



HAL
open science

Étude d'un biais prosodique précoce : le cas de la loi iambo-trochaïque

Nawal Abboub

► **To cite this version:**

Nawal Abboub. Étude d'un biais prosodique précoce : le cas de la loi iambo-trochaïque. Psychologie. Université Sorbonne Paris Cité, 2015. Français. NNT : 2015USPCB001 . tel-01865814

HAL Id: tel-01865814

<https://theses.hal.science/tel-01865814>

Submitted on 2 Sep 2018

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

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



Ecole Doctorale ED261:
« Cognition, Comportement, Conduites Humaines »

Étude d'un biais prosodique précoce :
Le cas de la loi iambo-trochaïque

Early prosodic biases: The iambic trochaic law

Nawal Abboub
Thèse de doctorat de Sciences Cognitives

Présentée et soutenue publiquement le vendredi 25 septembre 2015 devant un jury composé de

Marina Nesper
Fabrice Wallois
Hellmuth Obrig
Judit Gervain
Thierry Nazzi

Professeure
MCU-PH
Professeur
Chargée de recherche
Directeur de recherche

Rapporteur
Rapporteur
Examineur
Directrice de thèse
Directeur de thèse

Résumé

Le but de cette thèse a été d'explorer certains des mécanismes relatifs à la perception de la prosodie linguistique. Cet intérêt provient du fait que la prosodie, qui est portée dans le signal acoustique par les indices d'intensité, durée et pitch, pourrait guider les nourrissons dans l'apprentissage des mots et de leur ordre dans les phrases. Nous avons donc étudié un mécanisme qui pourrait sous-tendre ces capacités prosodiques précoces : la loi iambo-trochaïque (iambic-trochaic law : ITL).

L'ITL (Woodrow, 1909; Hayes, 1995; Nespors et al., 2008) est un des principes d'organisation de la perception auditive qui a été proposé comme ayant un rôle important non seulement dans la perception chez l'adulte, mais aussi pendant l'acquisition du langage chez le nourrisson. L'ITL postule une préférence générale pour le regroupement des éléments sonores (syllabes, notes musicales...) en paires, en fonction des variations d'intensité, de tonalité (pitch) ou de durée de ces éléments : les sons forts ou aigus marquent le début de groupes de pattern trochaïque (fort-faible ou aigu-grave), alors que les sons longs marquent la fin de groupes de pattern iambique (court-long). Toutefois, les langues diffèrent dans leur façon d'utiliser ces indices acoustiques pour marquer la prééminence prosodique dans les mots et dans les syntagmes. Par exemple, le français ne possède pas d'accent lexical mais un accent au niveau du syntagme où la dernière syllabe est allongée, ce qui crée un contraste court-long (pattern iambique basé sur la durée). À l'inverse, en anglais ou en allemand, l'accent lexical est en général sur la première syllabe, qui est plus forte et/ou plus aigüe, ce qui induit un pattern trochaïque basé sur l'intensité ou le pitch. L'environnement linguistique pourrait donc interagir avec l'ITL.

Cette thèse s'articule en deux axes. D'une part, dans une tâche de segmentation / reconnaissance de paires de syllabes, nos données montrent que des sensibilités à l'ITL sont présentes chez l'adulte (Exp. 1) et le nourrisson de 7,5 mois (Exp. 2), français et allemands. De faibles différences cross-linguistiques entre les deux groupes sont retrouvées pour le groupement basé sur l'intensité. D'autre part, à l'aide de la NIRS (Near InfraRed Spectroscopy), nous avons mesuré les réponses d'activité cérébrale chez des nouveau-nés et mis en évidence que des sensibilités à l'ITL sont présentes à la naissance et déjà influencées par la langue entendue *in utero* (Exp. 3-6). Pour finir, nous avons exploré de quelle manière l'environnement linguistique influence les compétences de discrimination de patterns rythmiques d'accentuation des mots chez des nourrissons de 10 mois (Exp. 7). Nos résultats révèlent que des nourrissons bilingues apprenant le français et une langue avec un accent lexical variable sont capables de discriminer des patterns d'accentuation contrairement aux monolingues français. Pris ensemble, nos résultats contribuent à la compréhension de l'origine développementale de l'ITL, de sa modulation par l'environnement linguistique et du développement langue-spécifique du traitement et des préférences rythmiques.

Mots clefs : Perception du langage, prosodie, acquisition du langage, nourrissons, nouveau-nés, bilinguisme, NIRS

Abstract

The goal of this doctoral dissertation was to explore the mechanisms underlying linguistic prosodic perception. Prosody is carried in the speech signal by a number of acoustic cues, including duration, intensity and pitch. Importantly for language acquisition, prosody could help infants learn words and word order in their native language. Therefore, we studied a mechanism that could support these early prosodic abilities: the iambic/trochaic law (ITL).

The ITL (Woodrow, 1909; Hayes, 1995; Nespor et al., 2008) is a mechanism that organizes auditory perception and was proposed to have an important role not only in adult speech perception but also in language acquisition in infancy. The ITL states that sounds (e.g., syllables, musical notes, etc.) contrasting in intensity/pitch form pairs with initial prominence, i.e., a trochaic pattern (strong-weak or high-low), and those contrasting in duration form pairs with final prominence, i.e., an iambic pattern (short-long). However, languages differ in how these acoustic cues mark prosodic prominence both at the level of words and of phonological phrases. For example, French has no lexical stress but has phrase-final stress, the last syllable of the phrase being lengthened, creating a short-long pattern (duration-based iambic pattern). Conversely, in English or in German, lexical stress is usually on the first syllable, which has higher intensity and/or pitch (intensity- or pitch-based trochaic pattern). Listeners' language background is therefore likely to interact with the ITL bias.

This thesis is divided into two main parts. First, in a segmentation / syllable pair recognition task, we found that sensitivity to the ITL was present in French and German adults (Exp. 1) and 7.5-month-old infants (Exp. 2). We found weak cross-linguistic differences between the two language groups for the intensity grouping in adults and infants. Secondly, using NIRS (Near InfraRed Spectroscopy), we measured cortical responses in newborns and demonstrated that sensitivity to the ITL was present at birth and were already influenced by the language the infants heard in utero (Exp. 3-6). Finally, we observed how language background influences the ability to discriminate lexical stress patterns in 10-month-olds (Exp. 7). Our findings show that bilingual infants simultaneously learning French and a language with variable lexical stress were able to discriminate stress patterns whereas monolingual French infants could not.

Taken together, our results contribute to a better understanding of the developmental origin of the ITL, its modulation by linguistic experience, and language-specific processing and rhythmic preferences.

Mots clefs : language perception, prosody, language acquisition, infants, newborns, bilingualism, NIRS

Liste des publications

Abboub N., Bijeljac-Babic R., Serres J & Nazzi T. (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, 132, 111-120.

Abboub N., Nazzi T. & Gervain J. (soumis). Rhythmic grouping at birth.

Remerciements

Il est loin d'être usuel d'avoir la possibilité de remercier toutes les personnes qui ont permis que ce projet voit le jour, alors je vais tâcher de n'oublier personne.

Dans un premier temps, j'aimerais remercier mes deux directeurs de thèse d'avoir été si complémentaires pendant ces quatre années, d'avoir relu et commenté ce manuscrit autant de fois. Je voudrais tout d'abord remercier Judit, celle qui m'a ouvert les portes de ce laboratoire, merci de m'avoir transmis cette passion des nouveau-nés, de m'avoir donné beaucoup d'autonomie dans mes projets, de m'avoir soutenue, rassurée et fait confiance. Tu m'as tout au long de ces quatre ans donné la possibilité de m'épanouir dans mes projets de recherches. Merci d'avoir toujours été emballée par nos résultats de ces tout petits bébés et d'être toujours aussi enthousiaste ! Merci aussi pour ta disponibilité (à toute heure du jour et de la nuit ;), ta patience, ton dynamisme et pour ta positivité qui te caractérise si bien. Merci pour ces bons moments passés en dehors du labo, notamment ceux de Trieste qui feront partie des meilleures conférences que j'ai faites. Je suis très fière d'avoir été une de tes premières thésardes !! Si Judit m'a ouvert les portes, Thierry, tu m'as donné toutes les clefs. Je ne pourrais probablement pas décrire avec les bons mots de tout ce que tu m'as apporté pendant ces quatre ans. Je mesure quelle chance j'ai eu de t'avoir comme directeur de thèse, peu de gens sont aussi disponibles, accessibles, et aussi à l'écoute que tu ne l'as été avec moi, une neurobiologiste un peu perdue dans un laboratoire de psycholinguistique. Tu as cette rare capacité à guider sans imposer, à questionner sans mettre en difficulté, à être exigeant mais tolérant, à être soucieux de tes thésards et avoir toujours cette touche d'humour qui fait toute la différence. Merci mille fois pour toutes ces discussions et d'avoir pris autant de temps à me transmettre ta passion pour la psycholinguistique et de ces détails si importants.

De ces 4 années de thèse à vos côtés, je ne vous remercierai jamais assez pour m'avoir tant apporté et j'ai pris énormément de plaisir à travailler avec vous deux !

J'aimerais également remercier Marina Nespor, Fabrice Wallois et Hellmuth Obrig d'avoir accepté de faire partie de mon jury de thèse. Je suis très heureuse et honorée de votre présence. Merci d'avoir pris le temps de lire et commenter ce travail ainsi que de votre participation à ma soutenance.

J'aimerais remercier tous les membres du LPP, tout particulièrement l'équipe Parole et celle du labo bébé dont l'atmosphère de travail a largement contribué à la réussite de mes projets. Merci à Viviane et Sylvie pour leur efficacité et leur grand soutien (merci pour ces centaines de listes de bébés). Merci à Lucie d'avoir été d'une si grande aide les samedis bondés. Merci à Lionel d'avoir été une aide aussi précieuse et efficace. J'aimerais remercier Josette pour m'avoir transmis sa passion pour les tout-petits, merci également à Ranka, la petite maman de notre équipe. Merci de nous avoir chouchoutés avec tes gâteaux et tes bons petits plats. Et merci pour cette collaboration qui m'a permis de publier un joli papier :) Merci à Anjali pour m'avoir aidée et soutenue dans mes premiers pas de thésarde, étudiant cette loi au nom un peu « barbare ». Merci d'avoir été si tolérante avec moi pour mon anglais si rudimentaire. Merci à Henny pour sa gentillesse et ses retours toujours pertinents. Merci à Silvana et Arnaud pour leur humour et bonne humeur !

I am also very grateful to Barbara Höhle and Natalie for this wonderful collaboration!

Et ENFIN, Merci à la TEAM 608, la meilleure du monde, je ne sais pas comment j'aurais fait sans vous, merci de votre soutien, de votre bonne humeur (ou pas ;)), des samedis plein à craquer au labo bébé, des conférences, des soirées raclettes, maki et tant d'autres !! Je tiens à remercier tout d'abord Camillia avec qui l'aventure NIRS a commencé (et les galères de l'hôpital avec :)), celle qui m'a mis dans le bain de la recherche dès mon arrivée au LPP, merci d'avoir pris ce temps pour moi. Merci à Laurianne, notre wonderwoman du labo, celle qui gère tout, sans que l'on comprenne comment, merci pour tous tes conseils, ton aide, ta disponibilité. Merci d'avoir relu des bouts de cette thèse, avec bienveillance. Tu feras une carrière brillante j'en suis convaincue !! Merci à Léo pour m'avoir fait découvrir la culture japonaise, finalement pas si éloignée de la culture kabyle ;). Merci pour ta disponibilité et pour avoir été le seul à remarquer mes nouvelles chaussures/sacs/robes/bijoux. Un remerciement tout particulier à mon âme-sœur de l'ITL, Carline, mon GPS-GFP (GPS-Gluten Free Poisseuse) celle qui a partagé avec moi l'amour de la loi iambo-trochaïque. Merci pour tous ces bons moments passés à tes côtés, merci pour ton rire si, hum, spécial et pour ce voyage à NYC pré-rédaction qui nous a fait un bien fou. Merci pour être un génie de word, un jour tu m'expliqueras ;) Et enfin ma loulou, merci pour être toujours en décalage (Hein ??), pour tout ton soutien, toujours rassurante et disponible, merci d'avoir été là pendant ces 4 ans surtout ces derniers mois, rien n'aurait été pareil sans toi et merci MERCI de n'avoir jamais utilisé ta batte contre moi ;). Une pensée et un grand merci aux anciennes Nayeli, LRF et aux dernières arrivées Cécile et Elena de votre soutien et bonne humeur, une très bonne relève de la 608 ((YEAH BABY) ;))

Merci à ceux qui sont loin mais qui restent si proches. Une pensée à Abdelmalek Sayad à qui je dois ma première rencontre avec le monde de la recherche et des corrections incessantes. A ma Lola, ma parisienne exilée en Alberta, malgré la distance, on a réussi tu vois :)), merci pour tes mots doux et ton soutien. Merci à Catherine pour tes conseils toujours si avisés et ton humour qui te caractérise bien. Merci à Anne pour ton univers si riche et pour cette bonne humeur qui ne te quitte jamais !

Merci à Alexandre pour tout ton soutien, tous ces beaux/bons moments de complicité et ce super we à NYC.

J'aimerais aussi remercier la Saint Charles pour m'avoir faite rencontrer des gens merveilleux, mon coach Bertrand, JB, mon hamster, Julie, mon rayon de soleil, Thomas, mon hipster incompris, Floriane, l'irremplaçable, Innée, la chercheuse de soleil éthiopienne, et bien d'autres. Une pensée pour Marie qui a relu entièrement cette thèse, d'avoir été là, merci pour tous ces bons moments passés ensemble et ton écoute si précieuse.

Et enfin j'aimerais remercier le trio et maintenant le quatuor de choc, celui qui m'a permis de m'évader de ma thèse et qui a subi toutes mes NIMBUSADES !! Je ne sais pas comment je pourrais vous remercier de tous ces moments aussi mémorables les uns que les autres qui ont rythmé ces dernières années ! Emilie, ma deuxième grande sœur, merci d'être aussi têtue que moi, aussi disponible et d'être une créatrice d'expression hors norme ! Merci de m'avoir fait découvrir cette merveilleuse chanson des corons ma Billy Joe Picarde ;) Merci à Alexandra pour ton répertoire aussi illimité de blagues, tu me surprendras toujours. Merci de tout votre soutien les girls et merci de m'avoir supportée quand j'étais si insupportable !

Merci à mon père pour m'avoir permise de m'évader de ma thèse avec ces expositions magnifiques, théâtres et ces jours passés à la maison de Corcelles dans notre région d'adoption. Merci pour ces bons petits plats et m'avoir chouchoutée l'été dernier.

Des milliards de merci à ma sœur Schéhérazade, mon double, ma meilleure amie, ma Miss Tick, celle qui m'a soutenue, même quand je ne le méritais pas. Je ne sais pas si je pourrais un jour trouver les bons mots pour te dire combien je t'aime... Merci d'exister et d'avoir été là dans tous ces merveilleux moments et aussi dans les plus durs. Merci d'avoir été si tolérante et patiente avec moi ces derniers mois. Merci d'avoir toujours les mots pour me faire rire et d'être si forte. Je suis tellement fière d'avoir une sœur aussi merveilleuse que toi !

Merci à mon petit frère Hicham pour tout, mais vraiment tout. Merci pour ton soutien, ton aide, ton intelligence, de m'avoir supportée ces derniers mois et d'avoir tout fait. Merci d'avoir été si patient et disponible et d'avoir fait preuve d'autant de maturité quand j'en manquais désespérément. Merci pour tous tes précieux conseils. Je t'aime tellement fort mon fakir !! Je suis extrêmement fière de l'homme que tu es devenu, et que tu es en passe de devenir.

Et je garde ce dernier paragraphe pour celle sans qui rien n'aurait été possible : ma mère, mon soleil. Merci de nous avoir transmis ta force, ton rire et ces valeurs de travail, courage et de pugnacité. Merci d'avoir été là dans tous ces moments difficiles mais aussi pour tous les bons. Merci pour toute l'énergie et le temps que tu as laissé dans notre éducation. Merci pour ton soutien indéfectible même quand tu étais épuisée. Aucun mot ne pourrait décrire tout l'amour que tu nous as donné et d'avoir créé une famille aussi unie. Je ne mesure probablement pas encore tout ce que tu nous as transmis mais je sais qu'aucune maman n'est aussi merveilleuse et courageuse que toi, un modèle difficilement égalable ! Une dernière chose : merci d'avoir toujours raison Monkey :)))

TABLE DES MATIERES

1 INTRODUCTION GENERALE	1
1.1 Aperçu des étapes précoces de l'acquisition du langage : la spécialisation pour la langue maternelle.....	6
1.1.1 Développement de la perception phonétique.....	6
1.1.2 Développement de la perception prosodique.....	7
1.2 Organisation de la prosodie linguistique	9
1.2.1 Définition.....	9
1.2.2 Le rôle de l'information prosodique.....	11
1.2.3 Hiérarchie prosodique.....	12
1.3 Un mécanisme d'apprentissage puissant : le <i>prosodic bootstrapping</i>	14
1.3.1 Le <i>prosodic bootstrapping</i>	16
1.3.2 Quelles sont les sensibilités précoces qui soutiennent l'hypothèse du <i>prosodic bootstrapping</i> ?.....	17
1.4 Acquisition des propriétés prosodiques et leur rôle dans le <i>bootstrapping</i> du vocabulaire et de la grammaire	27
1.4.1 Le problème de la parole continue.....	27
1.4.2 Le rôle de l'accentuation lexicale : trouver et représenter les mots	28
1.4.3 Le rôle de la proéminence dans les syntagmes : trouver l'ordre des mots.....	33
1.5 La loi iambo-trochaïque (ITL).....	38
1.5.1 Définition.....	38
1.5.2 Modulation par l'environnement linguistique.....	39
1.5.3 L'ITL chez l'adulte	40
1.5.4 L'ITL chez les nourrissons.....	41
1.6 Objectif et structure de la thèse	43
2 ITL et segmentation chez l'adulte et le nourrisson	45
2.1 CHAPITRE 1 : Comment l'ITL guide le processus de segmentation chez l'adulte ?....	46
2.1.1 Introduction.....	46
2.1.2 Expérience 1.....	53
2.1.3 Points clefs.....	76
2.2 CHAPITRE 2 : Comment l'ITL guide le processus de segmentation chez le nourrisson de 7,5 mois ?.....	77
2.2.1 Introduction.....	77
2.2.2 Expérience 2.....	81
2.2.3 Points clefs.....	94
2.3 Conclusion.....	95
3 ITL et préférences de patterns rythmiques chez le nourrisson.....	97
3.1 CHAPITRE 3 : L'ITL à la naissance.....	98
3.1.1 Introduction : Pourquoi utiliser l'imagerie optique pour tester l'ITL?.....	98
3.1.2 Les origines de la loi iambo-trochaïque.....	108
3.1.3 Expérience 3 : Contraste de durée chez des nouveau-nés monolingues français..	113
3.1.4 Expérience 4 : Contraste d'intensité chez des nouveau-nés monolingues français	120
3.1.5 Expérience 5 : Contraste de pitch chez des nouveau-nés monolingues français....	123
3.1.6 Expérience 6 : Contraste de pitch chez les nouveau-nés bilingues français-autre langue	126
3.1.7 Discussion.....	129
3.1.8 Points clefs.....	135

3.2	CHAPITRE 4 : Chez des nourrissons français monolingues ou bilingues de 10 mois	136
3.2.1	<i>Introduction</i>	137
3.2.2	<i>Expérience 7</i>	141
3.2.3	<i>Discussion</i>	148
3.2.4	<i>Points clefs</i>	152
4	DISCUSSION GENERALE	153
4.1	Résumé des résultats	155
4.1.1	<i>L'ITL guide le processus de segmentation chez l'adulte et le nourrisson</i>	155
4.1.2	<i>L'ITL et préférences de pattern rythmique chez le nourrisson</i>	156
4.2	Les limites dans la comparaison de nos études	158
4.2.1	<i>Une bonne méthode pour comparer adultes et nourrissons ?</i>	158
4.2.2	<i>De faibles différences cross-linguistiques ?</i>	159
4.2.3	<i>Quel niveau prosodique module l'ITL ?</i>	159
4.2.4	<i>L'impact du matériel et de la méthode sur les préférences rythmiques</i>	161
4.2.5	<i>Capacité auditive et seuil de discrimination</i>	161
4.2.6	<i>Saillance acoustique</i>	162
4.3	L'ITL : vers une théorie unifiée	163
4.3.1	<i>Durée</i>	163
4.3.2	<i>Intensité</i>	164
4.3.3	<i>Pitch</i>	166
4.4	Mécanismes sous tendant l'ITL	167
4.4.1	<i>Un héritage évolutif : contraintes perceptives</i>	167
4.4.2	<i>Des régularités prosodiques</i>	168
4.5	Bases cérébrales de l'ITL	170
4.5.1	<i>Partage de compétence</i>	170
4.5.2	<i>Les compétences des nouveau-nés bilingues : le rôle de l'expérience prénatale</i>	172
4.6	Perspectives	174
4.6.1	<i>Interaction des indices</i>	174
4.6.2	<i>Les valeurs de différences</i>	174
4.6.3	<i>Le choix des langues dans les études cross-linguistiques</i>	175
4.6.4	<i>Privation sensorielle précoce : surdité et implant cochléaire</i>	175
4.7	Conclusions	176
	Références	177
	Annexes	196
	Annexe 1 : Étude Pilote Expérience 2	197
	Annexe 2 : Pré-analyse Expérience 2	199
	Annexe 3 : Abboub, Bijeljac-Babic, Serres & Nazzi (2015)	200
	Annexe 4 : Abboub, Nazzi & Gervain (manuscrit soumis)	Erreur ! Signet non défini.

TABLE DES FIGURES

<i>Figure 1.1 : La chronologie du développement de la perception de la parole. Cette figure représente les changements majeurs qui apparaissent dans la perception de la parole chez des nourrissons, typiques pendant la première année de vie. (d'après Kuhl, 2004).</i>	8
<i>Figure 1.2: Informations prosodiques divisées par informations acoustiques (signal) et informations phonético-phonologiques (perception).</i>	10
<i>Figure 1.3: Hiérarchie prosodique selon Nespors et Vogel (1989), du rang le plus bas au plus élevé.</i>	13
<i>Figure 1.4 : Procédure de segmentation : (a) le signal de parole ne présente pas de frontières de mots apparentes, (b) alors qu'à l'écrit des espaces sont présents entre les mots (d'après Kuhl, 2004).</i>	27
<i>Figure 1.5 : Schéma de la loi iambo-trochaïque pour les trois indices acoustiques (durée, intensité, pitch).</i>	39
<i>Figure 2.1: Un extrait de la phase de familiarisation de l'étude de Bion et al., 2011, pour les trois conditions (a) condition contrôle, (b) condition pitch, (c), condition durée</i>	51
<i>Figure 2.2 : Répartition des participants dans les 4 conditions (durée, intensité, pitch et contrôle) avec le schéma des deux type de rampe pour chaque langue (L1). L'augmentation graduelle de la différence de durée, intensité et pitch commence soit sur la troisième syllabe (position initiale de la deuxième paire) ou soit sur la quatrième syllabe (position finale de la deuxième paire).</i>	58
<i>Figure 2.3 : Prédiction de la loi Iambo-trochaïque. (1) : condition durée : groupement iambique ; (2) : condition intensité et (3) : condition pitch (4) : groupement trochaïque ; condition contrôle : aucune prédiction de groupement.</i>	60
<i>Figure 2.4 : Score moyen (en %) de réponses "oui", pour les items adjacents (réponse correcte) et non-adjacents (réponse incorrecte) dans les quatre conditions (contrôle, durée, intensité et pitch).</i>	61
<i>Figure 2.5 : Score moyen de bonnes réponses (réponses « oui ») pour les items trochaïques et iambiques dans les conditions durée, intensité et pitch.</i>	63
<i>Figure 2.6 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition durée en fonction du type de rampe (3^{ème} ou 4^{ème} syllabe), séparé par langue maternelle.</i>	65
<i>Figure 2.7 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition intensité en fonction du type de rampe (3^{ème} ou 4^{ème} syllabe), séparés par langue maternelle.</i>	66
<i>Figure 2.8 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition du pitch en fonction du type de rampe.</i>	68
<i>Figure 2.9 : Cabine HPP (à gauche) et box de l'expérimentateur (à droite).</i>	84
<i>Figure 2.10: Temps d'orientation de regard pour les items iambiques et items trochaïques, en fonction de la condition (contrôle, durée, intensité et pitch) et de la langue maternelle des nourrissons (allemand, français).</i>	86
<i>Figure 2.11 Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition durée, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe).</i>	88
<i>Figure 2.12: Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition intensité, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe), puis par groupe de langue maternelle (français et allemand).</i>	89
<i>Figure 2.13 : Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition pitch, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe), puis par groupe de langue maternelle (français et allemand).</i>	90
<i>Figure 3.1 : (a) Propriétés optiques des tissus, (b) Trajectoire de la lumière proche infra-rouge traversant le scalp et le crâne pour atteindre le cerveau. (c) Schéma représentant l'absorption des différentes lumières et les composants de dispersion (les concentrations sont représentées par le degré de couleur) à l'intérieur du tissu (Tachtsidis, 2010).</i>	102
<i>Figure 3.2 : Une réponse réponse hémodynamique typique (HRF) chez l'adulte. La stimulation est délivrée à t=0. La réponse commence souvent avec une légère augmentation en déoxyHb, suivie par une augmentation d'oxyHb, puis par une diminution de concentration en déoxyHb (mesurée ici en unité arbitraire). Le pic du signal apparaît quelques secondes après le début du stimulus, puis revient à la normale (baseline) (Gervain et al., 2011, Figure 3).</i>	103
<i>Figure 3.3 : Design expérimental des Expériences 1-4. Chaque rectangle représente un bloc, contenant une série de 6 paires d'un type de groupement, ces paires étant séparées par de courtes pauses.</i>	115
<i>Figure 3.4 : Photo d'un nouveau-né portant le bonnet avec les optodes placées préférentiellement sur les aires temporo-pariétales. La disposition des optodes insérées dans le bonnet permet d'enregistrer simultanément l'activité des 24 canaux.</i>	116
<i>Figure 3.5 : Photod'un nouveau-né endormi dans son berceau et du dispositif juste avant le début de l'expérimentation, portant le bonnet avec les capteurs. Ces optodes sont reliés par des fibres optiques à la machine NIRSCOUT NIRx.</i>	117

<i>Figure 3.6 : Résultats des grandes moyennes de l'Expérience 3 (Condition durée, monolingues). Les canaux sont placés selon le même placement que sur la Figure 3.10. L'axe des x représente le temps en secondes ; l'axe des y représente les concentrations en mmol × mm. Le rectangle sur l'axe des x représente le temps de stimulation. Les lignes continues rouge et bleu dans le graphique représentent les changements de concentration de l'oxyHb et deoxyHb pour les séquences avec le groupement en désaccord avec l'ITL (trochaïque). Les lignes rouge et bleu avec des tirets représentent les concentrations de l'oxyHb et deoxyHb pour les séquences avec le groupement en accord avec l'ITL (iambique). Les lignes rouge et bleu en pointillés représentent les concentrations de l'oxyHb et deoxyHb pour le groupement sans contraste. *** : p < .001 ; ** : p < .01 ; * : p < .05 ; + : p < .01, sans correction.</i>	119
<i>Figure 3.7 : Résultats des grandes moyennes de l'Expérience 4 (condition intensité, monolingues). Toutes les conventions du graphiques sont identiques à la Figure 3.6.</i>	122
<i>Figure 3.8 : Résultats des grandes moyennes de l'Expérience 5 (condition pitch, monolingues). Toutes les conventions du graphique sont identiques à la Figure 3.6.</i>	125
<i>Figure 3.9 : Résultats des grandes moyennes de l'Expérience 6 (Condition pitch, bilingues). Toutes les conventions du graphique sont identiques à la Figure 6.</i>	128
<i>Figure 3.10 : Carte statistique (valeurs de p) comparant les réponses aux séquences : en accord avec ITL vs baseline, en désaccord avec ITL vs baseline et en accord avec ITL vs désaccord avec ITL. Pour chaque canal, les valeurs de p ont un code couleur : p < .001, bordeaux; p < .01, rouge ; p < .05, orange, p < .1, jaune, sans correction.</i>	130
<i>Figure 3.11: Moyenne des temps d'orientation (et les erreurs standard) pour les patterns d'accentuations familiers et nouveaux, séparés par groupe linguistique : français monolingues, français dominant bilingues (60 à 70% de français) et les français non dominant bilingues (30 à 50% de français), ** p < ,01</i>	148
<i>Figure 4.1 : Diagramme schématisant la relation fonctionnelle entre le traitement linguistique et le traitement syntaxique musical.</i>	172

TABLE DES TABLEAUX

<i>Tableau 2-1 : Détails des informations des participants par groupe.</i>	53
<i>Tableau 2-2 : Valeurs des trois indices acoustiques pour chaque type de séquence de familiarisation et pour les items tests.</i>	55
<i>Tableau 2-3 : Analyse globale sur les 4 conditions (durée, intensité, pitch et contrôle).</i>	62
<i>Tableau 2-4 : Test t sur les 4 conditions (durée, intensité, pitch et contrôle), par rapport au hasard (50%).</i>	62
<i>Tableau 2-5 : Analyse sur les 3 conditions (durée, intensité, et pitch).</i>	64
<i>Tableau 2-6 : Analyse pour la condition intensité, séparée par la langue maternelle.</i>	67
<i>Tableau 2-7 : Récapitulatif des analyses par condition chez tous les participants.</i>	68
<i>Tableau 2-8 : Détails des informations des participants par groupe.</i>	81
<i>Tableau 3-1 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FRD}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0].</i>	120
<i>Tableau 3-2 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FRD}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste].</i>	123
<i>Tableau 3-3 : Résultats des tests de Student pour l'oxyHb etdDeoxy uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FRD}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste].</i>	126
<i>Tableau 3-4 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FRD}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste].</i>	129
<i>Tableau 3-5 : Caractéristiques des nourrissons bilingues (% d'exposition au français, deuxième langue) et performances individuelles (temps d'orientation du regard en secondes pour les patterns d'accentuations familiers ou nouveaux à la phase de test et les scores de différence), en fonction du sous-groupe (français dominant versus français non-dominant).</i>	144
<i>Tableau 3-6 : Analyses acoustiques (durée, intensité et pitch) des voyelles contenues dans les stimuli utilisés dans notre étude (issue de Skoruppa et al. 2009) et celle de Bijeljac-Babic et al. (2012).</i>	145
<i>Tableau 4-1 : Résumé des résultats</i>	157

This task intrigues people because it's about us. But that doesn't make it a scientific question. It may be important for us to know where we came from, but if we can't answer that question scientifically, we can't answer it. If you want to tell stories, well then, tell stories."

Noam Chomsky

1 INTRODUCTION GENERALE

Une des compétences qui distinguent l'humain des autres espèces est la capacité de percevoir et de produire un signal acoustique aussi complexe que celui de la parole. Comme le code génétique, la structure de la parole est hiérarchique, générative, récursive et presque illimitée. Bien que l'ADN soit composé uniquement de 4 types de nucléotides, il est responsable de l'émergence de structures extrêmement complexes et d'une variété d'organismes et d'espèces quasi-illimitée. Le langage aussi possède un nombre limité de sons, mais rend possible une infinité de phrases. La question de l'acquisition du langage chez les humains fait partie des grandes problématiques qui ont été largement étudiées ces dernières décennies (Hauser, Chomsky, & Fitch, 2002).

Le langage est un code complexe où différentes informations sont disponibles, de la sémantique à la pragmatique et de la phonologie à la syntaxe. Il est en effet primordial, pour qu'un être humain puisse maîtriser une langue, de comprendre le sens des mots ainsi que l'organisation de ces mots au sein de la phrase. Cette complexité rend extrêmement difficile la compréhension de l'acquisition du langage. Quelles sont les unités qui transmettent les informations linguistiques pertinentes aux enfants ? Comment décodent-ils ces informations et surtout à partir de quel moment vont-ils y être sensibles ?

La perception de la parole par le nourrisson commence avant même sa naissance. La parole fait partie de l'environnement auditif du fœtus dès que son oreille est assez mature -dans le dernier trimestre de la grossesse- pour analyser ce signal acoustique (Pujol, Lavigne-Rebillard, & Lenoir, 1998). L'acquisition du langage chez l'enfant semble être rapide et réalisée sans apparente difficulté. Il est d'ailleurs de moins en moins rare que certains jeunes enfants grandissent dans des environnements multilingues (Grosjean, 2010). Ces enfants apprennent deux, voire trois langues, simultanément, sans difficultés et en quelques années, à l'inverse d'un adulte apprenant une seconde langue. Cette différence entre les enfants et les adultes a poussé de nombreux scientifiques à essayer de comprendre quels sont les « outils » dont disposeraient les enfants pour apprendre avec autant de facilité leur(s) langue(s) maternelle(s). De puissants mécanismes doivent être à l'origine de l'apprentissage d'un code aussi complexe que le langage. Mais d'où proviennent-ils ? Sont-ils présents dès la naissance ? Se développent-ils pendant les premiers mois de vie ?

Certaines écoles de pensées empiristes s'enracinant dans les écrits d'Aristote dans son traité *de l'âme* (III, 4, § 9, 430 avant JC), ou encore de Locke (1690), ont réduit

à peu de choses les capacités des nouveau-nés. Leurs comportements étaient définis comme de simples réflexes primaires, archaïques, complètement reproductibles à d'autres espèces, ou encore considérés comme aspécifiques à l'espèce humaine. Les capacités initiales des humains ont été ignorées au profit de l'étude, plus facile à mettre en œuvre, du traitement du langage chez des adultes possédant un système plus « mature ». Les nouveau-nés étaient considérés comme sourds et aveugles, ne disposant à la naissance, que de ces fameux réflexes archaïques pour appréhender leur environnement. Ces traditions empiristes prônaient notamment un apprentissage global et universel (Skinner, 1957). A l'opposé, depuis Platon (383 avant JC) et Descartes (1637), des théories plus innéistes ressurgissent sous l'influence de nombreux philosophes, linguistes et généticiens (Chomsky, 1957; Hauser, Chomsky, et al., 2002; Lenneberg, 1967).

Les recherches actuelles ont montré que, non seulement les nouveau-nés sont loin d'être aveugles, et peuvent percevoir des différences très subtiles dans leur environnement visuel (Izard, Sann, Spelke, & Streri, 2009; Pascalis, De Schonen, Morton, Deruelle, & Fabre-Grenet, 1995), mais également loin d'être sourds, capables de discriminer des sons très proches à la naissance, voire même avant (Kisilevsky et al., 2009; Mehler et al., 1988; Nazzi, Bertoncini, & Mehler, 1998; Nazzi, Floccia, & Bertoncini, 1998). De plus, bien que leur cerveau ait une taille bien inférieure à celui d'un adulte, il n'est pas aussi immature qu'il n'y paraît. De très récentes données issues de nouvelles méthodes d'imagerie cérébrale ont montré que les nouveau-nés possèdent une organisation cérébrale proche de celle de l'adulte ainsi que des réseaux de traitement du langage très complexes (Dehaene-Lambertz, Hertz-Pannier, Dubois, & Dehaene, 2008; Dubois et al., 2009, 2015). Ces capacités ont d'ailleurs été retrouvées très récemment chez de grands prématurés (Mahmoudzadeh et al., 2013). Bien loin, donc, de la fameuse *tabula rasa*.

Ce sont, entre autres, les théories chomskyennes (1957), associées à celle de Lenneberg (1967), qui ont persuadé un grand nombre de scientifiques d'explorer la notion que l'esprit humain ou le cerveau serait déjà équipé à la naissance de modules spécialisés pour l'acquisition du langage, appelé par Chomsky le LAD (pour Language Acquisition Device). Chomsky a proposé un cadre particulier, celui de Principe et Paramètres (P&P). Cette théorie nativiste postule que, lors de l'acquisition du langage,

interviennent, d'une part, des mécanismes innés qui guident l'apprentissage s'appliquant à toutes les langues (« principe ») et, d'autre part, des « paramètres » distinguant-guidant les choix entre plusieurs structures possibles. La formulation de cette théorie P&P (1980) a grandement guidé les sciences cognitives dans leur effort de comprendre de quelle manière les humains arrivent à apprendre leur langue maternelle aussi facilement et rapidement, malgré la pauvreté du signal linguistique. En effet, Chomsky part de l'observation que les enfants sont souvent confrontés à des phrases incomplètes, erronées, ou encore ne rencontrent jamais d'exemples négatifs. C'est donc sur une base d'informations partielles et peu fiables que les enfants acquièrent, sans difficulté majeure, les propriétés grammaticales de leur langue maternelle. C'est du constat de cette pauvreté du stimulus que Chomsky a proposé une alternative aux théories empiristes (Chomsky, 1965, 1975). P&P propose que les jeunes enfants naissent avec un savoir universel, notamment de la grammaire (UG, *Universal Grammar*). D'ailleurs, Chomsky écrit récemment (2008):

“The [P&P] approach suggests a framework for understanding how essential unity might yield the appearance of the limitless diversity that was assumed not long ago for language (as for biological organisms generally)... the approach suggests that what emerged, fairly suddenly, was the generative procedure that provides the principles, and that diversity of language results from the fact that the principles do not determine the answers to all questions about language, but leave some questions as open parameters.”

L'approche P&P suggère un cadre pour comprendre de quelle manière l'unité fondamentale pourrait prendre l'apparence de cette diversité illimitée qui a été présumée jusqu'à peu pour le langage (comme pour les organismes biologiques)... Cette approche suggère que ce qui a émergé, assez soudainement, est une procédure générative fournissant des principes, et que la diversité des langues résulte du fait que ces principes ne fournissent pas toutes les réponses à toutes les questions à propos du langage, mais laissent certaines questions comme des paramètres ouverts.

Ces propositions ont eu une très grande influence dans le domaine de l'acquisition du langage. Les théoriciens générativistes ont mis en évidence le fait que

l'environnement linguistique est trop appauvri pour que les nourrissons puissent apprendre la grammaire de leur langue maternelle par un mécanisme d'apprentissage classique. Ces théories générativistes, à l'inverse des théories classiques (Skinner, 1957) prônant un apprentissage non spécifique, peuvent expliquer pourquoi seulement les humains sont capables d'acquérir la grammaire. Par ailleurs, comment expliquer que les enfants acquièrent la grammaire, et plus généralement les propriétés de leur langue maternelle, avec la même rapidité quelle que soit la langue apprise, sachant qu'à l'heure actuelle plus de 6000 langues (Crystal, 1997) sont parlées à travers le monde ? En effet vers 6-7 ans, les enfants maîtrisent la quasi-totalité des règles de leur langue maternelle et peuvent produire et comprendre la plupart des structures langagières.

Si l'espèce humaine possède des mécanismes spécifiques pour apprendre le langage, quels sont-ils ? Comment ces mécanismes se développent-ils et comment interagissent-ils les uns avec les autres et dans quel ordre interviennent-ils ?

Les études expérimentales sur le développement du langage ont mis en évidence un nombre considérable de capacités langagières très précoces, dont certaines seraient déjà présentes à la naissance. Par ailleurs, suite à l'émergence de cette théorie, de nombreuses régularités présentes dans les langues ont pu être mises en évidence. Ces observations suggèrent donc que des mécanismes généraux sont impliqués dans l'apprentissage du langage. Dans cette thèse, nous allons justement contribuer à l'effort de comprendre quels sont les mécanismes dont pourraient disposer les nourrissons à la naissance pour apprendre leur(s) langue(s) maternelle(s).

1.1 Aperçu des étapes précoces de l'acquisition du langage : la spécialisation pour la langue maternelle

Apprendre une langue nécessite de déterminer quels sont les indices importants, contenus dans le signal de parole, qui véhiculent une information linguistique. De nombreuses recherches ont montré que, dans les premiers mois qui suivent la naissance des nourrissons, une spécialisation pour les sons de la langue apprise émerge.

