Jacquet-pfau (Eds.), *Autour du mot : pratiques et compétences*. (p. 201–222). Limoges, France: Lambert-Lu.
- Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal Top-Down Signals Increase Coupling of Auditory Low-Frequency Oscillations to Continuous Speech in Human Listeners. *Current Biology*, 25(12), 1649–1653.
- Park, H., Kayser, C., Thut, G., & Gross, J. (2016). Lip movements entrain the observers' low-frequency brain oscillations to facilitate speech intelligibility. *eLife*, 5, 1–17.
- Pasdeloup, V. (2004). Le rythme n'est pas élastique : étude préliminaire de l'influence du débit de parole sur la structuration temporelle. In *Actes des 25èmes Journées d'Etudes sur la Parole 2004* (pp. 397–400). Fès, Maroc.
- Pasdeloup, V. (2005). Figures et fond dans la scène prosodique: leur résistance face aux variations du débit de parole. In *Symposium Interface Discours-Prosodie-IDP05*. Aix-en-Provence, France.
- Pasquini, E. S., Corriveau, K. H., & Goswami, U. (2007). Auditory processing of amplitude envelope rise time in adults diagnosed with developmental dyslexia. *Scientific Studies of Reading*, 11(3), 259–286.
- Patel, A. D. (2007). *Music, Language, and the Brain*. Oxford, UK: Oxford University Press.
- Patel, A. D., Peretz, I., & Tramo, M. (1998). Processing Prosodic and Musical Patterns : A Neuropsychological Investigation, 144(61), 123–144.
- Paus, T., Zijdenbos, A., Worsley, K., Collins, L. D., Blumenthal, J., Giedd, J. N., ... Evans, A. C. (1999). *Structural Maturation of Neural Pathways in Children and Adolescents :*

- In Vivo Study. Science*, 283.
- Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, 3.
- Peelle, J. E., Eason, R. J., Schmitter, S., Schwarzbauer, C., & Davis, M. H. (2010). Evaluating an acoustically quiet EPI sequence for use in fMRI studies of speech and auditory processing. *NeuroImage*, 52(4), 1410–1419.
- Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex*, 23(6), 1378–1387.
- Peelle, J. E., & Sommers, M. S. (2015). Prediction and constraint in audiovisual speech perception. *Cortex*, 68, 169–181.
- Peelle, J. E., & Wingfield, A. (2005). Dissociations in Perceptual Learning Revealed by Adult Age Differences in Adaptation to Time-Compressed Speech. *Journal of Experimental Psychology. Human Perception and Performance*, 31(6), 1315–1330.
- Pefkou, M., Arnal, L. H., Fontolan, L., & Giraud, A.-L. (2017). Theta- and beta-band neural activity reflect independent syllable tracking and intelligibility of time-compressed speech. *The Journal of Neuroscience*, 37(33), 7930–7938.
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for Infant-Directed Over Speech: Evidence From 7-Week-Old Infants. *Infant Behavior and Development*, 15, 325–345.
- Pellegrino, F., Coupé, C., & Marisco, E. (2011). A Cross-Language Perspective on Speech Information Rate. *Language*, 87(3), 539–558.
- Peña, M., & Melloni, L. (2012). Brain Oscillations during Spoken Sentence Processing. *Journal of Cognitive Neuroscience*, 24(5), 1149–1164.
- Penfield, W., & Rasmussen, T. (1952). *The cerebral cortex in man*. New York, NY: Macmillan.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., ... Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758–4763.
- Peretz, I., & Coltheart, M. (2003). Modularity of music processing, 6(7), 688–691.
- Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, 7(30), 1–10.
- Peretz, I., Vuvan, D., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Royal Society Publishing*, B 370.
- Pérez, A., Carreiras, M., Gillon Dowens, M., & Duñabeitia, J. A. (2015). Differential oscillatory encoding of foreign speech. *Brain and Language*, 147, 51–57.
- Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics*, 22(3), 171–198.
- Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B., & Ostry, D. J. (2004). Baby hands that move to the rhythm of language: hearing babies acquiring sign languages babble silently on the hands. *Cognition*, 93(1), 43–73.
- Pickering, M. J., & Garrod, S. (2006). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3), 105–110.
- Pike, K. L. (1945). *The Intonation of American English*. Ann Arbor: University of Michigan

- Press.
- Pindzola, R. H., Jenkins, W. M., & Lokken, K. J. (1989). Speaking rates of young children. *Language, Speech, and Hearing Services in Schools, 20*, 133–138.
- Pisoni, D. B. (1977). Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops. *The Journal of the Acoustical Society of America, 61*(5), 1352–1361.
- Planchou, C. (2014). *Traitement auditifs non verbaux et troubles du développement du langage oral : perception et production musicales*.
- Plante, E., Swisher, L., & Vance, R. (1989). Anatomical correlates of normal and impaired language in a set of dizygotic twins. *Brain and Language, 37*(4), 643–655.
- Plante, E., Swisher, L., Vance, R., & Rapcsak, S. (1991). MRI findings in boys with specific language impairment. *Brain and Language, 41*(1), 52–66.
- Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquière, P., & Wouters, J. (2012). Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. *Ear and Hearing, 33*(1), 134–143.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows : cerebral lateralization as asymmetric sampling in time. *Speech Communication, 41*, 245–255.
- 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.
- Poldrack, R. A., Wagner, A. D., Prull, M. W., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1999). Functional Specialization for Semantic and Phonological Processing in the Left Inferior Prefrontal Cortex 1. *NeuroImage, 10*, 15–35.
- Port, R. F. (2003). Meter and speech. *Journal of Phonetics, 31*, 599–611.
- Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2012). Neural entrainment to rhythmically presented auditory, visual, and audio-visual speech in children. *Frontiers in Psychology, 3*.
- Power, A. J., Mead, N., Barnes, L., & Goswami, U. (2013). Neural entrainment to rhythmic speech in children with developmental dyslexia. *Frontiers in Human Neuroscience, 7*(7).
- Preis, S., Jäncke, L., Schittler, P., Huanq, Y., & Steinmetz, H. (1998). Normal intrasylvian anatomical asymmetry in children with developmental language disorder. *Neuropsychologia 1998; 36: 849–55. Neuropsychologia, 36*, 849–855.
- Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*, 62–88.
- Przybylski, L., Bedoin, N., Krifi-Papoz, S., Herbillon, V., Roch, D., Léculier, L., ... Tillmann, B. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology, 27*(1), 121–31.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nature Reviews. Neuroscience, 11*(5), 351–60.
- Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., 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), 7865–70.
- Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on