1.1.1 Développement de la perception phonétique

Avant de se spécialiser pour les sons de leur langue maternelle, ils sont capables, dès la naissance, de discriminer catégoriellement une grande variété de phonèmes présents dans les langues du monde (Cheour-Luhtanen et al., 1995; Dehaene-Lambertz & Pena, 2001; Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Toutefois, chaque langue du monde utilise un inventaire de sons spécifiques. Les jeunes nourrissons vont, sous l'effet de l'expérience linguistique, maintenir (et développer) des sensibilités pour les contrastes utilisés dans leur langue maternelle et montrer un déclin de ces sensibilités pour des contrastes non-natifs. Ce phénomène de spécialisation pour les phonèmes appartenant spécifiquement à la langue apprise a été défini comme l'adaptation perceptuelle. Une des premières études qui a étudié ce phénomène a montré que des nourrissons de 6/8 mois exposés à l'anglais discriminaient deux types de « d » qui, en hindi, mais pas en anglais, permettent de contraster le sens des mots (Werker, Gilbert, Humphrey, & Tees, 1981). En revanche, les nourrissons de 10/12 mois exposés à l'anglais n'étaient plus capables de discriminer ce contraste, alors que les nourrissons exposés à l'hindi maintenaient cette sensibilité (Werker & Tees, 1984). De nombreuses autres recherches ont mis en évidence des phénomènes similaires chez des nourrissons exposés à d'autres langues avec des pertes de sensibilité à des contrastes non natifs et des gains de sensibilité pour des contrastes natifs (Kuhl et al., 2006; Maurer & Werker, 2014; Narayan, Werker, & Beddor, 2010). A la naissance, les nourrissons sont donc capables de discriminer de nombreux contrastes phonétiques, mais, dans la deuxième partie de la première année, l'effet de l'expérience langagière va moduler la sensibilité à ces distinctions phonétiques.

Néanmoins, les nourrissons ne doivent pas apprendre uniquement quels sont les phonèmes utilisés pour créer un contraste, mais doivent également déterminer le rôle

des informations suprasegmentales de leur langue maternelle. Et ce sont ces informations prosodiques qui vont nous intéresser dans cette thèse.

1.1.2 Développement de la perception prosodique

Les enfants doivent apprendre les propriétés sonores ou acoustiques de leur langue maternelle au-delà de l'identité du phonème ; par exemple, les tons lexicaux, la longueur des voyelles et l'accentuation peuvent être utilisés pour distinguer des items lexicaux (Cutler & Norris, 1988). A la naissance, les nourrissons semblent avoir commencé à apprendre certaines de ces propriétés: les nouveau-nés préfèrent écouter de la parole plutôt que des sons non-linguistiques (Vouloumanos & Werker, 2007), la voix de leur mère plutôt que celle d'une autre femme (Decasper & Fifer, 1980; Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld, 1978), ainsi que leur langue maternelle plutôt qu'une langue étrangère (Byers-Heinlein, Burns, & Werker, 2010; Mehler et al., 1988; Moon, Panneton Cooper, & Fifer, 1993). Les nouveau-nés sont donc capables de discriminer des sons sur la base de leur prosodie.

Des recherches se sont aussi intéressées à la trajectoire développementale de la perception de ces structures suprasegmentales, par exemple celle réalisée par Mattock & Burnham (2006). Ces chercheurs ont observé les capacités de discrimination de contrastes de tons lexicaux thaï ou de leurs analogues non-linguistiques chez des nourrissons entre le 6^{ème} et 9^{ème} mois de vie, exposés à l'anglais ou au chinois. Pour les tons lexicaux, les résultats montrent que les capacités de discrimination ont commencé à décliner au 9^{ème} mois chez les nourrissons exposés à l'anglais, alors que la discrimination reste stable chez les nourrissons exposés au chinois. Pour les analogues non-linguistiques, les capacités de discrimination restent constantes entre 6 et 9 mois dans les deux groupes linguistiques. Ces résultats suggèrent que les capacités de discrimination de contrastes tonals déclinent si cette dimension prosodique n'est pas utilisée dans la langue maternelle. Une étude similaire, mais explorant cette fois-ci la perception d'un contraste de longueur de voyelles, qui est un contraste phonémique en japonais, a également montré un gain de sensibilité chez des nourrissons japonais (Sato, Sogabe, & Mazuka, 2010b). Ces auteurs rapportent que des nourrissons entre 4 et 7 mois de vie ne sont pas capables de discriminer des voyelles courtes des voyelles longues, mais dès 9 mois cette capacité semble émerger.

Prises ensemble, ces recherches montrent que pendant la première année de vie, l'expérience linguistique façonne la perception des sons à différents niveaux et celle-ci devient spécifique à la langue maternelle (Figure 1.1).

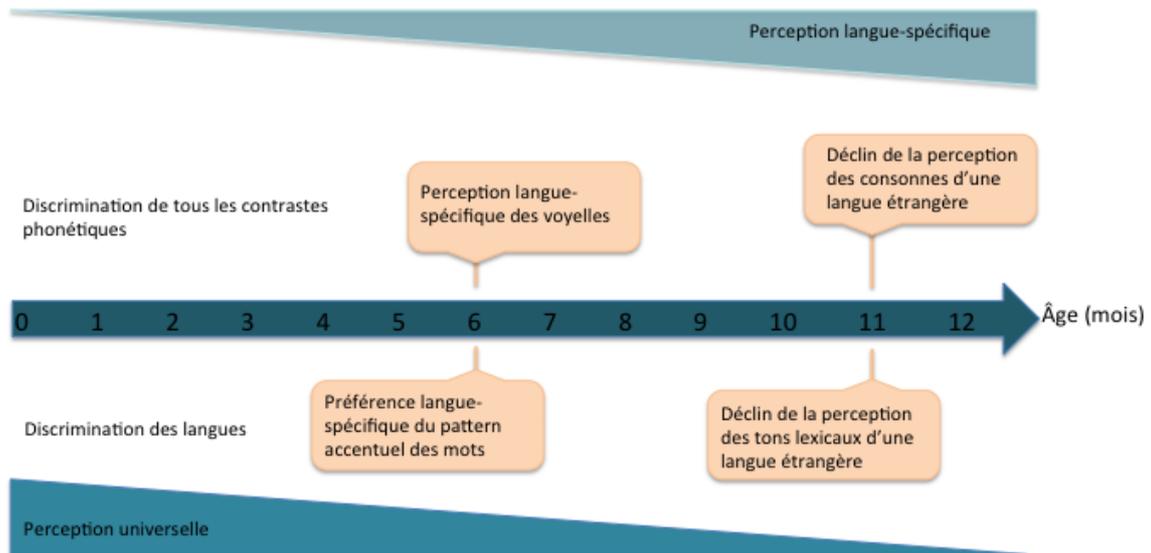


Figure 1.1 : La chronologie du développement de la perception de la parole. Cette figure représente les changements majeurs qui apparaissent dans la perception de la parole chez des nourrissons, typiques pendant la première année de vie. (d'après Kuhl, 2004).

Dans cette thèse, nous allons voir que les indices suprasegmentaux ont un rôle majeur dans l'apprentissage du langage. Par exemple, dans certaines langues comme l'anglais, les adultes comme les nourrissons se réfèrent à la syllabe accentuée pour localiser les frontières de mots ou pour en comprendre le sens, ce qui est un pré-requis pour l'acquisition du vocabulaire (Cooper, Cutler, & Wales, 2002; Cutler, Mehler, Norris, & Segui, 1986; Small, Simon, & Goldberg, 1988; Soto Faraco, Sebastián, & Cutler, 2001). Nous allons également décrire l'importance du rôle de ces indices pour apprendre l'ordre basique des mots, notamment en permettant aux enfants d'acquérir le paramètre de l'ordre de la tête (Christophe, Nespors, Teresa Guasti, & Van Ooyen, 2003; Mazuka, 1996; Nespors, Guasti, & Christophe, 1996). C'est pourquoi, dans cette thèse, nous allons nous intéresser exclusivement à la prosodie linguistique, puisqu'elle contient des indices acoustiques pouvant guider les nourrissons dans l'apprentissage des mots ou encore de la grammaire de leur langue maternelle. Nous allons nous intéresser aux mécanismes qui pourraient sous-tendre ces capacités.

Pour comprendre dans quel cadre s'insèrent nos questions de recherche, nous allons dans ce chapitre introductif définir, dans un premier temps, le terme de prosodie linguistique, en prenant en compte différentes considérations théoriques. Ensuite, nous introduirons la littérature sur le rôle de la prosodie dans le traitement du langage, notamment en ce qui concerne les capacités précoces de traitement, puis l'acquisition des propriétés prosodiques de la langue maternelle. Nous finirons sur la description d'un phénomène perceptuel qui pourrait sous-tendre ces mécanismes d'apprentissage : la loi iambo-trochaïque (ITL).

1.2 Organisation de la prosodie linguistique

La prosodie est un aspect clef du langage parlé. Elle possède une multitude de fonctions dans la communication quotidienne, puisqu'elle permet notamment de déterminer le sens d'un mot, de signaler le type d'intention de communication, tout comme lors de l'écoute ou de la lecture de cette phrase : « Et si on mangeait les enfants ? » par rapport à celle-ci : « Et si on mangeait, les enfants ? » ; la pause à l'oral, représentée par la virgule à l'écrit, permet de désambiguïser cette question surprenante dans le premier cas. La prosodie est définie par des variations de durée, d'intensité ou encore de hauteur (pitch) des éléments. Depuis quelques dizaines d'années, la prosodie, et notamment son rôle dans le développement du langage, a longuement été étudiée, au départ par les phonologues, puis par des psycholinguistes, psychologues et, plus récemment, par des neuroscientifiques. Dans le paragraphe suivant, nous allons décrire la prosodie présente dans le langage parlé, en décrivant dans un premier temps les propriétés acoustiques et phonologiques qui définissent la structure prosodique du langage. Ensuite nous développerons son rôle dans l'acquisition du langage.

1.2.1 Définition

Du grec *pros-ôdia* (pros : en cela ; ode : chant) et traduit littéralement par le latin *ac-centus* (accent tonique), la prosodie fait référence à la rythmique et aux propriétés mélodiques du langage parlé. La prosodie est une branche de la phonologie qui concerne l'étude des unités linguistiques à un niveau suprasegmental ; en d'autres termes, des unités plus grandes que le phonème (encore appelé segment). La parole n'est pas produite de manière isochrone et monotone ; les mots dans les phrases sont

prononcés selon un rythme, une intonation, une accentuation, et sont regroupés dans des unités prosodiques.

Les informations prosodiques du signal de parole d'une langue donnée peuvent être décrites selon ses caractéristiques acoustiques et phonologiques, reflétant respectivement ses propriétés physiques et ses représentations internes (Warren, 1999).

Le langage parlé est modulé selon trois paramètres physiques principaux : la fréquence fondamentale (perçue comme la hauteur ou le pitch), la durée ou l'extension dans le temps des sons (perçue comme la longueur des segments et pauses) et l'amplitude des sons (perçue comme la force ou l'intensité). Voir Figure 1.2.

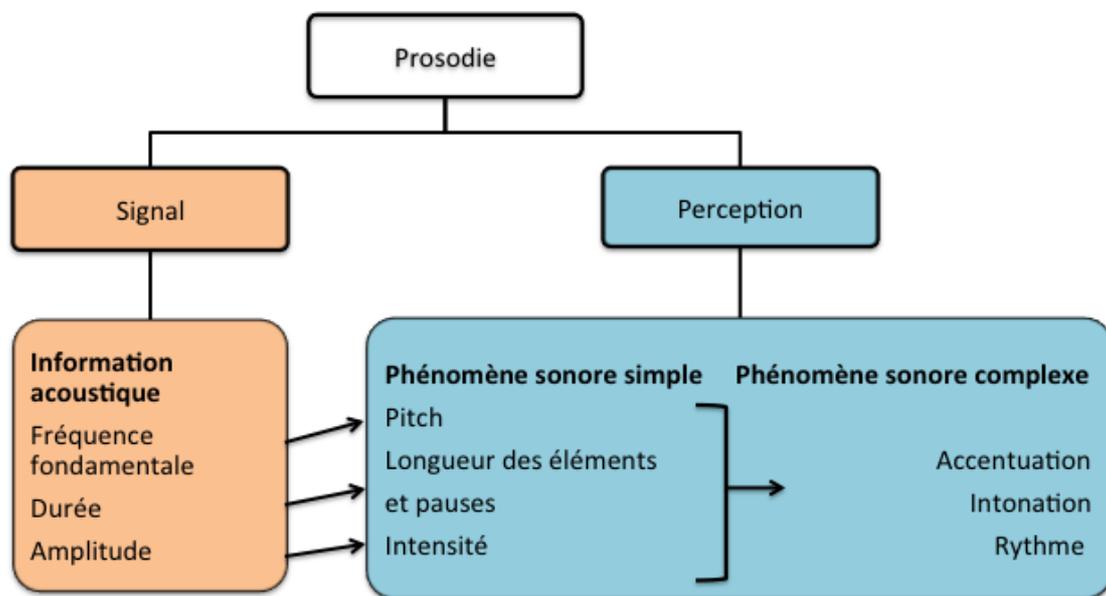


Figure 1.2: Informations prosodiques divisées par informations acoustiques (signal) et informations phonético-phonologiques (perception).

Chez le locuteur adulte, les processus permettant la transition du niveau acoustique vers celui de la perception sont loin d'être triviaux. Premièrement, il n'y a pas de lien linéaire entre les indices acoustiques et la fonction phonologique (Lehiste, 1970). Par exemple, une valeur de pitch de 200Hz va être perçue comme un son aigu ou haut, si la phrase est prononcée par la voix d'un homme, alors qu'elle sera perçue comme étant grave si elle est prononcée par un enfant. De plus, la durée perçue d'un mot ne dépend pas seulement de la véritable longueur de celui-ci, mais aussi de sa position dans la phrase et également de la vitesse de débit de la parole. La perception de ces indices acoustiques par le système perceptuel humain est donc déterminée par divers

facteurs comme le contexte ou par l'interaction avec de nombreux autres indices acoustiques.

1.2.2 Le rôle de l'information prosodique

La combinaison de ces caractéristiques acoustiques simples du signal sonore va être à l'origine de la perception de phénomènes acoustiques complexes comme l'intonation, l'accentuation et le rythme. Ces phénomènes sonores ont des fonctions linguistiques (Crystal, 1997), comme par exemple spécifier le mode de la phrase (question, demande, affirmation), définir une information de structure (accentuation d'une information importante dans une phrase) et marquer les structures syntaxiques. Par ailleurs, le même indice acoustique peut marquer différents phénomènes phonologiques ; par exemple, les changements de pitch contribuent au contour intonatif de la phrase, mais ils ont également un rôle crucial lorsque l'on veut accentuer une information.

Dans certaines langues, si le locuteur veut accentuer une information dans une phrase (accentuation du syntagme), il utilise une augmentation d'intensité, mais aussi de pitch (par exemple en anglais « *Who likes Mary? **Bill** likes Mary* »). L'accentuation peut aussi porter une information linguistique au niveau du mot. En effet, dans certaines langues, l'accentuation peut être la seule information disponible qui permette la différenciation dans la signification d'un mot (pour une revue de la littérature, voir Cutler, 2005). De plus, la présence ou la place de l'accentuation dans un mot peut parfois donner une information à propos de la catégorie syntaxique de celui-ci. En effet, les mots de contenus sont souvent accentués à l'inverse des mots de fonction et, en anglais, les verbes sont souvent accentués sur la seconde syllabe, alors que les noms sont souvent accentués sur la première syllabe (Kelly, 1996). Par exemple, en anglais : DIS¹count : réduction vs disCOUNT : ignorer ; en espagnol : SABana : draps saBAna : savane ; allemand : UMfahren : renverser vs umFAHREN : conduire.

Le phénomène prosodique du rythme ou de la métrique est caractérisé par l'alternance de ces variations d'accentuation (durée, intensité et pitch, Fraise, 1974, 1982). Des études sur la définition de la métrique rythmique des langues du monde ont proposé que le rythme dépende de l'alternance des voyelles et des consonnes dans le signal, notamment par les proportions occupées par les voyelles (Ramus et al. 1999).

¹ Les majuscules représentent les syllabes accentuées.

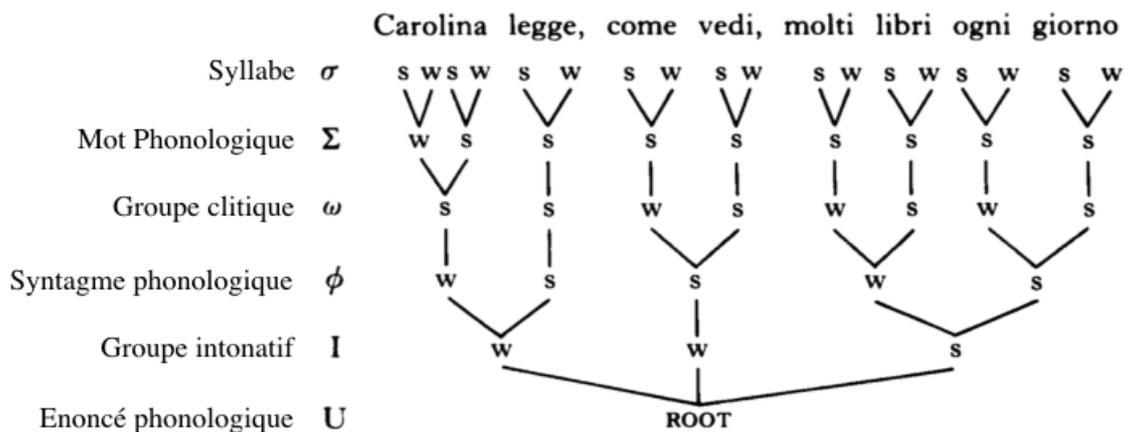
Par exemple, les mots en anglais ou en allemand présentent une alternance de syllabes accentuées (fortes) et non accentuées (faibles), comme c'est le cas dans des mots polysyllabiques : E-lec-TRI-ci-TY et PO-ssi-BI-li-TY. Les langues romanes (français, italien, espagnol) possèdent une alternance basée sur la syllabe (chaque syllabe alterne à un rythme relativement constant). D'autres langues (japonais, tamoul) utilisent la more, qui est une unité sub-syllabique. Une more peut correspondre à une syllabe courte/faible ou à une partie de syllabe ; ainsi, par exemple, le mot Nippon (« Japon ») est composé de deux syllabes, mais de quatre mores ni-p-po-n.

L'alternance de ces patterns d'accentuations, d'intonations ou rythmiques permet de structurer le signal de parole. La parole ne se résume donc pas à une séquence de mots produite linéairement et de manière monotone. D'ailleurs, celle-ci est organisée selon des domaines prosodiques particuliers qui lient les unités fondamentales du langage parlé entre elles (Nespor et al., 2008; Nespor & Vogel, 1986; Selkirk, 1984).

1.2.3 Hiérarchie prosodique

La plupart des théories phonologiques de la structure prosodique sont issues de la *strict layering hypothesis*, premièrement suggérée par Pierrehumbert (1980) et employée ensuite par de nombreux théoriciens (Hayes, 1989; Nespor & Vogel, 1986; Selkirk, 1984). La phonologie prosodique s'intéresse à l'organisation de la parole au niveau suprasegmental. Selon la *strict layering hypothesis*, les catégories ou domaines prosodiques sont organisés hiérarchiquement en couche, avec les domaines de haut niveau divisés exhaustivement en domaines de plus bas-niveau. Longuement décrits par Nespor et Vogel (1986), les éléments suprasegmentaux forment une hiérarchie organisée en couches, établissant un lien étroit entre l'organisation de la phrase et le signal acoustique de la parole. Ces auteurs postulent que chacun de ces niveaux est caractérisé par la proéminence prosodique relative des constituants. Chaque constituant qui porte la proéminence est défini comme fort et les autres composants sont définis, par contraste, comme faibles. Ces éléments forts reçoivent l'accentuation primaire (*primary stress*) et les éléments faibles sont soit non-accentués, soit porteurs d'une accentuation secondaire (*secondary stress*). A chaque niveau de la hiérarchie prosodique, des patterns prosodiques sont récurrents. La proéminence peut être en position initiale ou finale et va être influencée par les besoins des informations structurales, pour indiquer l'intérêt de ce constituant.

Ces unités prosodiques se définissent les unes par rapport aux autres, comme des niveaux emboîtés. Nespor et Vogel (1986) ont proposé 7 niveaux - du plus grand au plus petit- ou du plus saillant au moins saillant (voir Figure 1.3).



« Caroline lit, comme tu vois, beaucoup de livres par jour »

Figure 1.3: Hiérarchie prosodique selon Nespor et Vogel (1989), du rang le plus bas au plus élevé.

Suivant cette hiérarchie, chaque élément d'un niveau est composé d'au moins un élément du niveau inférieur. *L'énoncé phonologique* est l'unité la plus grande : « Apparemment, les petits enfants sont allés à la plage ». Elle peut être divisée en *groupes intonationnels* : [apparemment] [les petits enfants sont allés à la plage]. Puis, ces groupes intonationnels peuvent être constitués eux-mêmes de *syntagmes phonologiques* : [apparemment], [les petits enfants] [sont allés] [à la plage]. Ces syntagmes phonologiques peuvent être divisés en *groupes clitiques* : [apparemment] [les petits] [enfants] [sont] [allés] [à la plage] qui, eux-mêmes, peuvent être constitués d'un ou plusieurs *mots phonologiques*. Ces mots phonologiques (nom, verbe, adjectif) peuvent aussi être associés à leurs morphèmes grammaticaux, comme [les petit]. Les mots phonologiques sont divisés en *pièdes*, puis *syllabes*. Les unités prosodiques d'un même niveau sont toutes du même type et suivent une relation linéaire – non récursive. Ces règles de non-récursivité vont contraster les structures syntaxiques, autrement dit une unité ne peut être divisée qu'en constituants d'un type bien déterminé. Par exemple, un mot est composé d'au moins un pied et ce pied est composé d'au moins une syllabe.

Dans cette hiérarchie, les unités prosodiques sont étroitement liées aux fonctions grammaticales. En résumé, l'énoncé phonologique correspond à une phrase complète et les groupes intonationnels peuvent correspondre à des propositions syntaxiques. Puis les syntagmes phonologiques sont constitués de mots de contenu avec leurs mots de fonction associés. Et enfin, les mots phonologiques correspondent aux mots lexicaux. Ces diverses unités, séparées par des frontières prosodiques, reflètent donc la structure syntaxique de la phrase. Toutefois la correspondance n'est pas toujours stricte.

D'un point de vue acoustique, l'accentuation (proéminence) est marquée par des changements en intensité, pitch et durée. Ces trois indices principaux contribuent à ce que l'on définit comme la perception de l'accentuation ou la proéminence des syllabes. Les frontières des unités prosodiques sont majoritairement marquées par des indices acoustiques tel que des allongement finaux et des augmentations ou des baisses de pitch (Beckman & Pierrehumbert, 1986; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). A mesure que la taille des constituants diminue, les frontières prosodiques sont de moins en moins saillantes acoustiquement.

Des études ont montrés que ces indices de frontières prosodiques sont d'autant plus marqués dans la parole adressée à l'enfant (IDS, *Infant Directed Speech*) (Fisher & Tokura, 1996). Comment ces indices acoustiques vont-ils être perçus par les nourrissons et comment ces indices vont-ils les aider à acquérir leur lexique et l'ordre rudimentaire des mots ?

1.3 Un mécanisme d'apprentissage puissant : le *prosodic bootstrapping*

Un des mécanismes les plus impressionnants chez les très jeunes enfants est leur capacité à apprendre leur langue maternelle rapidement et sans apparente difficulté, alors qu'un adulte apprenant une deuxième ou une troisième langue a besoin de beaucoup plus de temps et d'efforts. De plus, un adulte n'atteint que rarement le même niveau de compétence qu'un enfant. La compréhension de certains mots apparaît dès 6 mois (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999, 2012), et les premiers mots vont être prononcés à partir de la fin de la première année de vie. Vers 18 mois, 75% des enfants comprennent environ 150 mots et peuvent en produire 50, et ce chiffre ne va faire qu'augmenter à mesure de leur développement (Fenson et al., 2000). Les

nourrissons vont aussi apprendre rapidement l'ordre selon lequel il faut organiser les mots dans les phrases, sans avoir aucune connaissance explicite de la grammaire. L'ordre des mots est une des premières propriétés syntaxiques que les nourrissons acquièrent. Ils ne font que très rarement d'erreurs sur l'ordre basique des mots lorsqu'ils commencent à prononcer leurs premières combinaisons de mots, entre 18 et 24 mois (Bloom, 1970; Meisel, 1992; Pinker, 1999). Par ailleurs, avant même de pouvoir les produire, les enfants réagissent différemment à l'écoute de phrases, selon si l'ordre des mots est correct ou non (Golinkoff, Jacquet, Hirsh-Pasek, & Nandakumar, 1996). Une des hypothèses principales soutient que les nourrissons acquièrent l'ordre des mots à un stade pré-lexical, donc bien avant l'acquisition du lexique (Gervain, Nespor, Mazuka, Horie, & Mehler, 2008; Gervain & Werker, 2013; Mazuka, 1996; Nespor et al., 2008).

Quels sont les mécanismes, présents dès la naissance, qui sous-tendent ces remarquables capacités ? Comment les enfants peuvent-ils lier cette chaîne continue de sons avec des représentations linguistiques encore très peu définies ?

Afin d'interpréter la phrase « Marie a cassé le vase », l'enfant doit identifier l'événement ou l'action décrite (ici : casser), les participants de l'action (Marie, vase) et identifier quel rôle chaque participant a joué (Marie est celle qui a cassé le vase et non l'inverse). Ce *linking problem* a reçu une attention considérable, à la fois par les théoriciens qui ont essayé de caractériser les liens entre la sémantique et la syntaxe, et par des psychologues spécialisés dans le développement et intéressés par la manière dont les enfants découvrent ces liens (Pinker, 1984, 1987). La question de quand et comment l'enfant commence à identifier et à extraire les informations nécessaires à l'apprentissage de la grammaire a été largement étudiée et formulée comme le *bootstrapping problem* (Pinker, 1987). De nombreuses solutions ont été proposées. En effet, de nombreux auteurs soulignent la nécessité de l'existence de connaissances domaine-spécifiques innées et de contraintes d'apprentissage. En fonction des types d'informations sur lesquels les enfants vont s'appuyer pour construire de nouvelles représentations linguistiques, plusieurs hypothèses de *bootstrapping* (ou « initialisation ») ont été émises. Par exemple, Pinker (1984) a proposé le *semantic bootstrapping*, suggérant qu'initialement les enfants résoudre le problème d'identification des catégories syntaxiques et des relations syntaxiques en lien avec leurs connaissances sémantiques des mots. Le *syntactic bootstrapping* (Gleitman, 1990)

propose que les enfants utilisent l'information syntaxique pour acquérir le sens des mots, notamment des verbes. D'autre part, le *distributional bootstrapping*, aussi appelé *correlational bootstrapping*, concerne la sensibilité aux informations distributionnelles dans le signal de la parole pour l'accès aux catégories syntaxiques (Maratsos & Chalkley, 1980). Dans cette lignée, Gervain (2008) propose le *frequency-based bootstrapping*, selon lequel la fréquence de distribution et l'ordre relatif des mots de fonction et de contenus permettent d'acquérir l'ordre basique des mots de la langue apprise. Enfin, le *prosodic bootstrapping* (Gleitman & Wanner, 1982, Nespor et al., 2008) propose que les indices suprasegmentaux, présents dans la parole, guideraient les nourrissons pour apprendre les mots et l'ordre basique des mots. Nous allons décrire en détails ce mécanisme dans la partie suivante.

1.3.1 Le prosodic bootstrapping

La théorie du *bootstrapping* prosodique, aussi appelée *bootstrapping from the signal*, propose que l'apprenant, pour extraire l'information concernant l'organisation syntaxique d'une langue donnée, se sert des informations prosodiques présentes dans le signal de parole. Gleitman et Wanner (1982) ont été les pionniers de l'hypothèse du *bootstrapping* prosodique. Ils proposent une série de mécanismes pour expliquer ces phénomènes, notamment l'étroit lien entre les indices de surface (suprasegmentaux) présents dans le signal de parole (signal acoustique) et les structures abstraites (catégories linguistiques et les règles qui les gouvernent) sous-jacentes. L'idée est que ces mécanismes aideraient les enfants à traiter le signal linguistique en utilisant ou en s'appuyant sur ces indices acoustiques de surface, qui les informeraient sur la structure syntaxique en signalant des frontières de mots et des constituants grammaticaux (Selkirk 1984 ; Nespor & Vogel 1986). Les nourrissons pourraient donc s'aider des variations de durée, intensité et de pitch, puisqu'ils sont étroitement liés aux frontières des unités syntaxiques.

Toutefois, les frontières syntaxiques ne sont pas toutes marquées par des frontières prosodiques, alors que toutes les frontières prosodiques sont des frontières syntaxiques. C'est pourquoi ces indices vont aider les nourrissons à organiser ce signal de parole continue, en permettant de définir la structure de la phrase, pour ensuite leur permettre d'apprendre les règles syntaxiques qui gouvernent le langage (voir aussi (Gleitman, Gleitman, Landau, & Wanner, 1988; Morgan, Meier, & Newport, 1987)

Le développement de ces théories a été repris et compilé par Morgan & Demuth (1996). Les hypothèses émises concernant le *prosodic bootstrapping* reposent sur 3 affirmations fondamentales :

- Des indices acoustiques fiables de la structure syntaxique sont présents dans la parole.
- Les nourrissons sont sensibles à ces indices et peuvent les exploiter lors du traitement de la parole.
- Les unités extraites vont aider les enfants dans la détection des relations entre les constituants de leur langue maternelle.

Ces auteurs formulent alors plus généralement que la prosodie constitue un signal qui peut agir comme un déclencheur et guider les enfants à apprendre la syntaxe et le lexique de leur langue maternelle. En effet, ces auteurs concluent que le fait de considérer le signal de parole comme contenant des indices nécessaires pour initialiser l'acquisition du langage (phonétique, phonotactique, prosodique) est essentiel pour comprendre les premières étapes de son acquisition. Une grande partie de ces hypothèses a été testée par de nombreux travaux expérimentaux, qui vont être présentés dans la section suivante.

Afin d'étudier le point de départ de ces compétences, nous allons décrire de quelle manière les nourrissons détectent ces indices prosodiques et à partir de quel moment ils y sont sensibles. Puis, nous décrirons comment les patterns accentuels au niveau des mots et des syntagmes sont importants dans l'acquisition des mots et de l'ordre des mots dans certaines langues.

1.3.2 Quelles sont les sensibilités précoces qui soutiennent l'hypothèse du *prosodic bootstrapping* ?

1.3.2.1 La nature du signal de parole in utero

L'exposition à la langue maternelle commence bien avant la naissance. Dès 25 semaines de gestation, la cochlée est fonctionnelle (Pujol et al., 1998), des potentiels évoqués auditifs commencent à apparaître et à être détectés à la fois dans le tronc cérébral et dans le cortex auditif, et vont devenir de plus en plus nombreux au fur et à mesure du développement cérébral. Vers 35 semaines de gestation, la biomécanique de

la cochlée, permettant une spécificité fréquentielle du codage des signaux acoustiques, est quasiment mature (Hall, 2000). L'analyse fine de sons de parole par les fœtus est alors possible, et ces sons commencent à être traités par le cortex auditif. Or, ces mécanismes sont encore peu connus, au vu de la difficulté d'accès au fonctionnement cérébral des fœtus. Toutefois, l'environnement auditif intra-utérin est très riche et complexe ; comment les fœtus font-ils pour reconnaître la parole dans un environnement où d'autres sons sont également présents, comme ceux provenant du système maternel cardiovasculaire, gastro-intestinal ou encore respiratoire ?

Une équipe de recherche a analysé les caractéristiques spectrales des enregistrements de l'environnement intra-utérin. Ils ont révélé que la voix maternelle reste un des signaux les plus saillants dans le signal acoustique transmis aux fœtus, notamment par sa double transmission, auditive et haptique. En effet, la voix maternelle est transmise par les tissus biologiques, majoritairement par la conduction des os, puis à travers le liquide amniotique, pour finir à l'oreille interne. De plus, lors de la transmission de la voix maternelle, celle-ci est amplifiée par la vibration interne du larynx et du diaphragme (Granier-Deferre & Busnel, 2011; Querleu, Renard, Versyp, Paris-Delrue, & Crèpin, 1988).

Des analyses spectrales ont également indiqué que, de la voix maternelle, seules les caractéristiques prosodiques étaient préservées : l'information rythmique et les contours de pitch (Spence & Decasper, 1987). En d'autres termes, les tissus biologiques feraient acte de filtres passe-bas, ne laissant passer qu'une bande de fréquences, celles inférieures à 300-400Hz, rendant pratiquement imperceptibles les variations phonémiques, ou plus largement segmentales (Querleu, Renard, Versyp, Paris-Delrue, & Vervoot, 1988). L'information de parole n'est donc pas transmise dans toute son intégralité et ce point tient une grande importance dans notre recherche.

1.3.2.2 Sensibilité aux indices prosodiques chez les fœtus

Bien avant la naissance, le fœtus est capable de reconnaître des sons de parole, et est sensible aux variations prosodiques. De nombreuses recherches ont montré, à l'aide de capteurs du rythme cardiaque fœtal, que les fœtus, dans le dernier trimestre de grossesse, sont capables de distinguer la musique de la parole (Granier-Deferre, Bassereau, Ribeiro, Jacquet, & Decasper, 2011; Kisilevsky, Hains, Jacquet, Granier-Deferre, & Lecanuet, 2004), leur langue maternelle d'une langue inconnue (Kisilevsky et al., 2009), la voix de leur mère de celle d'une étrangère (Kisilevsky et al., 2003), et,

très récemment, de reconnaître la voix du père, toujours en préférant la voix maternelle (Lee & Kisilevsky, 2014). Ces études suggèrent que, dès le début du troisième trimestre de grossesse, le système auditif est assez mature pour distinguer, notamment sur la base des indices prosodiques, deux types de sons avec une préférence pour les stimuli familiers. Par ailleurs, plus les fœtus se rapprochent du terme, plus ils sont en mesure de discriminer de fines différences acoustiques, comme des différences d'intensité, de fréquence, de pitch et de spectre, qui sont nécessaires pour percevoir les contours mélodiques (Decasper, Lecanuet, Busnel, Granier-Deferre, & Maugeais, 1994; Granier-Deferre, Ribeiro, Jacquet, & Bassereau, 2011; Hepper & Shahidullah, 1994; Lecanuet & Schaal, 1996). Ils peuvent aussi différencier des notes musicales de basse versus haute fréquence (Lecanuet, Granier-Deferre, Jacquet, & DeCasper, 2000). Ces nombreuses études ont donc montré des sensibilités très précoces aux propriétés prosodiques de la parole, avant même la naissance.

Ces préférences pour la voix maternelle, ou encore la langue maternelle, avant la naissance, reflètent non seulement la capacité à discriminer et à traiter ces stimuli auditifs, mais également la mise en mémoire de certaines caractéristiques auxquelles les fœtus ont été exposés au cours de la grossesse. Ces capacités reflètent alors, probablement, la présence de puissants mécanismes d'apprentissage associés à un système auditif fonctionnel. Ces études suggèrent qu'il existe donc un apprentissage pendant la vie fœtale. Une récente recherche semble montrer qu'une exposition prénatale à certaines mélodies induirait des représentations neurales à long terme (Partanen, Kujala, Tervaniemi, & Huotilainen, 2013; Partanen, Kujala, Naatanen, et al., 2013), mais de nombreuses questions restent ouvertes.

1.3.2.3 Sensibilité aux indices prosodiques dans les premiers mois de vie

1.3.2.3.1 Au niveau de la rythmicité générale des langues

Les recherches chez les nouveau-nés ont mis en évidence des capacités très précoces de discrimination et de préférence sur la base des indices prosodiques de la parole. A la naissance, les nourrissons préfèrent écouter leur langue maternelle plutôt qu'une langue étrangère. En effet, une des études pionnières sur les capacités langagières précoces a présenté à des nouveau-nés français des phrases prononcées en français ou en russe par une locutrice bilingue (Mehler et al., 1988). Les réponses des nouveau-nés ont été enregistrées avec une technique comportementale (HAS, *High*

Amplitude Sucking) mesurant leur taux de succion (qui a été utilisé dans de nombreuses études). Les résultats montrent que ces nouveau-nés sont capables de discriminer leur langue maternelle (le français) d'une langue étrangère (russe). Quelques années plus tard, ces travaux ont été reproduits chez des nouveau-nés exposés à d'autres langues (Byers-Heinlein et al., 2010; Moon et al., 1993), confirmant que, dès la naissance, les nourrissons sont capables de distinguer leur langue maternelle d'une langue étrangère. Mais sur quels paramètres se basent-ils pour discriminer-préférer ces différents stimuli ?

Afin de comprendre comment les nouveau-nés traitent ces séquences de sons de parole et quels indices sont nécessaires pour discriminer leur langue maternelle d'une langue étrangère, les chercheurs ont dégradé les stimuli utilisés dans l'expérience précédente avec un filtre passe bas (< 400Hz, Mehler et al., 1988). En dégradant ces séquences de parole naturelle, la plupart des informations phonétiques ou encore phonotactiques sont éliminées, alors que les informations prosodiques sont largement préservées et disponibles. Ce filtrage, même s'il est loin d'être parfait, permet d'isoler le rythme et l'intonation. Les auteurs ont présenté ces stimuli dégradés à des nouveau-nés et les résultats de cette étude montrent que les nouveau-nés sont toujours capables de discriminer deux langues avec des stimuli filtrés passe bas. Ces travaux suggèrent donc que les nouveau-nés utilisent les indices prosodiques pour discriminer les langues, peut-être le rythme. D'autres travaux ont exploré cette hypothèse rythmique en présentant d'autres paires de langues. Les nouveau-nés français se sont montrés capables de discriminer l'anglais du japonais, ou encore le néerlandais du japonais, mais pas l'anglais du néerlandais (Nazzi, Bertoncini, et al., 1998; Ramus, Hauser, Miller, Morris, & Mehler, 2000). Ces résultats suggèrent qu'ils sont capables de discriminer deux langues inconnues si elles sont rythmiquement différentes, mais qu'ils échouent si elles appartiennent à la même classe rythmique. En effet, le néerlandais et l'anglais font partie des langues accentuelles, alors que le japonais est une langue moraique et le français est une langue syllabique. Une étude a élargi ces recherches en essayant de déterminer quels indices prosodiques permettent aux nouveau-nés de discriminer les langues (Ramus, 2002). En préservant les indices rythmiques de la parole mais pas l'intonation, les nouveau-nés sont toujours capables de discriminer le néerlandais du japonais. Ces données montrent bien que les propriétés rythmiques de la parole sont détectées dès la naissance par les bébés et que ce sont ces indices qui leur permettent de discriminer les langues.

Cependant, ces capacités de discrimination des langues ne sont pas spécifiquement humaines. De nombreuses recherches ont montré, même chez des espèces assez éloignées de la nôtre, la capacité de discriminer les langues en fonction de leur rythme, chez le singe (Ramus et al., 2000) ou encore chez le rat (Toro, Trobalon, & Sebastián-Gallés, 2003). Ces travaux suggèrent ainsi que ces capacités de discrimination d'indices prosodiques pourraient être reliées à des capacités perceptives générales du système auditif partagées avec d'autres espèces.

Nous avons vu qu'à la naissance les nouveau-nés montrent des compétences générales de traitement de la parole sur la base du rythme. Comment se développent ces compétences ?

Des études ont montré que la sensibilité à certaines propriétés prosodiques des langues devient de plus en plus fine. En effet, des nourrissons américains anglophones sont capables de discriminer, à partir de 5 mois, différents dialectes de la même langue (anglais britannique versus anglais américain), tout en continuant de faire la distinction entre les langues appartenant à différentes classes rythmiques (anglais britannique versus japonais) (Nazzi, Jusczyk, & Johnson, 2000). Des résultats similaires sont trouvés avec des monolingues espagnols ou catalans qui sont capables de distinguer ces deux langues, rythmiquement similaires, à 4 mois (Bosch & Sebastián-Gallés, 1997). Très récemment, une étude avec des nourrissons monolingues basques ou espagnols, et des bilingues à forte dominance basque ou espagnole, ont trouvé des résultats similaires, cette fois-ci chez des nourrissons encore plus jeunes de 3,5 mois (Molnar, Gervain, & Carreiras, 2014).

Ces recherches montrent que les nourrissons ont des préférences initiales et se spécialisent, partant de capacités globales de traitement de la prosodie à la naissance vers une perception plus fine de l'organisation de la structure prosodique de leur langue maternelle. Ces sensibilités sont un pré-requis pour percevoir les changements des patterns rythmiques qui peuvent indiquer les frontières des unités prosodiques de la parole et donc acquérir leur langue maternelle.