- spoken-word perception. *Phonetica*, 62(1), 55–67.
- Ramus, F., Hauser, M. D., Miller, C. A., Morris, R. D., & Mehler, J. (2000). Language Discrimination by Human Newborns and by Cotton-Top Tamarin Monkeys. *Science*, 288(5464), 349–351.
- Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265–292.
- Rapin, I., & Allen, D. A. (1983). Developmental language disorders: Nosologic considerations. In U. Kirk (Ed.), *Neuropsychology of language, reading, and spelling* (p. 155–184). New York, NY: Academic Press.
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271(1–2), 16–25.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–24.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of ““ what ”” and ““ where ”” in auditory cortex. *PNAS*, 97(22).
- Reilly, S., Bishop, D. V. M., & Tomblin, B. (2014). Terminological debate over language impairment in children: Forward movement and sticking points. *International Journal of Language and Communication Disorders*, 49(4), 452–462.
- Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., ... Wake, M. (2014). Specific language impairment: A convenient label for whom? *International Journal of Language and Communication Disorders*, 49(4), 416–451.
- Rice, M. L. (2015). Specific Language Impairment, Nonverbal IQ, Attention-Deficit/Hyperactivity Disorder, Autism Spectrum Disorder, Cochlear Implants, Bilingualism, and Dialectal Variants: Defining the Boundaries, Clarifying Clinical Conditions, and Sorting Out Causes. *Journal of Speech, Language, and Hearing Research*, 59, 122–132.
- 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.
- Richardson, U., Thomson, J. M., Scott, S. K., & Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia*, 10(3), 215–233.
- Ridouane, R., Meynadier, Y., & Fougeron, C. (2009). *La syllabe : objet théorique et réalité physique*. *Faits de langues*, 37, La parole : origines, développement et structures (p. 213–234). Paris, France: Ophrys.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188–194.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11(4), 264–274.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, 15(8), 1261–9.
- Rodd, J. M., Longe, O. A., Randall, B., & Tyler, L. K. (2010). The functional organisation of

- the fronto-temporal language system: Evidence from syntactic and semantic ambiguity. *Neuropsychologia*, 48(5), 1324–1335.
- Röder, B., Stock, O., Neville, H., Bien, S., & Rösler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: a functional magnetic resonance imaging study. *NeuroImage*, 15(4), 1003–1014.
- Rogalsky, C., Love, T., Driscoll, D., Anderson, S. W., & Hickok, G. (2011). Are mirror neurons the basis of speech perception? Evidence from five cases with damage to the purported human mirror system. *Neurocase*, 17(2), 178–187.
- Romero, L., Walsh, V., & Papagno, C. (2006). The neural correlates of phonological short-term memory: a repetitive transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, 18(7), 1147–55.
- Roncaglia-Denissen, M. P., Schmidt-Kassow, M., & Kotz, S. A. (2013). Speech rhythm facilitates syntactic ambiguity resolution: ERP evidence. *PloS One*, 8(2).
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions: Biological Sciences*, 336(1278), 367–373.
- Rossi, M. (1989). «De la quiddité des variables». In H. Meloni (Ed.), *Actes du séminaire Variabilité et spécificité du locuteur : Etudes et Applications* (p. 11–31). Marseille, France.
- Roy, A. C., Craighero, L., Fabbri-Destro, M., & Fadiga, L. (2008). Phonological and lexical motor facilitation during speech listening: A transcranial magnetic stimulation study. *Journal of Physiology Paris*, 102, 101–105.
- Sallat, S., & Jentschke, S. (2015). Music Perception Influences Language Acquisition – Melodic and Rhythmic-Melodic Perception in Children with Specific Language Impairment. *Behavioural Neurology*, 2015: 606470.
- Sammler, D., Grosbras, M. H., Anwender, A., Bestelmeyer, P. E. G., & Belin, P. (2015). Dorsal and ventral pathways for prosody. *Current Biology*, 25(23), 3079–3085.
- Sato, M., Tremblay, P., & Gracco, V. L. (2009). A mediating role of the premotor cortex in phoneme segmentation. *Brain and Language*, 111(1), 1–7.
- Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.-S., ... Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105(46), 18035–18040.
- Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, 34(7), 1015–1022.
- Scharinger, M., Bendixen, A., Trujillo-Barreto, N. J., & Obleser, J. (2012). A sparse neural code for some speech sounds but not for others. *PloS One*, 7(7).
- Schelstraete, M.-A., & Collette, E. (2012). Retard de langage et dysphasie: question de diagnostic. *Langage et Pratiques*, 50, 48–68.
- Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, 1060, 219–230.
- Schmidt-Kassow, M., Heinemann, L. V., Abel, C., & Kaiser, J. (2013). Auditory-motor synchronization facilitates attention allocation. *NeuroImage*, 82, 101–106.
- Schomers, M. R., Kirilina, E., Weigand, A., Bajbouj, M., & Pulvermüller, F. (2015). Causal