1.3.2.3.2 Au niveau des unités prosodiques majeures : groupes intonationnels et syntagmes

Une des études pionnières qui a mesuré les sensibilités des nourrissons aux corrélats acoustiques des frontières prosodiques au niveau des groupes intonationnels (*clauses*) a utilisé la méthode de regard préférentiel (HPP), chez des nourrissons de 7 à 10 mois américains anglophones. Ils ont présenté aux nourrissons des passages de parole naturelle (anglais) avec des pauses artificielles, où les pauses étaient insérées à l'exacte frontière des groupes intonationnels (position congruente) ou au milieu (position non-congruente). Les auteurs trouvent que les nourrissons écoutent plus longtemps les passages où la position des pauses est congruente, plutôt que non-congruente (Hirsh-Pasek, Nelson, Jusczyk, Druss, & Kennedy, 1987; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989).

Des travaux sur les sensibilités des nourrissons aux frontières des groupes intonationnels se sont intéressés à la spécificité des résultats précédents. En effet, est-ce que les mêmes effets seraient retrouvés avec des passages de parole naturelle, mais prononcée dans une langue étrangère ? Ainsi, en utilisant la même méthode (HPP) et procédure, les auteurs montrent que des nourrissons de 4,5 mois exposés à l'anglais sont sensibles aux corrélats acoustiques des frontières prosodiques, même si les stimuli sont présentés dans une langue qui leur est étrangère, le polonais (Jusczyk, Hohne, & Mandel, 1995; Jusczyk, 1989). En revanche, à 6 mois, les nourrissons anglophones montrent une préférence uniquement quand ils sont présentés à des stimuli prononcés dans leur langue maternelle. Mais sur quels indices exactement les nourrissons se sont-ils basés ? Les auteurs ont essayé de répondre à cette question en présentant aux nourrissons de 6 mois anglophones, des passages prononcés en polonais, mais ayant subi un filtrage passe bas, éliminant ainsi la possibilité que des indices phonémiques ou phonotactiques du polonais puissent expliquer ce résultat. Des résultats similaires ont été trouvés, les nourrissons préférant les passages avec les insérées aux positions congruentes. Mais ces études ont été réalisées avec des nourrissons exposés à l'anglais ; retrouve-t-on les mêmes résultats dans d'autres langues ?

Des études similaires avec des nourrissons exposés à d'autres langues montrent que cette capacité émerge à la même période et n'est donc pas spécifique des nourrissons anglais : les nourrissons allemands préfèrent également les passages où les pauses ont été insérées aux frontières congruentes à ceux où elles ont été insérées à des

positions non congruentes (Schmitz, 2008). Par contre, ces effets sont retrouvés un peu plus tardivement chez les nourrissons japonais, autour de 8 mois (Hayashi & Mazuka, 2002). Par conséquent, cette capacité à reconnaître les frontières prosodiques dans la parole continue est une compétence qui semble être partagée entre des nourrissons exposés à des langues différentes. Jusczyk et al (1995) suggère que ces résultats montrent que les nourrissons sont sensibles dès 4,5 mois aux indices des frontières prosodiques et graduellement se spécialisent, vers 6 mois, pour les caractéristiques prosodiques de leur langue maternelle. Donc la perception des groupes intonationnels peut être influencée en partie par les propriétés prosodiques de la langue maternelle et ce assez tôt dans le développement.

D'autres recherches ont essayé de comprendre comment la présence de frontières prosodiques affecte la segmentation de la parole continue. Une étude par Nazzi et al. (2000) a exploré, dans une tâche de segmentation, comment les nourrissons de 6 mois exposés à l'anglais reconnaissent des séquences de mots dans des passages, lorsque ces séquences ont été présentées en cohérence avec les frontières naturelles des groupes intonationnels, ou à cheval entre deux syntagmes ex : "rabbits eat leafy vegetables" ou à cheval "... rabbits eat. Leafy vegetables ...". Les résultats montrent que la présentation des mots en cohérence avec les frontières prosodiques aide les nourrissons à mieux mémoriser la séquence pendant la phase de familiarisation et facilite la reconnaissance de cette séquence dans la phase de test. Or, dans cette étude, les nourrissons ont été familiarisés avec des séquences cibles pré-extraites et ont été testés sur leurs capacités à détecter ces séquences quand elles étaient imbriquées dans des passages, ce qui a pu faciliter la tâche. Une étude, utilisant un dessin expérimental comparable à l'étude de Nazzi et al. (2000), a cette fois-ci présenté, en phase de familiarisation ou en test, des passages, ce qui rendait la tâche un peu plus complexe (Soderstrom, Nelson, & Jusczyk, 2005). Ces auteurs trouvaient les mêmes résultats. Ces données ont mis en évidence que les nourrissons sont hautement sensibles à l'information prosodique des groupes intonationnels et que cela influence leur mémorisation des séquences de mots. Des recherches chez des nourrissons plus jeunes ont montré qu'ils étaient aussi capables d'exploiter l'information prosodique des groupes intonationnels. Les auteurs ont observé que des nourrissons de 2 mois exposés à l'anglais sont capables de percevoir les corrélats prosodiques des groupes intonationnels, et sont sensibles aux changements phonétiques ou d'ordre de mots à

l'intérieur de ces groupes (Mandel, Jusczyk, & Nelson, 1994; Mandel, Kemler Nelson, & Jusczyk, 1996). Mais comment perçoivent-ils les frontières à l'intérieur de ces groupes intonationels ?

Des études réalisées chez des nourrissons de 9 mois américains anglophones, utilisant la même technique que Hirsh-Pasek et al (1987), ont inséré des pauses dans des passages, soit avant le verbe (frontière entre le sujet et le syntagme verbal), soit après le verbe (dans le syntagme verbal). Les auteurs observent des préférences pour les pauses congruentes au niveau des frontières des syntagmes (Gerken, Jusczyk, & Mandel, 1994; Jusczyk et al., 1992). Ils retrouvent également ce pattern de résultats si les stimuli langagiers ont subi un filtrage passe bas, ce qui renforce l'idée que les indices prosodiques ont un rôle très important chez les nourrissons pour détecter les frontières de ces unités à cet âge.

Ces résultats ont été généralisés à des nourrissons plus jeunes, à l'aide d'une procédure de segmentation en HPP, en présentant à des nourrissons de 6 ou 9 mois des passages contenant des syntagmes phonologiques cibles lors d'une phase de familiarisation (Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003). Puis, en phase test, les auteurs leur ont fait écouter des passages contenant les syntagmes présentés en familiarisation (congruents) ou une séquence phonétiquement identique, mais à cheval sur la frontière prosodique (non-congruente). Les résultats montrent que les nourrissons de 6 mois préfèrent écouter les passages avec les syntagmes congruents, plutôt que les syntagmes non-congruents. Mais cette préférence n'a pas été répliquée dans un second set de stimuli. Des analyses acoustiques du matériel utilisé entre les deux expériences révèlent de plus grandes différences de durée (comme l'allongement final) et de pitch dans la première expérience par rapport à la seconde. Les auteurs en concluent que les nourrissons de 6 mois sont sensibles aux frontières prosodiques qui coïncident avec les frontières syntaxiques. Ces travaux ont été également répliqués chez des nourrissons exposés à d'autres langues, comme le hollandais (Johnson & Seidl, 2008). Ces résultats montrent que les nourrissons, dès 6 mois, sont sensibles aux frontières des syntagmes. Par ailleurs, l'identification des mots, comme appartenant à un syntagme, est un pré-requis pour ensuite acquérir les régularités grammaticales des syntagmes et de trouver les mots.

1.3.2.3.3 Au niveau des frontières des mots

Dès la naissance, les nourrissons présentent des sensibilités aux frontières des mots dans les syntagmes. Une recherche a présenté à des nouveau-nés des dissyllabes « mati » dans « panorama **typique** » versus « mathé**mat**icien ». Ils ont enregistré les réponses à l'aide de la technique de succion non-nutritive (HAS) et montrent que les nouveau-nés perçoivent les corrélats acoustiques des frontières prosodiques au niveau des syntagmes phonologiques (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe, Mehler, & Sebastián-Gallés, 2001). Les nouveau-nés perçoivent également les contours de pitch de mots japonais (Nazzi, Floccia, et al., 1998), suggérant ainsi une sensibilité très précoce au contour de pitch. De plus, une étude, explorant la perception des nouveau-nés aux patterns accentuels des mots, a mis en évidence que, lors de la présentation de mots multi-syllabiques différant sur leur accentuation (MAma versus maMA), les nouveau-nés italiens étaient capables de discriminer les deux patterns d'accentuations (Sansavini, Bertoncini, & Giovanelli, 1997). Ces données établissent que les nourrissons possèdent dès la naissance des sensibilités prosodiques aux indices prosodiques de frontières de mots, ou marquant des différences d'accentuation lexicale entre mots.

Discussion intermédiaire

L'ensemble de ces recherches montre que, dès la naissance, les nourrissons perçoivent les indices de frontières prosodiques qui marquent les unités linguistiques comme les allongements finaux, les montées ou descentes de pitch. Toutefois, ces études n'explorent pas le rôle spécifique de l'information prosodique dans le traitement de la parole. Comment les nourrissons s'aident-ils de ces indices pour encoder les mots, trouver leur ordre dans les phrases ? Des sensibilités initiales aux indices prosodiques sont donc présentes très tôt dans le développement. Par conséquent, la sensibilité au rythme de la parole, ainsi qu'aux frontières des unités prosodiques, pourrait permettre aux nourrissons d'être informés sur les unités rythmiques ainsi que les patterns accentuels présents dans leur langue maternelle, donc sur les unités qui sont pertinentes dans la langue en acquisition. Ces frontières prosodiques vont donc permettre aux nourrissons de commencer à mettre en place et à détecter des propriétés prosodiques de leur langue maternelle et à initialiser l'acquisition de la syntaxe et du lexique.

1.4 Acquisition des propriétés prosodiques et leur rôle dans le *bootstrapping* du vocabulaire et de la grammaire

1.4.1 Le problème de la parole continue

Dès la naissance, un des grands problèmes auxquels font face les nourrissons pour trouver les mots dans les phrases est la continuité du signal. En effet, les mots sont rarement produits en isolation (Brent & Siskind, 2001). De plus, lorsque les mères veulent apprendre un mot nouveau à leur enfant, il est rare qu'elles le présentent en isolation (Woodward & Aslin, 1990, 1991, 1992, cité dans Aslin, Woodward, LaMendola, & Bever, 1996). Comment les enfants font-ils alors pour apprendre les mots aussi rapidement ? Alors que les adultes vont pouvoir utiliser leur lexique pour trouver les mots dans les phrases, c'est loin d'être encore le cas chez les nourrissons dans la première année de vie. A l'inverse de l'écrit, où des espaces sont présents et visibles par le lecteur, il est impossible à l'oral de « voir » ces espaces, ni de les entendre (Figure 1.4).

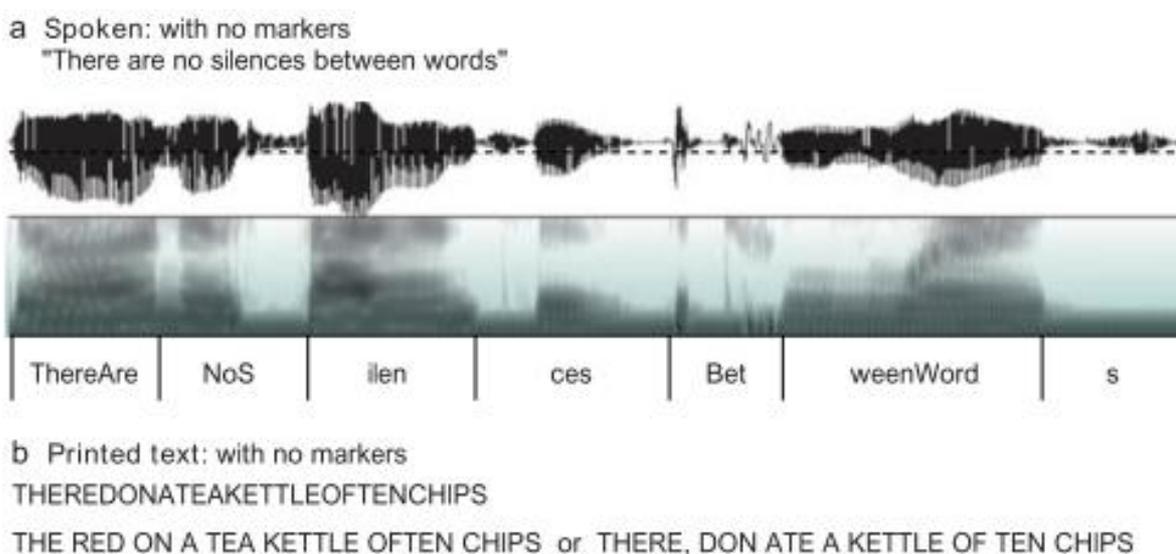


Figure 1.4 : Procédure de segmentation : (a) le signal de parole ne présente pas de frontières de mots apparentes, (b) alors qu'à l'écrit des espaces sont présents entre les mots (d'après Kuhl, 2004).

Les enfants doivent donc extraire les mots dans ce flux continu de parole afin de constituer leur lexique. De nombreuses hypothèses ont été proposées concernant les mécanismes sous-jacents à ces capacités, impliquant de nombreux indices du signal de parole sur lesquels les nourrissons pourraient s'appuyer pour segmenter la parole. De

nombreux indices ont été identifiés comme ayant un rôle pour aider les nourrissons à segmenter la parole en mots : prosodiques (Curtin, Mintz, & Christiansen, 2005; Jusczyk, Houston, & Newsome, 1999; Nazzi, Iakimova, Bertoncini, Fredonie, & Alcantara, 2006; Nishibayashi, Goyet, & Nazzi, 2015), statistiques (Johnson & Jusczyk, 2001; Nazzi, Mersad, Sundara, Iakimova, & Polka, 2014; Saffran, Newport, & Aslin, 1996; Thiessen & Saffran, 2003), phonotactiques (Gonzalez-gomez & Nazzi, 2013; Mattys, Jusczyk, Luce, & Morgan, 1999; Mattys & Jusczyk, 2001b), de coarticulation (Johnson & Jusczyk, 2001), allophoniques (Jusczyk, Hohne, & Bauman, 1999; Mattys & Jusczyk, 2001a).

Selon Mattys, White et Melhorn (2005), alors que les adultes vont essentiellement s'aider des indices lexicaux et segmentaux (ex : contraintes phonotactiques, coarticulation) pour segmenter, les nourrissons vont s'aider dans un premier temps des indices prosodiques qui ne requièrent pas de connaissance spécifiquement langagière. Il apparaît donc qu'au début de l'acquisition du langage, un des indices qui apparaît comme l'un des plus cruciaux dans le début du processus de segmentation correspond aux indices de la prosodie linguistique. De quelle manière ces indices vont-ils aider les enfants à segmenter la parole ?

1.4.2 Le rôle de l'accentuation lexicale : trouver et représenter les mots

1.4.2.1 Les indices disponibles

De récentes recherches ont mis en évidence, au moins dans certaines langues, que l'information prosodique est cruciale pour segmenter la parole en mots. Étant donné qu'il n'y a pas de marqueurs acoustiques clairs des frontières de mots, la segmentation des mots est loin d'être triviale. Quels sont les indices qui aident les nourrissons pour trouver les mots ?

Comme nous l'avons vu dans la section 2.2., les mots contiennent des syllabes accentuées et non accentuées et les patterns rythmiques qui en découlent pourraient aider les nourrissons à encoder les mots (Gleitman & Wanner, 1982). Or les langues diffèrent selon ce paramètre. Ces différences cross-linguistiques ont d'importantes implications dans le rôle de la perception de l'accentuation, pour l'accès à la forme sonore des mots et donc pour l'accès au lexique (Cutler & Norris, 1988; Cutler, 1975). En effet, de nombreux travaux ont présenté de solides preuves sur les procédures de

segmentation, qui seraient largement influencées par la langue maternelle des individus testés. Des études ont montré que ces procédures semblent reposer sur l'unité rythmique de la langue apprise, par exemple : accentuelle chez les anglais (Cutler et al., 1986), syllabique chez les français (Cutler et al., 1986; Mehler, Dommergues, Frauenfelder, & Segui, 1981) et moraïque chez les japonais (Cutler et al., 1986; Otake, Hatano, Cutler, & Mehler, 1993). Mais qu'en est-il pour les nourrissons ? Des hypothèses ont été proposées, postulant que les stratégies dans les processus de segmentation seraient partagées entre les adultes et les nourrissons (Nazzi et al., 2006).

1.4.2.2 Acquisition des patterns prosodiques des mots

Les nourrissons commencent à segmenter et à extraire les mots dans la parole continue à partir de la deuxième moitié de la première année de vie. En effet, les premières études ont été réalisées en utilisant la technique HPP chez des nourrissons exposés à de l'anglais américain et âgés de 7,5 mois ou de 10,5 mois. Les résultats montrent que, dès 7,5 mois, les nourrissons sont capables d'extraire de la parole continue des mots monosyllabiques (Jusczyk & Aslin, 1995) et certains types de mots bi-syllabiques (Johnson & Jusczyk, 2001; Jusczyk, Houston, et al., 1999). Dans cette dernière étude, les nourrissons de 7,5 mois reconnaissent les mots dissyllabiques trochaïques (KINGdom ou DOCTOR) dans la parole continue, mais pas les mots iambiques (guiTAR ou surPRISE). Ce n'est qu'à 10,5 mois que les nourrissons sont capables d'extraire les mots iambiques. Le pattern accentuel trochaïque semble donc être traité/reconnu plus tôt au cours du développement que le pattern iambique, suggérant que les nourrissons anglophones segmentent d'abord sur la base de l'unité trochaïque. Des patterns de résultats similaires ont été trouvés chez les nourrissons acquérant le néerlandais et l'allemand (langues très similaires à l'anglais sur cette dimension) de 7,5 mois à 9 mois (Höhle, 2002; Houston, Jusczyk, Kuijpers, Coolen, & Cutler, 2000; Kooijman, Hagoort, & Cutler, 2009).

Ces études réalisées chez les anglais, allemands et hollandais montrent donc un clair avantage des unités trochaïques dans les étapes précoces de la segmentation de la parole. De nombreuses hypothèses ont été émises concernant l'origine de ces compétences, et comment elles émergent.

Certains auteurs proposent que ce biais trochaïque est inné et que les nourrissons commencent à segmenter en extrayant les unités trochaïques indépendamment de la langue apprise (Allen & Hawkins, 1978). Une seconde possibilité est que les

nourrissons commencent par faire des analyses de distribution de patterns rythmiques dans le signal de la parole, ou à extraire et généraliser les patterns rythmiques des mots prononcés isolément. Étant donné que 90% des mots de contenus (substantif) sont accentués sur la première syllabe en anglais (Cutler & Carter, 1987), extraire les unités trochaïques serait donc une procédure de segmentation efficace dans cette langue. Par conséquent, les nourrissons exposés à l'anglais seraient sensibles au pattern prédominant des mots de leur langue. Une troisième alternative dans la lignée de la théorie du *prosodic bootstrapping* (Gleitman & Wanner, 1982) et de la stratégie de segmentation métrique trouvée chez les adultes (Cutler et al., 1986; Cutler & Norris, 1988) a été proposée par Nazzi et al. (2006) : l'hypothèse d'initialisation rythmique. Cette hypothèse propose que le processus de segmentation précoce repose sur l'unité rythmique de la langue apprise. Cette prédiction est en accord avec les données présentées ci-dessus pour les anglais, allemands et néerlandais. Ceci a également été confirmé par des récents travaux chez les nourrissons français (Nazzi et al., 2006, 2014; Nishibayashi et al., 2015) et chez des nourrissons espagnols (Bosch, Figueras, Teixidó, & Ramon-Casas, 2013; Pons & Bosch, 2010), qui utilisent l'unité de ces langues, la syllabe, pour segmenter précocement. Ainsi, le biais trochaïque observé chez les anglais, allemands et néerlandais ne correspondrait pas à l'acquisition du pattern accentuel prédominant des mots, ni à une compétence innée générale, mais à l'acquisition de l'unité rythmique de ces langues avant le début des capacités de segmentation, donc avant 7-8 mois.

Toutefois, ces études ne répondent pas clairement à la question de l'origine et des mécanismes sous-tendant l'émergence de ce biais prosodique. Il est donc difficile de déterminer exactement quels sont les indices sur lesquels les nourrissons s'appuient et, surtout, comment ces compétences évoluent pendant la première année de vie. Une alternative récente a émergé, soulignant le rôle des caractéristiques physiques des syllabes accentuées et non accentuées pouvant sous-tendre ces processus : la loi Iambo-trochaïque (ITL) que nous allons décrire en détails dans la partie suivante.

C'est également à cette période que des biais de préférence pour certains types de pattern rythmique émergent. Une des premières études à avoir observé ces préférences a été réalisée par une équipe américaine (Jusczyk, Cutler, & Redanz, 1993). Ces chercheurs ont présenté à des nourrissons de 9 mois exposés à l'anglais américain des listes de mots dissyllabiques ayant soit un pattern trochaïque (fort-faible), soit un pattern iambique (faible-fort). Les résultats montrent que les nourrissons regardent plus

longtemps les listes de mots trochaïques que les listes de mots iambiques. Cependant, aucune préférence n'est retrouvée à 6 mois. De plus, des résultats similaires sont obtenus lorsque ces mêmes stimuli ont été filtrés par un filtre passe-bas, préservant leurs caractéristiques prosodiques. De nombreuses répliques de ces travaux ont suivi, avec différents types de stimuli (Echols, Crowhurst, & Childers, 1997; Turk, Jusczyk, & Gerken, 1995). Un biais prosodique, ici trochaïque, est donc retrouvé chez des nourrissons de 9 mois. Toutefois, ces études n'ont été réalisées que chez des nourrissons exposés à l'anglais, qui est une langue accentuelle à prédominance trochaïque, ce qui limite l'interprétation de ces données. Est-ce seulement un biais général pour un pattern fort-faible ? Est-il spécifique de la langue apprise ?

Une recherche en potentiels évoqués (ERPs) démontre que des nourrissons français et allemands de 4 mois traitent différemment des dissyllabes (/baba/) ayant un pattern trochaïque versus iambique (Friederici, Friedrich & Christophe, 2007). Il semble donc que les nourrissons de 4 mois soient capables de discriminer les patterns d'accentuations, et le fait que les nourrissons allemands et français traitent différemment les patterns trochaïques et iambiques suggère une spécialisation précoce pour le pattern rythmique de leur langue. Parallèlement, Höhle et al. (2009) ont étudié chez des nourrissons exposés au français ou à l'allemand, âgés de 4 ou de 6 mois, la perception d'un non-mot dissyllabique (/gaba/) possédant un pattern soit trochaïque, soit iambique, cette fois-ci en HPP (Headturn Preference Paradigm). Ces auteurs ne trouvent pas de réponse préférentielle chez les nourrissons français de 6 mois, quel que soit le type de séquences, même s'ils peuvent discriminer les 2 patterns accentuels. Par contre, les bébés allemands présentent, dès 6 mois, une réponse préférentielle aux séquences trochaïques. A ces différences cross-linguistiques s'ajoute le fait que les capacités de discrimination accentuelle semblent décroître à partir de 10 mois chez les nourrissons français. En effet, les nourrissons de 10 mois monolingues français ont besoin d'un temps plus long de familiarisation pour discriminer les stimuli (Bijeljic-Babic, Serres, Höhle, & Nazzi, 2012) par rapport aux nourrissons de 6 mois (Höhle et al., 2009). Par ailleurs, les travaux de Mattock & Burnham (2006) montrent un déclin de sensibilité aux dimensions prosodiques qui ne sont pas contrastives dans la langue maternelle, sur cette même période comprise entre 6 et 9 mois. Il est donc également possible que les résultats de Höhle et al. (2009) et de Friederici et al. (2007) montrent que les nourrissons allemands ont acquis le biais trochaïque et/ou que les enfants français ont perdu en partie ce biais prosodique. Prises ensemble, ces recherches

montrent qu'en apprenant leur langue maternelle, les nourrissons se spécialisent et maintiennent des sensibilités à détecter différents patterns accentuels.

Les recherches réalisées chez les nourrissons français montrent qu'ils sont capables de discriminer les patterns accentuels, suggérant ainsi que les nourrissons français sont loin d'être sourds à l'accentuation. Toutefois, ces études ont été réalisées dans un contexte de faible variabilité segmentale. D'ailleurs, des études à la naissance rapportent déjà la capacité des nourrissons à discriminer les mots selon leurs patterns accentuels (Nazzi, Floccia, et al., 1998; Sansavini et al., 1997). Par contre, dès lors que la variabilité segmentale des stimuli est plus importante, des différences apparaissent entre les nourrissons exposés à des langues maternelles ayant des systèmes d'accentuation lexicale différents. Des recherches ont montré que les nourrissons de 8-9 mois, exposés à l'espagnol ou à l'anglais (langues à accentuation lexicale variable), sont capables de discriminer les patterns accentuels dans une liste de mots qui varie, contrairement aux nourrissons français (une langue sans accentuation lexicale) du même âge (Skoruppa et al., 2009; Skoruppa, Cristià, Peperkamp, & Seidl, 2011). Ces résultats suggèrent que les nourrissons exposés à une langue à accentuation lexicale variable ont la capacité d'encoder les patterns accentuels en présence de variabilité segmentale, mais pas les nourrissons français. Ces observations posent la question de savoir comment cette capacité se développe chez des nourrissons bilingues entendant le français et une langue possédant une accentuation lexicale variable. En effet, comment ces enfants traitent-ils l'accentuation dans un contexte qui implique d'extraire et de généraliser les patterns d'accentuation ?

Nous avons vu à travers l'ensemble de ces études que les nourrissons apprennent donc, dans la première année de vie, les propriétés prosodiques des mots de leur langue maternelle. Ces indices prosodiques vont les guider dans l'apprentissage des mots, mais ils ont également un rôle essentiel dans l'apprentissage des règles rudimentaires de grammaire, comme l'ordre des mots.

1.4.3 Le rôle de la proéminence dans les syntagmes : trouver l'ordre des mots

1.4.3.1 Les indices disponibles

Les nourrissons, dans la première année vie, comme nous l'avons vu précédemment, ne possèdent pas encore un lexique à proprement parlé. Afin de comprendre le sens d'une phrase, comment vont-ils faire pour identifier les propriétés syntaxiques des mots et se représenter leur ordre ? Pour ce faire, plusieurs indices seraient à leur disposition dans la parole, notamment des indices statistiques (*bootstrapping* fréquentiel) et prosodiques (*bootstrapping* prosodique). Nous allons nous intéresser dans cette thèse exclusivement aux indices prosodiques.

Il a ainsi été proposé que les indices de proéminence prosodique présents dans le signal de parole pourraient guider les nourrissons dans leur acquisition de certaines propriétés syntaxiques de leur langue maternelle, notamment l'ordre des mots. Un des paramètres qui pourrait permettre aux enfants d'apprendre l'ordre des mots de leur langue maternelle est celui de la tête-complément.

Selon Chomsky (1981), le paramètre de direction de tête permet de distinguer deux catégories de grammaire. En effet, dans la plupart des langues, les têtes des unités se trouvent systématiquement avant leur complément, ou après leur complément. De même, les langues se distinguent en fonction du placement de la préposition principale avant ou après les prépositions subordonnées. En d'autres termes, l'ordre des mots est relié à la direction de la tête: dans un premier cas, les verbes (V) précèdent les objets (O), donc la direction est à droite (VO), comme dans plupart des langues indo-européennes : anglais, français ou encore grec. Dans un second cas, les objets (O) précèdent les verbes (V), direction à gauche (OV), comme en turc, hindi ou japonais (Dryer, 1992). Par exemple, en français, dans le groupe verbal « mange une pomme », le verbe manger, qui correspond à la tête lexicale du constituant syntaxique, est au début du groupe. Par contre, en japonais, le pattern est opposé : dans *ringo-wo taberu* (« pomme mange »), le verbe est à la fin du groupe. Il a été mis en évidence que les langues VO ou OV suivent généralement le même ordre que d'autres catégories de syntagmes (prépositions et substantifs, conjonctions et subordonnées, ... ; Greenberg 1963; Dryer 1992). Par exemple, le japonais est OV, avec les mots de fonction en position finale (*Tokyo kara* : de Tokyo), alors que le français est VO, avec les mots de fonction en position initiale (*de Paris*; Gervain et al., 2008). Bien qu'elles soient peu

nombreuses, il existe des langues mixtes, où les deux ordres sont possibles, comme l'allemand (Nespor et al., 2008; Smith, 1995). Il existe donc des variations spécifiques pour une langue donnée dans la détermination de l'ordre des mots. Sur quels indices les nourrissons vont-ils s'appuyer pour apprendre l'ordre des mots ?

Des travaux soutiennent l'idée que la proéminence donne un indice de l'ordre des mots d'une langue (Christophe & Dupoux, 1996; Christophe, Guasti, Nespor, Dupoux, & Van Ooyen, 1997; Mazuka, 1996; Nespor et al., 2008, 1996). Selon cette théorie, les nourrissons peuvent découvrir le paramètre de direction de tête sur la base des indices prosodiques présents dans le signal de parole. En effet, ces travaux postulent que la proéminence prosodique dans les groupes phonologiques ou l'accentuation des mots corrèle avec l'ordre des mots. De très récentes recherches rapportent des mesures systématiques de la correspondance acoustique de la proéminence (Nespor et al., 2008). Ces recherches montrent que l'ordre VO est accompagné d'un allongement final du syntagme, alors que l'ordre OV est relié à une augmentation du pitch et de l'intensité au début du syntagme. Ces résultats établissent que la réalisation acoustique de la proéminence change en fonction de l'ordre des mots. En turc par exemple, qui est une langue OV, la proéminence est à gauche, en début de syntagme, et elle est marquée par une fréquence fondamentale (pitch) et une amplitude (intensité) plus hautes sur l'objet. En français, qui est VO, la proéminence est à droite et elle est marquée par une plus longue durée. Dans des langues mixtes comme l'allemand, la localisation de la proéminence dans le syntagme varie en fonction de l'ordre de celui-ci.

Le langage parlé contient donc des variations prosodiques systématiques basées sur un certain nombre d'indices acoustiques. Ce lien étroit entre le signal de la parole et l'organisation syntaxique a permis de faire émerger de nombreuses hypothèses concernant les mécanismes impliqués dans l'acquisition du langage, notamment dans celle de l'ordre des mots. Comment les enfants font-ils pour exploiter ces indices dans le signal de parole ?

1.4.3.2 L'acquisition de l'ordre des mots.

Cette hypothèse de *bootstrapping* prosodique au niveau de la proéminence des syntagmes est soutenue par de nombreux travaux montrant que les nourrissons sont sensibles à ces différences prosodiques. En effet, nous avons vu précédemment que, dès la naissance, voire même avant, les nourrissons sont sensibles aux indices prosodiques qui indiquent les frontières des constituants syntaxiques majeurs. Dès 2 mois, les

nourrissons sont capables de discriminer des prosodies VO et OV. En effet, une étude a présenté à des nourrissons français de 6 à 12 semaines des phrases qui ont été resynthétisées pour enlever toutes les informations phonétiques, permettant ainsi d'isoler les caractéristiques prosodiques (Christophe, Nespoulet, et al., 2003). En mesurant les réponses des nourrissons à l'aide de la succion non-nutritive, les résultats montrent que les nourrissons sont capables de discriminer des syntagmes français (VO) de syntagmes turcs (OV), exclusivement sur la base de la proéminence prosodique. Des résultats similaires ont été obtenus par une équipe de recherche allemande, montrant que, pour le cas des langues mixtes comme l'allemand, des nourrissons de 14 mois s'appuient également sur les indices prosodiques pour discriminer les syntagmes allemands, qui diffèrent seulement par la position de la tête (Bion, Höhle, & Schmitz, 2007). En revanche, ces études ne permettent pas de savoir sur quels indices acoustiques s'appuient les nourrissons pour discriminer les patterns prosodiques des syntagmes, étant donné que tous les indices acoustiques étaient présents simultanément. Les préférences des nourrissons peuvent donc être dues à des variations de durée, d'intensité ou de pitch. Encore une fois, il est impossible de pouvoir déterminer le rôle respectif de chaque indice. Comment font les nourrissons pour détecter ces différences de proéminence prosodique ?

Des études ont essayé de présenter en isolation certains indices prosodiques en les couplant à des indices fréquentiels. Celle qui a été menée chez des nourrissons français de 8 mois a utilisé la technique du regard préférentiel (HPP), testant les représentations rudimentaires de l'ordre des mots (Bernard & Gervain, 2012). Pour cela, les auteurs ont familiarisé les nourrissons avec une langue artificielle. Cette phase de familiarisation était de deux types : soit les indices fréquentiels et prosodiques étaient en conflit (unités fréquentes de durée plus longue), soit ils étaient cohérents (unités fréquentes de durée plus courte). Les résultats ont montré que les nourrissons préfèrent la grammaire où la fréquence des mots et leur durée étaient en cohérence (par ailleurs également cohérente avec l'organisation de leur langue maternelle). En d'autres termes, ils préfèrent les séquences où les mots fréquents (correspondant aux mots de fonction du français) sont en position initiale, et les mots non-fréquents (correspondant aux mots de contenu) en position finale. En somme, les nourrissons de 8 mois exposés au français préfèrent les items structurés dans le même ordre que le français (Verbe-Objet), par opposition aux items structurés dans un ordre opposé (Objet-Verbe, comme en japonais par exemple), seulement lorsqu'il y a une cohérence entre la fréquence des mots et leur

prosodie (basée sur la durée). Ces recherches montrent donc que, dès 8 mois, les nourrissons français ont fixé le paramètre tête-complément de leur langue maternelle et sont sensibles aux patterns prosodiques de celle-ci.

Des résultats similaires ont été obtenus chez des nourrissons exposés à des langues différant par l'ordre des mots (Gervain & Werker, 2013). Les nourrissons bilingues sont également capables d'exploiter les informations prosodiques ajoutées aux indices fréquentiels. En effet, des nourrissons bilingues exposés à une langue OV (japonais, hindi, punjabi, farsi ou coréen) et une langue VO (anglais) préfèrent écouter les séquences où les mots fréquents sont en position initiale quand la séquence suit le pattern typique des langues VO (avec la durée comme indice prosodique), mais ils préfèrent les séquences où les mots fréquents sont en position finale quand ils sont exposés à une séquence qui suit le pattern typique des langues OV (avec le pitch comme indice prosodique). Ces travaux suggèrent que des nourrissons bilingues sont sensibles aux patterns prosodiques des syntagmes auxquels ils sont exposés et peuvent exploiter ces indices prosodiques relatifs à l'ordre des mots de leur(s) langue(s) maternelle(s). Toutefois, bien que ces deux dernières études aient isolé les indices acoustiques, des indices fréquentiels y étaient couplés, pouvant ainsi être une information supplémentaire.

L'ensemble de ces travaux suggère que les nourrissons de 8 mois peuvent potentiellement s'appuyer sur les patterns prosodiques pour apprendre certaines propriétés syntaxiques de leur langue maternelle, notamment l'ordre basique des mots.

But de la thèse

Des études ont montré que les éléments contrastant en intensité ou en pitch sont groupés par paires, avec l'élément le plus aigu/fort sur la partie initiale, et les éléments contrastant en durée sont groupés par paires, avec l'élément le plus long sur la partie finale (Nespor et al., 2008; Woodrow, 1909). Ce principe a été énoncé comme la loi iambo-trochaïque (ITL) par Hayes (1995). Étant donné qu'une intensité plus forte et/ou pitch plus élevé, ainsi qu'un allongement de durée, corrélerent avec le pattern accentuel des syllabes dans les mots et le pattern accentuel au niveau des syntagmes, il est possible que l'ITL joue un rôle dans le groupement des syllabes en mots, et le groupement des mots en syntagmes. Cette hypothèse est d'autant plus probable que le pitch, la durée et l'intensité sont des indices acoustiques universaux de la prosodie linguistique ; les préférences peuvent alors émerger de la manière dont notre système perceptuel perçoit et organise les variations de ces corrélats acoustiques de l'accentuation. L'ITL pourrait donc guider le processus de segmentation de telle façon que les séquences contrastant en intensité et/ou pitch seraient groupées de façon trochaïquement et les séquences contrastant en durée seraient groupées de façon iambique.

Or, les études précédentes ne nous renseignent pas sur les capacités des nourrissons à percevoir le groupement de syllabes en paires en fonction de la durée, de l'intensité ou encore du pitch, puisque toutes ces dimensions étaient mélangées. Dans cette thèse, nous allons essayer de comprendre quels sont les indices acoustiques qui pourraient être impliqués dans l'acquisition du biais trochaïque en explorant la question du groupement perceptuel et l'application de la loi iambo-trochaïque pour chaque indice. Mais quelle est cette loi et quelle est son origine ?

1.5 La loi iambo-trochaïque (ITL)

1.5.1 Définition

Nous avons vu que les nourrissons sont sensibles aux indices prosodiques de la parole dès la naissance. Des biais dans le traitement et lors de l'acquisition des patterns prosodiques de la langue maternelle semblent être présents pendant la première année de vie du nourrisson. Ces biais prosodiques vont leur permettre d'apprendre les mots et l'ordre rudimentaire des mots de leur langue maternelle. Un biais perceptuel qui peut essayer d'expliquer ou de répondre à cette question est celle du groupement rythmique. Les questions centrales de cette thèse concernent la manière dont ces biais prosodiques peuvent émerger.

Les premières observations sur le groupement perceptuel ont été réalisées lors de la segmentation d'une séquence auditive, composée de sons identiques, en groupant les éléments. Ce phénomène peut être illustré par de nombreux exemples de la vie quotidienne, notamment lors de la perception d'une séquence de sons identiques, comme une suite de paires de sons comme le pin-pon d'une sirène d'ambulance ou encore le tic-tac d'une horloge (Bolton, 1894). Or, en fonction de l'indice acoustique variant dans la séquence, le groupement induit ne va pas suivre la même organisation. Ces principes de groupement ont été proposés dans un premier temps pour la perception auditive de tons musicaux (Bolton, 1894; Cooper & Meyer, 1960; Woodrow, 1951) : les séquences alternant en intensité sont groupées avec l'élément le plus fort en position initiale (groupement trochaïque) et les séquences variant en durée sont groupées avec l'élément le plus long en position finale (groupement iambique). Ce biais perceptif (voir Figure 1.5) a ensuite été étendu à des sons langagiers et définit comme la loi Iambo-trochaïque (ITL : *Iambic Trochaic Law*) par Hayes (1995). En effet, l'ITL a d'abord été proposée pour expliquer les différentes manières dont les mots peuvent être accentués. Dans la production de la parole, l'ITL a été prise en compte pour expliquer la position de l'accentuation à travers les langues. Si l'accentuation est en position finale, alors elle est souvent marquée par une augmentation de la durée. Si elle est en position initiale, elle est alors marquée par une augmentation d'intensité.

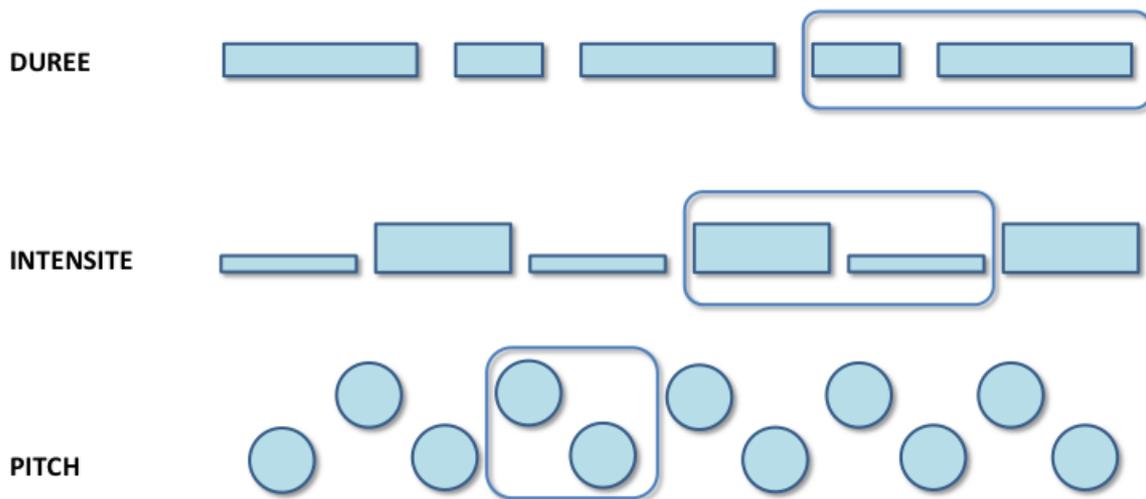


Figure 1.5 : Schéma de la loi iambo-trochaïque pour les trois indices acoustiques (durée, intensité, pitch).

1.5.2 Modulation par l'environnement linguistique

Cependant, comme nous l'avons vu dans ce chapitre introductif, chaque langue possède des propriétés prosodiques intrinsèques. Des recherches ont montré que l'accentuation dans les langues au niveau des mots et des syntagmes repose sur des différences de durée, d'intensité et de pitch (Beckman & Pierrehumbert, 1986; Delattre, 1966; Dogil & Williams, 1999; Fry, 1955; Jessen, Marasek, Schneider, & Classen, 1995). Toutefois, les langues diffèrent notamment dans leur façon d'utiliser ces indices pour signaler l'accentuation des mots et les relations hiérarchiques entre les mots, notamment entre le français et l'allemand.

En effet, au niveau du mot, l'allemand est prosodiquement organisé en pieds trochaïques. La majorité des mots allemands dissyllabiques ont une accentuation initiale (Féry, 1998; Wiese, 1996) qui est marquée par une augmentation de la durée et de l'intensité de la voyelle. Les contrastes de durée sont souvent plus faibles que ceux de l'intensité (Delattre, 1966; Dogil & Williams, 1999; Jessen et al., 1995). Les syllabes accentuées sont parfois accompagnées d'un accent de pitch (Jessen et al., 1995). En revanche, le français n'a pas d'accentuation lexicale.

Au niveau du syntagme, l'accentuation n'est pas fixe en allemand, mais dépend de la position de la syllabe accentuée du mot (Atterer & Ladd, 2004; Dogil & Williams, 1999; Féry, Hörnig, & Pahaut, 2011), et chaque syntagme est organisé autour d'un accent de pitch qui est associé à l'accentuation lexicale (Féry et al., 2011). Ces accents

de pitch sont souvent marqués par des contours montants (Grabe, 1998), mais dans les phrases déclaratives, ils sont descendants et en positions finales (Dogil & Williams, 1999). De plus, leur structure varie en fonction du contexte pragmatique, discours, etc. (Féry & Kügler, 2008). La durée a un rôle mineur dans l'accentuation du syntagme (Féry et al., 2011). En allemand comme en français, l'intensité est le marqueur le moins pertinent de l'accentuation du syntagme. Cependant, l'interaction entre intensité et pitch est différente dans les deux langues. En allemand, l'intensité et la F0 marquent généralement ensemble la syllabe accentuée (plus haute et plus forte), alors qu'en français l'intensité et la F0 peuvent être dissociées (Nespor et al., 2008; Vaissière & Michaud, 2006). De plus, en français, « l'accentuation » au niveau du syntagme est systématiquement en position finale (Delattre, 1966; Jun & Fougeron, 2002). Cette accentuation principale est acoustiquement marquée par la durée ainsi qu'un mouvement de pitch : la syllabe finale est toujours plus longue et est marquée par une augmentation ou une baisse de la fréquence fondamentale (Féry et al., 2011; Jun & Fougeron, 2002; Michelas & D'Imperio, 2010). Il peut y avoir de manière optionnelle une accentuation secondaire, qui peut apparaître sur la partie initiale du syntagme qui contient plus de deux syllabes (et rarement dans les syntagmes avec seulement deux mots de contenus). Cette accentuation secondaire est marquée par une augmentation de F0 sans allongement (Jun & Fougeron, 2002) et le pic n'est pas associé à une syllabe en particulier, mais peut apparaître sur la première, seconde ou plus rarement troisième syllabe du syntagme (Jun & Fougeron, 2002; Vaissière & Michaud, 2006; Welby & Loevenbruck, 2006). L'intensité n'est pas un indice pertinent pour marquer l'accentuation au niveau du syntagme (Delattre, 1966).

Ces particularités semblent nous montrer une utilisation des signaux acoustiques différente en fonction des langues. Dans ce contexte, la perception de la loi iambotrochaïque (ITL) est-elle semblable pour un locuteur germanophone ou francophone? Si elle diffère, se posent alors différentes problématiques : comment ? De quelle manière? À partir de quel âge ? Ces différences se limitent-elles au domaine du langage, ou influencent-elles aussi la perception musicale ?

1.5.3 L'ITL chez l'adulte

Une étude récente suggère que le traitement du rythme est déterminé par l'ITL d'une façon qui n'est pas spécifique à l'environnement linguistique (Hay & Diehl,

2007). En effet, lors de l'écoute de chaînes de syllabes /ga/ répétées, variant en intensité ou en durée (le pitch étant stable dans les deux types de séquences), les sujets adultes, américains et français, perçoivent de la même manière le groupement : trochaïque pour les variations d'intensité, iambique pour les variations de durée. Cependant, des effets cross-linguistiques sont retrouvés par certains auteurs lors de l'écoute de sons alternant soit en durée, soit en intensité, soit en pitch : entre des japonais et des anglais (Iversen, Patel, & Ohgushi, 2008) et entre des allemands et des français (Bhatara, Boll-Avetisyan, Unger, Nazzi, & Höhle, 2013). Ces données semblent remettre en question l'universalité de la perception acoustique du groupement suivant l'ITL. Cependant, ces études ont été réalisées avec des stimuli différents : des sons linguistiques pour Hay & Diehl (2007) et Bhatara et al. (2013) et des sons non-linguistiques pour Iversen et collaborateurs (2008). C'est pourquoi la question de différences possibles dans la perception de son de parole (Hay & Diehl) et des sons non-linguistiques (Iversen et al., 2008) se pose. De plus, des différences cross-linguistiques de traitement ont été trouvées pour des sons linguistiques entre adultes français et allemands (Bhatara et al., 2013), mais essentiellement pour des stimuli langagiers complexes (combinaison de différentes syllabes) et pas si les séquences présentées sont constituées d'une seule syllabe répétée, comme dans l'étude de Hay & Diehl. Par ailleurs, les auteurs ont présenté les séquences en mélangeant au sein des mêmes blocs de test les deux types de variations : intensité et durée dans l'Exp. 1; pitch et durée dans l'Exp. 2. Des résultats différents ont été obtenus pour la durée dans ces deux expériences, suggérant des effets modulateurs du contexte. Ces résultats révèlent que la perception universelle de l'ITL pourrait donc dépendre de la complexité du matériel utilisé, de la procédure ou encore de la langue apprise.

Par ailleurs, des questions concernant le développement de ces biais se posent. C'est pourquoi de nombreux chercheurs se sont intéressés à l'étude de la perception de l'ITL chez le nourrisson, afin de comprendre comment se met en place cette compétence, si elle est innée et/ou si elle est modulée par l'environnement linguistique.

1.5.4 L'ITL chez les nourrissons

La première étude sur l'ITL chez les nourrissons révèle que, dès 8 mois, les bébés traitent l'allongement de durée comme un indice de fin d'unité, alors que l'intensité n'est pas perçue comme un indice qui permet de grouper les sons (Trainor &

Adams, 2000). Des études plus récentes, portant sur l'écoute de sons variant en termes de pitch et de durée chez des nourrissons italiens de 7 mois, observent des effets de groupement pour le pitch, mais pas pour la durée (Bion, Benavides-varela, & Nespor, 2011). D'autres études n'ont testé que la durée chez des nourrissons de 7-8 mois et trouvent un effet cross-linguistique : les nourrissons anglais perçoivent le groupement selon l'ITL pour la durée, mais pas les japonais (Yoshida et al., 2010). Une récente étude montre des patterns de résultats similaires pour l'indice d'intensité chez des nourrissons de 9 mois bilingues espagnol-basque (Molnar, Lallier, & Carreiras, 2014) ; toutefois, en fonction de la dominance à l'espagnol ou au basque, des résultats différents sont retrouvés pour la durée.

D'une part, l'ensemble de ces travaux révèle des sensibilités très précoces à l'ITL, émergeant autour de 7-9 mois de vie, mais les résultats diffèrent d'une étude à l'autre. D'autre part, des modulations cross-linguistiques, en termes de sensibilité du groupement ITL, existent pour certains indices et dans certaines langues, mais il est encore difficile de comprendre de quelle manière. En effet, ces études diffèrent en termes de matériel utilisé : alors que Trainor & Adams (2000), Yoshida et al. (2010) et Molnar et al. (2014) ont utilisé des tons, Bion et al. (2011) ont utilisé des syllabes. De plus, les langues auxquelles ont été exposés les nourrissons, ainsi que le dessin expérimental utilisé, ne sont pas identiques d'une étude à l'autre. Toutes ces différences sont autant de facteurs qui ne permettent pas de trouver de consensus, tant sur la période d'émergence du biais que sur sa modulation par l'environnement linguistique. A ce jour, aucune étude n'a étudié les trois dimensions acoustiques dans une même expérience et en comparant des enfants exposés à des langues différentes.

1.6 Objectif et structure de la thèse

La première partie de cette thèse va donc explorer la présence de l'ITL ainsi que sa modulation par l'environnement linguistique chez l'adulte (Chapitre Expérimental 1, Exp. 1) et le nourrisson de 7,5 mois (Chapitre Expérimental 2, Exp. 2). Nous allons pour la première fois étudier les trois indices acoustiques séparément. Nous utiliserons un matériel linguistique que nous présenterons dans une tâche similaire de groupement/segmentation aux deux groupes linguistiques (allemand versus français) et aux deux groupes d'âge (adultes et nourrissons de 7,5 mois).

De plus, la seule étude qui s'est intéressée à des nourrissons plus jeunes a porté sur des nourrissons de 5-6 mois et n'a pas mis en évidence de perception de l'ITL. D'autres tranches d'âges, chez des nourrissons plus jeunes, notamment de 0 à 4 mois, n'ont pas encore été testées expérimentalement, ce qui laisse des questions ouvertes sur la trajectoire développementale précoce du biais prosodique. Il n'est pas impossible que la perception de l'ITL suive une trajectoire développementale en forme de U, avec un stade initial universel qui disparaîtrait avant l'âge de 5 mois, pour se réorganiser et réapparaître comme une conséquence de l'expérience linguistique à partir de 7-8 mois. Des expériences à la naissance sont donc nécessaires pour clarifier ces hypothèses. C'est ce sur quoi porte le Chapitre Expérimental 3, lequel explore les capacités perceptives à la naissance : sont-elles déjà structurées par la loi iambo-trochaïque à la naissance, pour la durée, l'intensité et le pitch (Exp. 3-5) ? De quelle manière peuvent-elles déjà être modulées par l'expérience prénatale (Exp. 6) ?

Pour finir, après avoir fait la revue des processus sous-tendant les biais prosodiques dans le traitement du langage, nous allons explorer le développement de préférences de traitement pour certains patterns rythmiques/accentuels, en fonction de l'exposition linguistique reçue pendant les 10 premiers mois de vie (Exp. 7). Le dernier chapitre expérimental consistera donc à étudier le développement de ces biais de préférence, cette fois-ci dans un contexte où tous les indices acoustiques seront présents, chez des nourrissons de 10 mois monolingues français et des bilingues exposés au français et à une langue à accentuation lexicale variable.

2 ITL ET SEGMENTATION CHEZ L'ADULTE ET LE NOURRISSON

2.1 CHAPITRE 1 : Comment l'ITL guide le processus de segmentation chez l'adulte ?

2.1.1 Introduction

2.1.1.1 *La loi iambo-trochaïque (ITL)*

Comme nous l'avons vu dans le chapitre introductif, la loi iambo-trochaïque (ITL) stipule que lors de l'écoute de séquences d'éléments contrastant en intensité, les sons sont automatiquement groupés trochaïquement (fort-faible), alors que des sons contrastant en durée sont groupés iambiquement (court-long) (Hayes, 1995; Woodrow, 1909). Très récemment, des travaux ont suggéré que le pitch pourrait avoir un rôle similaire à l'intensité (Bion et al., 2011; de la Mora, Nespor, & Toro, 2013; Nespor et al., 2008).

Les études chez l'adulte validant la formulation originelle de l'ITL (durée et intensité) ont été réalisées dans différents contextes. En effet, deux types de recherches ont étudié la perception de l'ITL : soit pour mettre en évidence la présence de l'ITL (anglais: Bolton, 1894, Trainor & Adams, 2000, Woodrow, 1909; italien : Bion et al., 2011; espagnol : de la Mora et al., 2013), soit pour étudier sa modulation en fonction de la langue maternelle des participants (Bhatara, Boll-Avetisyan, Agus, Höhle, & Nazzi, sous presse; Bhatara et al., 2013; Hay & Diehl, 2007; Iversen et al., 2008; Kusumoto & Moreton, 1997). Les résultats issus de ces travaux ont révélé des modulations dans la perception de l'ITL qui peuvent être expliquées en partie par la langue maternelle des locuteurs, mais également en fonction du matériel ainsi que des méthodes utilisées. Dans ce qui suit nous allons discuter des facteurs affectant la perception de l'ITL en détaillant ce que l'on sait pour chacun de ces trois indices.

2.1.1.2 *Modulation par l'environnement linguistique*

Seules 5 études cross-linguistiques entre différents locuteurs ont été menées et elles différaient sur le type de matériel utilisé (linguistique versus non linguistique). Les quatre études qui ont été réalisées avec des stimuli non-linguistiques ont révélé des disparités entre les locuteurs. Alors que des différences cross-linguistiques concernant les performances du jugement rythmique ont été mises en évidence entre anglais et japonais (Iversen et al., 2008; Kusumoto & Moreton, 1997) ou entre allemands et

français (Bhatara et al., 2013, sous presse), aucune n'a été retrouvée entre des anglais et des français (Hay & Diehl, 2007).

Dans l'étude de Hay & Diehl (2007), des locuteurs français ou anglais ont écouté des séquences de sons, soit non-linguistiques (comme dans Iversen et al., 2008), soit linguistiques, qui variaient soit en intensité, soit en durée. Les séquences de sons non-linguistiques consistaient en une répétition de bips complexes et celles linguistiques étaient composées de la répétition d'une seule syllabe synthétisée /ga/. Tous les participants, quelle que soit leur langue (anglais ou français) et quel que soit le type de matériel (linguistique ou non linguistique), ont eu des performances semblables. Ils rapportaient, lors de l'écoute de séquences alternant en intensité, qu'elles étaient composées de paires trochaïques (fort-faible), et lors de l'écoute des séquences variant en durée qu'elles étaient composées de paires iambiques (court-long). Ces données indiquent que, dans certaines conditions expérimentales, l'ITL ne serait pas modulée par la langue maternelle, tout du moins ici entre le français et l'anglais.

Bhatara et al. (2013) ont proposé que ce manque de variation cross-linguistique puisse être en partie relié au matériel et la procédure utilisée par Hay & Diehl (2007). En effet, dans l'expérience de Hay & Diehl (2007), des séquences composées d'une seule syllabe ou d'un seul bip répété, séparées par des pauses, ont été présentées aux participants. De plus, leurs stimuli étaient présentés par bloc, avec le premier variant en durée et le deuxième variant en intensité, ce qui a pu diminuer la difficulté de la tâche. Or, il est probable que traiter des séquences de stimuli relativement plus complexes, en l'occurrence de nature langagière, implique une demande plus importante en terme de ressource attentionnelle et de mémorisation. Il est donc possible, tout du moins dans ce type de tâche plus complexe, qui demande de traiter les corrélats acoustiques de l'accentuation, que les participants puissent faire appel à des représentations linguistiques internes et non à un traitement purement acoustique (Dupoux et al., 1997, 2001). Ce manque de différence cross-linguistique trouvée par Hay & Diehl (2007) peut donc refléter un mécanisme reposant sur des compétences plus générales. Donc, si la complexité acoustique des stimuli est plus élevée, alors des effets plus clairs de l'expérience de la langue maternelle peuvent être observés. Bhatara et al. (2013) ont donc présenté des séquences composées de 16 syllabes différentes (issues de la combinaison de différentes consonnes et voyelles), agencées sans pauses, ainsi que de différents types de variations de durée et d'intensité (Exp.1) ou de durée et de pitch

(Exp.2). Ces différents paramètres pourraient permettre d'avoir un matériel plus écologique possédant un niveau de complexité plus élevé que celui de Hay et Diehl, forçant les participants à traiter les informations à un niveau plus linguistique, plus abstrait et moins acoustique.

Une autre possibilité, non exclusive à la première, qui pouvait expliquer ce manque de différence cross-linguistique dans l'étude de Hay & Diehl (2007) est que le français et l'anglais ont un ordre des mots similaire. Bhatara et al. (2013) ont choisi de comparer des locuteurs allemands et français, puisque ces deux langues diffèrent au niveau du mot et également au niveau de l'ordre des mots dans les syntagmes. A la vue de ces différences prosodiques, les français et les allemands pourraient donc avoir des performances différentes lors de la perception de l'ITL.

Ainsi, en présentant des séquences linguistiquement plus complexes à des locuteurs allemands et français, Bhatara et collègues (2013) ont observé des différences claires entre les deux groupes linguistiques dans la perception du groupement rythmique. Par rapport aux français, les allemands ont plus souvent rapporté des réponses trochaïques lors de l'écoute de séquences variant en intensité et en pitch, et des réponses iambiques lors de l'écoute des séquences qui variaient en durée. Ces résultats mettaient en évidence un clair avantage des allemands dans les performances de jugements de l'ITL par rapport aux français. Par ailleurs, une étude complémentaire utilisant cette fois-ci des stimuli non-linguistiques complexes retrouvaient les mêmes résultats entre des français et des allemands (Bhatara et al., en révision). Ces études remettent donc en question la position de Hay & Diehl (2007) sur un simple biais général auditif.

Comme nous l'avons vu dans l'introduction générale, la modulation de l'ITL a été reliée à deux niveaux prosodiques : lexical ou syntagmatique. Une des hypothèses qui a été proposée par Bhatara et al. (2013), concerne l'existence patterns trochaïques au niveau du mot et de la réalisation de l'accentuation du syntagme par le pitch et l'intensité qui existe en allemand, mais pas en français. Cette différence pourrait expliquer la perception plus stable de l'intensité et du pitch chez les locuteurs allemands. De plus, la tâche utilisée dans cette étude était basée sur la catégorisation prosodique des séquences, ce qui a pu rendre la tâche plus difficile pour les français. En effet, de nombreux travaux ont montré que les adultes français, et même les nourrissons,

ont des difficultés à percevoir l'accentuation lexicale marquée justement par l'augmentation de l'intensité ou encore du pitch (Abboub, Bijeljac-Babic, Serres, & Nazzi, 2015; Bijeljac-Babic et al., 2012; Dupoux et al., 2001, 1997; Skoruppa et al., 2009). Concernant la durée, Bhatara et collaborateurs (2013) ont proposé que la manière dont ils ont présenté les séquences (en faisant varier des différences faibles de durée par séquence) ait pu avantager les allemands, puisque les variations prosodiques dans cette langue sont plus faibles qu'en français. Ces résultats ont donc mis pour la première fois en évidence l'impact de la langue maternelle sur la perception de l'ITL entre les allemands et les français.

Si Bhatara et al. (sous presse, 2013) ont mis en évidence l'apparition de différences cross-linguistiques, mais seulement pour un matériel acoustiquement complexe entre des allemands et des français, deux autres études sur la perception du groupement rythmique ont trouvé des différences pour un matériel plus simple mais avec des langues prosodiquement plus éloignées (Iversen et al., 2008; Kusumoto & Moreton, 1997). Ils ont présenté à des participants anglais et japonais des séquences de tons non-linguistiques complexes. Dans ces études, chaque séquence alternait selon le pitch ou la durée (Kusumoto & Moreton, 1997) et l'intensité ou la durée (Iversen et al., 2008). Les participants devaient indiquer s'ils avaient perçu les séquences selon un groupement trochaïque ou un groupement iambique. Les auteurs ont rapporté que, quelle que soit leur langue maternelle, les participants ont perçu lors de l'écoute de séquences variant en intensité un pattern trochaïque (fort-faible), donc en accord avec l'ITL. Cependant, pour la durée, des paires iambiques (court-long) comme énoncées par l'ITL n'ont été perçues que par les anglais. Des réponses plus hétérogènes ont été retrouvées chez les japonais, avec une tendance à percevoir un groupement trochaïque, allant cette fois-ci à l'inverse des prédictions de l'ITL. Pour le pitch, Kusumoto & Moreton (1997) ont rapporté seulement des tendances de groupement trochaïque chez les anglais et certains locuteurs japonais (kansai), mais aucune pour les locuteurs japonais tokyoïtes. Ils ont suggéré que le pitch serait un indice moins stable pour grouper les sons (Rice, 1992). Pour la durée, les auteurs ont expliqué que cette modulation cross-linguistique pouvait être due principalement à des différences prosodiques au niveau de l'ordre des mots dans le syntagme et également au niveau du mot. L'anglais possède, au niveau du mot, une accentuation lexicale marquée par la durée, l'intensité et le pitch, et généralement sur la première syllabe des mots (Cutler,

1986), alors qu'en japonais les accents de pitch ne sont pas marqués par la durée ou l'intensité et sont généralement situés sur la partie finale des mots (Beckman, 1986). De plus, au niveau de l'ordre des mots, l'anglais a une structure syntaxique VO (verbe-objet), alors que le japonais suit un pattern OV (objet-verbe). Ces deux facteurs pourraient ainsi moduler le jugement du groupement rythmique, notamment celui de la durée.

Un autre facteur qui a pu influencer le jugement du groupement rythmique est la manière dont les auteurs ont présenté les séquences de sons. Dans l'expérience de Bhatara et collaborateurs, les auteurs ont présenté les indices en mélangeant les types de séquence (Exp.1 durée et intensité ; Exp.2 durée et pitch) et n'ont pas obtenu exactement le même type de résultats pour la durée dans l'Expérience 1 et 2. En effet, pour la durée, les allemands ont rapporté des réponses moins stables dans l'Expérience 2 et les français ne montraient plus d'effet de groupement. Le contexte de présentation semble donc influencer la perception du groupement rythmique.

Une étude très récente a exploré le rôle de l'ITL dans la perception du langage avec une nouvelle tâche de groupement (Bion et al., 2011). Ces auteurs ont étudié deux indices séparément, le pitch et la durée, et les ont présentés à des adultes et des nourrissons de 7,5 mois italophones. Ils ont choisi de ne pas tester l'intensité, en se justifiant par le fait que des recherches montraient que les nourrissons étaient sensibles aux variations de pitch et de durée, mais pas d'intensité (Singh, White, Morgan, & White, 2008). La méthode et le matériel de l'étude de Bion et al. (2011) diffèrent sur plusieurs aspects des précédentes. Ils ont basé leur tâche sur la mémorisation d'une séquence, dans laquelle ils ont manipulé les indices prosodiques. A l'inverse de Bhatara et al. (2013) et Hay & Diehl (2007), ils ont séparé la présentation de chacun de ces indices pour les étudier isolément. Les participants ont entendu des séquences de familiarisation pouvant être de trois types, variant soit en durée, soit en pitch, soit n'ayant pas de modulation prosodique. Ces séquences étaient composées de 10 paires de syllabes, séparées par des pauses entre chaque syllabe (voir Figure 2.1). Après cette phase de familiarisation (1mn d'écoute), ils ont mesuré comment les participants ont retenu cette séquence, si certaines combinaisons ont été mieux mémorisées que d'autres. Les participants ont mieux mémorisé les paires qui suivaient un pattern trochaïque (aigu-grave) lors de l'écoute de la séquence variant en pitch, et un pattern iambique (court-long) lors de l'écoute de la séquence variant en durée. Ces résultats mettent donc

en évidence une perception du groupement rythmique basée sur le pitch et la durée chez des locuteurs italiens.

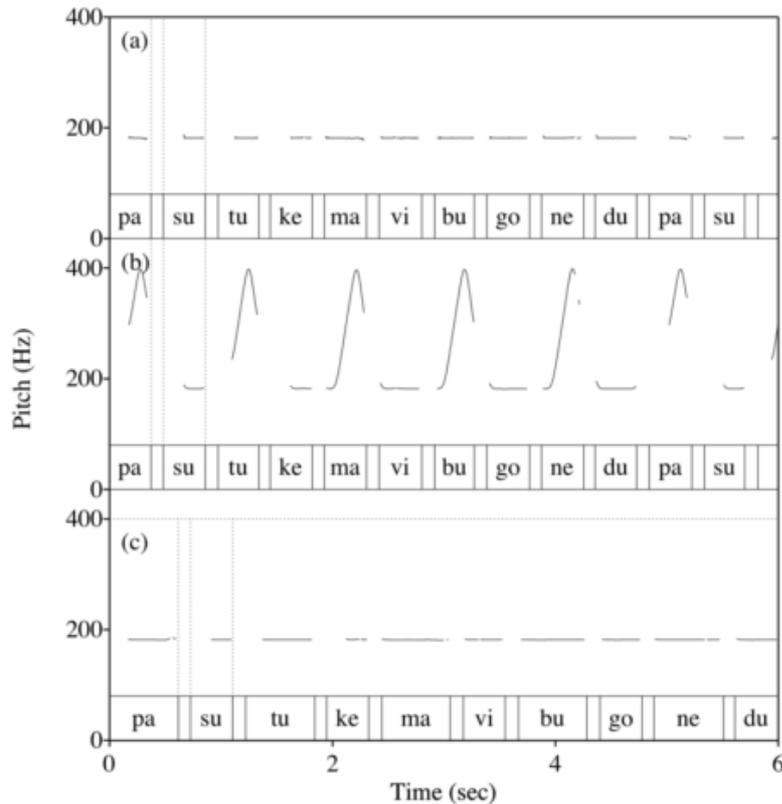


Figure 2.1: Un extrait de la phase de familiarisation de l'étude de Bion et al., 2011, pour les trois conditions (a) condition contrôle, (b) condition pitch, (c), condition durée

Dans les études précédentes (Bhatara et al. 2013 ; sous presse ; Hay & Diehl, 2007 ; Iversen et al., 2008), les auteurs ont demandé aux participants si la séquence présentée était de type iambique ou trochaïque, ce qui a pu influencer les réponses, étant donné qu'on leur demandait explicitement une catégorisation prosodique. Ici Bion et al. (2011) ont mesuré indirectement la perception du groupement en observant si les sujets mémorisaient mieux les unités linguistiques suivant le groupement prédit par l'ITL que celles qui ne le suivaient pas. Cette méthode permet de mesurer plus implicitement le groupement prosodique. D'ailleurs, de nombreuses études utilisant des tâches de segmentation en vision, ont montré que la mémorisation d'évènements est une mesure implicite robuste du processus de segmentation (Kurby & Zacks, 2008; Swallow & Zacks, 2008). De plus, les travaux de Dupoux et al. (1997, 2001) rapportent que des tâches faisant appel à la mémoire de travail sur l'étude de la perception de

l'accentuation permettent d'avoir des mesures plus fiables. L'utilisation de ce type de mesures indirectes permet aussi de comparer au mieux les résultats des adultes et des nourrissons, ce qui est impossible à faire avec la tâche de Bhatara et al. (2013) ou Hay & Diehl (2007), qui demande un jugement catégorique clair. Pour des raisons évidentes, il est impossible de recueillir ce type de réponse chez des nourrissons.

Ces observations prises ensemble suggèrent que, non seulement la langue maternelle du participant, mais également le contexte de présentation des indices et la méthode, peuvent moduler la stabilité du jugement rythmique.

2.1.1.3 But et hypothèses

L'objectif de notre étude va être d'évaluer comment l'ITL reflète un biais général auditif qui peut être modulé par l'expérience linguistique. En se référant aux expériences réalisées sur la perception de l'ITL, nous avons décidé de comparer des locuteurs natifs allemands et français. Comme nous l'avons décrit dans le chapitre introductif, le français et l'allemand diffèrent principalement au niveau de la prosodie des mots. L'allemand possède une accentuation lexicale alors que ce n'est pas le cas du français (Féry et al., 2011). De plus, ces deux langues sont différentes au niveau des syntagmes ; l'ordre des mots est mixte en allemand et peut suivre à la fois l'ordre VO et OV, alors que le français est VO (Nespor et al. 2008). Notre étude va donc s'intéresser à la perception de l'ITL en séparant les indices acoustiques et à l'aide d'une méthode plus adaptée pour comparer les performances des adultes et des nourrissons.

Comme Bion et al. (2011), nous voulions présenter le même matériel à des nourrissons et avons donc adopté leur méthode. Toutefois, nous avons décidé de modifier certains paramètres de leur méthode. Premièrement, nous avons modifié le nombre de syllabes, puisque nous voulions avoir un matériel très similaire entre les adultes et les nourrissons (Bion : 10 pour les adultes et 6 pour les nourrissons, notre étude : 8 pour les adultes et 6 pour les nourrissons). Deuxièmement, étant donné que la complexité acoustique semble être un paramètre important pour faire émerger des différences cross-linguistiques (Bhatara et al., 2013, sous presse), nous avons retiré les pauses entre les syllabes et augmenter le nombre de syllabes. Troisièmement, nous avons ajouté la condition de l'intensité, afin de pouvoir tester tous les indices énoncés par l'ITL. Ces changements nous donnent ainsi la possibilité d'étendre les travaux de Bion et al. (2011), réalisés chez des italiens, à des locuteurs de langue maternelle différente (allemands et français), mais aussi ceux de Bhatara et al (2013), en étudiant

séparément les indices acoustiques avec une méthode plus adaptée pour comparer nos deux groupes.

Sur la base des résultats obtenus par Bion et al. (2011) et Bhatara et al. (2013, sous presse), si les adultes ont perçu l'ITL, alors ils grouperont les séquences de syllabes qui variaient en durée de façon iambique, et les séquences qui variaient en intensité ou en pitch de façon trochaïque. De plus, si notre matériel et notre tâche sont assez complexes, des différences cross-linguistiques émergeront. Selon les résultats obtenus par Bhatara et al (2013), des jugements plus stables de l'ITL seront retrouvés chez les allemands.

2.1.2 Expérience 1

2.1.2.1 *Méthode*

2.1.2.1.1 Participants

Nous avons recruté 64 participants adultes monolingues français par l'intermédiaire du RISC (Relais d'Information sur les Sciences de la Cognition) et d'affiches au sein de l'Université Paris Descartes. De plus, 64 monolingues allemands ont également été testés à Potsdam en Allemagne, via un recrutement sur le campus de l'Université de Potsdam. Dans chaque groupe de langue, 16 participants ont été inclus pour chaque condition (détails dans Tableau 2-1). Leur statut de monolingue a été vérifié à l'aide d'un questionnaire (basé sur le LEAP-Q ;Marian, Blumenfeld, & Kaushanskaya, 2007).

Tableau 2-1 : Détails des informations des participants par groupe.

Condition	Monolingues français				Monolingues Allemands			
	Durée	Intensité	Pitch	Contrôle	Durée	Intensité	Pitch	Contrôle
Femme	9	11	11	12	13	12	15	14
Age moyen (ans)	27.3	26.7	28.7	32.7	22,4	22,3	24,1	24,4
Écart type (ans)	7.3	7.9	9.1	9.5	3,4	3,3	8,5	5,2
Intervalle (ans)	[20-44]	[18-45]	[19-47]	[20-46]	[18-30]	[17-28]	[18-51]	[20-37]

2.1.2.1.2 *Stimuli*

Pour la phase de familiarisation, nous avons synthétisé une séquence composée de 8 syllabes dont les variations prosodiques entre syllabes consécutives différaient en fonction de la condition. Elles alternaient en durée (courte – longue) dans la condition (1), en intensité (forte - faible) dans la condition (2), en pitch (grave – aigu) dans la condition (3). Enfin, dans la condition contrôle (4), toutes les syllabes avaient les mêmes durées, intensités et pitch.

Le choix de construction et de la nature de ce matériel s'est basé sur différentes études utilisant également un langage artificiel (Bhatara et al., 2013; Tyler & Cutler, 2009). La séquence est composée de huit syllabes CV différentes, obtenues en combinant quatre voyelles /e/, /i/, /o/, /a/ à huit consonnes /f/, /n/, /g/, /p/, /z/, /r/, /v/, /t/. Cet ensemble de phonèmes a été sélectionné pour plusieurs raisons. Premièrement, il fait partie de l'inventaire des phonèmes en commun entre le français et l'allemand. Ainsi, bien que certains phonèmes ne soient probablement pas perçus de manière similaire, la variabilité segmentale des stimuli devrait être identique pour les deux groupes. Deuxièmement, les voyelles et consonnes varient en terme de traits phonologiques (voyelles : arrondissement, hauteur et place ; consonnes : voisement, manière et place d'articulation). Troisièmement, nous avons, pour éviter les mots monosyllabiques, choisi des consonnes peu fréquentes dans le lexique du français (LEXIQUE, New et al., 2001) et de l'allemand (CELEX, Baayen et al., 1995). Ces syllabes ont été agencées de manière à éviter de former des mots bi-syllabiques allemands ou français, ou encore de contenir des syllabes adjacentes avec la même voyelle, soit la séquence : **/fo na gi pe za ro ve ti /**.

2.1.2.1.2.1 *Construction de la séquence de familiarisation*

Cette séquence a été répétée 20 fois (environ 1 minute). Nous avons synthétisé cette séquence de familiarisation à l'aide de MBROLA (Dutoit, Pagel, Pierret, Bataille, & van der Vrecken, 1996), dans deux versions, une avec une voix féminine allemande (de5) et une avec une voix féminine française (fr4).

Nous avons ensuite manipulé les caractéristiques acoustiques de cette séquence en fonction de la condition, à l'aide de PRAAT (Boersma & Weenink, 2010). Les valeurs de durée, d'intensité ou de pitch (Tableau 2-2) ont été choisies sur la base de

celles utilisées dans la littérature. Ainsi, nous nous sommes principalement basés sur l'étude de Bion et al. (2011) et de Bhatara et al. (2013).

Tableau 2-2 : Valeurs des trois indices acoustiques pour chaque type de séquence de familiarisation et pour les items tests.

	Syllabe	PITCH	INTENSITE	DUREE
FAMILIARISATION	1. DUREE	200Hz	70 dB	260-460 ms
	2. INTENSITE	200Hz	66-74 dB	360 ms
	3. PITCH	200- 420 Hz	70 dB	360 ms
	4. CONTROLE	200Hz	70dB	360ms
TEST	Items test	200Hz	70dB	360ms

Effectivement, le choix de ces valeurs repose à la fois sur des mesures acoustiques spécifiques de l'accentuation de l'allemand et du français (Bhatara et al., 2013), mais également sur les choix faits dans la procédure originale de Bion et al. (2011), afin de rester au plus proche de cette dernière pour nous permettre de comparer nos résultats. Toutefois, nous avons décidé de retirer les pauses entre les syllabes (à l'inverse de Bion et al., 2011 et Hay & Diehl, 2007, mais comme Bhatara et al., 2013). De ce fait, toutes les syllabes sont co-articulées. Nous avons fait ce choix, puisque l'étude de Bhatara et al. (2013) montre que la complexité acoustique portée par la variabilité segmentale des stimuli utilisés et l'absence de pauses entre les syllabes font émerger des différences cross-linguistiques.

2.1.2.1.2.2 Début de la séquence de familiarisation

Afin d'éviter que les participants ne se basent sur la première paire de syllabes pour grouper prosodiquement la séquence de familiarisation, nous nous sommes basés sur le dessin expérimental de Bion et collaborateurs (2011). Nous avons donc utilisé la

même rampe qui consiste en une augmentation graduelle des contrastes de durée, intensité ou de pitch, implémentée à partir du début de chaque séquence. C'est ainsi que dans la condition durée, les deux premières syllabes ont des durées égales (200ms), puis la différence de durée entre les syllabes augmente au fur et à mesure, par palier de 20ms, à partir de la troisième syllabe. Donc, à chaque syllabe impaire une augmentation de 20 ms est effectuée, pour atteindre un maximum de 460ms. Autrement dit, la troisième syllabe a une durée de 220ms, la quatrième 200, la cinquième 240ms, la sixième 200 et ainsi de suite, jusqu'à ce que la syllabe impaire atteigne 460ms. Dans la condition intensité, la séquence de familiarisation commence par deux éléments de la même intensité, puis à partir de la troisième syllabe, l'intensité augmente de 1 dB à chaque syllabe impaire, en partant d'une base de 66dB pour atteindre 74 dB. Enfin, pour la condition pitch, les deux premières syllabes ont un pitch de 200Hz, puis le pitch augmente par palier de 20 Hz dès la troisième syllabe sur chaque syllabe impaire, pour atteindre 420Hz. Dans la condition contrôle, aucune variation de durée, intensité ou de pitch n'a été réalisée ; toutes les syllabes ayant donc les mêmes caractéristiques prosodiques.

Cependant, Woodrow rapporte déjà en 1909 que le jugement de groupement rythmique dépend crucialement du premier son entendu. En effet, le phénomène de groupement d'items est connu pour s'appliquer à différents types de séquence. De très anciens travaux (Bolton, 1894) montrent que des tons identiques apparaissant à un temps régulier sont également perçus par paire (effet du tic-tac de l'horloge). Un exemple connu illustrant ce phénomène est celui du numéro de téléphone où la mémorisation d'une suite de numéros n'est pas encodée de manière isochrone, mais par groupe de numéros (Thrope & Rowland, 1965). Ces observations révèlent donc que, même sans aucune modulation prosodique pour des tons ayant les mêmes propriétés acoustiques, un phénomène de groupement dès le début de la séquence est présent.

Une centaine d'années plus tard, Hay and Diehl (2007) retrouvent ce problème méthodologique dans différentes études pilotes réalisées sur la perception de l'ITL. Ils rapportent en effet que, quel que soit le son entendu en premier, les participants le reportent comme le début du groupe rythmique. Ils ont alors tenté de neutraliser cet effet, en appliquant un masquage du début de séquence. Le masquage consiste à augmenter graduellement l'intensité, à partir du début de la séquence, en parallèle de la présence d'un bruit blanc dont l'intensité diminue progressivement. La fin de la séquence n'est pas masquée, étant donné que des études pilotes n'ont reporté que de très

faibles effets. De plus, ils ont ajouté un facteur de contrebalancement : la moitié de leurs séquences commencent par un stimulus plus intense/fort, ou plus long, et l'autre moitié par un stimulus plus faible ou plus court. Sans ajouter de masquage, nous avons décidé de contrebalancer également le début de la rampe en commençant à augmenter les différences soit sur la 3^{ème} syllabe (impaire) comme dans Bion et al. (2011), soit sur la 4^{ème} syllabe (paire) voir Figure 2.2.

Ce contrebalancement de la rampe sur la 3^{ème} ou 4^{ème} syllabe pourrait affecter la perception du groupement. Comme nous l'avons décrit plus haut, la séquence peut être groupée par paire à partir du premier son entendu (groupement par défaut). Si le groupement par défaut coïncide avec le groupement de syllabe prédit par l'ITL, les adultes devraient avoir de meilleures performances. Donc, si la rampe commence à partir de la 3^{ème} syllabe, alors la deuxième paire (du groupement par défaut) est trochaïque et cela coïncide avec l'ITL pour les conditions intensité et pitch. Si la rampe commence à partir de la 4^{ème} syllabe, alors la deuxième paire (par défaut) est iambique, cela coïncide avec l'ITL pour la condition durée. La position de la rampe est donc un facteur susceptible d'interagir avec le groupement basé sur l'ITL.

Conditions		Séquence de familiarisation							L1	N			
DUREE	Rampe	3 ^{ème} syllabe	fo	na	GI	pe	ZA	ro	VE	ti	FO	Français	8
		4 ^{ème} syllabe	fo	na	gi	PE	za	RO	ve	TI	fo	Allemands	8
	Rampe	3 ^{ème} syllabe	fo	na	GI	pe	ZA	ro	VE	ti	FO	Français	8
		4 ^{ème} syllabe	fo	na	gi	PE	za	RO	ve	TI	fo	Allemands	8
PITCH	Rampe	3 ^{ème} syllabe	fo	na	GI	pe	ZA	ro	VE	ti	FO	Français	8
		4 ^{ème} syllabe	fo	na	gi	PE	za	RO	ve	TI	fo	Allemands	8
	Rampe	3 ^{ème} syllabe	fo	na	GI	pe	ZA	ro	VE	ti	FO	Français	8
		4 ^{ème} syllabe	fo	na	gi	PE	za	RO	ve	TI	fo	Allemands	8
CONTROLE	fo na gi pe za ro ve ti fo							Français	8				
	fo na gi pe za ro ve ti fo							Allemands	8				

Figure 2.2 : Répartition des participants dans les 4 conditions (durée, intensité, pitch et contrôle) avec le schéma des deux type de rampe pour chaque langue (L1). L'augmentation graduelle de la différence de durée, intensité et pitch commence soit sur la troisième syllabe (position initiale de la deuxième paire) ou soit sur la quatrième syllabe (position finale de la deuxième paire).