- influence of articulatory motor cortex on comprehending single spoken words: TMS evidence. *Cerebral Cortex*, 25, 3894–3902.
- Schön, D., Gordon, R., Campagne, A., Magne, C., Astesano, C., Anton, J.-L., & Besson, M. (2010). Similar cerebral networks in language, music and song perception. *NeuroImage*, 51(1), 450–61.
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1).
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., Puce, A., Program, S., & Hall, A. S. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, 12(3), 106–113.
- 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, 172–176.
- Schwartz, J.-L., Basirat, A., Ménard, L., & Sato, M. (2012). The Perception-for-Action-Control Theory (PACT): A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, 25(5), 336–354.
- Schwartz, J.-L., Berthommier, F., & Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2), 69–78.
- Scott, S. K. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123(12), 2400–2406.
- Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, 123, 2400–2406.
- Scott, S. K., & Johnsrude, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, 26(2), 100–107.
- Segui, J., Dupoux, E., & Mehler, J. (1991). The role of the syllable in speech segmentation, phoneme identification, and lexical access. In *Cognitive models of speech processing* (p. 263–280). Cambridge, MA : MIT-Press.
- Segui, J., Frauenfelder, U. H., & Mehler, J. (1981). Phoneme monitoring, syllable monitoring and lexical access. *British Journal of Psychology*, 72, 471–477.
- Serniclaes, W., Sprenger-Charolles, L., Carré, R., & Demonet, J.-F. (2001). Perceptual Discrimination of. *Journal of Speech Language and Hearing Research*, 44, 384–399.
- 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, 336–361.
- Shahin, A. J., Trainor, L. J., Roberts, L. E., Backer, K. C., & Miller, L. M. (2010). Development of auditory phase-locked activity for music sounds. *Journal of Neurophysiology*, 103(1), 218–29.
- Shannon, R. V, Zeng, F.-G. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234), 303–4.
- Skeide, M. A., Brauer, J., & Friederici, A. D. (2016). Brain Functional and Structural Predictors of Language Performance. *Cerebral Cortex*, 26(5), 2127–2139.
- Skipper, J. I., Goldin-Meadow, S., Nusbaum, H. C., & Small, S. L. (2007). Speech-associated gestures, Broca's area, and the human mirror system. *Brain and Language*, 101(3), 260–277.
- Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor

- cortical activation during speech perception. *NeuroImage*, 25, 76–89.
- Slater, J., Tierney, A., & Kraus, N. (2013). At-Risk Elementary School Children with One Year of Classroom Music Instruction Are Better at Keeping a Beat. *PLoS ONE*, 8(10), 1–9.
- Snow, E., & Ferguson, C. A. (1977). *Talking to children: language input and acquisition*. Cambridge, MA : MIT-Press.
- Snowling, M., Bishop, D. V. M., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry*, 41, 587–600.
- Soltész, F., Szücs, D., Leong, V., White, S., & Goswami, U. (2013). Differential entrainment of neuroelectric delta oscillations in developmental dyslexia. *PloS One*, 8(10), 1–11.
- Soriano-Mas, C., Pujol, J., Ortiz, H., Deus, J., López-Sala, A., & Sans, A. (2009). Age-related brain structural alterations in children with specific language impairment. *Human Brain Mapping*, 30(5), 1626–1636.
- Stark, R. E., & Montgomery, J. W. (1995). Sentence processing in language-impaired children under conditions of filtering and time compression. *Applied Psycholinguistics*, 16(2), 137–154.
- Stern, D. N. (1974). The Goal and Structure of Mother-Infant Play. *Journal of the American Academy of Child Psychiatry*, 13(3), 402–421.
- Stetson, R. H. (1951). *Motor Phonetics: a Study of Speech Movements in Action*. Amsterdam, Netherland: North-Holl.
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*, 111(4), 1872–1891.
- Stevens, K. N., & Blumstein, S. E. (1978). Invariant cues for place of articulation in stop Consonants. *Journal of Acoustic Society of America*, 64, 1358–1368.
- Stevens, K. N., & Klatt, D. (1974). Role of formant transitions in the voiced-voiceless distinction for stops. *The Journal of the Acoustical Society of America*, 55(3), 653–9.
- Strauß, A., & Schwartz, J.-L. (2017). The syllable in the light of motor skills and neural oscillations. *Language, Cognition and Neuroscience*, 32(5), 562–569.
- Strauß, A., Wöstmann, M., & Obleser, J. (2014). Cortical alpha oscillations as a tool for auditory selective inhibition. *Frontiers in Human Neuroscience*, 8(350), 1–7.
- Streeter, L. A., & Nigro, G. N. (1979). The role of medial consonant transitions in word perception. *Journal of Acoustic Society of American*, 65, 1533–1541.
- Sturm, J. A., & Seery, C. H. (2007). Speech and Articulatory Rates of School-Age Children in Conversation and Narrative Contexts. *Language Speech and Hearing Services in Schools*, 38, 47–59.
- Takano, T., & Ogawa, T. (1998). Characterization of developmental changes in EEG-gamma band activity during childhood using the autoregressive model. *Acta Paediatrica Japonica*, 40, 446–452.
- Tallal, P. (2003). Language learning disabilities: integrating research approaches. *Current Directions in Psychological Science*, 271, 81–84.
- Tallal, P., Miller, S. L., Bedi, G., Byrna, G., Wang, X., Nagarajan, S. S., ... Merzenich, M. M. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science*, 271, 81–84.

- Tallal, P., & Piercy, M. (1973a). Defects of Non-Verbal Auditory Perception in Children with Developmental Aphasia. *Nature*, *241*(5390), 468–469.
- Tallal, P., & Piercy, M. (1973b). Developmental aphasia: impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia*, *11*, 189–398.
- Tallal, P., & Piercy, M. (1975). Developmental aphasia: the perception of brief vowels and extended stop consonants. *Neuropsychologia*, *13*(1), 69–74.
- Tallal, P., Stark, R. E., & Curtiss, B. (1976). Relation between speech perception and speech production impairment in children with developmental dysphasia. *Brain and Language*, *3*, 305–317.
- Tallon-baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation, *3*(4), 151–162.
- Tallon-baudry, C., Bertrand, O., & Fischer, C. (2001). Oscillatory synchrony between human extrastriate areas during visual short-term memory maintenance. *The Journal of Neuroscience*, *21*, 1–5.
- Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., ... Wartenburger, I. (2009). Sensitivity of Newborn Auditory Cortex to the Temporal Structure of Sounds. *Journal of Neuroscience*, *29*(47), 14726–14733.
- Telkemeyer, S., Rossi, S., Nierhaus, T., Steinbrink, J., Obrig, H., & Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near-infrared spectroscopy and EEG study. *Frontiers in Psychology*, *2*.
- Tettamanti, M., Alkadhi, H., Moro, A., Perani, D., Kollias, S., & Weniger, D. (2002). Neural correlates for the acquisition of natural language syntax. *NeuroImage*, *17*(2), 700–709.
- 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*(19737), 1–6.
- Thomson, J. M., Fryer, B., & Maltby, J. (2006). Auditory and motor rhythm awareness in adults with dyslexia, *29*(3), 334–348.
- 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), 120–9.
- Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience*, *8*.
- Tillmann, B. (2012). Music and Language Perception: Expectations, Structural Integration, and Cognitive Sequencing. *Topics in Cognitive Science*, *4*(4), 568–584.
- Tomblin, B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research*, *40*, 1245–1260.
- Trehub, S. E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development*, *47*, 466–472.
- Tuller, B., & Kelso, J. A. (1990). Phase Transitions in Speech and their Perceptual Consequences. In *Attention et performance* (p. 429–451). Paris : Editions Odile Jacob.
- Tzourio-Mazoyer, N., Heim, A., Zilbovicius, M., Gerard, C., & Mazoyer, B. (1994). Abnormal regional CBF response in left hemisphere of dysphasic children during a language task. *Pediatric Neurology*, *10*(1), 20–26.
- Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., & Singer, W. (2010). Neural

- synchrony and the development of cortical networks. *Trends in Cognitive Sciences*, 14(2), 72–80.
- Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: the procedural deficit hypothesis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 41(3), 399–433.
- Valero-Cabré, A., Pascual-Leone, A., & Coubard, O. A. (2011). Transcranial magnetic stimulation (TMS) in basic and clinical neuroscience research. *Revue Neurologique*, 167(4), 291–316.
- Van Son, R. J. J. H., & Pols, L. C. W. (1996). A comparison between the acoustics of vowel and consonant reduction. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 20, 13-25.
- Van Son, R. J. J. H., & Pols, L. C. W. (1999). An acoustic description of consonant reduction. *Speech Communication*, 28, 125–140.
- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences*, 102(4), 1181–1186.
- Vander Ghinst, M., Bourguignon, M., Op de Beeck, M., Wens, V., Marty, B., Hassid, S., ... de Tiège, X. (2016). Left Superior Temporal Gyrus Is Coupled to Attended Speech in a Cocktail-Party Auditory Scene. *The Journal of Neuroscience*, 36(5), 1596–1606.
- Vanvooren, S., Poelmans, H., Hofmann, M., Ghesquière, P., & Wouters, J. (2014). Hemispheric asymmetry in auditory processing of speech envelope modulations in prereading children. *The Journal of Neuroscience*, 34(4), 1523–1529.
- Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The Brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience*, 2, 229–239.
- Vargha-Khadem, F., Gadian, D. G., Copp, A., & Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience*, 6, 131–138.
- Vargha-Khadem, F., Watkins, K. E., Price, C. J., Ashburner, J., Alcock, K. J., Connelly, A., ... Passingham, R. E. (1998). Neural basis of an inherited speech and language disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 95(21), 12695–700.
- Vaxelaire, B. (1993). *Etude comparée des effets des variations de débit -lent, rapide- sur les paramètres articulatoires, à partir de la cinéradiographie (sujets français)*. Strasbourg II.
- Watkins, K. E., Dronkers, N. F., & Vargha-Khadem, F. (2002a). Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain*, 125(3), 452–464.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: the role of Broca's area. *Journal of Cognitive Neuroscience*, 16(6), 978–987.
- Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, 41(8), 989–994.
- Watkins, K. E., Vargha-Khadem, F., Ashburner, J., Passingham, R. E., Connelly, A., Friston, K. J., ... Gadian, D. G. (2002b). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain*, 125(3), 465–478.
- Weiss, A. H., Granot, R. Y., & Ahissar, M. (2014). The enigma of dyslexic musicians. *Neuropsychologia*, 54(1), 28–40.
- Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in