2.1.2.1.2.3 Items de test

Pour la phase de test, trois types de stimuli ont été synthétisés avec MBROLA, encore une fois avec la voix allemande de5 ou française fr4. Tous ces stimuli sont composés de paires de syllabes sans modulation prosodique. Ces items ont donc tous la même durée, intensité et pitch (voir Tableau 2-2). Toutes les syllabes présentes dans les paires sont les mêmes que celle de la phase de familiarisation. Les paires sont combinées de différentes façons, soit elles étaient adjacentes, soit non-adjacentes dans la phase de familiarisation. Pour les paires adjacentes, la première possibilité de combinaison correspond à celle où, dans la phase de familiarisation la syllabe longue, forte ou aiguë était en position initiale et par conséquent la syllabe courte, faible ou grave était en position finale (items trochaïques). La seconde possibilité correspond à celle où dans la phase de familiarisation, la syllabe courte, faible ou grave était en position initiale et par conséquent la syllabe longue, forte ou aiguë était en position

finale (items iambiques). Enfin, la dernière combinaison possible correspond à des paires de syllabes qui étaient non-adjacentes dans la phase de familiarisation (items non-adjacents) : /fope/, /give/, /vegi/, /tiza/, /naro/, /pefo/, /rona/, /zati/.

Tous les participants ont entendu les mêmes items de test (4 trochaïques, 4 iambiques et 8 non-adjacents), quelle que soit la phase de familiarisation entendue, et ceci avec la même voix que celle entendue dans la phase de familiarisation (de5 ou fr4). De cette manière, les différences de performances qui pourraient être observées ne pourraient être attribuées qu'au type de familiarisation entendue.

2.1.2.1.3 Procédure

Les participants étaient assis confortablement dans une pièce silencieuse et les stimuli étaient présentés à l'aide de PsyScope X (disponible sur <http://psy.ck.sissa.it/>), sur un Macbook Pro portable. A Potsdam les stimuli étaient présentés à l'aide d'un casque AKG K 55 et à Paris avec un casque Sennheiser HD558. Pour chaque langue, nous avons présenté aux quatre groupes de participants l'une des quatre conditions : durée, intensité, pitch ou contrôle. Chaque condition correspondait à une phase de familiarisation spécifique. La phase de test était exactement la même pour les quatre conditions. Pendant la phase de familiarisation, les participants avaient pour instruction d'écouter attentivement la séquence qui leur était présentée. Après la phase de familiarisation, 48 items tests au total leur étaient présentés. Il y avait 3 types d'items tests [iambique (4 paires), trochaïque (4 paires), non-adjacent (8 paires)] ; chaque type étant répété 3 fois dans trois blocs consécutifs, pour un total de 48 essais. Les items étaient présentés dans un ordre aléatoire. Dans chaque condition, les participants ont entendu les stimuli soit avec la voix allemande (de5), soit avec la voix française (fr4) ; c'est à dire que la moitié des effectifs allemands ou français a entendu la voix allemande et l'autre moitié a entendu la voix française. De plus, ils ont pu avoir une rampe commençant sur la 3^{ème} ou la 4^{ème} syllabe (voir Figure 1.2 pour la répartition des participants dans les sous-groupes).

La tâche des sujets était de répondre à la question « Avez-vous perçu cette paire comme adjacente pendant la phase de familiarisation ? ». Une bonne réponse correspondait à répondre « oui » si la paire était adjacente en phase de familiarisation (qu'elle soit iambique ou trochaïque) et à répondre « non » si la paire était non-adjacente en phase de familiarisation.

2.1.2.1.4 Prédictions

Si les sujets peuvent mémoriser l'ordre des syllabes entendues pendant la phase de familiarisation, leurs performances devraient être supérieures à 50 %. Concernant l'ITL, notre hypothèse est que, si les participants sont sensibles au biais de groupement rythmique, ils auront de meilleurs scores de bonnes réponses pour les items iambiques que trochaïques dans la condition durée et pour les items trochaïques que iambiques dans les conditions intensité et pitch (Figure 2.3). De plus, comme mentionné plus haut, ces effets pourraient être modulés par la correspondance ou non entre le groupement basé sur l'ITL et le groupement basé sur un découpage de paires de syllabes à partir de la première syllabe. Dans ce cas-là, nous devrions voir des différences de performance en fonction de la rampe.

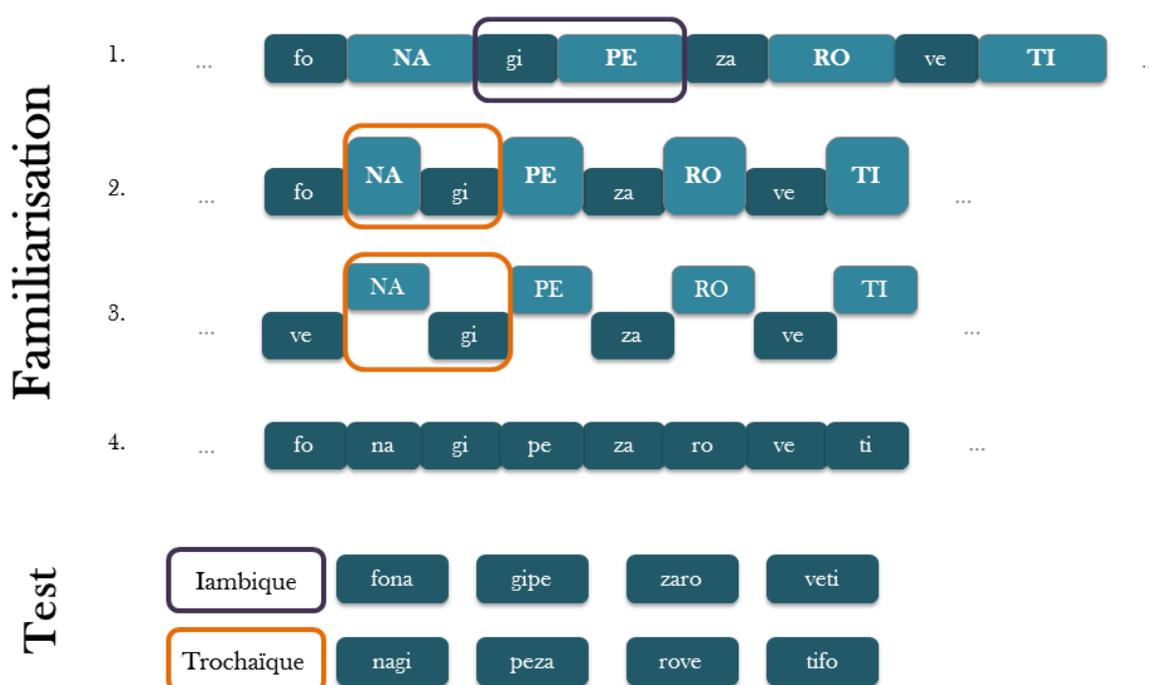


Figure 2.3 : Prédiction de la loi Iambo-trochaïque. (1) : condition durée : groupement iambique ; (2) : condition intensité et (3) : condition pitch (4) : groupement trochaïque ; condition contrôle : aucune prédiction de groupement.

2.1.2.2 Résultats

2.1.2.2.1 Les participants ont-ils retenu l'ordre des éléments de la séquence ?

Les réponses « oui » pour les items qui sont adjacents et non-adjacents dans la phase de familiarisation ont été sommées et transformées en % pour chaque sujet et c'est cette mesure qui a défini notre variable dépendante dans nos analyses. ; sachant que la réponse « oui » correspondait à une bonne réponse si la paire était adjacente en phase de familiarisation (qu'elle soit iambique ou trochaïque), mais correspondait à une mauvaise réponse si la paire était non-adjacente en phase de familiarisation. Le score moyen de bonnes réponses (%) a donc été calculé pour chaque groupe et condition (durée, intensité, pitch et contrôle). Les résultats sont illustrés dans la Figure 2.4.

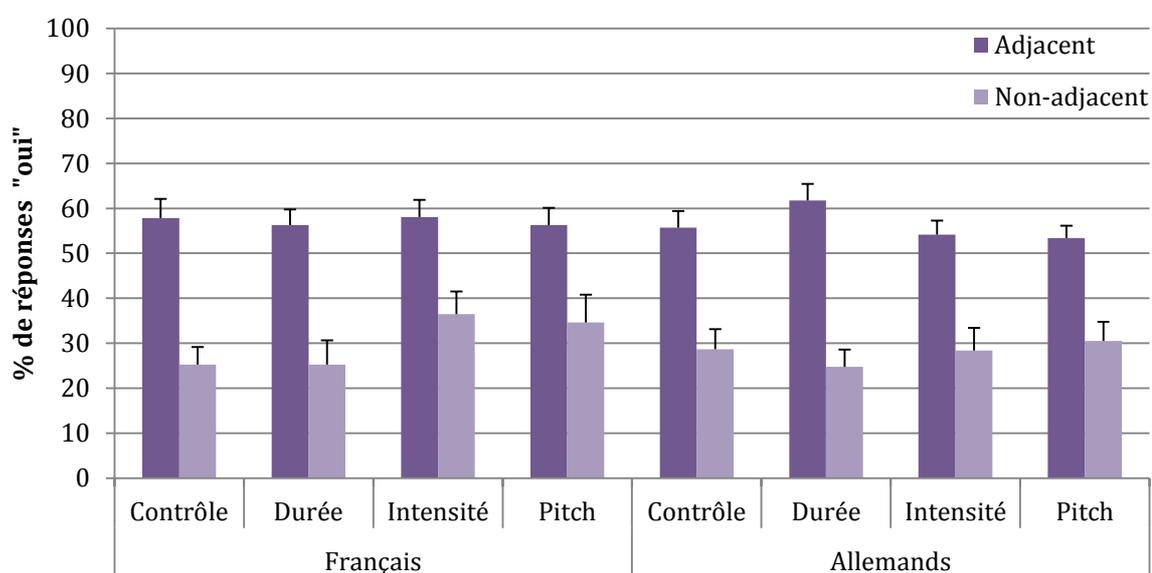


Figure 2.4 : Score moyen (en %) de réponses "oui", pour les items adjacents (réponse correcte) et non-adjacents (réponse incorrecte) dans les quatre conditions (contrôle, durée, intensité et pitch).

Une ANOVA à mesures répétées a été réalisée avec comme facteur intra-sujet le type d'item (non-adjacent, adjacent) et comme facteurs inter-sujets la condition (durée, intensité, pitch et contrôle) et la langue maternelle (français, allemand). Les résultats révèlent un effet principal d'item [$F(1, 120) = 37.26, p < .00001, \eta^2_p=0.24$]. Aucun autre effet ni interaction n'atteint la significativité ; voir Tableau 2-3. Les participants discriminent donc bien les items adjacents des items non-adjacents et répondent correctement plus souvent lorsque les items sont adjacents (57%) dans la séquence de

familiarisation que lorsque les items y sont non-adjacents (29%). Par ailleurs, cet effet ne diffère ni en fonction de la condition, ni de la langue.

Tableau 2-3 : Analyse globale sur les 4 conditions (durée, intensité, pitch et contrôle).

Adjacent vs Non adjacent			
Contrastes	ddl (x)	F (x, 120)	p
Langage	1	0,15	0,70
Condition	3	1,98	0,12
Type d'item	1	37,26	<0,00001
Langage × Condition	3	0,42	0,74
Type d'item × Langage	1	0,48	0,49
Type d'item × Condition	3	0,27	0,84
Type d'item × Langage × Condition	3	0,70	0,56

De plus, des tests t ont montré que les participants ont détecté les paires non-adjacentes différemment du hasard, quelle que soient la condition et leur langue (Tableau 2-4). Les adultes français comme allemands ont donc mémorisé l'ordre des syllabes présentes dans la séquence de familiarisation et présentent les mêmes performances quelle que soit la langue maternelle, nous permettant ainsi d'établir que la procédure utilisée est pertinente.

Tableau 2-4 : Test t sur les 4 conditions (durée, intensité, pitch et contrôle), par rapport au hasard (50%).

Langue	Paires	Non-adjacentes	
	Condition	p	t(15)
Français	Contrôle	<0,00001	6,333
	Durée	<0,001	4,594
	Intensité	0,017	2,68
	Pitch	0,025	2,492
Allemands	Contrôle	<0,001	4,747
	Durée	<0,001	6,626
	Intensité	0,001	4,305
	Pitch	<0,001	4,526

Or de quelle manière ont-ils encodé cette séquence de syllabes ? Étant donné que les paires adjacentes contiennent deux possibilités de patterns prosodiques : iambiques ou trochaïques, nous allons dans ce qui suit analyser ces patterns de façon séparée, dans les trois conditions où les propriétés prosodiques sont manipulées (durée, intensité et pitch) et en fonction des groupes. Cela nous donne la possibilité d'évaluer si

l'un des deux types d'items a été mieux retenu lors de la présentation de la séquence de familiarisation.

2.1.2.2.2 Groupement selon la loi lambo-trochaïque

2.1.2.2.2.1 Analyse globale

Les bonnes réponses (réponse « oui » car les syllabes sont adjacentes) pour les items qui ont une structure trochaïque (long-court/fort-faible/aigu-grave) ou iambique (court-long/faible-fort/grave-aigu) dans la phase de familiarisation ont été sommées pour chaque sujet et transformées en %. Le score moyen de bonnes réponses a été calculé pour chaque groupe et condition (durée, intensité, pitch). Les résultats sont illustrés dans la Figure 2.5.

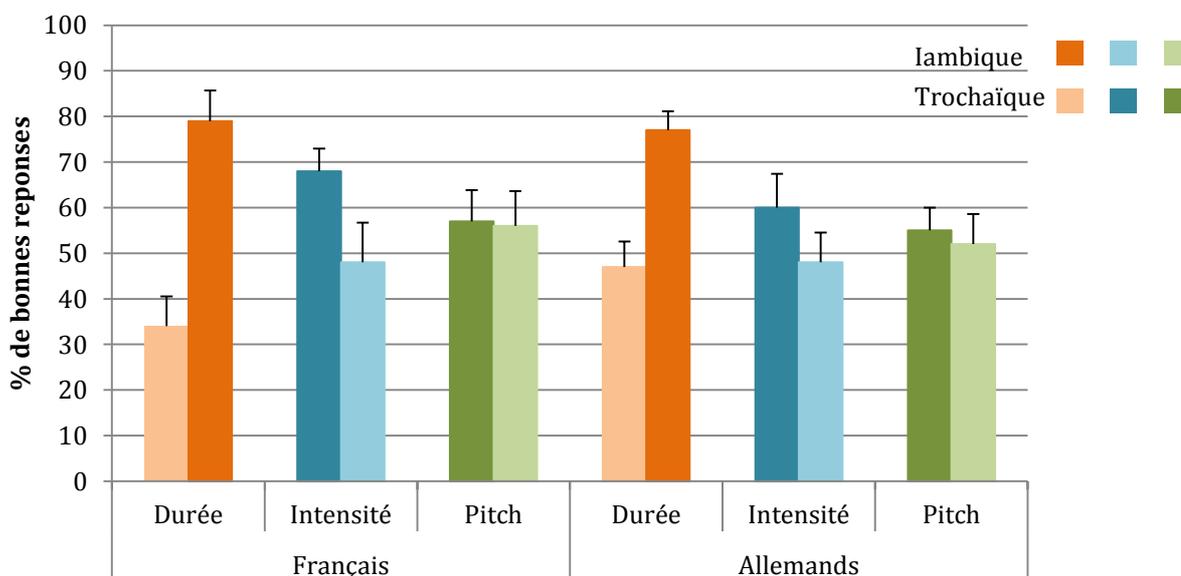


Figure 2.5 : Score moyen de bonnes réponses (réponses « oui ») pour les items trochaïques et iambiques dans les conditions durée, intensité et pitch.

Nous avons réalisé une ANOVA à mesures répétées sur le score moyen de bonnes réponses avec, pour facteur intra-sujet, le type d'items (iambique, trochaïque) et, comme facteurs inter-sujet, la langue maternelle et la condition. L'interaction entre le type d'items et la condition est significative [$F(2, 90) = 12.88, p < .0001, \eta^2_p = 0.12$]. Ces résultats montrent que les participants ont des scores de bonnes réponses qui diffèrent entre les items trochaïques et iambiques en fonction de la condition de familiarisation. Aucun autre effet ni interaction n'atteignent la significativité (voir Tableau 2-5). Afin de

comprendre plus en détails cette interaction entre la condition et le type d'items, nous avons analysé séparément chaque condition.

Tableau 2-5 : Analyse sur les 3 conditions (durée, intensité, et pitch).

Trochaïque versus Iambique			
Contrastes	ddl (x)	F (x, 90)	p
Langage	1	0,02	0,88
Condition	2	0,75	0,47
Type d'item	1	1,90	0,17
Langage × Condition	2	1,09	0,34
Type d'item × Langage	1	0,14	0,71
Type d'item × Condition	2	12,88	0,00001
Type d'item × Langage × Condition	2	0,55	0,58

2.1.2.2.2.2 Condition durée

Une ANOVA à mesures répétées a été réalisée sur le score moyen de bonnes réponses avec, en facteur intra-sujet, le type d'items (trochaïque ou iambique) et, en facteur inter-sujet, la langue maternelle et le type de rampe. Elle révèle un effet principal du type d'items [$F(1,28)=34.58$, $p<.0001$, $\eta^2_p=0.55$]. Les participants ont de meilleurs scores pour les items iambiques ($M_{\text{iambique}} = 75\%$) que pour les items trochaïques ($M_{\text{trochaïque}} = 40\%$) et ceci ne diffère pas en fonction de leur langue maternelle. Ainsi, les participants français ont de meilleurs scores pour les items iambiques ($M_{\text{iambique}} = 79\%$ vs $M_{\text{trochaïque}} = 34\%$), tout comme les participants allemands ($M_{\text{iambique}} = 77\%$ vs $M_{\text{trochaïque}} = 45\%$). Ces résultats sont illustrés dans la Figure 2.6. Aucun autre effet ni interaction ne sont significatifs (voir Tableau 2-7, colonne de gauche).

Pendant l'écoute de la séquence de familiarisation variant en durée, les participants allemands et français ont donc groupé les syllabes iambiquement, comme le prédisait la loi Iambo-trochaïque, sans aucune modulation de la langue maternelle, ni du type de rampe.

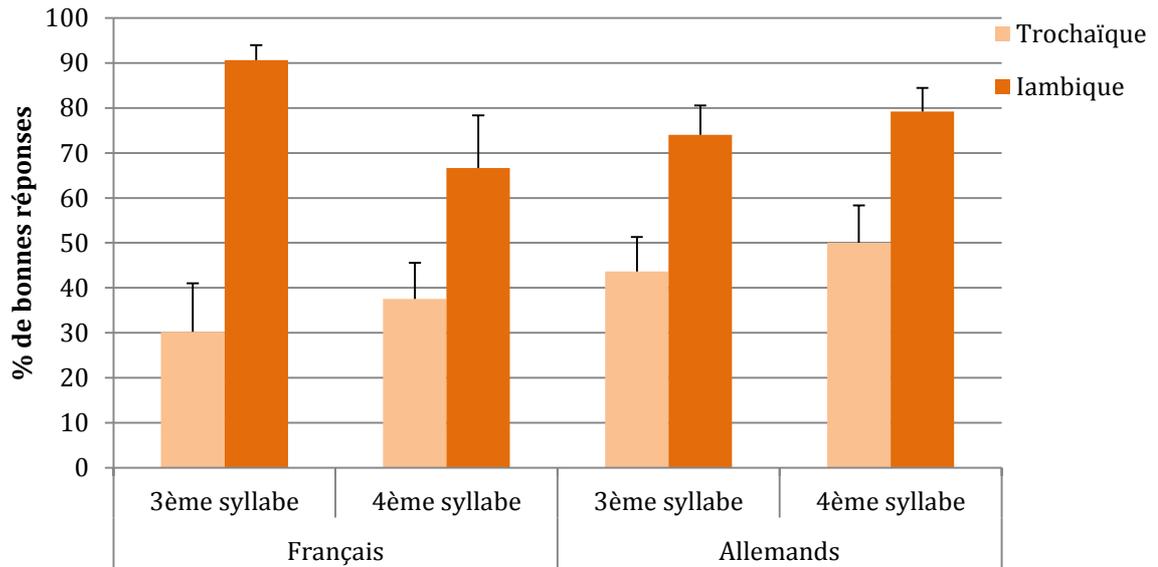


Figure 2.6 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition durée en fonction du type de rampe (3^{ème} ou 4^{ème} syllabe), séparé par langue maternelle.

2.1.2.2.2.3 Condition intensité

Une ANOVA à mesures répétées a été réalisée sur les scores moyens de bonnes réponses avec le facteur intra-sujet type d'items et les facteurs inter-sujets, langue maternelle et type de rampe. Les résultats révèlent un effet principal du type d'items [$F(1, 28) = 6.60, p = .016, \eta^2_p=0.19$]. Les sujets ont de meilleurs scores de bonnes réponses pour les items trochaïques ($M_{\text{Trochaïque}} = 64\%$) que les items iambiques ($M_{\text{Iambique}} = 48\%$). Toutefois, le type de rampe interagissait, d'une part, avec le type d'items [$F(1, 28) = 25.51, p < .00001, \eta^2_p=0.48$] et, d'autre part, avec la langue maternelle [$F(1, 28) = 6.6, p = .016, \eta^2_p=0.19$]. Aucun autre effet ou interaction n'étaient significatifs (voir Tableau 2-7, colonne centrale).

Afin de comprendre l'interaction entre le type de rampe et le type d'items, nous avons évalué l'effet du type d'items en fonction du type de rampe. Il apparaît que les adultes ont eu de meilleures performances pour les items trochaïques que les items iambiques si la rampe commençait par la 3^{ème} syllabe [$M_{\text{Trochaïque}} = 79\%$ vs $M_{\text{Iambique}} = 30\%$, $t(15) = -4.74, p < .001$]. Par contre, s'ils avaient entendu une séquence avec la rampe commençant par la 4^{ème} syllabe, aucune différence de performance n'était retrouvée ; une tendance de performance inverse étant même présente [$M_{\text{Iambique}} = 66\%$ vs $M_{\text{Trochaïque}} = 50\%$, $t(15) = -1.43, p = .087$].

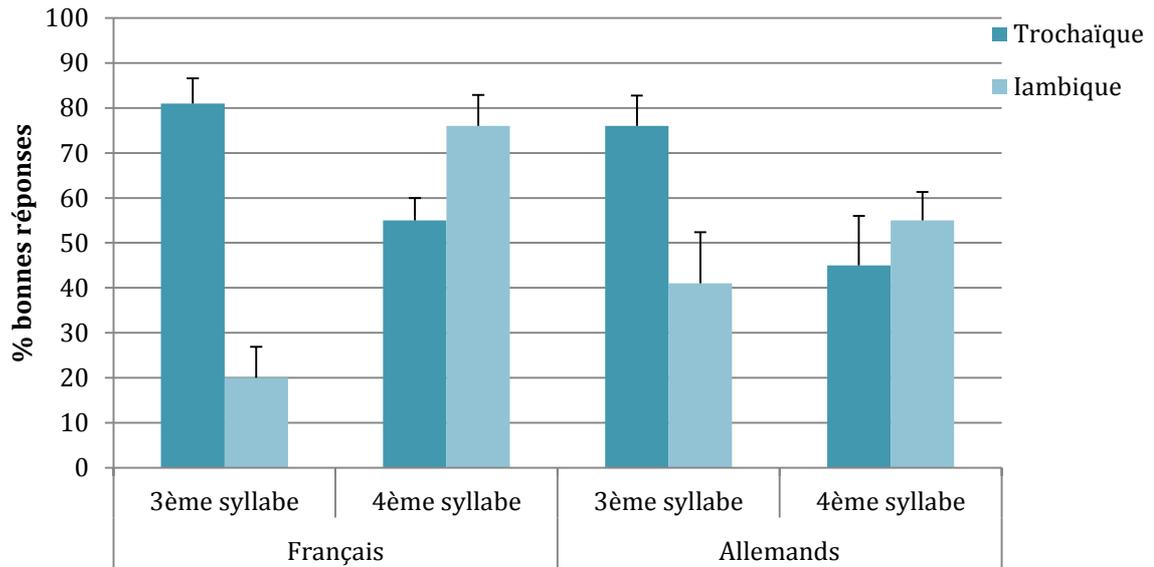


Figure 2.7 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition intensité en fonction du type de rampe (3^{ème} ou 4^{ème} syllabe), séparés par langue maternelle.

Afin de préciser l'interaction du type de rampe avec la langue maternelle, nous avons évalué l'effet de la rampe pour chaque groupe linguistique ; voir Figure 2.7. Les français ont de meilleurs scores de bonnes réponses pour les items trochaïques par rapport aux iambiques quand la séquence commence avec la rampe sur la 3^{ème} syllabe [$M_{\text{Trochaïque}} = 81\%$ vs $M_{\text{Iambique}} = 20\%$, $t(7) = -6.27$, $p < .001$]. Par contre, si la rampe commence sur la 4^{ème} syllabe, alors les français ont un pattern inverse de résultats ; ils ont de meilleures performances pour les items iambiques par rapport aux items trochaïques [$M_{\text{Iambique}} = 76\%$ vs $M_{\text{Trochaïque}} = 55\%$, $t(7) = -2.81$, $p = .013$]. Chez les allemands, aucune différence significative n'est mise en évidence, que ce soit pour la rampe sur la 3^{ème} syllabe ($M_{\text{Trochaïque}} = 76\%$ vs $M_{\text{Iambique}} = 41\%$) ou pour celle sur la 4^{ème} syllabe ($M_{\text{Trochaïque}} = 45\%$ vs $M_{\text{Iambique}} = 55\%$). Les résultats sont rapportés Tableau 2-6.

Tableau 2-6 : Analyse pour la condition intensité, séparée par la langue maternelle.

Langue	Trochaïque vs iambique	INTENSITE	
	Contraste	F(1,28)	p
Français	Type de rampe	4,89	0,04
	Type d'item	14,18	0,002
	Type d'item × Type de rampe	58,17	0,000002
Allemands	Type de rampe	1,90	0,19
	Type d'item	1,19	0,29
	Type d'item × Type de rampe	3,99	0,07

En résumé, ces résultats montrent que les adultes ont mieux mémorisé les items trochaïques que les iambiques, comme prédit par l'ITL. Toutefois, les participants ont eu tendance à grouper différemment les séquences en fonction du type de rampe. Ceci signifie que, quand le groupement trochaïque est aligné avec le groupement formé à partir de la première syllabe de la séquence de familiarisation (comme c'est le cas pour la séquence commençant par la 3^{ème} syllabe), alors les performances sont meilleures que quand les deux types de groupement ne sont pas alignés. De plus, le type de rampe n'affecte pas de la même manière le jugement du groupement rythmique suivant la langue maternelle des participants, puisque les français y sont plus sensibles que les allemands.

2.1.2.2.2.4 Condition Pitch

Une ANOVA à mesures répétées a été réalisée sur le score moyen de bonnes réponses avec, en facteur intra-sujet le type d'items et, en facteurs inter-sujet, la langue maternelle et la rampe. Les résultats révèlent une interaction entre la rampe et le type d'items [$F(1, 28) = 7.54, p = .01$]. Lorsque la rampe commence sur la 3^{ème} syllabe, les participants ont tendance à avoir de meilleurs scores pour les items trochaïques par rapport aux items iambiques [$M_{\text{Trochaïque}} = 63\%$ vs $M_{\text{Iambique}} = 41\%$; $t(15) = 1.98, p = .06$]. Par contre, si la rampe commence sur la 4^{ème} syllabe, ils ont une tendance inverse [$M_{\text{Iambique}} = 67\%$ vs $M_{\text{Trochaïque}} = 49\%$; $t(15) = -1.3, p = .06$]. Ces effets ne sont cependant que des tendances et sont indépendants de la langue maternelle des participants. Aucun autre effet ou interaction n'atteignent la significativité (voir Tableau 2-7, colonne de droite). Les résultats sont illustrés dans la Figure 2.8.

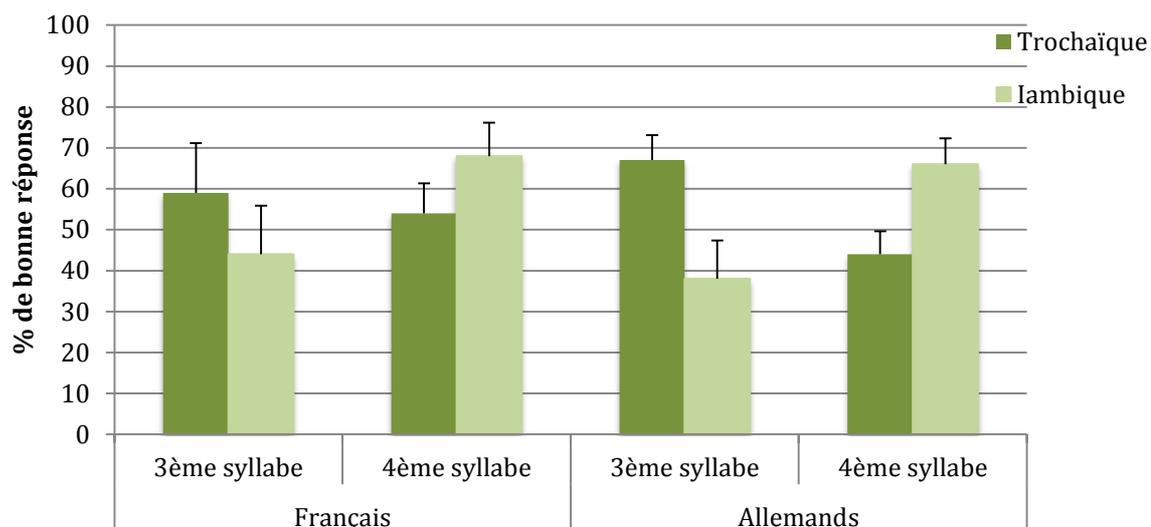


Figure 2.8 : Scores moyens de bonnes réponses pour les items trochaïque et iambique dans la condition du pitch en fonction du type de rampe.

Ces données montrent que, dans cette tâche, la performance des adultes est influencée par la caractéristique de la rampe pour grouper la séquence. En résumé, aucun effet de groupement n'est mis en évidence, suggérant l'instabilité de la perception de l'ITL basé sur le pitch, influencé par la rampe.

Tableau 2-7 : Récapitulatif des analyses par condition chez tous les participants.

Trochaïque versus Iambique	DUREE		INTENSITE		PITCH	
	F(1, 28)	p	F(1,28)	p	F (1,28)	p
Contraste						
Langage	1,16	0,29	0,73	0,40	0,37	0,55
Type de rampe	0,07	0,80	0,55	0,46	1,61	0,21
Type d'item	34,58	<0,00001	6,69	0,02	0,10	0,75
Langage × Type de rampe	1,92	0,18	6,60	0,02	0,52	0,48
Type d'item × Langage	1,42	0,24	0,38	0,54	0,03	0,86
Type d'item × Type de rampe	1,63	0,21	25,51	<0,0001	7,53	0,01
Type d'item × Langage × Type de rampe	1,42	0,24	2,07	0,16	0,56	0,46

2.1.2.3 Discussion

Dans cette étude, nous avons étudié la perception de la loi iambo-trochaïque dans le cadre d'une tâche de mémorisation de séquences linguistiques. Nous avons exploré de quelle manière les adultes français et allemands ont mémorisé et traité des séquences composées de syllabes alternant en durée, intensité ou pitch, ainsi qu'en l'absence de modulation prosodique. Ont-ils groupé ces séquences selon l'ITL ? Leurs réponses différaient-elles en fonction de leur langue maternelle ? Nos résultats confirment les travaux antérieurs sur le groupement prosodique, révélant des effets universaux, mais également des différences cross-linguistiques.

2.1.2.3.1 Perception de la Loi Iambo-Trochaïque (ITL) : validation de l'effet originel

Nos résultats confirment la formulation originelle de la perception du groupement rythmique proposée par Woodrow (1909), ensuite reprise et énoncée comme Loi Iambo-Trochaïque (ITL) par Hayes (1995). Les participants ont segmenté ces séquences linguistiques différemment en fonction de l'indice acoustique manipulé. En effet, les participants se souvenaient mieux des paires qui finissaient par une syllabe longue lors de l'écoute de séquences alternant en durée et de celles qui commençaient par une syllabe forte lors de l'écoute de séquences alternant en intensité. Ainsi, un allongement de syllabe a été traité comme un indice indiquant la position finale (groupement iambique) et une syllabe plus forte est perçue comme indiquant la position initiale (groupement trochaïque). Les données des locuteurs français et allemands confirment donc les prédictions de l'ITL, trochaïque pour l'intensité et iambique pour la durée. Ces résultats confirment que la perception du groupement rythmique repose sur des principes universaux, s'ajoutant ainsi au travail originel de Woodrow (1909) et à ceux plus récents de Hay & Diehl (2007), Bion et al. (2011) et Bhatara et al. (2013).

Par contre, nos résultats sur le groupement basé sur le pitch, proposé plus récemment par Nespors (2008) et soutenu par les études de Bion et al. (2011) chez des locuteurs italiens et de la Mora et al. (2013) chez des locuteurs espagnols ne sont pas aussi clairs. En utilisant une méthode similaire à Bion et al. (2011), nous n'avons pas mis en évidence de groupement stable trochaïque basé sur le pitch. Nos résultats chez des locuteurs français et allemands ne généralisent donc pas les travaux obtenus chez des italiens (Bion et al., 2011) et chez des espagnols (de la Mora et al., 2013), mais ils sont en accord avec ceux de Bhatara et al. (2013) pour les français, et de Kusumoto et

Moreton (1997) pour les anglais et japonais. Ces derniers ont aussi observé des résultats peu consistants de groupement basé sur le pitch, suggérant qu'il ne serait pas un aussi bon indice que l'intensité et la durée pour grouper les sons selon l'ITL, tout du moins chez ces locuteurs. On notera aussi que nous ne trouvons pas de préférence de groupement trochaïque pour le pitch chez les allemands, alors que Bhatara et collaborateurs (2013) montraient qu'ils étaient sensibles à ce type de groupement de la même manière qu'à l'intensité. Ceci nous laisse penser que le matériel et les procédures utilisées dans ces études pourraient être en partie responsables de ces différences d'effet.

Un des facteurs qui a joué sur nos résultats est notamment le type de début de séquence. Là où Bion et al. (2011) n'avait qu'un seul type de début de séquence, nous avons décidé de contrebalancer ce facteur (voir les raisons présentées plus haut dans la section méthode). Ce dernier élément est apparu comme un modulateur important de la perception du groupement rythmique, ce dont nous allons discuter dans la section suivante.

2.1.2.3.2 Effets du type de séquence

Nos résultats montrent que le jugement rythmique peut être dépendant des caractéristiques de la rampe et que celle-ci a eu un poids important dans la manière dont les sujets ont mémorisé certaines séquences de familiarisation. Nous avons retrouvé ce phénomène pour le pitch et l'intensité, mais pas pour la durée.

La rampe que nous avons utilisée est la même que celle utilisée par Bion et al. (2011), qui consistait en une augmentation graduelle de la différence de durée/intensité/pitch. Nous avons décidé d'augmenter pas à pas la différence, soit à partir de la 3^{ème} syllabe, soit à partir de la 4^{ème} syllabe. Cet effet du type de rampe met en évidence une combinaison de deux phénomènes ; le premier concerne l'effet groupement de l'ITL, le second suggérant un groupement de paires de syllabes -induit « par défaut »- à partir de la première syllabe entendue de la séquence de familiarisation. En fonction des indices acoustiques, le mécanisme de groupement prosodique pourrait être modulé par le groupement « par défaut ». Si les deux facteurs jouent dans le même sens (ce qui est le cas pour la rampe commençant sur la 3^{ème} syllabe pour les conditions intensité et pitch), alors les participants se souviennent mieux des items trochaïques. Par contre, si les facteurs ne jouent pas dans le même sens

(pour la rampe commençant sur la 4^{ème} syllabe), alors les participants ne se souviennent pas plus souvent des items trochaïques que des items iambiques.

On voit que l'effet de variation d'intensité ou de pitch est modulé par le type de la rampe. Toutefois, si nous retrouvions un effet significatif du type d'items pour l'intensité, aucun effet global n'était présent pour le pitch, alors que deux récentes études ayant un dessin expérimental très proche du nôtre trouvaient un effet (Bion et al., 2011 ; de la Mora et al., 2013). Cependant, dans ces deux études, les auteurs n'avaient pas contrebalancé le facteur de type de rampe. En effet, le contraste de pitch des syllabes commençait toujours sur la 3^{ème} syllabe, donnant un groupement trochaïque allant dans le même sens que le groupement par défaut à partir de la première syllabe. Donc l'utilisation d'un seul type de début de familiarisation dans les deux études précédentes, notamment celui favorisant le groupement trochaïque, a dû les aider à segmenter la séquence plus facilement et à trouver un effet pour le pitch. Mais cet effet résulte au moins en partie de l'indice supplémentaire du groupement par défaut. Si nous restreignons nos analyses à cette même rampe, nous retrouvons des résultats similaires à ces deux études (Bion et al., 2011 ; de la Mora et al. 2013), puisque nous retrouvons des tendances avec 8 participants seulement par langue. Certes il ne s'agit que de tendances, mais il est possible qu'en ayant le même effectif de participants (16 sujets) cette tendance puisse devenir significative. Les résultats clairs obtenus dans les deux études précédentes pour le groupement basé sur le pitch ne reflètent pas un groupement si stable pour le pitch chez l'adulte, tout du moins dans ce type de tâche.

Si dans ces deux dernières études, un seul type de rampe a pu favoriser la perception du groupement rythmique, dans d'autres études qui ont tenté de contrôler ce paramètre, de résultats faibles, voir nuls, ont été retrouvés. C'est le cas de l'étude de Trainor et Adams (2000), qui a utilisé une méthode différente de la nôtre (tâche de détection de silence), sans masquage ni rampe, mais avec le même facteur de contrebalancement pour le début de la séquence. Leurs résultats de locuteurs anglophones montrent un effet de groupement pour la durée, mais pas pour l'intensité. Ils ont interprété ces résultats comme signifiant le rôle modéré de l'intensité comme marqueur de frontière prosodique, interagissant probablement avec d'autres indices comme la durée ou le pitch. Cependant, dans leur interprétation, ces auteurs n'ont pas exploité le fait qu'ils ont obtenu un très fort effet du début de la séquence pour l'intensité, qui établit que ce facteur de contrebalancement est un modulateur de l'effet de l'ITL. Nos résultats vont dans le même sens que Trainor & Adams (2000). Il est

possible que la variation d'intensité seule n'ait pas un rôle aussi pertinent dans le groupement rythmique, et donc que l'intensité *per se* ne joue pas. Par ailleurs, comme notre étude, ils ne retrouvent pas d'effet de rampe pour la durée. Les auteurs suggèrent que la durée serait un indice plus important dans la segmentation de la séquence de syllabe. En effet, il est possible que l'absence d'effet de rampe soit dû au fait que les anglophones, les allemands et les français ont une représentation plus robuste de la durée, en lien notamment avec l'allongement systématique de la dernière syllabe des syntagmes dans ces trois langues (Cutler, 2005; Delattre, 1966; Dogil & Williams, 1999; Jessen et al., 1995). D'ailleurs, de récentes études chez l'adulte confirmaient l'importance de cet indice (Langus, Marchetto, Bion, & Nespors, 2012; Tyler & Cutler, 2009).

Des études ultérieures seront nécessaires pour mieux comprendre ce phénomène de biais de début de séquence, son impact en fonction des indices et surtout pour contrôler très attentivement ce facteur lors de prochaines expériences sur l'ITL.

2.1.2.3.3 Effet cross-linguistique

La seule différence cross-linguistique retrouvée dans cette étude concerne la différence de sensibilité des participants au type de rampe pour la condition de l'intensité. Les résultats du pitch et de la durée ne révèlent par contre aucune interaction significative avec la langue maternelle. Deux questions se posent alors. Pourquoi les français et les allemands présentent-ils des sensibilités différentes au type de rampe ? De plus, pourquoi retrouve-t-on cet effet seulement pour l'intensité ?

Nous avons observé que les français sont très affectés par la rampe lors du processus de mémorisation de la séquence, alors que les allemands semblent être moins affectés. Il est probable que, lors de l'écoute de la séquence d'intensité commençant par la rampe qui ne concorde pas avec l'ITL (4ème syllabe), les français aient une perception moins stable et inversent le groupement de syllabes au cours de l'écoute. Il est donc possible que cette sensibilité au type de la rampe reflète une différence de stabilité des représentations internes, ici des variations d'intensité. En effet, il est possible que les allemands traitent mieux l'intensité et, donc, sont moins sensibles à l'effet de la position de la rampe. D'ailleurs, si des différences significatives en terme de stabilité du jugement rythmique n'ont pas été mises en évidence dans l'étude de Hay et Diehl (2007), ils reportaient toutefois que la perception de l'ITL semblait être plus stable chez les anglophones que chez les francophones. Une seule étude cross-

linguistique, utilisant un matériel linguistique et comparant français et allemands, a révélé pour la première fois un clair avantage des allemands par rapport au français (Bhatara et al., 2013). Bhatara et collaborateurs (2013) ont interprété leurs résultats, notamment le fait que la perception des allemands soit plus stable que celle des français, de part de nombreuses différences du système prosodique de ces deux langues, comme par exemple le fait que l'allemand possède une accentuation lexicale au niveau du mot, mais pas le français (voir section introduction pour plus de détails). Cela pourrait en partie expliquer pourquoi les français inversent le groupement en fonction de la rampe, alors que les allemands y sont moins sensibles.

Par ailleurs, nos résultats ont révélé qu'effectivement nos deux groupes sont sensibles au phénomène de groupement perceptif établi par l'ITL pour la durée et l'intensité. Une de nos hypothèses principales était que la langue maternelle pouvait moduler ce biais. Alors que dans l'étude de Bhatara et collaborateurs (2013), les allemands étaient clairement plus sensibles à l'ITL que les français, nos résultats semblent montrer des différences cross-linguistiques beaucoup moins importantes. A quoi cela peut-il être dû ?

Une première possibilité concerne la manière dont nous avons présenté ces indices. Dans notre étude, les participants n'ont écouté qu'un seul type d'indice acoustique, soit la durée, l'intensité ou le pitch. A l'inverse, Bhatara et al. (2013) ont présenté soit des séquences alternant en durée et mélangées avec des séquences variant en intensité (Exp. 1), soit des séquences alternant en durée et mélangées avec des séquences variant en pitch (Exp. 2). Leurs résultats révèlent de cruciales différences entre les deux expériences. Dans la première expérience, les deux groupes perçoivent le groupement en accord avec l'ITL, révélant un large avantage des allemands. Cependant, quand la durée et le pitch sont couplés, les français sont incapables de grouper les deux types de séquences et les allemands rapportent des réponses un peu moins stables pour la durée par rapport à la première expérience. Ces résultats suggèrent que le fait de présenter deux types de séquences en même temps a un effet sur les stratégies de groupement basées sur chacun des deux indices présentés. De manière parallèle, une étude récente utilisant une méthode reliée également à la mémoire à court-terme, consistant à faire écouter des items trochaïques ou iambiques à des locuteurs anglophones afin qu'ils les mémorisent, montre le même type d'effet (Morgan, Edwards, & Wheeldon, 2014). En effet, dans une première expérience, les mots sont prononcés par une locutrice anglophone avec tous les indices acoustiques présents

(durée, intensité et pitch). Les participants se souviennent alors mieux des items trochaïques que des items iambiques. Par contre dans une deuxième expérience, après manipulation de l'intensité de sorte qu'elle soit identique entre les items trochaïques et iambiques, les sujets ne pouvant s'appuyer que sur le pitch et la durée, ils se souviennent autant des items trochaïques que des items iambiques. Ces observations reflètent donc à la fois la nécessité d'une cooccurrence d'indices acoustiques, mais également une compétition inter-indices pour percevoir le groupement rythmique.

Ces deux dernières expériences ont été réalisées chez des adultes allemands (Bhatara et al., 2013) et anglais (Morgan et al., 2014), deux langues possédant une accentuation lexicale. Nous savons que la perception de cette accentuation implique de traiter simultanément plusieurs indices acoustiques (Cutler, 2005; Delattre, 1966; Dogil & Williams, 1999; Jessen et al., 1995). Alors que le français, comme vu précédemment, n'a pas de stress lexical, son « accentuation » est au niveau du syntagme, en position finale (Jun & Fougeron, 2002; Goedemans and van der Hulst, 2011; Nespor et al., 2008). De plus, comme nous l'avons vu dans l'introduction, le pitch est souvent associé à l'intensité en allemand, alors qu'en Français ces indices peuvent être indépendants. Ces réalisations acoustiques marquées par un seul indice (français) ou avec une cooccurrence de deux indices (allemand) peuvent donc favoriser les français et désavantager les allemands. De plus, notre tâche ne demandait pas explicitement de catégoriser l'item en fonction de son pattern prosodique, comme le demandait Bhatara et al. (2013) ou Hay & Diehl (2007). Ceci pouvait rendre la tâche plus difficile pour les français par rapport aux allemands, étant donné l'absence d'accentuation en français (Dupoux et al., 1997). Par contre, dans notre tâche, les participants étaient testés sur leur capacité à mémoriser des items et à rapporter si ces items étaient présents ou non, en ne faisant appel à aucune représentation prosodique, directement du moins.

Il est probable que les spécificités ci-dessus aient joué sur le fait que nous retrouvons peu de différences cross-linguistiques dans notre expérience. Des études ultérieures testant différentes combinaisons d'indices seraient utiles pour comprendre comment le groupement rythmique peut être perçu efficacement en fonction de sa cohérence avec la structure prosodique de la langue maternelle. Ceci permettra de déterminer quelles combinaisons d'indices sont plus avantageuses pour une langue donnée.

Un deuxième facteur a pu conduire à l'observation de ces faibles différences cross-linguistiques. Dans l'étude de Bhatara et collaborateurs (2013), réalisée également

avec des adultes français et allemands, 16 syllabes avaient été utilisées. Elles avaient été agencées sans pauses, avec différents types de variations de durée et d'intensité (Exp.1) ou de durée et de pitch (Exp.2). Ces différents paramètres permettaient d'avoir un matériel possédant un niveau de complexité linguistique plus élevé que dans les études précédentes sur l'ITL, forçant les participants à traiter l'information à un niveau plus abstrait et moins acoustique. Jusqu'alors, la seule étude ayant étudié l'ITL cross-linguistiquement avaient utilisé des stimuli moins complexes, constitués de la répétition d'une syllabe ou d'un son non-linguistique, agencés avec des pauses (Hay & Diehl, 2007). Ces auteurs ne rapportaient aucune différence cross-linguistiques entre les français et les américains anglophones. Or, dans notre étude, nous avons utilisé 8 types de syllabes différentes, alors que Bhatara et al. (2013) en présentaient 32 par séquence (chaque syllabe étant présentée une fois de façon accentuée, une fois non). Notre matériel possédait donc deux fois moins de variabilité segmentale que Bhatara et al. (2013), mais 8 fois plus que Hay et Diehl (2007). De plus, la variation des contrastes en durée, intensité ou pitch était fixe dans notre étude (passée la rampe initiale), alors que dans l'étude de Bhatara et al. (2013), elle était variable suivant les essais. Ces différences ont probablement limité la complexité linguistique de notre matériel et a fortiori l'émergence de différences entre nos groupes. Des études complémentaires seront nécessaires pour comprendre à partir de quel niveau de complexité ces différences apparaissent et ce en fonction du type de langues contrastées. De plus, un nombre plus important de participants pour ce type d'études pourrait être nécessaire.

Notre étude a confirmé la perception de l'ITL pour la durée et l'intensité, mais pas pour le pitch chez des locuteurs adultes français et allemands. Les résultats sont d'ailleurs plus consistants pour la durée que l'intensité. De plus, elle a révélé de faibles différences cross-linguistiques. Dans le prochain chapitre, nous allons nous intéresser au développement de ces préférences. Comment ces capacités émergent-elles et comment l'environnement linguistique pourrait interagir très précocement avec ces préférences de groupement ? Quelques récentes recherches ont étudié ces questions à partir de 5-6 mois de vie, cependant les résultats de ces études peinent à se mettre en accord quant à la période d'émergence du biais prosodique et ce en fonction de l'indice. De plus, la manière dont l'environnement linguistique module ces capacités perceptuelles précoces reste encore peu connue. Notre étude va apporter de nouvelles données sur ces questions.

2.1.3 Points clefs

- La perception de la loi iambo-trochaïque basée sur la durée et l'intensité est confirmée par notre étude pour des locuteurs allemands et français.
- Nous ne généralisons pas les travaux réalisés chez des locuteurs italiens (Bion et al. 2011) et espagnols (de la Mora et al. 2013) pour le groupement basé sur le pitch à des locuteurs français et allemands.
- De faibles différences cross-linguistiques ont été trouvées.
- La caractéristique des premiers éléments biaise le jugement du groupement rythmique pour l'intensité et le pitch.
- Une rampe d'augmentation, pas à pas, de la différence d'intensité ou de pitch n'est pas une méthode valide pour masquer le début de la séquence.

2.2 CHAPITRE 2 : Comment l'ITL guide le processus de segmentation chez le nourrisson de 7,5 mois ?

2.2.1 Introduction

La première étude qui a montré une sensibilité au groupement rythmique est celle de Trainor et Adams (2000), chez des nourrissons de 8 mois apprenant l'anglais. Ces nourrissons ont été exposés à une séquence de tons complexes (non-linguistiques), alternant en durée (Expérience 1) ou en intensité (Expérience 2) dans une tâche de détection de silence. Les résultats montrent une sensibilité précoce à la perception du groupement basé sur la durée, mais pas sur l'intensité. Les auteurs interprètent ces résultats selon l'idée que le groupement basé sur la durée et l'intensité ne suit pas la même trajectoire développementale. Cette hypothèse semble être soutenue par les travaux de Saffran et Hay (2012) chez des nourrissons de 6 mois exposés à l'anglais ; toutefois les auteurs rapportent des résultats inverses. En effet, en présentant des séquences de syllabes variant selon la durée ou l'intensité, ils observent une sensibilité au groupement pour l'intensité, mais pas pour la durée (Expérience 3). Ce n'est que chez les nourrissons de 9 mois que les effets sont retrouvés pour l'intensité et la durée (Exp. 1). On pourrait penser que cette inversion d'effet pourrait être due au type de matériel utilisé : alors que Trainor et Adams ont utilisé des tons complexes, Saffran et Hay (2012) ont utilisé des syllabes. Or, Saffran et Hay ont aussi réalisé une expérience chez les 9 mois avec cette fois-ci des tons (Exp. 2). Bien que les effets soient un peu plus faibles, ils retrouvent le même pattern de résultats que lorsqu'ils ont présenté les séquences de syllabes. Par conséquent, il est peu probable que cette différence entre les études soit liée à la nature du matériel en lui-même. Une autre possibilité repose sur le fait que Saffran et Hay ont couplé des indices prosodiques avec des indices statistiques (probabilité de transition), ce qui a pu interférer avec la perception de l'ITL. Ces résultats semblent néanmoins confirmer des différences d'émergences de la perception du groupement basé sur l'intensité et celui basé sur la durée, suggérant que ces indices ne seraient pas traités de la même façon.

Des différences de trajectoires développementales ressortent aussi de l'étude de Bion et al. (2011). Alors que jusqu'à présent les études concernaient la durée et l'intensité, ces auteurs ont présenté pour la première fois des séquences variant en pitch ainsi qu'en durée. Bion et collaborateurs (2011) ont familiarisé les nourrissons à une

séquence de 6 syllabes pendant 3 minutes consécutives. Ils ont testé ensuite leurs préférences pour les items (paire de syllabes) qui étaient trochaïques (long-court, aigu-grave) ou iambiques (court-long, grave-aigu) pendant la phase de familiarisation. Les auteurs reportent une préférence pour les items trochaïques (effet de familiarité) pour la séquence alternant en pitch, mais aucune pour la durée. Les auteurs interprètent cette différence en proposant des trajectoires développementales différentes entre les indices. Par ailleurs, des données obtenues chez le rat montrent des résultats similaires avec une sensibilité au groupement de pitch et non à la durée (de la Mora et al., 2013). Les auteurs proposent que la perception du groupement basé sur le pitch résulte probablement d'un processus perceptuel automatique puissant (universel), alors que celui sur la durée pourrait être lié à l'expérience linguistique, et donc émergerait plus tardivement dans le développement.

Seulement deux études se sont intéressées à l'impact de l'environnement linguistique sur l'ITL, en comparant directement des nourrissons exposés à des langues maternelles différentes. La première étude a voulu répliquer et étendre des travaux réalisés chez l'adulte (Iversen et al., 2008), à des nourrissons de 5-6 mois et de 7-8 mois, exposés soit à l'anglais, soit au japonais (Yoshida et al., 2010). Dans cette étude, les auteurs ont présenté une séquence de familiarisation composée de tons alternant en durée pendant 2 minutes. Pendant la phase de test, les auteurs ont mesuré la préférence des nourrissons pour des séquences de tons présentant soit un groupement iambique, soit un groupement trochaïque. Les auteurs n'ont pas testé l'intensité et le pitch. Les résultats montrent que seuls les nourrissons de 7-8 mois exposés à l'anglais traitent les séquences iambiques et trochaïques différemment (préférant les séquences trochaïques). Les nourrissons japonais ne montrent pas de préférence. Les auteurs suggèrent que les nourrissons apprenant l'anglais dès 7-8 mois sont capables de segmenter les séquences variant en durée selon un groupement iambique et interprètent cette préférence pour le groupement trochaïque comme un effet de nouveauté. Ces résultats étendent donc l'étude originale chez les adultes américains anglophones et japonais, confirmant l'impact de l'environnement linguistique sur le groupement ITL basé sur la durée, ici très précocement, puisque cette modulation est déjà présente à 7-8 mois. Or, ces auteurs n'ont pas présenté de séquence contrôle en phase de familiarisation, laissant ouverte la possibilité que cette préférence pour les items trochaïques en phase test reflète le biais trochaïque qui émerge également à cette période chez les nourrissons américains anglophones (Jusczyk et al., 1993) et non pas une sensibilité à l'ITL (Thiessen &

Saffran, 2007). De plus, la procédure utilisée est peut-être trop complexe pour des nourrissons aussi jeunes et l'absence d'effet ne résulte pas nécessairement d'un manque de sensibilité à cet indice.

La deuxième étude a utilisé la même procédure que Yoshida et al. (2010) avec une familiarisation plus courte (1mn30s), chez des bilingues dominants basques ou espagnols de 9 mois et, cette fois-ci, les auteurs ont testé la durée et l'intensité, mais pas le pitch (Molnar, Lallier, et al., 2014). Ils retrouvent comme Yoshida et al. (2010) une différence cross-linguistique pour la durée. En effet, les bilingues dominants espagnols avaient une préférence pour les séquences trochaïques, alors qu'aucune préférence significative n'était trouvée pour les bilingues dominants basques. Les auteurs expliquent la préférence pour les séquences trochaïques comme reflétant un effet de nouveauté (similaire à Yoshida et al. 2010). Toutefois, Molnar et al. (2014) reportent dans une analyse de corrélation chez les nourrissons bilingues dominants basques que, plus l'exposition au basque est importante, plus la préférence pour les séquences iambiques semble apparaître. Par contre, aucune différence cross-linguistique n'est retrouvée pour l'intensité, les deux groupes de nourrissons montrant une claire préférence pour les séquences trochaïques (effet de familiarité). Ces résultats répliquent donc dans le groupe de bilingues dominants espagnols ceux de Yoshida et al. (2010), chez des nourrissons plus âgés et dans une autre langue.

Toutes les études citées précédemment (voir Tableau 2-8) suggèrent qu'entre 7 et 9 mois des préférences de groupement émergent. Cependant, le nombre d'études est relativement limité et les résultats divergent concernant les questions de quand et comment les préférences sont observées. Par ailleurs, même s'il apparaît clairement que des modulations, liées à la nature de l'exposition linguistique, sont présentes très tôt dans le développement, cela reste encore peu compris. Les différences semblent être retrouvées principalement sur la durée dans les deux seules études cross-linguistiques réalisées chez des nourrissons. Pour certains auteurs, ces études suggèrent que le groupement basé sur la durée résulterait plus vraisemblablement de l'expérience linguistique, alors que le groupement basé sur le pitch et l'intensité serait régi par des mécanismes perceptifs auditifs plus généraux. En effet, certains auteurs proposent que le groupement de la durée serait dépendant des propriétés prosodiques de la langue au niveau du syntagme (Gervain & Werker, 2013; Iversen et al., 2008; Yoshida et al., 2010) ou encore au niveau du mot (Bhatara et al., sous presse, 2013). Toutefois, Yoshida et al. (2010) n'ont testé qu'un seul indice (durée) chez des nourrissons de 7-

8mois et leur familiarisation durait 2mn, alors que l'étude de Molnar et al. (2014) a été réalisée chez des nourrissons plus âgés, de 9 mois, (durée et intensité) avec une familiarisation plus courte (1mn30s). Donc ces patterns de résultats sont encore difficiles à comprendre. De plus, aucune étude n'a testé les trois indices en même temps, dans deux groupes de nourrissons du même âge exposés à des langues différentes, avec la même procédure.

Notre expérience a pour but de répondre aux questions suivantes : quand ces préférences de groupement apparaissent-elles, ont-elles des trajectoires développementales différentes et sont-elles précocement modulées par l'environnement linguistique. Nous avons décidé de nous baser sur le dessin expérimental de Bion et al. (2011) afin de tester des nourrissons du même âge (7,5mois), mais exposé soit au français, soit à l'allemand.

Tableau 2-8 : Tableau récapitulatif des études ITL

Etude	Langue		Stimulus	Age	Tâche	Durée de la familiarisation	Préférence	Effet
Trainor & Adams, 2000	anglais	Monolingues	Tons complexes	8 mois	Détection de silences			nouveauté
Yoshida et al. 2010	anglais	Monolingues	Tons complexes	5-6 mois	Segmentation	2 mn	aucune	-
				7-8 mois	Segmentation	2 mn	trochaïque	nouveauté
	japonais	Monolingues		5-6 mois	Segmentation	2 mn	aucune	-
				7-8 mois	Segmentation	2 mn	tendance iambique	familiarité
Bion et al. 2011	italien	Monolingues	Syllabes	7-8 mois	Segmentation	3 mn	aucune	-
Hay & Saffran, 2012	anglais	Monolingues	Syllabes	6 mois	Segmentation	3 mn	aucune	-
				9 mois	Segmentation	3 mn	iambique	familiarité
			Tons	9 mois	Segmentation	3 mn	tendance iambique	familiarité
Molnar et al. 2014	espagnol & basque	Bilingues espagnol dominant	Tons	9 mois	Segmentation	1,5 mn	trochaïque	nouveauté
	espagnol & basque	Bilingues basque dominant	Tons	9 mois	Segmentation	1,5 mn	tendance iambique	familiarité
Trainor & Adams, 2000	anglais	Monolingues	Tons complexes	8 mois	Détection de silences			-
Hay & Saffran, 2012	anglais	Monolingues	Syllabes	6 mois	Segmentation	3 mn	trochaïque	familiarité
				9 mois	Segmentation	3 mn	trochaïque	familiarité
			Tons	9 mois	Segmentation	3 mn	trochaïque	familiarité
Molnar et al. 2014	espagnol & basque	Bilingues espagnol dominant	Tons	9 mois	Segmentation	1,5 mn	trochaïque	familiarité
	espagnol & basque	Bilingues basque dominant	Tons	9 mois	Segmentation	1,5 mn	trochaïque	familiarité
Bion et al. 2011	italien	Monolingues	Syllabes	7-8 mois	Segmentation	3 mn	trochaïque	familiarité

Notre hypothèse est que si la perception de l'ITL est déjà présente à 7 mois et n'est pas encore modulée par les propriétés de la langue maternelle (ITL universel), alors ces enfants seront capables de segmenter les séquences alternant en pitch et en intensité trochaïquement, et les séquences en durée iambiquement. Si par contre, l'ITL est déjà modulée par l'environnement linguistique, alors des différences comme celles observées chez les adultes pourraient être déjà présentes. Notre étude chez les adultes a

révélé des différences cross-linguistiques de stabilité du jugement rythmique pour l'intensité, mais aucune n'est trouvée pour la durée ou le pitch.

2.2.2 Expérience 2

2.2.2.1 *Méthodes*

2.2.2.1.1 Participants

Nous avons testé 160 nourrissons monolingues de 7.5 mois nés à terme, apprenant le français à Paris (France) et l'allemand à Potsdam (Allemagne). 20 nourrissons par condition expérimentale et par langue maternelle (allemand, français) ont été inclus (durée, intensité, pitch, contrôle). 45 nourrissons ont été exclus pour cause de pleurs et/ou de fatigue, ou pour ne pas avoir fini la passation.

Tableau 2-9 : Détails des informations des participants par groupe.

Condition	Monolingues Français				Monolingues Allemands			
	Durée	Intensité	Pitch	Contrôle	Durée	Intensité	Pitch	Contrôle
Filles	12	10	9	8	10	12	10	10
Nourrissons exclus	3	6	6	4	5	7	10	4
Age moyen (mois) Écart type	7,5 (0,3)	7,4 (0,3)	7,5 (0,3)	7,5 (0,2)	7,6 (0,3)	7,4 (0,3)	7,2 (0,2)	7,5 (0,4)
Intervalle (ans)	7,1-8,0	7,0-7,9	7,1-8,0	7,1-7,8	7,1-8,0	7,0-8,0	7,1-7,8	7,0-8,3

2.2.2.1.2 Stimuli

Initialement, nous voulions utiliser le même matériel que pour l'expérience chez les adultes. Toutefois, des études pilotes ont indiqué que ce matériel n'était pas adapté pour des nourrissons aussi jeunes ; aucun résultat significatif n'ayant été obtenu (voir Annexe 1). L'augmentation du nombre de syllabes par rapport à Bion et al (2011) a pu rendre la séquence plus difficile à mémoriser (Bion : 6 versus stimuli de notre étude adulte : 8). Nous avons donc décidé de simplifier le matériel en nous rapprochant du matériel initial de Bion et al. (2011) : pour la phase de familiarisation, la séquence était composée de 6 syllabes séparées par des pauses de 100ms (au lieu de 8 sans pauses chez

l'adulte), dont les variations prosodiques entre syllabes consécutives différaient en fonction de la condition (comme pour les adultes). Elles alternaient en durée (courte-longue), en intensité (forte-faible) ou en pitch (grave-aigu). Enfin, dans la condition contrôle, toutes les syllabes avaient les mêmes durée, intensité et pitch.

Les stimuli étaient composés d'une séquence de 6 syllabes CV différentes, obtenues en combinant six voyelles /a/, /e/, /i/, /o/, /u/, /y/ avec six consonnes /f/, /n/, /g/, /p/, /r/, /z/. Cet ensemble de phonèmes a été sélectionné avec la même stratégie que pour l'étude adulte.

2.2.2.1.2.1 Construction de la séquence de familiarisation

Ces syllabes ont été agencées de manière à éviter de former des mots bisyllabiques allemands ou français, soit la séquence : / **na zu gi pe fy ro** /. Contrairement à l'étude chez les adultes, nous avons rétabli les pauses de 100ms entre les syllabes qui étaient initialement présentes dans l'étude de Bion et al. (2011). Cette séquence a été répétée 66 fois (environ 3 minutes). Nous avons synthétisé cette séquence de familiarisation à l'aide de MBROLA, dans deux versions, une avec la voix féminine allemande de5 et l'autre avec la voix féminine française fr4. Les mêmes valeurs de variations acoustiques que dans l'expérience adulte ont été appliquées à ces nouvelles séquences de familiarisation (Exp. 1, Tableau 2-2).

2.2.2.1.2.2 Début de la séquence de familiarisation

Nous avons observé dans notre étude chez les adultes que les participants pouvaient, en fonction de l'indice, se baser sur les propriétés de la première paire de syllabes pour grouper la séquence de familiarisation (effet de la rampe). De ce fait, en plus de la rampe, qui consistait en une augmentation graduelle des contrastes de durée, d'intensité ou de pitch, implémentée à partir de la 3^{ème} syllabe ou la 4^{ème} syllabe, nous avons ajouté un bruit blanc au début de la séquence. Nous nous sommes basés sur le dessin expérimental de Bhatara et collaborateurs (2013) qui, eux aussi, avaient utilisé cette méthode de masquage, avec une combinaison de bruit blanc sur les 4 premières répétitions de la séquence (24 syllabes)

2.2.2.1.2.3 Items de test

Pour la phase de test, deux types de stimuli ont été synthétisés avec MBROLA, encore une fois avec la voix allemande de5 ou française fr4. Tous ces stimuli étaient

composés de paires de syllabes sans modulation prosodique. Les syllabes de ces items ont donc eu toutes la même durée, intensité ou pitch (Exp. 1, Tableau 2-2). Toutes les syllabes présentes dans les paires étaient les mêmes que celle de la phase de familiarisation. Les syllabes ont été combinées de différentes façons ; la première possibilité de combinaison était celle où, dans la phase de familiarisation, la syllabe longue, forte ou aiguë était en position initiale et, par conséquent, la syllabe courte, faible ou grave était en position finale (items trochaïques). La seconde possibilité était celle où, dans la phase de familiarisation, la syllabe courte, faible ou grave était en position initiale et, par conséquent, la syllabe longue, forte ou aiguë était en position finale (items iambiques). Nous avons fait le choix de ne garder que les deux types d'items principaux (trochaïques et iambiques), étant donné le faible nombre d'essais tests possible dans ce type d'études avec des nourrissons.

Chaque essai est composé d'un des 6 items, répétés 16 fois (14,5s). Tous les participants ont entendu les mêmes items (3 trochaïques, 3 iambiques, répétés 2 fois, pour un total de 12 essais), quelle que soit la phase de familiarisation entendue, et ceci avec la même voix que celle entendue dans la phase de familiarisation (soit de5 ou fr4).

2.2.2.1.3 Procédure

Nous avons utilisé la procédure de regard préférentiel (*Headturn Preference Procedure*, HPP) (D. G. K. Nelson et al., 1995). Les nourrissons étaient assis confortablement sur les genoux d'un de leurs parents au centre d'une cabine insonorisée. Devant le nourrisson, une lumière verte était fixée sur le panneau central. Deux autres lumières rouges étaient fixées de chaque côté, sur des panneaux latéraux, à la même hauteur que la lumière verte. Les haut-parleurs étaient cachés derrière les panneaux sous chaque lumière rouge (Figure 2.9).

L'expérimentateur était placé en dehors de la cabine insonorisée et observait les mouvements du nourrisson sur un écran télévisé relayant en direct la vidéo de l'enfant, via une caméra cachée, fixée sous la lumière verte. L'expérimentateur contrôlait la présentation des stimuli audio et le clignotement des lumières par le biais d'un boîtier à 3 boutons relié à l'ordinateur. Un bouton servait à déclencher l'essai et deux autres à mesurer le temps de regard dirigé vers chaque lumière rouge. L'ordinateur enregistrait la durée de chaque regard de l'enfant, indiqué par les pressions des boutons de l'expérimentateur. L'expérimentateur et le parent portant l'enfant portaient un casque

audio par lequel était diffusée une bande sonore qui masquait les stimuli joués dans la cabine.

Pendant la phase de familiarisation, le nourrisson a écouté en continu, provenant des deux haut-parleurs simultanément (SONY xs-F1722), via un amplificateur audio (Marantz PM4000), une des quatre séquences de familiarisation (durée, intensité, pitch ou contrôle). Pendant que le nourrisson écoutait la séquence de familiarisation, les lumières vertes et rouges clignotaient en fonction du comportement du nourrisson et indépendamment du son joué. Pendant la phase de test, le nourrisson a écouté 12 essais (3 trochaïques et 3 iambiques, répétés 2 fois), chaque essai étant composé d'un des six items répété 16 fois (environ 14s).



Figure 2.9 : Cabine HPP (à gauche) et box de l'expérimentateur (à droite).

Chaque essai commençait par la lumière verte clignotante, afin de centrer l'attention du nourrisson. Dès lors qu'il avait regardé cette lumière verte, l'expérimentateur déclenchait le clignotement d'une des lumières rouges sur un des côtés, en appuyant sur le bouton central du boîtier. Quand le nourrisson se tournait d'au moins 30° vers la lumière rouge, l'expérimentateur déclenchait, via un des boutons latéraux du boîtier, le stimulus audio test joué par le haut-parleur fixé en dessous de la lumière clignotante. Si le nourrisson arrêta de regarder la lumière pendant plus de 2 secondes alors que le son était joué, la lumière rouge s'éteignait, le son s'arrêtait et la lumière verte centrale recommençait à clignoter signifiant le début d'un nouvel essai. Par contre, si l'enfant détournait le regard de la lumière de moins de 2s et revenait sur la cible, alors le son continuait d'être joué. Cependant, le temps passé à « regarder

ailleurs » était automatiquement soustrait des temps d'orientation pour cet essai. Donc, le temps maximum d'orientation du regard pour un essai correspondait à la durée du fichier son (ici 14,5s). Un PC (DELL Optiplex) a permis de présenter les stimuli et d'enregistrer les données.

Nous avons présenté à quatre groupes de nourrissons, pour chaque langue maternelle, l'une des quatre conditions : durée, intensité, pitch ou contrôle. Chaque condition correspondait à une phase de familiarisation spécifique. Cependant, la phase de test était exactement la même pour les quatre conditions. Les 12 items étaient présentés dans un ordre aléatoire pour chaque nourrisson et l'ordre était contrebalancé d'un nourrisson à l'autre. Dans chaque condition, les nourrissons ont entendu soit les stimuli avec la voix allemande de5, soit avec la voix française fr4. De plus, dans chaque condition, la moitié des enfants ont entendu une séquence dont la rampe commence sur la 3^{ème} syllabe et l'autre moitié sur la 4^{ème}.

2.2.2.2 Résultats

Nous avons décidé de restreindre nos analyses aux temps de regard moyen des trois premiers essais trochaïques et des trois premiers essais iambiques. Nous justifions ce choix par le fait que notre phase test contenait 12 essais, rendant ainsi notre phase test assez longue, après 3 minutes d'écoute en continu de la séquence de familiarisation. Les temps de regards de la deuxième moitié de l'expérience étaient faibles, reflétant une perte ou un manque de concentration (pour l'analyse voir Annexe 2).

2.2.2.2.1 Groupement selon la loi lambo-trochaïque

2.2.2.2.1.1 Analyse globale

Les temps d'orientation de regard pour les items qui avaient une structure trochaïque (long-court/fort-faible/aigu-grave) ou iambique (court-long/faible-fort/grave-aigu) dans la phase de familiarisation ont été moyennés pour chaque participant. Pour la condition contrôle, nous avons attribué le type « trochaïque » pour les paires de syllabes adjacentes paires (1-2) et le type « iambique » pour les paires de syllabes adjacentes impaires (2-3). Les temps moyens d'orientation du regard ont été calculés pour chaque groupe et condition (durée, intensité, pitch et contrôle). Les résultats sont illustrés dans la Figure 2.10.

Nous avons réalisé une ANOVA à mesures répétées sur les temps moyens d'orientation du regard, avec pour facteur intra-sujet le type d'items (iambique, trochaïque) et comme facteurs inter-sujet, la langue maternelle (français, allemand) et la condition (durée, intensité, pitch, contrôle). Un effet significatif du type d'items a été retrouvé [$F(1, 151) = 4.33, p = .04, \eta^2_p = .03$], ainsi qu'une interaction entre le type d'items et la condition [$F(3, 151) = 3.08, p = .03, \eta^2_p = .06$]. Ces résultats montrent que les nourrissons ont des temps d'orientation de regard qui diffèrent entre les items trochaïques et iambiques, en fonction de la condition entendue dans la phase de familiarisation. Des effets marginaux de la condition [$F(3, 151) = 2.4, p = .07, \eta^2_p = .05$] et de l'interaction entre la langue maternelle et la condition [$F(3, 151) = 2.1, p = .09, \eta^2_p = .04$] ont été également observés. Aucun autre effet ni interaction n'ont atteint la significativité. Afin de comprendre plus en détails l'interaction entre la condition et le type d'items, nous avons analysé séparément chaque condition.

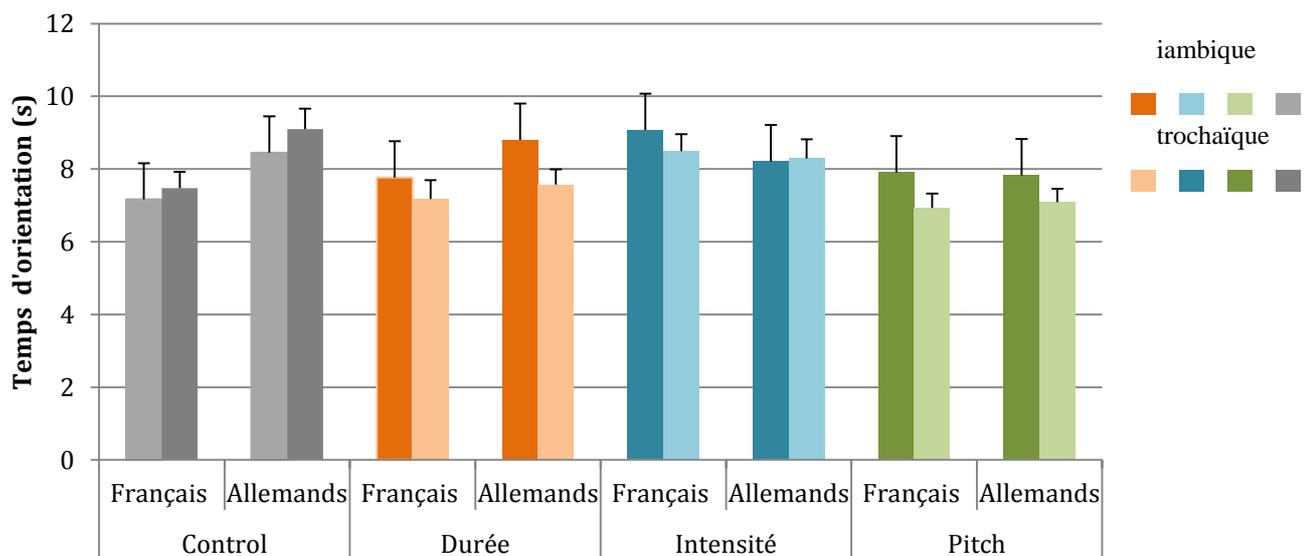


Figure 2.10: Temps d'orientation de regard pour les items iambiques et items trochaïques, en fonction de la condition (contrôle, durée, intensité et pitch) et de la langue maternelle des nourrissons (allemand, français).

2.2.2.2.1.2 Condition contrôle

Une ANOVA à mesures répétées a été réalisée sur les temps moyens d'orientation du regard, avec en facteurs le type d'items et la langue maternelle. Seul, l'effet principal de la langue maternelle [$F(1,38) = 5.19, p = .03, \eta^2_p = .12$] est

significatif. Les nourrissons allemands ont des temps de regard plus longs que les français. Mais aucune interaction significative n'est présente avec le type d'items.

Les nourrissons ne présentent donc pas de préférence pour un type de paires, suggérant qu'ils ont également mémorisé (ou non) celles-ci et il n'y a pas de biais dû au matériel utilisé.

2.2.2.2.1.3 Condition durée

Une ANOVA a été réalisée sur les temps moyens d'orientation du regard avec, en facteur intra-sujet, le type d'items et, en facteurs inter-sujet, la langue maternelle et le type de rampe. Elle révèle un effet principal du type d'items [$F(1,35) = 7.13$, $p = .01$, $\eta^2_p = .17$]. Les nourrissons avaient des temps d'orientation du regard plus longs pour les items iambiques ($M_{\text{iambique}} = 8.23$ s) que pour les items trochaïques ($M_{\text{trochaïque}} = 7.37$ s) et ceci ne différait pas en fonction de leur langue maternelle. Une interaction significative entre le type de rampe et le type d'items a été retrouvée [$F(1, 35) = 9.47$, $p = .004$, $\eta^2_p = .21$]. Aucun autre effet ou interaction n'était significatif.

Afin de comprendre l'interaction, nous avons évalué l'effet du type d'items en fonction du type de rampe. Il est apparu que les nourrissons ont des temps d'orientation du regard plus longs pour les items iambiques que trochaïques si la rampe commençait par la 4^{ème} syllabe ; voir Figure 2.11, [$M_{\text{iambique}} = 9.03$ s vs $M_{\text{Trochaïque}} = 7.27$ s, $t(20) = 4.02$, $p < .001$]. Par contre, s'ils avaient entendu une séquence dont la rampe commençait par la 3^{ème} syllabe, aucune différence de performance n'était retrouvée [$M_{\text{Trochaïque}} = 7.49$ s vs $M_{\text{iambique}} = 7,39$ s, $t(18) = -.25$, $p = .804$].

Les nourrissons allemands et français ont des temps de regard plus longs pour les items iambiques, ils sont donc capables d'exploiter la durée pour grouper les syllabes. Cet effet n'est pas modulé par la langue maternelle, mais influencé par le type de rampe. Ces résultats suggèrent qu'un indice additionnel, comme l'alignement de la rampe, favorise le groupement de la séquence de syllabes selon l'ITL.

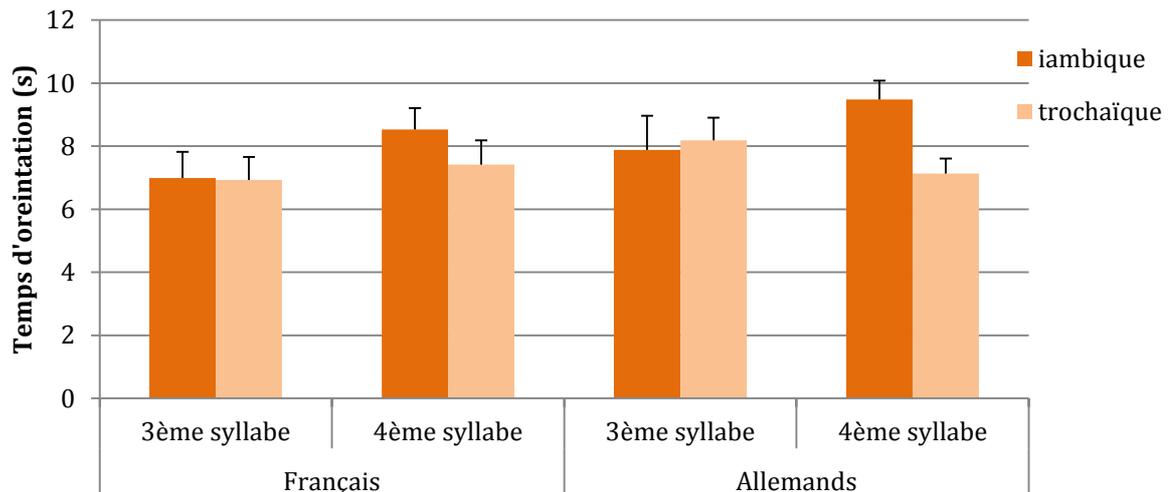


Figure 2.11 Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition durée, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe).

2.2.2.2.1.4 Condition intensité

Une ANOVA sur les temps moyens d'orientation du regard, avec les facteurs type d'items, langue maternelle et type de rampe, n'a révélé aucun effet ou interaction significatif. Seul le type de rampe interagit marginalement avec le type d'items et la langue maternelle [$F(1, 36) = 4.08$, $p = .051$, $\eta^2_p = 0.1$, $d = .67$]. Bien que cette triple interaction soit marginale, nous avons évalué l'effet du type d'items en fonction du type de rampe ; voir Figure 2.12. Il apparaît que si la rampe commence par la 3^{ème} syllabe, les nourrissons français regardent autant de temps les items iambiques que les items trochaïques [$M_{\text{Iambique}} = 8.89$ s vs $M_{\text{Trochaïque}} = 7.94$ s, $t(9) = 1.37$, $p = .204$], alors que les allemands préfèrent regarder les items trochaïques plutôt que les items iambiques [$M_{\text{Trochaïque}} = 8.89$ s vs $M_{\text{Iambique}} = 7.94$ s, $t(9) = -2.41$, $p = .039$]. Par contre, si la rampe commence par la 4^{ème} syllabe, aucune différence de temps d'orientation du regard n'est retrouvée, ni chez les français [$M_{\text{Iambique}} = 9.26$ vs $M_{\text{Trochaïque}} = 9.05$, $t(9) = .280$, $p = .786$], ni chez les allemands [$M_{\text{Iambique}} = 9.35$ vs $M_{\text{Trochaïque}} = 8.22$, $t(9) = 1.07$, $p = .313$].

En résumé, ces résultats montrent que les nourrissons ont eu tendance à grouper différemment les séquences en fonction du type de rampe et de la langue maternelle. Il apparaît que, quand la position de la rampe est aligné au le groupement trochaïque, les nourrissons allemands vont avoir tendance à préférer les items trochaïques ; cet effet n'est pas retrouvé chez les français. Par contre, si la rampe n'est pas alignée avec le

groupement de l'ITL, alors les temps de regard ne sont pas significativement différents entre les deux types d'items dans les deux groupes.