- audition: cognitive and clinical perspectives. *Frontiers in Psychology*, 2(73), 1–15.
- Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17(3), 323–333.
- Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Cerebral dominance for language function in adults with specific language impairment or autism. *Brain*, 131(12), 3193–3200.
- Whitford, T. J., Rennie, C. J., Grieve, S. M., Clark, R. C., Gordon, E., & Williams, L. M. (2007). Brain maturation in adolescence: Concurrent changes in neuroanatomy and neurophysiology. *Human Brain Mapping*, 28(3), 228–237.
- Wilson, S. M. (2009). Speech perception when the motor system is compromised. *Trends in Cognitive Sciences*, 13(8), 329–330.
- Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33, 316–325.
- 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.
- Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., & Fries, P. (2007). Modulation of Neuronal Interactions Through Neuronal Synchronization. *Science*, 316, 1609–1612.
- Woodruff Carr, K., White-Schwoch, T., Tierney, A., Strait, D. L., & Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences*, 111(40), 14559–14564.
- World Health Organization. (2005). ICD-10 International Statistical Classification of Diseases and Related Health Problems. 2, 10th revision.
- Worrell, G. A., Jerbi, K., Kobayashi, K., Lina, J. M., Zemann, R., & Le Van Quyen, M. (2012). Recording and analysis techniques for high-frequency oscillations. *Progress in Neurobiology*, 98(3), 265–278.
- Zatorre, R. J. (2003). Music and the brain. *Annals of the New York Academy of Sciences*, 999(1), 4–14.
- 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.
- 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.
- Zheng, X., & Pierrehumbert, J. B. (2010). The effects of prosodic prominence and serial position on duration perception. *The Journal of the Acoustical Society of America*, 128(2), 851–859.
- Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, 131(1), 3–29.
- Zion Golumbic, E. M., Cogan, G. B., Schroeder, C. E., & Poeppel, D. (2013). Visual Input Enhances Selective Speech Envelope Tracking in Auditory Cortex at a “Cocktail Party.” *Journal of Neuroscience*, 33(4), 1417–1426.
- Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., ... Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended

- speech at a “cocktail party.” *Neuron*, 77(5), 980–991.
- Zion Golumbic, E. M., 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), 151–61.
- Zoefel, B., & Vanrullen, R. (2015). Selective Perceptual Phase Entrainment to Speech Rhythm in the Absence of Spectral Energy Fluctuations. *Journal of Neuroscience*, 35(5), 1954–1964.

ANNEXE 1. Oscillations et perception de la parole chez les adultes : étude en MEG



Tracking of Speech Rhythm by Neuronal Oscillations: a MEG Study of Normal vs Fast Speech Processing

Ana-Sofia Hincapié^{1,2}, Hannu Laaksonen^{3,4}, Dimitri Bayle⁵, Hélène Guiraud³, Karim Jerbi^{1*}, Véronique Boulenger^{3*}

¹Psychology Department, University of Montreal, Quebec, Canada; ²Pontificia Universidad Católica de Chile, Santiago de Chile, Chile; ³Laboratory Dynamics of Language CNRS/Université Lyon2 UMR5596, Lyon, France; ⁴Lyon Neuroscience Research Center, DyCog team, Inserm U1028, CNRS UMR5292, Lyon, France; ⁵Université Paris Ouest Nanterre La Défense, Paris, France.

* co-last authors

ashincap@uc.cl, helene.guiraud@univ-lyon2.fr, karim.jerbi@umontreal.ca, veronique.boulenger@cnrs.fr



Take Home Message

Oscillatory brain activity displays task-specific coupling with the speech signal at frequencies that reflect syllable rate and pitch (f_0). Our findings provide further evidence for power modulations in auditory cortices during speech processing: desynchronization at the alpha band (8-10Hz) and synchronization at the high gamma band (60-90Hz).

Introduction

- Close correspondence between speech rhythm and cortical oscillations has been shown. As seen in Figure 1, evidence suggests that when the speech signal is perceived there is phase shift on the large-scale oscillatory networks involved in auditory cognition. This entrainment allows the brain to parse the acoustic signal into linguistic elements critical for language comprehension [1-3].
- Previous studies have shown coupling of ongoing oscillations in auditory regions to the amplitude envelope of the speech signal or to its fundamental frequency (f_0) [4-7].
- Alignment between oscillatory activity in auditory and premotor regions has further been reported in speech-relevant frequency bands [8-9], supporting recent dual-stream models of a functional role of the motor system in speech perception [10].

Objective: Investigate the effect of speech rate shifts (from normal to fast) on the properties of acoustic-cortical coupling (a) at syllable rate and (b) at the fundamental frequency of speech (pitch).

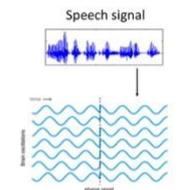


Figure 1. Brain Oscillations phase shift when speech signal is perceived

Methods

Brain activity was recorded using a 275-channel whole-head CTF MEG system. Data was acquired at the Lyon Neuroscience Research Center, France. The task was to listen to sentences from a list and to press a button with the left index finger whenever a filler trial was heard. Analysis were carried out using Matlab and Fieldtrip. Table 1 shows the characteristics of the auditory signals and the frequency bands chosen for the acoustic-cortical coherence analysis. Figure 2 summarizes the methodology used.

Condition	Mean Speech Rate (Syllables/s)	Frequency Band Speech Rate (Hz)	Mean f_0 (Hz)	Frequency Band f_0 (Hz)
Normal	6,76 ± 0,57	5,7 - 7,7	79 ± 2,3	77 - 81
Natural Fast	9,15 ± 0,60	8 - 10	88,3 ± 7,2	86,3 - 90,3
Time-Compressed	9,15 ± 0,60	8 - 10	79,5 ± 2,5	77 - 81

Table 1. Characteristics of the auditory signals and the frequency bands chosen for the analysis

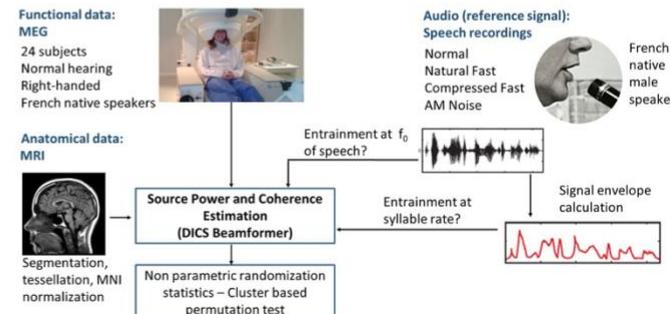


Figure 2. Methodology

Results

Tracking Speech Rate

Cortico-acoustic Coupling @ [5,7-7,7Hz] (syllabic rate for normal speech stimuli) Cortico-acoustic Coupling @ [8-10Hz] (Syllabic rate for fast speech stimuli)

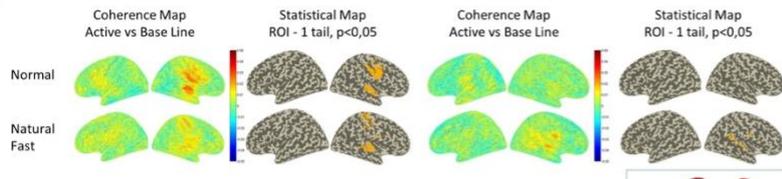


Figure 3. Tracking of speech rate

Power Modulations

Task-based Power (8-10Hz)

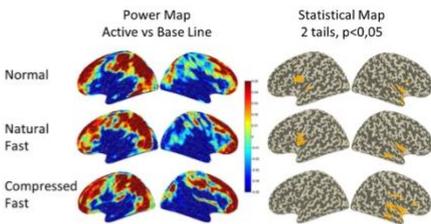


Figure 5. Power Modulations

Tracking of f_0

Cortico-acoustic Coupling @ [77-81Hz] (f_0 range for normal & compressed speech) Cortico-acoustic Coupling @ [86,3-90,3Hz] (f_0 range for fast natural speech)

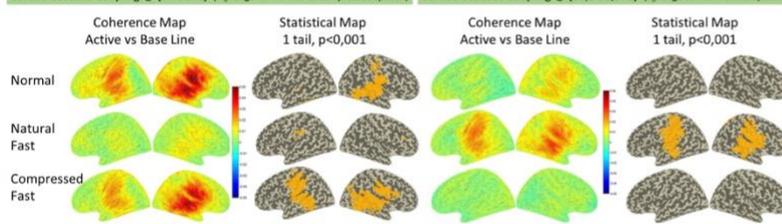


Figure 4. Tracking of fundamental frequency

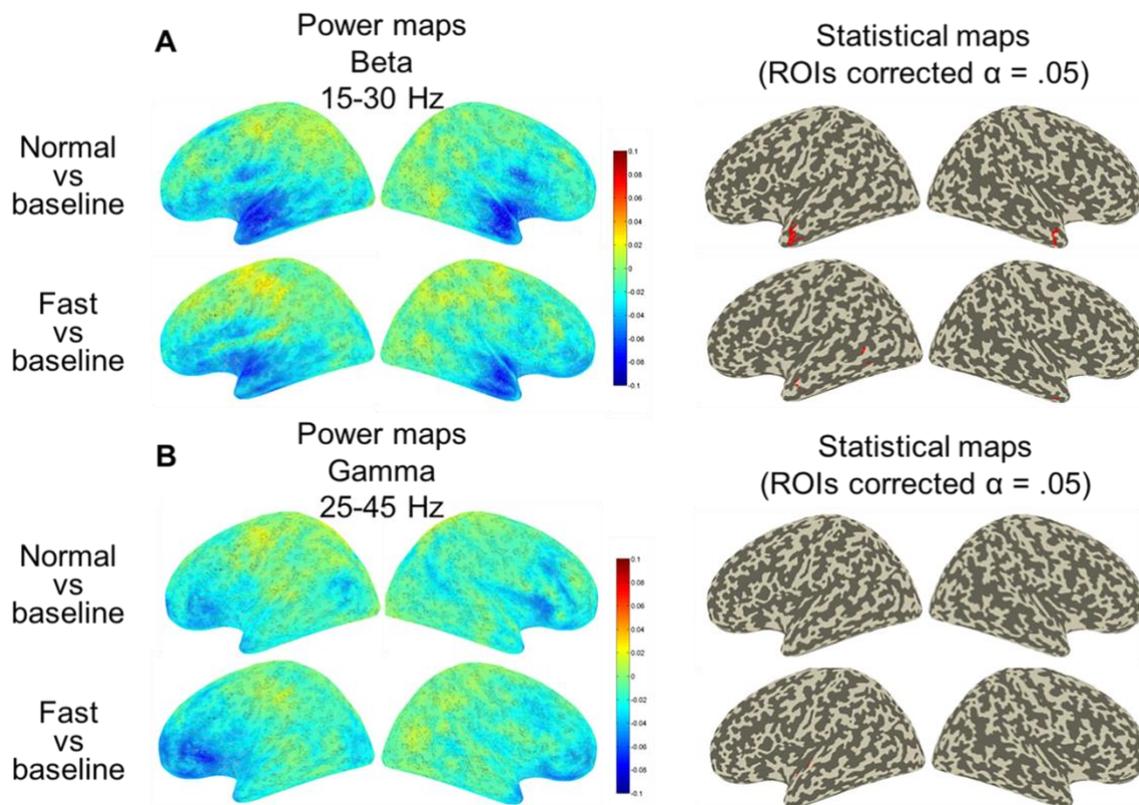
Discussion :

- Statistically significant increase of acoustic-cortical coherence in the right hemisphere, tracking syllabic rate. Motor activation in the natural fast condition at the normal speech rate might indicate recruitment of motor areas to facilitate comprehension (Figure 3).
- Statistically significant entrainment of cortical oscillations at f_0 , with shift in coupling frequency mirroring changes in pitch (Figure 4). Note that the compressed fast condition has approx the same mean f_0 as the normal condition (Table 1), which explains the coupling patterns.
- Desynchronization in the alpha band (8-10Hz) and increase in synchronization in the high gamma band (60-90Hz) (Figure 5).