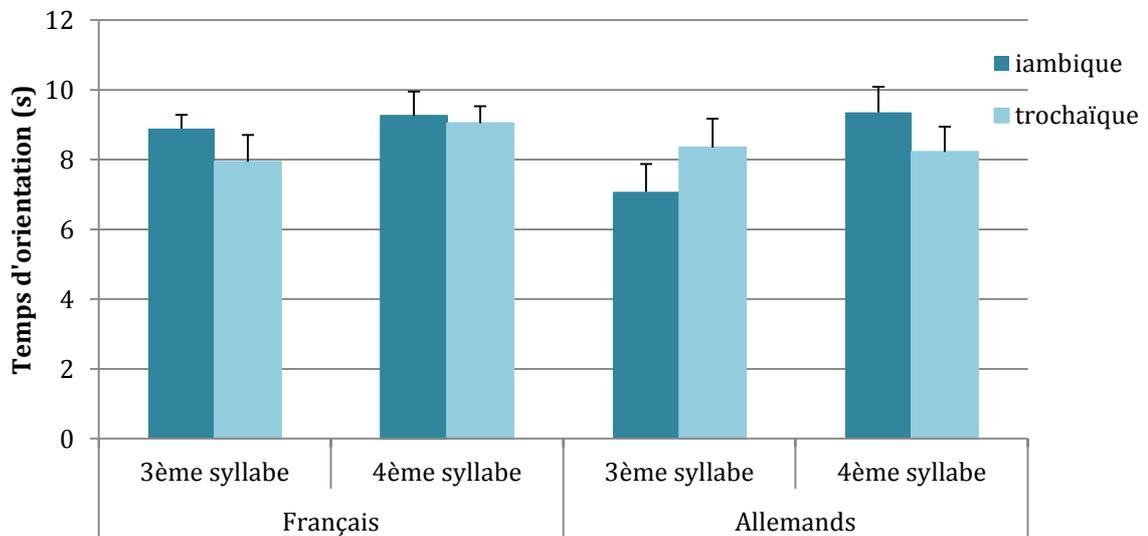


Figure 2.12: Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition intensité, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe), puis par groupe de langue maternelle (français et allemand).

2.2.2.2.1.5 Condition pitch

Une ANOVA sur les temps moyens d'orientation du regard, avec les facteurs type d'items, langue maternelle et type de rampe, révèle un effet principal du type d'items [$F(1,36) = 4.72, p = .036, \eta^2_p = .12$]. Les nourrissons regardent plus longtemps les items iambiques ($M_{iambique} = 7.87$ s) que les items trochaïques ($M_{trochaïque} = 7.01$ s) et ceci ne diffère pas en fonction de leur langue maternelle ; voir Figure 2.13. Aucun autre effet ou interaction n'est significatif.

Les nourrissons allemands et français ont des temps de regard plus longs pour les items iambiques, ils sont donc capables d'exploiter le pitch pour grouper les syllabes. Cet effet n'est pas modulé par la langue maternelle.

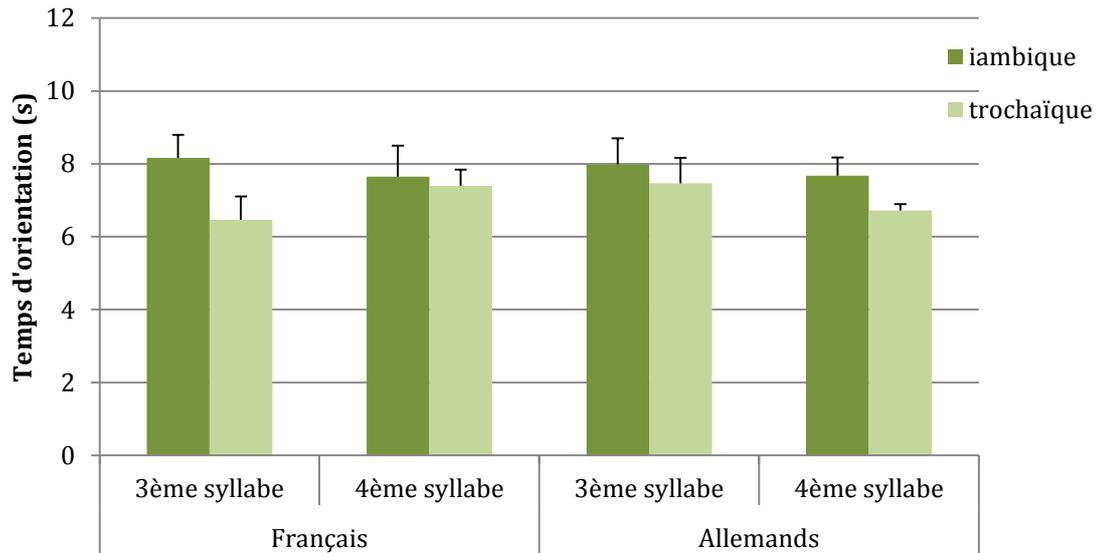


Figure 2.13 : Temps d'orientation du regard pour les items iambiques et items trochaïques pour la condition pitch, séparés par type de rampe (3^{ème} ou 4^{ème} syllabe), puis par groupe de langue maternelle (français et allemand).

2.2.2.3 Discussion

Dans cette étude, nous avons évalué la perception de l'ITL chez des nourrissons exposés à l'allemand ou au français lors de l'écoute de séquences linguistiques. Quelles étaient les préférences de groupement pour les trois indices et comment ont-elles pu être modulées par l'environnement linguistique ?

Dans nos deux groupes linguistiques, les nourrissons regardent plus longtemps, pour la durée et le pitch, les items tests qui ont été iambiques dans la phase de familiarisation. Notre étude montre ainsi que les nourrissons français et allemands de 7-8 mois traitent différemment les deux types d'items tests, suggérant qu'ils sont capables d'exploiter les variations de pitch et de durée pour segmenter la séquence de syllabes. Aucun effet n'est retrouvé pour l'intensité ; cependant, des différences cross-linguistiques semblent être présentes concernant leur sensibilité à la position de la rampe. Les résultats de cette étude amènent plusieurs interprétations possibles.

2.2.2.3.1 Durée

2.2.2.3.1.1 *Validation de l'effet*

Concernant le biais prosodique basé sur la durée, les études ont trouvé des patterns de préférences complètement opposés aux nôtres. En effet, les travaux

antérieures sur la durée chez des nourrissons américains anglophones de 7-8 mois (Yoshida et al., 2010) ou encore chez des nourrissons bilingues espagnols-basques (dominants espagnols) de 9 mois (Molnar, Lallier, et al., 2014), ont mis en évidence des préférences pour le groupement trochaïque. Les auteurs ont interprété ces résultats comme un effet de nouveauté selon le modèle de Hunter et Ames (1988). A l'inverse, nous avons trouvé une préférence pour les items iambiques lors de l'écoute de séquences variant en durée, que nous avons interprété comme un effet de familiarité. Une interprétation possible pourrait être que nos enfants ont eu plus de difficultés que ceux des deux autres études à traiter la durée. Une autre possibilité de cette inversion de patterns de résultats concerne la nature du matériel et des procédures utilisées. Alors que ces deux dernières études ont utilisé des répétitions de tons complexes (variabilité segmentale faible), nous avons utilisé différentes syllabes (combinaison de différentes voyelles et consonnes) présentant une variabilité segmentale plus importante. De plus, les nourrissons dans notre tâche devaient mémoriser une séquence de syllabes dans la phase de familiarisation, ce qui diffère beaucoup des dessins expérimentaux utilisés par Yoshida et al. (2010) et Molnar et al (2014). La comparaison entre les études antérieures est donc difficile, étant donnée la variabilité des méthodes utilisées.

2.2.2.3.1.2 Effet de la rampe

Nous avons essayé de masquer par un bruit blanc le début de la séquence mais celui-ci ne semble pas être suffisant pour éviter que les nourrissons groupent les syllabes avec les caractéristiques du premier son entendu. Puisque, sans bruit blanc, le même phénomène a été observé dans l'Expérience 1 de cette thèse et dans d'autres études ITL (Hay & Diehl, 2007; Trainor & Adams, 2000). En fonction du type de rampe, les nourrissons ne vont pas mémoriser la séquence de la même manière.

Pour la durée, si le groupement de paires par défaut dès la première syllabe entendue est aligné à celui de l'ITL, ce qui est le cas quand l'allongement est sur la 4^{ème} syllabe, alors les nourrissons sont capables de segmenter les séquences. Par contre, quand l'allongement est sur la 3^{ème} syllabe, le groupement par défaut n'est pas aligné et les nourrissons ne sont plus capables de segmenter. Le groupement par défaut agirait donc comme un facilitateur pour détecter le groupement rythmique selon l'ITL et, donc, pour segmenter efficacement la séquence. En revanche, si le groupement par défaut est opposé au groupement prosodique, alors les préférences vont disparaître, pouvant suggérer que le groupement par défaut a autant de poids que la variation de la durée. Ici

les deux indices de groupement (par défaut et prosodique) s'opposent et les nourrissons semblent être sensibles aux deux. Mais rappelons que sans la prosodie ils ne peuvent segmenter la séquence, comme nous l'indique la condition contrôle où aucun effet n'est retrouvé. Le groupement par défaut n'est donc pas suffisant pour qu'ils puissent grouper la séquence.

C'est probablement l'une des raisons pour laquelle Bion et al. (2011) n'ont pas trouvé d'effet pour la durée chez des nourrissons italiens, mais que nous en retrouvons chez les nourrissons allemands et français. Les auteurs ont justement allongé systématiquement la 3^{ème} syllabe. Dans cette condition de rampe, nous retrouvons exactement leur résultat : nous n'obtenons aucune préférence. Il est donc possible qu'en présentant l'allongement sur la 4^{ème} syllabe, Bion et collaborateurs (2011) auraient pu obtenir une préférence comme nous. Par ailleurs, l'expérience chez les rats a aussi utilisé le même type de procédure que Bion et al. (2011) et a obtenu les mêmes résultats que chez les nourrissons de 7 mois italiens. Il est donc probable que l'absence de résultats de Bion et al. (2011) sur le groupement basé sur la durée ne reflète pas une différence de trajectoire développementale entre la durée et le pitch, mais plutôt une différence de méthodologie. Toutefois, s'il a été démontrée la présence, chez les nourrissons monolingues anglais, bilingues espagnols-basques (espagnols dominants) et maintenant monolingues français et allemands, d'une sensibilité à la durée, pour d'autres nourrissons exposés à d'autres langues plus éloignées prosodiquement (japonais), aucune préférence n'a été retrouvée. Ces résultats suggèrent ainsi que d'autres facteurs interagissent avec ce phénomène, notamment la nature de l'exposition linguistique.

2.2.2.3.1.3 Pitch

Concernant le biais prosodique basé sur le pitch, Bion et al. (2011) retrouvent un effet de familiarité chez des nourrissons italiens, alors que nous trouvons un effet de nouveauté. Ici le matériel est sensiblement le même, ce qui indique que le type de matériel utilisé n'est pas le seul responsable de cet effet ; la langue maternelle à laquelle ont été exposés les nourrissons apparaît comme un facteur. Notre étude a été réalisée chez des nourrissons exposés à l'allemand et au français, qui sont des langues différentes de l'italien. D'autres études cross-linguistiques sur la perception du groupement basé sur le pitch sont nécessaires.

2.2.2.3.1.4 Intensité

Aucun effet principal n'est mis en évidence dans notre étude pour l'intensité. Notre étude est la première à avoir utilisé des syllabes pour tester le groupement rythmique (uniquement) pour cet indice, donc il est encore difficile d'interpréter ces résultats. En effet, la majorité des études ont utilisé des tons (Molnar, Lallier, et al., 2014; Trainor & Adams, 2000) ; une autre a présenté une grammaire artificielle contenant des indices de probabilité de transition (Hay & Saffran, 2012). Dans cette dernière étude, il a été montré que des nourrissons américains anglophones de 6 mois sont sensibles au groupement basé sur l'intensité. En effet, ils sont capables d'exploiter cet indice pour grouper les syllabes. Seulement nous ne pouvons pas déterminer si cet effet est lié à l'exposition à l'anglais, ou à une compétence universelle.

Or, dans notre étude, nous ne retrouvons pas d'effet principal dans les deux groupes linguistiques, mais un effet est présent et il est lié à la rampe chez les nourrissons allemands. Effectivement, nous retrouvons des effets par rapport au type de rampe pour l'intensité, mais ils ne sont pas tout à fait similaires à ceux trouvés pour la durée. Nos résultats montrent que, si la rampe est alignée au groupement de l'ITL dans la phase de familiarisation, alors les allemands ont tendance à préférer les items trochaïques (effet de familiarité) par rapport aux items iambiques, mais pas les français. Ici le groupement par défaut semble être un indice supplémentaire à l'indice d'intensité, chez les nourrissons allemands, pour détecter le groupement trochaïque. Il n'est pas surprenant de retrouver cette sensibilité chez les nourrissons allemands, et non chez les français, puisque l'allemand et l'anglais ont des prosodies très proches, mais éloignées du français. Ces langues (anglais et allemand versus français) diffèrent largement au niveau du mot, mais possèdent des similarités au niveau du syntagme (voir aussi chapitre introductif). C'est également à cette période et chez des nourrissons exposés à l'anglais et l'allemand qu'émerge le biais trochaïque (Höhle et al., 2009; Jusczyk et al., 1993).

Enfin, il est possible que des différences en termes de difficulté de traitement entre la durée et le pitch existent. Nous retrouvons un effet de familiarité pour la durée qui est modulé par l'effet de la rampe, alors que les résultats du pitch ne sont pas affectés par les caractéristiques de la rampe. Il est possible que traiter l'indice de pitch puisse appeler un mécanisme de traitement plus automatique, peut-être plus acoustique, alors que celui de la durée fait appel à un traitement plus linguistique, et moins

automatique, demandant un temps un peu plus long pour traiter la séquence. Cette hypothèse est également soutenue par une étude montrant que ce mécanisme de groupement du pitch est présent chez d'autres espèces comme le rat (de la Mora et al., 2013), suggérant ainsi la présence d'un mécanisme auditif perceptuel plus général, non spécifique au langage, pour le pitch, puisqu'il semble être présent même en l'absence d'expérience langagière. De futures recherches sont nécessaires pour comprendre à quel niveau les indices sont traités et à quels mécanismes ils font appel.

En conclusion, cette étude met en évidence, pour la première fois avec des stimuli linguistiques, des préférences de groupement pour le pitch et la durée, chez des nourrissons de 7 mois apprenant le français et l'allemand. Une différence marginale cross-linguistique est trouvée pour l'intensité. Ces mécanismes permettant de détecter le groupement rythmique sont donc en place très tôt et pourraient initialiser ainsi l'acquisition des mots ou encore de l'ordre des mots de leur langue maternelle.

2.2.3 Points clefs

- Des préférences de groupement rythmique sont observées pour la durée et le pitch chez des nourrissons allemands et français.
- Nous généralisons les travaux de Bion et al. (2011) chez des nourrissons italiens pour le groupement basé sur le pitch chez des nourrissons français et allemands.
- De faibles différences cross-linguistiques sont trouvées pour l'intensité.

2.3 Conclusion

Dans ces deux études, un matériel similaire a été utilisé entre les adultes et les nourrissons de 7,5 mois afin d'explorer la perception de l'ITL. Nous avons observé et mesuré de quelle manière les adultes et les nourrissons groupent des séquences de syllabes présentées avec une modulation prosodique. Nos résultats avec les adultes français et allemands confirment le groupement basé sur la durée et l'intensité, mais pas celui basé sur le pitch. De plus, des différences cross-linguistiques semblent être présentes entre nos deux groupes linguistiques pour l'intensité. Chez les nourrissons de 7,5 mois, un biais de groupement est retrouvé pour la durée et le pitch ; en revanche, aucun n'est mis en évidence pour l'intensité. Les nourrissons sont donc capables de traiter les modulations de pitch et de durée à 7,5 mois.

Est-il possible que cette sensibilité à l'ITL soit présente encore plus tôt dans le développement ? Ces résultats reflètent-ils une prédisposition pour traiter le groupement rythmique, ou l'exposition à la langue maternelle est-elle seule responsable du développement de ces compétences ? La majorité des études n'ont étudié ces capacités qu'à partir de 5-6 mois au plus tôt, ce qui nous pousse à nous questionner quant au point de départ, l'origine de ces sensibilités. Le chapitre 3 va donc explorer ces capacités très précoces de sensibilité à l'ITL.

3 ITL ET PREFERENCES DE PATTERNS RYTHMIQUES CHEZ LE NOURRISSON

3.1 CHAPITRE 3 : L'ITL à la naissance

Après avoir étudié le développement du biais iambo-trochaïque, en testant des nourrissons de 7 mois exposés au français et à l'allemand, nous allons maintenant nous intéresser à l'origine de celui-ci. Aucune étude sur la perception du biais n'a été réalisée à la naissance. Comment les nouveau-nés sont-ils équipés pour traiter des indices prosodiques à la naissance et plus précisément, sont-ils déjà sensibles au groupement prosodique. De plus, comme nous l'avons vu précédemment, ce phénomène peut être modulé par l'expérience linguistique. Est-il possible que l'expérience prénatale influence déjà la perception du groupement ? Ces questions restent largement inexplorées aujourd'hui encore.

Contrairement à la première partie où nous avons étudié le phénomène du groupement rythmique dans une tâche de segmentation, ici nous allons nous intéresser à la préférence de patterns rythmiques. Nous avons fait ce choix pour plusieurs raisons. La première est que nous étudions une population différente (des nouveau-nés) et que nous utilisons une technique expérimentale particulière (la NIRS, *near-infrared spectroscopy* : « Spectrométrie proche de l'infra-rouge »). Afin de s'adapter à ces deux contraintes, nous avons fait le choix de leur présenter des séquences où les sons étaient déjà groupés selon un pattern rythmique donné (iambique ou trochaïque).

Dans ce chapitre, nous allons justifier le choix de cette technique pour tester les préférences de groupement, puis nous allons présenter les quatre expériences nous permettant de répondre à nos deux questions principales : les nouveau-nés sont-ils sensibles à l'ITL et l'expérience linguistique (ici, essentiellement prénatale) module-t-elle déjà ce biais à la naissance ?

3.1.1 Introduction : Pourquoi utiliser l'imagerie optique pour tester l'ITL?

Ces dernières années ont été marquées par un avancement remarquable des diverses techniques d'imagerie cérébrale, notamment l'électroencéphalographie (EEG), l'imagerie par résonance magnétique (IRM), la tomographie par émission de positon (PET), la magnétoencéphalographie (MEG) ou encore l'imagerie optique aussi nommée spectroscopie proche infrarouge (*near-infrared spectroscopy*, NIRS). Ces techniques ont permis d'approcher de plus près l'organisation cérébrale, tant au point de vue anatomique que fonctionnel. Néanmoins, de nombreuses techniques comme l'IRM

ou la PET restent difficilement applicables aux études développementales, en particulier aux plus jeunes nourrissons, car elles présentent des contraintes (immobilité, usage de radio traceur) et des inconvénients (bruits), rendant leur usage courant peu pratique chez ces participants vulnérables. En revanche, la NIRS présente plusieurs avantages pratiques, notamment son utilisation silencieuse, ainsi que des contraintes de mouvement beaucoup moins strictes qu'en EEG ou en IRM. Son utilisation est alors tout à fait adaptée à de très jeunes populations, notamment dans des études utilisant des stimuli auditifs.

3.1.1.1 Un bref historique

Les travaux pionniers de l'application de la NIRS chez l'humain remontent à 1977 (Jöbsis, 1977; Jöbsis-vanderVliet, 1999). Au départ, la NIRS a été utilisée pour comprendre l'oxygénation cérébrale, que ce soit chez l'adulte, l'enfant ou encore le nouveau-né, en recherche expérimentale ou en clinique. Entre 1993 et 1998, de nombreux groupes de recherches ont fait émerger de nouvelles possibilités permettant d'accéder aux activations cérébrales de manière totalement non-invasive chez l'adulte (Chance, Zhuang, UnAh, Alter, & Lipton, 1993; Hoshi & Tamura, 1993; Kato, Kamei, Takashima, & Ozaki, 1993; Villringer, Planck, Hock, Schleinkofer, & Dirnagl, 1993) et également chez le nourrisson (Meek et al., 1998).

Depuis plus de 20 ans, de nombreuses études en NIRS se sont intéressées aux zones dédiées au traitement du langage chez l'adulte, avant de s'intéresser au très jeune enfant. Les recherches fondatrices en NIRS concernant la perception du langage chez l'adulte commencent en 1998, suite à l'observation d'activités cérébrales bilatérales dans les régions préfrontales et frontales (Fallgatter & Strik, 1998; Sakatani, Chen, Lichty, Zuo, & Wang, 1999; Sakatani, Xie, Lichty, Li, & Zuo, 1998), enregistrées avec des machines n'utilisant encore que deux canaux. Cette même année, des activations dans l'aire de Broca pendant une tâche langagière ont été trouvées à l'aide de la première machine commercialisée (Hitachi ETG-100), cette fois-ci avec 24 canaux (Watanabe et al., 1998). A partir de ces recherches innovantes, de robustes activations corticales vont ensuite être observées dans des régions classiques du langage aussi bien chez l'adulte, l'enfant ou encore le nouveau-né, avec des machines de plus en plus sophistiquées.

Depuis son introduction dans la recherche neurocognitive développementale il y a 16 ans (Meek et al., 1998), la NIRS a apporté de nombreux éléments à nos

connaissances du cerveau en développement, notamment sur les bases cérébrales sous-tendant l'acquisition du langage (Gervain, Macagno, Cogoi, Peña, & Mehler, 2008; Gervain et al., 2011; Minagawa-Kawai, Cristià, & Dupoux, 2011; Peña et al., 2003; Vanderwert & Nelson, 2014).

3.1.1.2 Principes généraux

De même que l'IRM, l'imagerie optique est une mesure métabolique, donc indirecte, de l'activité neuronale. Lors d'une stimulation, une demande importante d'oxygène provenant des cellules nerveuses est entraînée. Cette dépense énergétique locale, via le couplage neuro-vasculaire, est régulée en apportant une quantité d'oxygène bien plus que nécessaire dans cette région spécifique. C'est ce surplus d'oxygène qui va être détecté et mesuré par la NIRS.

La méthode la plus courante dans la recherche en développement est l'utilisation de machines à onde continue (*continuous wave*, CW). Notre machine repose sur les mêmes principes que ces machines, excepté que c'est une onde pulsée qui a été utilisée au cours de cette recherche. Notre machine utilise une lumière LED (*light emitting diode*), émettant dans le proche infrarouge (*near infrared*, NIR). Cette lumière est émise dans une bande étroite autour de deux longueurs d'ondes spécifiques. Elle est utilisée pour mesurer « in vivo » les changements de concentrations de certains chromophores. Ces chromophores ont des groupes d'atomes qui absorbent la lumière incidente et sont responsables de l'aspect coloré des molécules. Ils sont présents dans les tissus biologiques et sont étroitement reliés à l'activité neuronale, comme l'oxyHb, déoxyHb ou encore le cytochrome c-oxydase (Jöbsis, 1977; Obrig & Villringer, 2003). Ceci est différent de l'IRM, dont le signal BOLD reflète uniquement la déoxyHb.

Quand cette lumière traverse un milieu homogène (non biologique), la dispersion ou la transmission dépendant des propriétés physico-chimiques du milieu (coefficient d'absorption (ϵ_λ), concentration (c), ...) sont négligeables. En effet, la majorité de la lumière est absorbée (A) suivant la loi de Beer-Lambert (avec I = intensité de lumière transmise après avoir traversé le milieu, I_0 = intensité de lumière incidente, λ = longueur d'onde utilisée et l = distance parcourue):

$$A = -\log\left(\frac{I}{I_0}\right) = c \times \epsilon_\lambda \times l \quad (1)$$

Le cerveau n'étant pas un milieu homogène (car biologique), il est indispensable de prendre en compte la complexité des différents tissus que la lumière doit traverser. En effet, les tissus biologiques comme la peau, le crâne et le cerveau étant des milieux où la dispersion de la lumière est importante, cette loi ne peut s'appliquer *stricto-sensu* et doit être modifiée afin de tenir compte de la dispersion (G), ainsi que du trajet non linéaire de la lumière dans ces milieux biologiques (differential path length factor, DPF):

$$-\log\left(\frac{I}{I_0}\right) = (c \times \varepsilon_\lambda \times l \times DPF) + G \quad (2)$$

Les estimations assez précises de la dispersion et du trajet non-linéaire de la lumière à travers ces milieux permettent donc de calculer les valeurs relatives (et non absolues) de concentrations en oxyHb et déoxyHb, et ainsi d'évaluer les changements d'oxygénation du sang (cf. Figure 3.1).

Dans la bande de fréquence proche de l'infrarouge (650nm – 1000nm), l'eau est à son niveau d'absorption le plus bas, ce qui rend les tissus biologiques, comme la peau ou encore les os, relativement transparents à cette lumière. Cette bande de fréquence est donc bien adaptée pour notre mesure. C'est grâce à de nombreuses recherches, encore actuelles (Gervain et al., 2011), que ces deux types de longueurs d'ondes ont été sélectionnés pour avoir le signal avec la meilleure qualité possible.

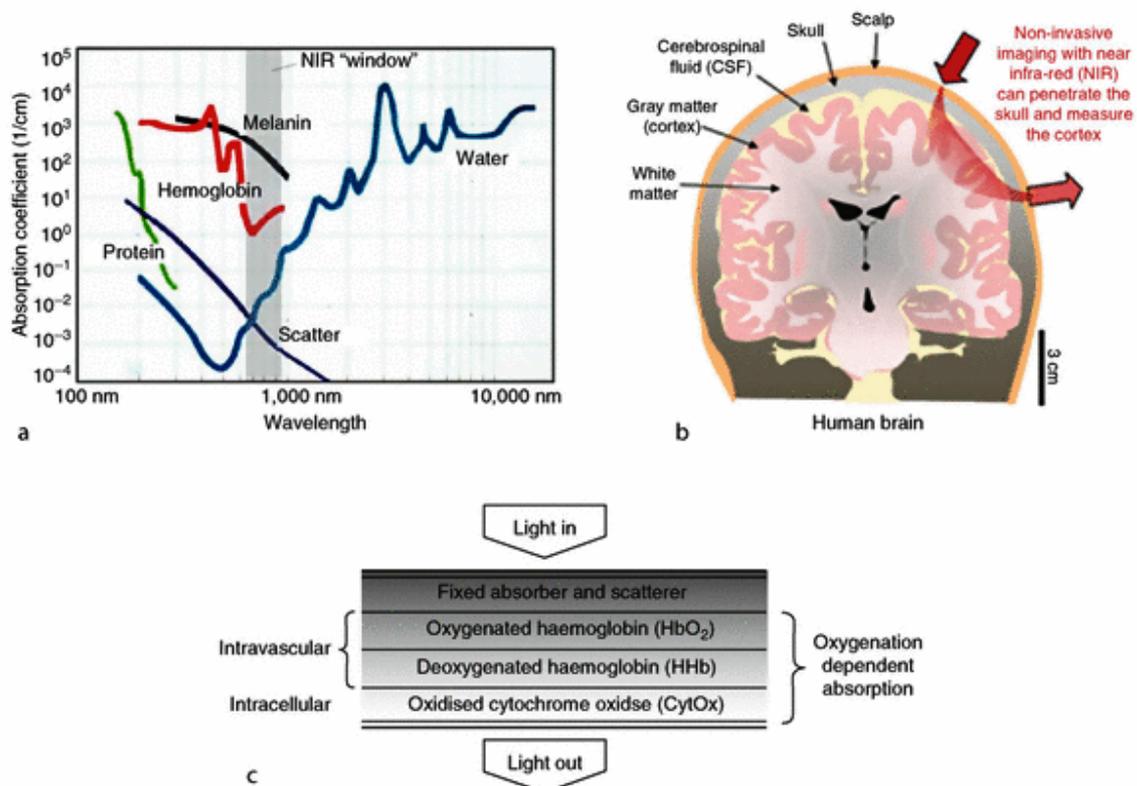


Figure 3.1 : (a) Propriétés optiques des tissus, (b) Trajectoire de la lumière proche infra-rouge traversant le scalp et le crâne pour atteindre le cerveau. (c) Schéma représentant l'absorption des différentes lumières et les composants de dispersion (les concentrations sont représentées par le degré de couleur) à l'intérieur du tissu (Tachtsidis, 2010).

3.1.1.3 Paramètres physiologiques

Les systèmes d'onde continue (CW) à onde laser, ou pulsée avec les LEDs (comme notre machine), permettent de mesurer les changements relatifs d'oxyHb et de déoxyHb comme une mesure indirecte de l'activité cérébrale. L'augmentation d'activité cérébrale demande des ressources métaboliques adaptées à celle-ci, notamment en oxygène et glucose. Ces changements s'accompagnent d'une vasodilatation locale transitoire, qui augmente le débit sanguin donc l'oxygénation des tissus, entraînant un apport d'oxygène en excès. Ce phénomène métabolique est appelé découplage neurovasculaire, permettant de réguler localement les besoins métaboliques, en contrôlant et adaptant la concentration en oxygène via le débit sanguin.

La réponse métabolique associée à ce phénomène est appelée réponse hémodynamique (cf. Figure 3.2). Elle apparaît plusieurs secondes après la stimulation. Les mécanismes exacts sous-tendant la réponse hémodynamique ne sont pas encore

parfaitement compris, et d'autant moins dans le cerveau en développement. Les études s'accordent sur le fait que cette réponse serait plus lente chez les nourrissons (Meek, 2002). De plus, il est possible que les réponses hémodynamiques relativement atypiques que l'on observe chez ces derniers soient liées à l'immaturation de la vascularisation, donc du découplage neuro-vasculaire, ou aux propriétés de l'hémoglobine fœtale, ou encore à l'immaturation du cerveau (myélinisation, etc.). Ces hypothèses doivent être explorées afin de comprendre la signification physiologique de ces réponses atypiques.

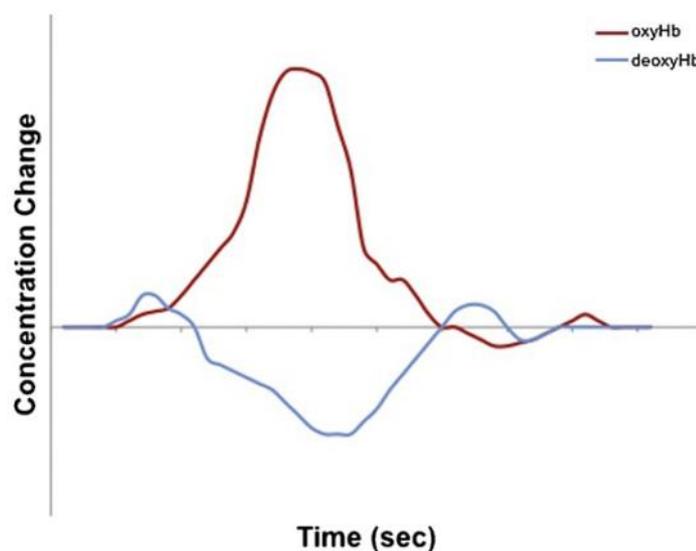


Figure 3.2 : Une réponse hémodynamique typique (HRF) chez l'adulte. La stimulation est délivrée à $t=0$. La réponse commence souvent avec une légère augmentation en déoxyHb, suivie par une augmentation d'oxyHb, puis par une diminution de concentration en déoxyHb (mesurée ici en unité arbitraire). Le pic du signal apparaît quelques secondes après le début du stimulus, puis revient à la normale (baseline) (Gervain et al., 2011, Figure 3).

3.1.1.4 Avantages de la NIRS

En NIRS, aucun champ magnétique ni onde de radiofréquence n'est utilisé, ni aucune injection de traceur ou de substance radioactive dans le sang, ce qui en fait un outil sans risque. La réponse hémodynamique est mesurée à la surface du crâne de façon totalement non-invasive et silencieuse, par la simple émission de lumière traversant différentes couches de tissus biologiques (peau, crâne, liquide céphalo-rachidien, etc.) avant d'atteindre le cortex. La mesure conjointe des changements de concentrations d'oxyHb et de déoxyHb a l'avantage de fournir des données physiologiquement pertinentes sur les corrélats métaboliques de l'activité cérébrale, puisque la somme

d'oxyHb et de déoxyHb correspond au volume sanguin régional cérébral (rCBV). Rappelons que le signal BOLD mesuré par l'IRM reflète exclusivement la déoxyHb (en perdant ses molécules d'oxygène, l'hémoglobine devient paramagnétique et donc sensible au champ magnétique), ce qui ne permet pas l'évaluation du débit sanguin global.

Par rapport à l'EEG, la localisation de la NIRS est bien plus précise. En effet, l'EEG enregistre les courants électriques (mesure de différence de potentiel électrique) et le problème inverse associé à la localisation des sources des potentiels électriques provenant du crâne ne permet pas une localisation précise. Cependant la NIRS reste bien loin de la résolution anatomique de l'IRM. Nous utilisons un bonnet ajustable en fonction de la taille du périmètre crânien, mais nous manquons d'images structurelles pour nous guider dans le placement des optodes de manière standardisée pour tous les bébés.

De plus, lorsque l'on compare la NIRS à d'autres techniques d'imagerie cérébrale, on constate qu'elle présente des avantages spécifiques à l'investigation de populations développementales. Les nouveau-nés ainsi que les nourrissons ont une peau plus fine ainsi qu'un crâne moins épais, la lumière atteint donc des régions plus profondes (10-15 mm chez le nouveau-né pour 3-5mm chez l'adulte, avec une distance de séparation source-détecteur typique de 3cm). Les jeunes enfants ont aussi généralement moins de cheveux ce qui améliore le contact entre la tête et les optodes, réduisant également les artefacts et le taux de signal réfléchi.

Contrairement à l'IRM qui utilise des bobines de gradient, induisant un bruit très important, la NIRS est totalement silencieuse, ce qui permet de présenter des stimuli auditifs dans des conditions de silence, avantage précieux pour l'étude de la perception précoce de la parole. Elle est également moins sensible aux mouvements par rapport à l'IRM ou à l'EEG, il n'est donc pas nécessaire de maintenir la tête et/ou le corps du sujet dans une rigidité totale, autre avantage crucial avec les nourrissons. De plus, le coût d'une machine NIRS est nettement moins élevé que pour l'IRM, proche du coût d'un appareillage EEG.

Enfin, la NIRS permet une alternative intéressante aux méthodes comportementales, restant peu nombreuses chez les nouveau-nés. En effet, le rythme

biologique des enfants à la naissance, comprenant peu de phase d'éveil ainsi qu'un contrôle moteur encore peu développé, limite beaucoup l'expérimentation comportementale. La NIRS nous permet d'enregistrer l'activité des réponses neuronales, mesurant ainsi des capacités qui auraient pu rester indétectables dans le passé avec des méthodes comportementales. Par ailleurs, elle permet de présenter des stimuli auditifs même lors des phases de sommeil, étant donné que les aires auditives des jeunes nourrissons sont actives même lors du sommeil (Gervain, Macagno, et al., 2008; May, Byers-Heinlein, Gervain, & Werker, 2011; Peña et al., 2003). Cette possibilité est un avantage non négligeable au vu des caractéristiques de leur rythme biologique. De plus, cela nous permet d'avoir un temps d'acquisition plus long, et par conséquent d'obtenir un nombre suffisant d'essais par condition pour nous donner les moyens d'avoir assez de puissance statistique lors de l'analyse de nos résultats. En dernier lieu, la NIRS ne fournit pas des réponses uniquement binaires -oui/non- à nos questions expérimentales, mais nous renseigne aussi au niveau de la localisation des activités cérébrales.

Cependant la NIRS présente certaines limites. Elle présente un désavantage au niveau de la précision de la localisation des activités cérébrales, dû à sa faible résolution spatiale et au manque d'imagerie structurelle pour guider le placement des optodes. De plus, elle ne permet l'investigation que des couches superficielles du cerveau et exclut celle des structures plus internes. Ceci reste acceptable pour l'étude du langage et de l'audition, mais ne constitue pas un outil de choix pour d'autres domaines (émotions, mémoire, etc.).

3.1.1.5 La NIRS : un outil essentiel pour l'étude du développement de l'acquisition du langage.

Ces dernières années ont été marquées par les avancées remarquables de la NIRS. Un grand nombre de travaux en NIRS ont porté sur les corrélats neuronaux de la perception précoce du langage. A la naissance, des études montrent que les nourrissons sont capables de discriminer un signal de parole d'un signal non-linguistique. Ces travaux mettent également en évidence une spécialisation hémisphérique dans le traitement langagier (Bortfeld, Fava, & Boas, 2009; Peña et al., 2003). Depuis Broca (1861), de très nombreuses études chez l'adulte ont confirmé l'existence d'asymétries fonctionnelles et ce en fonction de la nature des stimuli auditifs (Hickok & Poeppel, 2007), amenant alors de nombreuses interrogations quant à l'origine de la latéralisation

du langage. De plus, ces capacités précoces pour traiter des stimuli linguistiques ou non-linguistiques semblent reposer sur un réseau complexe, très loin de l’immaturité que l’on a longtemps portée au système perceptif/sensoriel des nourrissons. Comment la NIRS a participé à éclaircir ces compétences précoces et les bases cérébrales qui y sont associées ?

Une des études pionnières sur les bases cérébrales de l’apprentissage du langage en neuro-imagerie fonctionnelle a été conduite en IRMf (Imagerie par Résonance Magnétique Fonctionnelle) avec des nourrissons âgés alors de 3 mois (Dehaene-Lambertz, Dehaene, & Hertz-Pannier, 2002). Pour la première fois, ces chercheurs ont comparé les activations cérébrales, notamment leur localisation, lors de l’écoute de phrases normales ou alors inversées (ces dernières constituant un signal non-linguistique). Leurs résultats révèlent une implication préférentielle de l’hémisphère gauche dans le traitement de stimuli langagiers, notamment dans les aires temporales et le gyrus angulaire ; un pattern d’activation relativement similaire à celui des adultes. Par ailleurs, une seconde étude, réalisée quelques années plus tard par la même équipe sur le même groupe d’âge, montre, en recomposant les séquences temporelles d’activation que le gyrus de Heschl est le premier locus à décharger au sein des régions périsylviennes. L’activité se propage ensuite aux aires plus postérieures et antérieures incluant jusqu’à l’aire de Broca. Ces résultats suggèrent non seulement une sensibilité précoce à des stimuli langagiers, mais également un engagement préférentiel de l’hémisphère gauche, ainsi que des structures complexes organisées hiérarchiquement au sein de la région périsylvienne (Dehaene-Lambertz et al., 2006).

Les récentes recherches en NIRS, comme celles décrites précédemment en IRMf, ont également observé ce type de réponse lors de l’écoute de la parole, soulignant un réseau précoce du traitement de la parole, déjà présent à la naissance. En effet, une étude pionnière en NIRS a montré, pour la première fois, que le cerveau des nouveau-nés répond différemment lorsqu’on leur présente leur langue maternelle normale versus inversée (Peña et al., 2003). De plus, cette étude retrouve également des activités cérébrales plus importantes dans les régions temporales périsylviennes de l’hémisphère gauche par rapport à celle de l’hémisphère droit, lors de la présentation de la parole normale comparée à la parole inversée. Pour la première fois, cette étude révèle que les nouveau-nés ont des patterns de réponses semblables à ceux observés chez l’adulte. Ceci suggère que les nouveau-nés ont déjà des structures cérébrales spécialisées à la naissance pour traiter le langage. Ces résultats étendent les nombreux travaux en IRMf

chez des nourrissons plus âgés (Dehaene-Lambertz et al., 2002). Toutefois, des différences d'activations cérébrales entre la parole normale versus inversée, notamment dans les lobes frontaux, sont retrouvées en IRMf, mais pas en NIRS. Ces différences peuvent être expliquées par le fait que l'âge des nourrissons n'est pas le même : des nouveau-nés ont été testés en NIRS alors que des nourrissons de 3 mois ont été recrutés pour l'étude en IRMf. Cette différence d'âge peut moduler les réponses obtenues, beaucoup d'études ayant révélé des changements cruciaux à ces périodes de développement, à la fois au niveau de la perception de la parole, mais également pour d'autres capacités cognitives (mémoire, attention, etc.). De plus, les nouveau-nés dans l'étude de Peña et collègues (2003) étaient endormis pour la plupart, alors que les nourrissons de Dehaene-Lambertz et al. (2002) étaient éveillés, et c'est chez eux que des réponses importantes sont retrouvées dans les lobes frontaux².