References

- [1] Ghazizadeh, O. (2011). "Linking speech perception and neurophysiology: speech decoding guided by cascaded oscillators locked to the input rhythm". *Frontiers in Psychology*, vol. 2, p. 130.
- [2] Giraud, A.L., Poeppel, D. (2012). "Cortical oscillations and speech processing: emerging computational principles and operations". *Nature Neuroscience*, vol. 15, pp. 511-517.
- [3] Poeppel, D., Davis, M.H. (2012). "Neural Oscillations Carry Speech Rhythm through to Comprehension". *Frontiers in Psychology*, vol. 3, p. 320.
- [4] Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., et al. (2001). "Speech comprehension is correlated with temporal response patterns recorded from auditory cortex". *Proceedings of the National Academy of Science USA*, vol. 98, pp. 13367-13372.
- [5] Poeppel, D., Gross, J., Davis, M.H. (2012). "Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension". *Cerebral Cortex*.
- [6] Bourguignon M, De Tielge X, de Beeck M, Ligtot N, Paquier P, Van Bogaert P, Goldman S, Hari R, Jousmäki V. (2013) The pace of prosodic phrasing couples the listener's cortex to the reader's voice. *Hum Brain Mapp.* Feb;34(2):314-26. doi: 10.1002/hbm.22402
- [7] Oumecké G, Suarez Garcia S, Bourguignon M, Wens V, Op de Beeck M, Marty B, Decouck N, Soncarriu MV, Goldman S, Jousmäki V, Van Bogaert P, Poeppel D, Ligtot N, de Beeck M, Ligtot N, Paquier P, Goldman S, Hari R, Jousmäki V. (2013) "Endogenous cortical rhythms determine cerebral specialization for speech perception and production". *Neuron*, vol. 56, pp. 1127-1134.
- [8] Morillon, B., Lehongre, K., Frackowiak, R.S., Ducours, A., Kleinschmidt, A., et al. (2010). "Neurophysiological origin of human brain asymmetry for speech and language". *Proceedings of the National Academy of Science USA*, vol. 107, pp. 18688-18693.
- [9] Rauschecker, J.P., Scott, S.K. (2009). "Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing". *Nature Neuroscience*, vol. 12, no. 6, pp. 718-724.

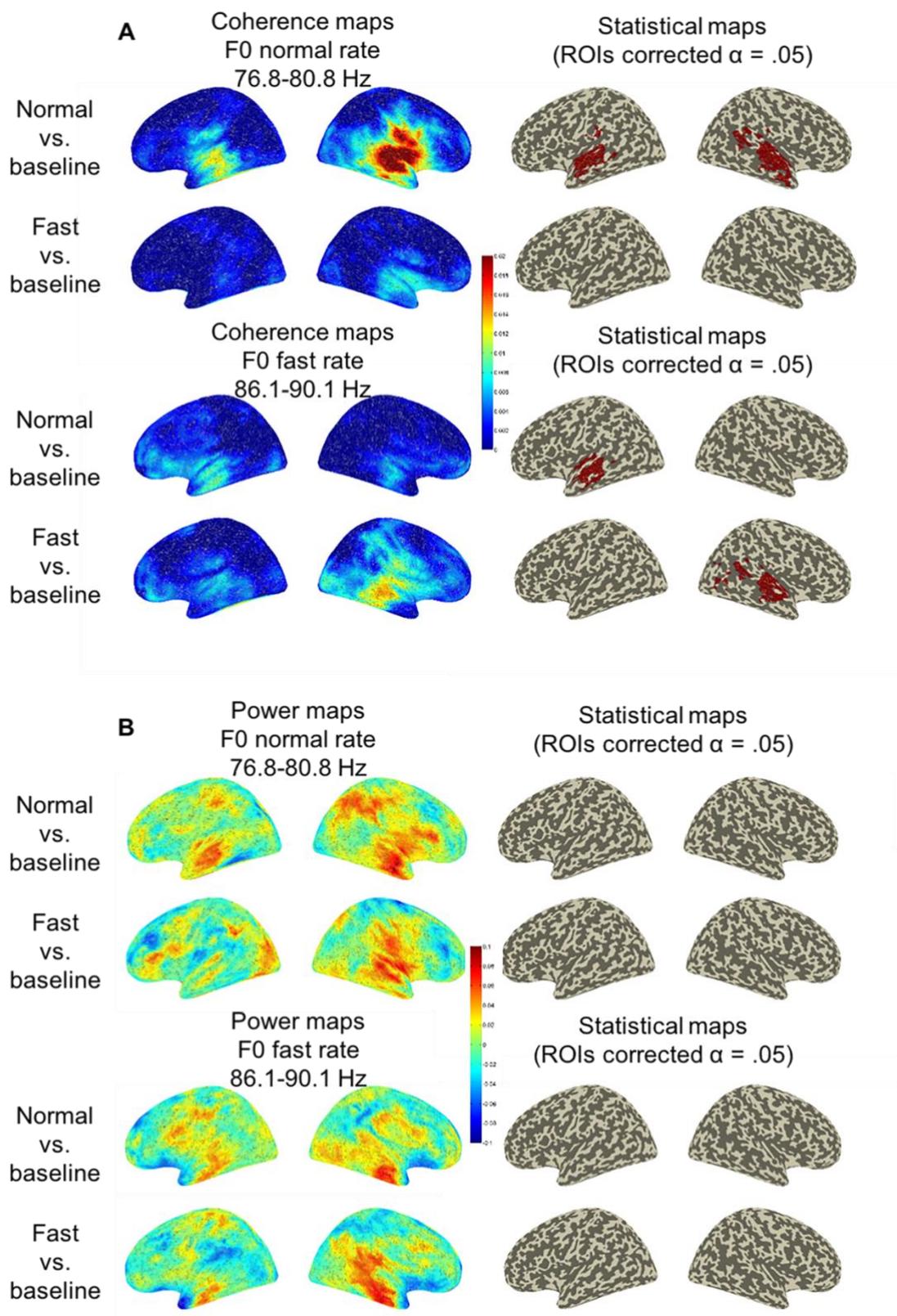
ANNEXE 2. Modulations de puissance dans les bandes beta et gamma basse chez les enfants neurotypiques



Cartes des modulations de puissance dans les bandes de fréquence (A) beta (15-30 Hz) et (B) gamma basse (25-45 Hz) dans les conditions de perception de parole à débit normal et rapide, comparé à la ligne de base. Les cartes statistiques correspondantes sont présentées sur les 4 régions d'intérêt (ROIs) après correction ($\alpha = .05$).

Une désynchronisation beta significative est observée chez ces enfants au niveau des régions temporelles antérieures bilatérales lors de la perception de parole à débit normal et de façon plus latéralisée à gauche lors de l'écoute de parole rapide au sein des régions temporelles antérieures et postérieures. En gamma, aucune modulation de puissance significative n'est observée.

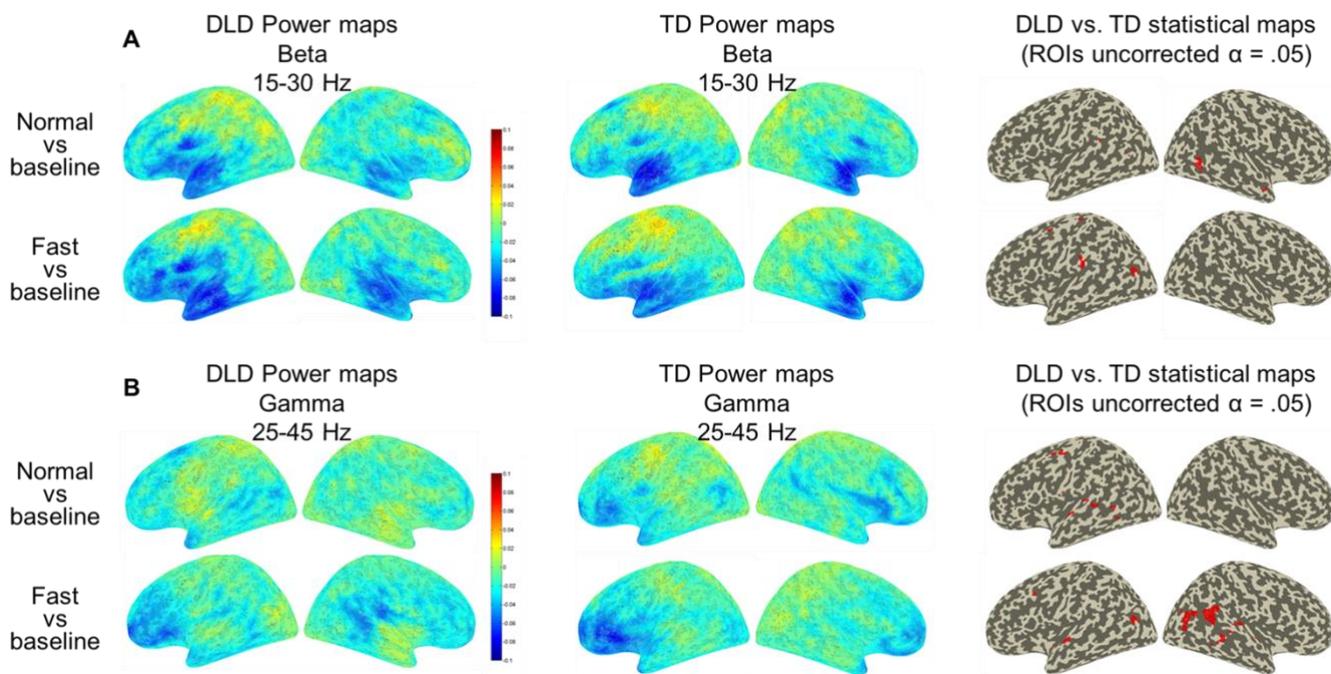
ANNEXE 3. Analyses de la fréquence fondamentale chez les enfants neurotypiques



Cartes de (A) cohérence et (B) modulations de puissance dans la bande de fréquence englobant la fréquence fondamentale moyenne (F0) de la parole à débit normal (haut, 76.8-80.8 Hz) et rapide (bas, 86.1-90.1 Hz) dans les conditions de parole à débit normal et rapide, comparé à la ligne de base. Les cartes statistiques correspondantes sont présentées pour les 4 ROIs après correction ($\alpha = .05$).

Une augmentation significative de la cohérence cortico-acoustique au sein des régions temporales bilatérales est observée lors de la perception de parole à débit normal dans la bande de fréquence englobant la F0 moyenne des phrases de cette condition (76.8-80.8 Hz). De la même manière, une augmentation significative de cohérence est observée dans ces mêmes régions, même si principalement latéralisées à droite, dans la bande de fréquence englobant la F0 moyenne des phrases de la condition rapide (86.1-90.1 Hz) lors de la perception de parole rapide. Un pic de cohérence à gauche semble également être observé lors de la perception des phrases à débit normal dans cette bande de fréquence.

ANNEXE 4. Modulations de puissance dans les bandes beta et gamma basse chez les enfants dysphasiques



Cartes des modulations de puissance dans les bandes (A) beta (15-30 Hz) et (B) gamma basse (25-45 Hz) dans les conditions de perception de parole à débit normal et rapide, comparé à la ligne de base. Les cartes statistiques correspondantes sont présentées pour les 4 ROIs sans correction ($\alpha = .05$). La significativité des résultats ne survit pas à la correction, cependant les résultats non corrigés semblent encourageants.