La NIRS n'est pas seulement réservée à l'étude des nouveau-nés, mais est également très utile à l'étude du langage chez des nourrissons plus âgés et pas uniquement en modalité auditive. Bortfeld et al. (2006) ont utilisé des séquences audiovisuelles chez des nourrissons de 6-9 mois. Ces auteurs montrent que les activations dans l'hémisphère gauche sont plus fortes lors de la présentation de la parole uniquement, puisque quand des animations visuelles sont aussi présentes, des réponses sont observées seulement dans les régions occipitales. Une extension de ces travaux a également confirmé que la région temporo-pariétale gauche est bien impliquée spécifiquement dans le traitement de sons langagiers (Bortfeld et al., 2009). Ces recherches soutiennent donc l'idée d'une latéralisation hémisphérique précoce, voire même à la naissance, pour traiter des sons langagiers préférentiellement à gauche.

Cependant, si certaines études retrouvent une latéralisation hémisphérique à la naissance, d'autres travaux ne convergent pas dans le même sens. Une étude en NIRS chez des nouveau-nés, explorant les caractéristiques de cette latéralisation, retrouve des activations bilatérales lors de la présentation de stimuli ayant des caractéristiques proches de celles retrouvées dans la parole, à savoir présentant des modulations rapides,

² Récemment, une étude japonaise a répliqué celle de Peña (2003) avec une machine plus sophistiquée pouvant enregistrer simultanément jusqu'à 72 canaux (Sato et al., 2011), testant également des nouveau-nés. Ils trouvent des profils de résultats similaires à Pena et al, avec une activation plus forte pour le japonais (langue native) dans l'hémisphère gauche par rapport au japonais inversé et l'anglais normal, mais une réponse bilatérale pour l'anglais normal et inversé. Des résultats plus faibles sont retrouvés dans une étude avec des finnois où cette fois-ci les auteurs ont présenté à des nouveau-nés comme stimuli contrôle de la musique (Kotilahti et al., 2010). Ces résultats semblent donc confirmer les travaux de Peña et collaborateurs.

alors qu'une latéralisation à droite émerge lors de la présentation de stimuli ayant des modulations plus lentes (Telkemeyer et al., 2009). De plus, May et collaborateurs (2011), présentant à des nouveau-nés des phrases prononcées dans leur langue maternelle (anglais) ou dans une langue étrangère (espagnol), ne retrouvent pas de latéralisation préférentielle à gauche. D'autres travaux, chez des nourrissons de 2 à 4 mois, rapportent également des réponses non latéralisées lors de la présentation de stimuli langagiers (Cristia et al., 2014; Minagawa-Kawai, van der Lely, et al., 2011; Taga & Asakawa, 2007). A l'inverse, Homae et al. (2006) trouvent des réponses très fortes, cette fois-ci dans l'hémisphère droit, pour des stimuli linguistiques portant un accent de pitch, chez des nourrissons japonais de 3 mois.

La NIRS est donc un outil qui a contribué largement à l'étude de la perception précoce de la parole et à son développement. Toutes les recherches présentées ci-dessus semblent donc suggérer que le type de matériel utilisé (langagier versus non-langagier) influence les réponses préférentielles dans l'un ou l'autre des hémisphères [pour une discussion plus détaillée, voir (Minagawa-Kawai, Cristia, et al., 2011)]. Ces questions mériteront d'être approfondies dans des recherches futures, et la NIRS semble être un outil indispensable pour y parvenir.

3.1.2 Les origines de la loi lambo-trochaïque

-- Cette partie est issue de la traduction libre de
Abboub N., Nazzi T. & Gervain J. (soumis). Rhythmic grouping at birth --

De nombreuses études ont montré que les nouveau-nés, voire même les fœtus, sont sensibles à des propriétés acoustiques spécifiques du langage. Sans reprendre dans les détails ce que nous avons abordé dans l'introduction, ils sont capables, à la naissance, de préférer un stimulus langagier à un stimulus non-linguistique (Vouloumanos & Werker, 2007), la voix de la mère à une voix étrangère (Decasper & Fifer, 1980; Mehler et al., 1978), leur langue maternelle à une langue étrangère de rythme différent (Mehler et al., 1988; Moon et al., 1993), ou de distinguer deux langues rythmiquement différentes qu'ils n'ont jamais entendues auparavant (Nazzi, Bertoncini, et al., 1998). Ils peuvent aussi détecter des indices acoustiques signalant les frontières de mots correspondant aux syntagmes prosodiques (Christophe et al., 1994). Les

capacités à percevoir et discriminer des séquences langagières sur la base de propriétés spécifiques, ici prosodiques - pour la majorité des études présentées - sont un important prérequis pour l'acquisition du langage chez les très jeunes enfants. Effectivement, c'est grâce à ces mécanismes de détection qu'ils vont réussir à organiser le flux continu de parole en unités linguistiques comme les mots ou encore repérer l'ordre des mots (pour une revue de la littérature, voir Gervain & Mehler, 2010).

La loi iambo-trochaïque (ITL) pourrait faire partie de ces biais perceptifs présents à la naissance guidant les nouveau-nés dans la préhension de leur environnement, en l'occurrence ici acoustique. Comme nous l'avons vu précédemment, percevoir des contrastes de durée, d'intensité ou de pitch, pourrait guider les nourrissons dans l'acquisition des mots ou de la structure grammaticale de leur langue maternelle. Selon certains auteurs, la perception de la loi iambo-trochaïque serait liée aux propriétés intrinsèques du système auditif. Notre système perceptif grouperait automatiquement, lors de l'écoute de séquences de sons, les éléments par paire, mais ne les grouperait pas de la même manière en fonction de l'indice acoustique. Si ces sons alternent en durée, alors l'élément le plus long sera perçu comme la fin de l'unité (groupement iambique), alors que s'ils alternent en intensité ou pitch l'élément le plus fort/haut sera perçu comme le début de l'unité (groupement trochaïque). Comme nous l'avons vu dans l'introduction générale, l'ITL a d'abord été proposée pour expliquer le groupement de séquences musicales (Bolton, 1894; Cooper & Meyer, 1960; Woodrow, 1951). Un exemple assez connu reste la sirène des pompiers, perçue par la majorité des personnes comme une séquence de paires de sons avec le premier son plus haut/aigu que le second. Le principe du groupement rythmique a ensuite été étendu aux régularités retrouvées dans les langues au niveau de la production, ainsi que dans la perception de la parole chez l'adulte (Bion et al., 2011; Hay & Diehl, 2007; Hayes, 1995; Nespore et al., 2008). Une des hypothèses relatives à l'ITL est qu'elle serait un mécanisme perceptif universel. Cette hypothèse est soutenue par des études qui ont comparé des locuteurs parlant deux langues rythmiquement différentes (anglais vs français) sans trouver aucune différence entre les groupes (Hay & Diehl, 2007). De plus, de récentes recherches ont mis en évidence chez des adultes et nourrissons italiens, langue dont la prosodie ne contient qu'une proportion faible de contrastes de pitch, ou encore chez des rats, sans aucune expérience langagière, une perception du groupement trochaïque basé sur le pitch (Bion et al., 2011; de la Mora et al., 2013). Ces résultats suggèrent que les préférences de groupement prosodique pourraient exister en absence de toute

expérience langagière et être partagées entre des humains et d'autres mammifères, renforçant le caractère universel de l'ITL.

Cependant, une récente alternative à ces hypothèses a été proposée, selon laquelle le groupement prosodique pourrait être au moins en partie influencé par l'expérience linguistique. De récentes recherches cross-linguistiques ont été réalisées entre des adultes japonais et anglais, montrant que, bien qu'ils aient des performances similaires pour grouper trochaïquement les séquences variant en intensité, ils diffèrent pour grouper iambiquement les séquences de durée (Iversen et al., 2008). Les deux langues diffèrent rythmiquement au niveau du syntagme, le japonais étant trochaïque (^*Tokyo ni*, Tokyo to, 'vers Tokyo', avec la proéminence marquée par un plus haut pitch sur le mot de contenu « Tokyo » sur la position initiale ; Gervain et Werker, 2013), alors que l'anglais est iambique (to *Ro:me*, avec la proéminence prosodique marquée par un allongement de la durée du mot de contenu « Rome » en position finale). Une autre étude cross-linguistique a comparé des adultes allemands et français. Les deux groupes sont sensibles à l'ITL, cependant les allemands montraient un effet plus stable dans la perception du groupement. De plus, l'effet basé sur le pitch n'était trouvé que chez les allemands, mais pas chez les français (Bhatara et al., 2013). Des résultats similaires ont été trouvés pour des stimuli non-linguistiques complexes (Bhatara, sous presse). Les auteurs de ces deux recherches expliquent ces différences par le fait que l'allemand possède au niveau du mot une accentuation lexicale, alors que ce n'est pas le cas du français. De plus, le français est iambique au niveau du syntagme, alors que l'allemand peut avoir les deux patterns rythmiques. Des études chez les nourrissons âgés de 7-8 mois japonais et anglais ont révélé des résultats assez similaires à ceux trouvés chez les adultes, avec notamment une préférence de groupement par la durée retrouvée uniquement chez les anglais (Yoshida et al., 2010). Une récente étude a mis en évidence que des nourrissons bilingues, espagnols dominants ou basques dominants, de 9-10 mois, montraient une perception du groupement stable pour l'intensité dans les deux groupes, mais des différences émergeaient pour la durée (Molnar, Lallier & Carreiras, 2014). Les auteurs reliaient ces résultats aux différences de propriétés prosodiques entre ces deux langues.

En conclusion, sur le peu d'études cross-linguistiques réalisées, il semblerait que les différences soient trouvées essentiellement pour la durée et le pitch, suggérant que l'environnement pourrait influencer très tôt les préférences de groupement et que ces indices acoustiques ne sont pas utilisés de la même manière en fonction de la langue

parlée et/ou entendue. Ces travaux amènent de nombreuses questions, parmi lesquelles : l'ITL est-elle présente à la naissance ? Et, si oui, est-elle modulée par l'environnement linguistique ? Notre étude a donc pour but de comprendre, dans un premier temps, si les nouveau-nés ont la capacité de détecter et de traiter différents types de pattern prosodique, ici de groupement. Deuxièmement, nous prendrons en compte l'environnement linguistique prénatal afin d'explorer l'impact potentiel très précoce de la langue maternelle.

Un point important de notre recherche concerne les bases cérébrales du groupement prosodique qui est un domaine encore totalement inexploré. Chez l'adulte, la perception et la compréhension du langage, incluant le traitement de la morphosyntaxe ainsi que de la sémantique, sont préférentiellement traitées dans l'hémisphère gauche, alors que la prosodie linguistique recrute un réseau plus dynamique dans l'hémisphère droit (Friederici, 2012; Hickok & Poeppel, 2007). Or, la latéralisation du traitement prosodique semble aussi dépendre de la valeur linguistique des stimuli utilisés et du contexte expérimental. C'est ainsi qu'une dominance hémisphérique gauche a été retrouvée si les indices prosodiques portent une valeur lexicale ou morphosyntaxique, comme les tons lexicaux pour les locuteurs de langues tonales (Gandour et al., 2004; Kreitewolf, Friederici, & von Kriegstein, 2014; Sato, Sogabe, & Mazuka, 2007, 2010a). Chez les très jeunes enfants, peu d'études ont examiné les bases neurales du traitement de la prosodie et aucune n'a spécifiquement étudié la perception du groupement prosodique. Les seules études en imagerie optique montrent, chez des nouveau-nés endormis et chez des nourrissons de 3 mois (Homae, Watanabe, Nakano, & Taga, 2007; Sato et al., 2010a; Telkemeyer et al., 2009), tout comme chez les enfants de 4 ans (Wartenburger et al., 2007), une dominance hémisphérique droite pour le traitement de la prosodie en général. Néanmoins, ces études ont examiné la perception de la prosodie en présentant des phrases entières et ont donc utilisé des stimuli dans toute leur complexité acoustique. De nombreuses incertitudes restent quant à la manière dont les jeunes enfants vont traiter les indices prosodiques en isolation (variation seulement en durée, seulement en intensité ou seulement en pitch) et quelles seront les bases cérébrales associées.

Pour répondre à nos questions expérimentales, nous avons fait écouter à trois groupes de nouveau-nés différents types de groupements variant selon les trois principaux indices prosodiques (durée, intensité, pitch). Ces trois premiers groupes (Expériences 1-3) sont constitués de nouveau-nés ayant entendus « in utero »

exclusivement du français, une langue qui utilise majoritairement dans sa prosodie des variations de durée, avec un allongement sur la dernière syllabe au niveau du syntagme (Dell, Hirst, & Vergnaud, 1984; Nespor et al., 2008). Dans un deuxième temps, nous avons évalué comment ces capacités pouvaient être modulées par l'environnement linguistique. Le dernier groupe (Expérience 4) est ainsi constitué de nouveau-nés ayant été exposés non seulement au français, mais également à une autre langue (voir détails dans la partie Méthode). Toutes ces autres langues possèdent des contrastes de pitch (plus important qu'en français) dans leur prosodie linguistique à différents niveaux (accentuation lexicale, tons lexicaux, réalisation de la proéminence des syntagmes, voir détails dans la discussion). Ces nouveau-nés ont donc eu pendant la grossesse, ainsi que dans les quelques heures de vie après l'accouchement, une exposition à des stimuli linguistiques qualitativement différents que ceux des trois groupes précédents.

Étant donné la nature du signal linguistique perçu, nous avons donc décidé d'utiliser la même procédure pour l'Expérience 4 que celle de l'Expérience 3, utilisant des contrastes de pitch. Le choix de cet indice repose sur deux raisons principales. Récemment, des études ont montré que le pitch pourrait être le plus fort/saillant des trois indices de groupement (Bion et al., 2011; de la Mora et al., 2013; Nespor et al., 2008). En effet, ces études ont montré que des nourrissons de 7 mois, italiens, langue dont le pitch n'est pas un des indices les plus utilisés dans la prosodie linguistique, utilisent le pitch mais pas la durée dans une tâche testant l'ITL (on notera toutefois que les Expériences 1 et 2 de cette thèse ne généralisent toutefois pas ce résultat au cas du français et de l'allemand). De même, les rats groupent des séquences linguistiques variant en pitch comme des unités trochaïques, mais ne semblent pas utiliser la durée pour le groupement iambique. Ces résultats suggèrent que le biais trochaïque basé sur le pitch pourrait être indépendant de la langue en acquisition et automatique. A l'inverse, l'utilisation de la durée serait probablement régie par l'expérience linguistique.

Dans notre étude, nous avons mesuré de quelle manière les nouveau-nés traitent et distinguent différents patterns prosodiques. La NIRS nous a permis d'enregistrer les réponses hémodynamiques de certaines aires corticales du cerveau pendant l'écoute de nos stimuli. Dans quatre expériences, nous avons ainsi présenté des séquences formées de paires de tons purs variant soit en durée (Expérience 1), soit en intensité (Expérience 2), soit en pitch (Expériences 3 & 4). Chaque indice prosodique a été présenté selon 3 conditions de groupement : en accord avec l'ITL (court-long, fort-faible, aigu-grave), en désaccord avec l'ITL (long-court, faible-fort, grave-aigu) et une condition sans

contraste où les éléments à l'intérieur des paires étaient identiques (durée, pitch et intensité égales).

Si les nouveau-nés utilisent l'alternance d'un indice prosodique pour grouper les sons, des différences dans l'amplitude des réponses et/ou dans leur localisation vont être trouvées pour les séquences de groupement en accord avec l'ITL et en désaccord avec l'ITL. Aucune étude n'ayant testé les bases cérébrales du groupement prosodique dans une population développementale, nous n'avons de ce fait pas de prédictions claires sur la localisation des réponses cérébrales. En effet, s'il a été trouvé que le traitement de la prosodie lors de l'écoute de phrases s'effectuerait préférentiellement dans l'hémisphère droit très tôt au cours du développement (Homae et al., 2007; Sato et al., 2010a; Telkemeyer et al., 2009), rien n'est aussi sûr pour des indices en isolation. En effet, les études précédentes ont toutes comparé des phrases dans leur entière complexité acoustique à des phrases où toutes les modulations prosodiques étaient absentes. La différence cruciale dans notre méthode est la conformité du groupement entendu à l'ITL, et porte donc sur son organisation séquentielle. Ces informations de structure et d'ordre sont connues pour être traitées préférentiellement par l'hémisphère gauche (Dehaene-Lambertz et al., 2002; Gervain, Berent, & Werker, 2012; Gervain, Macagno, et al., 2008). Il est donc envisageable que nous retrouvions soit une implication des deux hémisphères, soit un avantage pour l'hémisphère gauche.

Sachant l'importance des contrastes de durée en français, l'Expérience 1 présente des variations de durée à des nouveau-nés exposés à un environnement linguistique exclusivement français.

3.1.3 Expérience 3 : Contraste de durée chez des nouveau-nés monolingues français

3.1.3.1 Méthode

3.1.3.1.1 Participants

Dix-huit nouveau-nés, nés à terme (13 filles, 5 garçons, moyenne d'âge: 2,05 jours; étendue 1-4 jours) dans des familles parlant exclusivement français ont participé à cette troisième expérience. Sept nouveau-nés supplémentaires ont été testés mais exclus des analyses pour cause de mouvement (3), cheveux noirs très épais (2) ou parce qu'ils ont pleuré avant la fin de l'expérience (2).

Les nouveau-nés ont été recrutés à l'Hôpital Robert Debré à Paris, France. Pour être retenu dans notre étude ils doivent : être nés entre 38 et 42 semaines d'aménorrhées, avoir un poids de plus de 2600g à la naissance, avoir un score d'Apgar supérieur ou égal à 8 à la cinquième ainsi qu'à la dixième minute après la naissance, avoir un diamètre crânien compris entre 33,5 et 36 cm et être en bonne santé. Le déroulement de la grossesse ne devait également présenter aucune anomalie tout comme l'accouchement. Le conseil d'évaluation éthique pour les recherches en santé (CERES, Université Paris Descartes) a donné sa permission et a approuvé le dessin expérimental de l'expérience. Tous les parents ont signé un consentement de participation éclairé avant de commencer l'expérience. Des explications détaillées sur le déroulement de la passation sont données par l'expérimentateur avant le début de chaque expérience ; au moins un des deux parents est présent dans la pièce lors de la passation.

3.1.3.1.2 Matériel et design expérimental

Nous avons utilisé 8 paires différentes de tons purs (bip). Les deux tons d'une paire étaient contrastés uniquement en durée. L'un était court (250ms) et l'autre était long (500ms), mais ils avaient la même intensité (70dB) et le même pitch (un exemplaire des huit notes de l'échelle de la musique occidentale): une paire était constituée soit de deux notes A (220Hz), soit de deux notes B (247Hz), etc. Ces paires de tons étaient présentées selon 3 conditions (voir Figure 3.3). Dans la condition en accord avec l'ITL, les paires de tons suivaient un groupement prédit par l'ITL et qui se trouve également en français (court-long). Dans la condition en désaccord avec l'ITL, le groupement suivait le type non-prédit par l'ITL et n'existant pas en français (long-court). Dans la condition sans contraste, il n'y avait aucune variation prosodique, tous les indices (durée, intensité et pitch) ayant des valeurs égales.

Chaque condition était présentée en 8 blocs pour un total de 24 blocs (3×8), chaque bloc comprenant 6 paires de tons. Les blocs étaient séparés par des intervalles de silence variables en durée (26-32 s), afin de permettre de revenir à un état de baseline de la réponse hémodynamique entre chaque stimulation. A l'intérieur de chaque bloc, les paires de tons étaient séparées par un silence (480ms), conduisant à une durée entre 8 et 12s par bloc. Ces 24 blocs étaient présentés dans un ordre aléatoire (voir la Figure 3.3 pour le dessin expérimental), évitant qu'un même type de bloc ne se suive. L'ordre de présentation était randomisé entre chaque sujet.

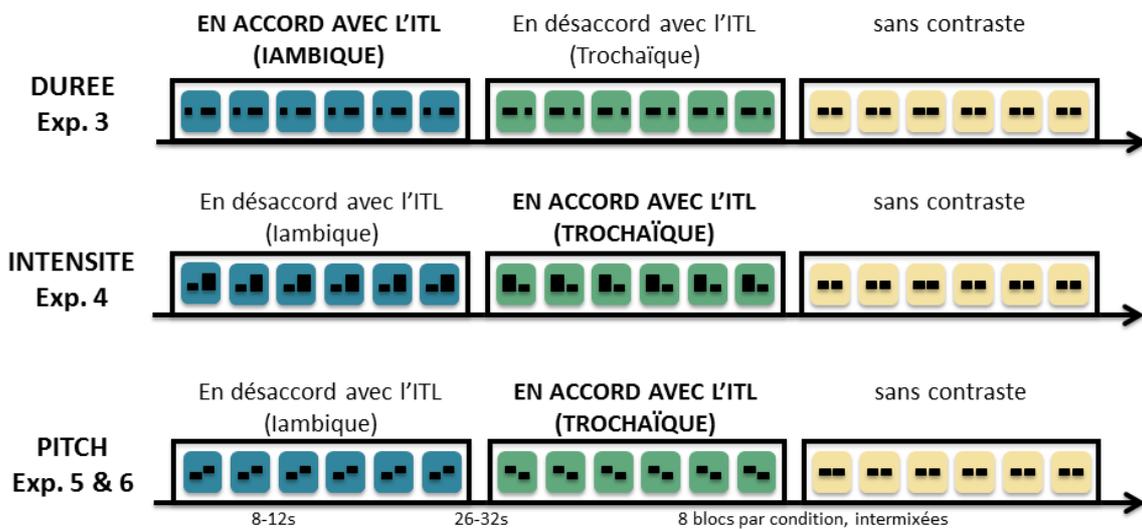


Figure 3.3 : Design expérimental des Expériences 3-6. Chaque rectangle représente un bloc, contenant une série de 6 paires d'un type de groupement, ces paires étant séparées par de courtes pauses.

3.1.3.1.3 Procédure

Les nouveau-nés ont été testés à l'aide d'une machine NIRx NIRScout 816 (USA), avec une distance source-détecteur de 3cm, 2 longueurs d'ondes pulsées (760 nm et 850 nm), un taux d'échantillonnage de 10,417 Hz. La machine utilise une lumière pulsée, émise par des LEDs. Le temps de la passation était d'environ 16-17 min. Les capteurs optiques (optodes) sont insérés bilatéralement sur un bonnet EEG composé d'un tissu semi-stretch, avec pour repère anatomique de surface : inion, nasion, vertex et les points bilatéraux péri-oculaires (Figure 3.4). Évidemment, ce bonnet possède des capteurs fixes et nous ne pouvions éviter la variabilité interindividuelle en terme de localisation anatomique due à la forme de la tête ou encore à sa taille, rendant la localisation au niveau du groupe approximative. Cependant, le placement des capteurs a été décidé afin de recouvrir les régions temporo-pariétales avec la plus grande probabilité possible, afin de contrôler au maximum ces variabilités. De plus, le bonnet est équipé de 8 sources et de 10 détecteurs, nous donnant la possibilité d'avoir 12 canaux par hémisphère, donc 24 au total.



Figure 3.4 : Photo d'un nouveau-né portant le bonnet avec les optodes placées préférentiellement sur les aires temporo-pariétales. La disposition des optodes insérées dans le bonnet permet d'enregistrer simultanément l'activité des 24 canaux.

La passation se déroulait dans une pièce du service de la maternité de l'hôpital Robert Debré à Paris. Nous avons expliqué aux parents le but de ces recherches ainsi que le déroulement de la passation. Au moins un des deux parents était présent lors de la passation. Après avoir recueilli les signatures du consentement de participation, nous installions le berceau en face des haut-parleurs à une distance de 1m, faisant un angle de 30° (Figure 3.5). Nous posions ensuite le bonnet toujours de la même manière pour tous les nouveau-nés testés. La majorité des nouveau-nés dormaient lors de la pose du bonnet ou étaient dans un état d'éveil calme afin d'éviter d'avoir un nombre trop important de mouvements. Pour préserver leur état de sommeil ou d'éveil calme et pour éviter toute interférence avec la lumière des LEDs, l'expérimentation se déroulait dans la pénombre (lumière et écrans d'ordinateurs éteints). Un Dell Optiplex PC 380 jouait les stimuli sous le contrôle d'E-Prime et un ordinateur portable HP enregistrait le signal NIRS.



Figure 3.5 : Photod'un nouveau-né endormi dans son berceau et du dispositif juste avant le début de l'expérimentation, portant le bonnet avec les capteurs. Ces optodes sont reliés par des fibres optiques à la machine NIRSCOUT NIRx.

3.1.3.1.4 Traitement de données et analyses

Les analyses ont été faites sur les concentrations d'oxyHb et de déoxyHb. Ces concentrations ont été calculées en appliquant la loi Beer Lambert modifiée, grâce aux valeurs d'intensité de lumière récupérées et enregistrées par la machine NIRS. Afin d'éliminer les bruits de hautes fréquences dus aux paramètres physiologiques comme le battement cardiaque ou encore les fluctuations du débit sanguin systémique, les données ont été filtrées entre 0,01 et 0,7 Hz. Les artéfacts de mouvements, identifiés et rejetés automatiquement, ont été définis de la manière suivante : les changements de concentration de $0,1 \text{ mmol} \times \text{mm}$ pendant 0,2 s ont été retirés en rejetant le bloc du canal où l'artéfact était présent. Pour les blocs non-rejetés, une baseline est linéairement ajustée entre la moyenne des 5 secondes qui précèdent le début de la stimulation et celle des 5 secondes après les 15 secondes qui suivent la fin de la période de repos (à partir de la fin du bloc). Pour chaque bloc, les données sont moyennées sur une fenêtre de 27s

en partant du début du bloc (12 s de stimulation + 15 s de retour à la baseline). Les nouveau-nés qui ont été inclus dans les analyses ont eu moins de 30% de données rejetées.

Afin d'identifier les régions montrant des activations significatives, nous avons considéré les données individuelles comme un effet aléatoire et avons réalisé des test de Student (test-t), canal par canal, entre nos trois conditions (en accord avec l'ITL, en désaccord avec l'ITL, sans contraste) versus la baseline (0), ainsi qu'entre les deux conditions de groupement ITL (en accord vs désaccord). Toutes les analyses statistiques portent sur l'oxyHb et la déoxyHb, mais nous ne rapporterons dans cette partie que les résultats sur l'oxyHb. Afin de corriger les comparaisons multiples, nous avons utilisé la correction False Discovery Rate (FDR) avec $q < .05$, (Benjamini & Hochberg, 1995; Genovese, Lazar, & Nichols, 2002; Singh & Dan, 2006).

3.1.3.2 Résultats et discussion

Les réponses ont été moyennées par canal et par condition et sont présentées dans la Tableau 3-1 et la Figure 3.6 pour l'oxyHb et la déoxyHb. Pour les comparaisons contre la baseline, les test-t (illustrés dans la Figure 3.10) pour l'oxyHb ont révélé des activations significatives ou marginales pour le groupement en désaccord avec l'ITL dans plusieurs canaux [CH 17 ($p < .01$); CH 1, 3, 4, 5, 9, 10, 24 ($p < .05$); CH 6, 16, 19 ($p < .1$), non corrigé], et un effet marginal pour la condition sans contraste dans le canal 11 ($p = .052$). Par ailleurs, les canaux présentant une activation significative étaient majoritairement dans l'hémisphère gauche.

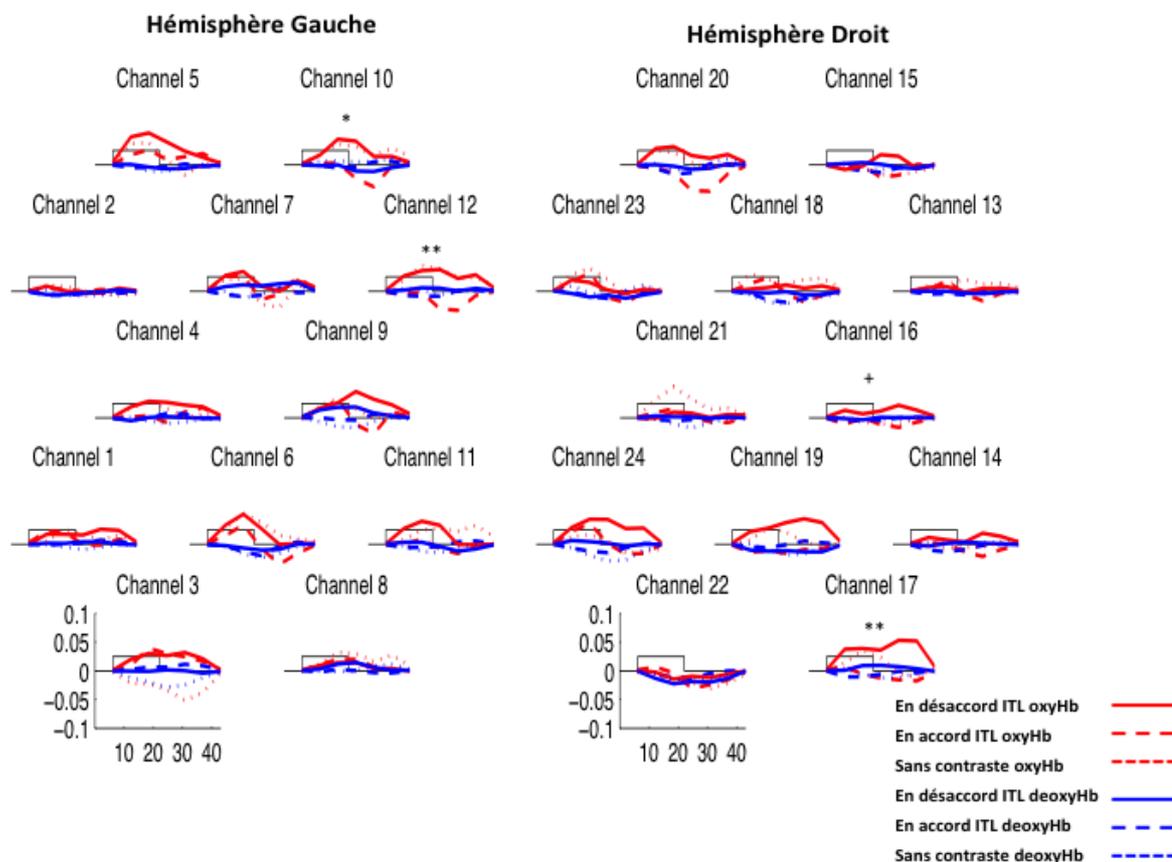


Figure 3.6 : Résultats des grandes moyennes de l'Expérience 3 (Condition durée, monolingues). Les canaux sont placés selon le même placement que sur la Figure 3.10. L'axe des x représente le temps en secondes ; l'axe des y représente les concentrations en $\text{mmol} \times \text{mm}$. Le rectangle sur l'axe des x représente le temps de stimulation. Les lignes continues rouge et bleu dans le graphique représentent les changements de concentration de l'oxyHb et de deoxyHb pour les séquences en désaccord avec l'ITL (trochaïque). Les lignes rouge et bleu avec des tirets représentent les concentrations de l'oxyHb et de deoxyHb pour les séquences avec le groupement en accord avec l'ITL (iambique). Les lignes rouge et bleu en pointillés représentent les concentrations de l'oxyHb et de deoxyHb pour le groupement sans contraste. *** : $p < .001$; ** : $p < .01$; * : $p < .05$; + : $p < .01$, sans correction.

En comparant les deux conditions principales (en accord versus en désaccord avec l'ITL), des différences significatives ou marginales apparaissaient dans plusieurs canaux [CH 12, 17 ($p < .001$); 10 ($p < .05$); 16 ($p < .1$)], encore une fois avec une implication plus importante de l'hémisphère gauche.

Ces résultats montrent que le cerveau des nouveau-nés distingue clairement les deux types de groupement basés sur des contrastes de durée. Par ailleurs, nous trouvons de plus fortes activations pour le groupement en désaccord avec l'ITL (trochaïque) et avec la structure prosodique du français (allongement à la fin des syntagmes). Nous interprétons cette augmentation d'activité comme un signe d'un plus grand effort pour

le traitement du groupement non familial et en désaccord avec l'ITL (Ding, Fu, & Lee, 2014). Ces résultats établissent clairement une préférence de groupement prosodique à la naissance. Cependant, rappelons que la durée est un indice acoustique qui joue un rôle majeur dans la proéminence prosodique du français. Par conséquent, ce résultat seul ne nous permet pas de distinguer entre un biais général auditif ou un effet de l'expérience prénatale avec la langue maternelle. Dans ce cadre, l'Expérience 4 va tester si les nouveau-nés monolingues français peuvent discriminer des séquences de groupement en accord ou en désaccord avec l'ITL, en utilisant cette fois-ci des contrastes d'intensité, jouant un rôle moins important dans la prosodie du français.

Tableau 3-1 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FDR}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0].

Mesure	OxyHb				DeoxyHb	
Conditions	A vs DA		DA vs O		SC vs O	
CH	p_{unc}	p_{FDR}	p_{unc}	p_{FDR}	P_{unc}	P_{FDR}
3			0,012	0,097	0,013	0,308
4			0,04	0,118		
5			0,044	0,118		
9			0,034	0,118		
10	0,011	0,089	0,023	0,118		
12	0,004	0,089	0,037	0,118		
16	0,05	0,302				
17	0,01	0,089	0,003	0,073		
24			0,03	0,118		

3.1.4 Expérience 4 : Contraste d'intensité chez des nouveau-nés monolingues français

3.1.4.1 Méthode

3.1.4.1.1 Participants

Dix-huit nouveau-nés, nés à terme (11 filles, 7 garçons, moyenne d'âge: 2,28 jours; étendue 1-4 jours) dans des familles parlant exclusivement français, ont participé à cette quatrième expérience. Les autres critères de sélection des enfants étaient les mêmes que pour l'Expérience 3. Huit nouveau-nés supplémentaires ont été testés mais exclus des analyses pour cause de mouvement ou réveil.

3.1.4.1.2 Matériel

Nous avons utilisé huit types de tons purs (notes A, B, C, D, E, F, G, AA). Ces tons purs étaient groupés par paires. A l'intérieur de chaque paire, les tons pouvaient contraster en intensité (70 vs 76dB), mais étaient identiques en durée (soit deux notes de 250ms ou de 500ms) et en pitch (deux mêmes notes, ex : A-A, B-B, C-C, D-D, E-E, F-F, G-G, AA-AA). Les séquences de paires de tons étaient présentées selon 3 conditions de groupement (voir Figure 3.3). Dans la condition en accord avec l'ITL, les paires de tons suivaient un groupement prédit par l'ITL (fort-faible). Dans la condition en désaccord avec l'ITL, le groupement suivait le type non prédit par l'ITL (faible-fort). Dans la condition sans contraste, il n'y avait aucune variation prosodique, tous les indices (durée, intensité et pitch) ayant des valeurs égales.

3.1.4.1.3 Design expérimental, procédure, traitement des données et analyses.

Le design expérimental, la procédure, le traitement des données et les analyses sont identiques à l'Expérience 3.

3.1.4.2 Résultats et discussion

La Figure 3.7 représente la grande moyenne des changements de concentration pour l'oxyHb et la deoxyHb pour tous les blocs et pour chaque pattern prosodique. Pour les comparaisons contre la baseline, des t-tests réalisés canal par canal pour l'OxyHb (illustrés dans la Figure 3.7, Tableau 3-2) montrent des activations significatives ou marginales dans plusieurs canaux : pour les séquences en accord avec l'ITL [CH 9 ($p < .01$) et CH 7, 16 ($p < .05$)], pour les séquences en désaccord avec l'ITL [CH 3, 4 ($p < .05$)] et pour les séquences sans contraste [CH 5 ($p < .05$), et CH 14, 21 ($p < .1$)]. Des comparaisons directes entre les conditions principales (en accord versus en désaccord avec l'ITL) ont été également réalisées et révèlent des différences significatives ou marginales dans les canaux 3 ($p < .05$) et 22 ($p < .1$), avec plus d'activation pour la condition en désaccord avec l'ITL.

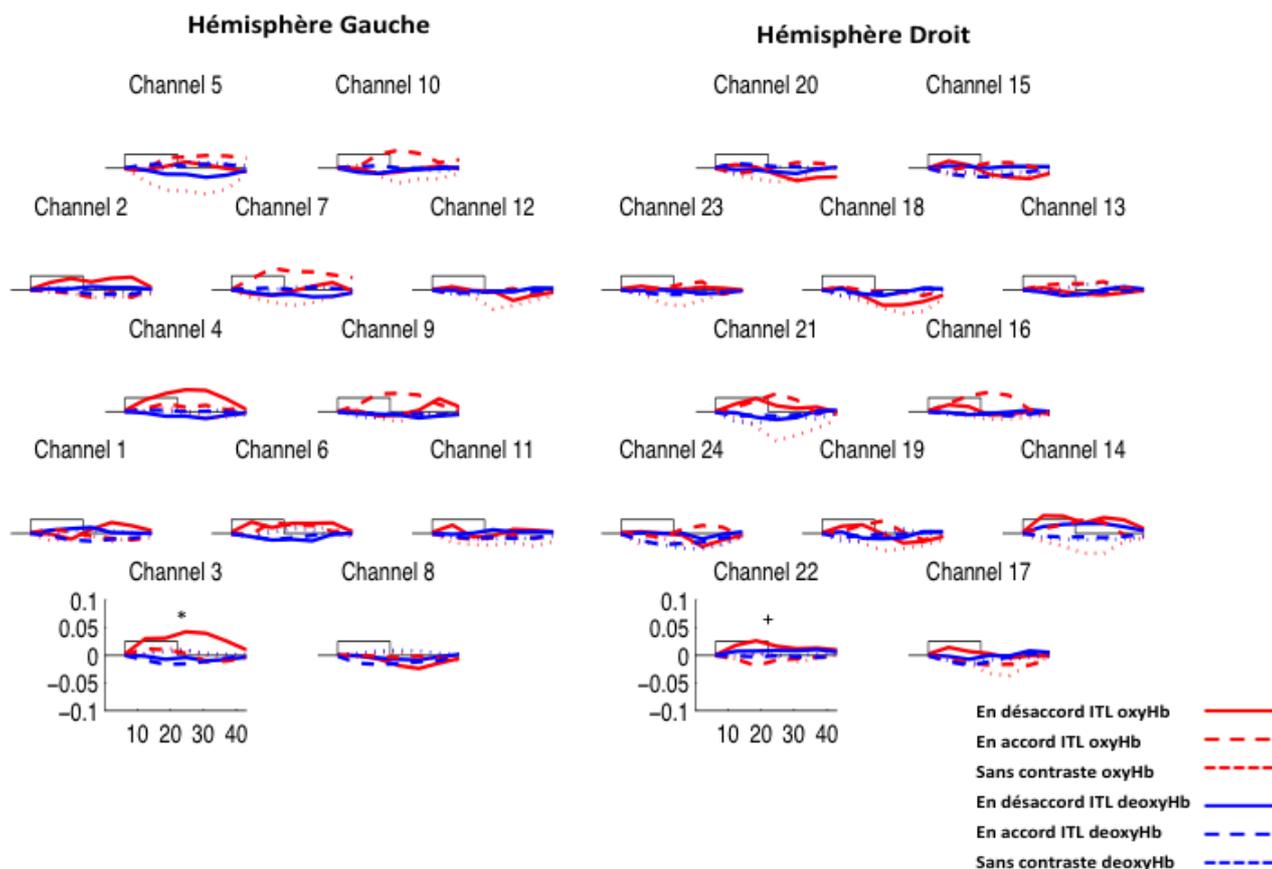


Figure 3.7 : Résultats des grandes moyennes de l'Expérience 4 (condition intensité, monolingues). Toutes les conventions du graphiques sont identiques à la Figure 3.6.

Les résultats de la perception du groupement basé sur l'intensité révèlent des activations globalement plus faibles que dans l'Expérience 3 (durée). Ce résultat peut être dû au fait que l'intensité joue un rôle moins important que la durée dans la prosodie du français. Mais il est à noter que lorsque des activations différentes sont trouvées entre les deux conditions principales, le groupement en désaccord avec l'ITL a une activation plus importante, ce qui confirme notre hypothèse que ce pattern de résultats pourrait être attribué à un effort plus important pour traiter ces séquences. Ceci suggère une certaine sensibilité à la perception de l'ITL, même pour une dimension acoustique comme l'intensité, un indice peu important dans la prosodie du français.

De récents travaux ont étendu l'étude de ces sensibilités du groupement rythmique à un autre indice acoustique : le pitch (Bhatara et al., sous presse, 2013 ; Bion et al., 2011 ; de la Mora et al., 2013). L'Expérience 5 a donc exploré comment le cerveau des nouveau-nés traite le groupement en accord versus en désaccord avec l'ITL

quand seul le pitch variait et qui, comme l'intensité, est moins important que la durée dans la prosodie du français.

Tableau 3-2 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FDR}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste].

Mesures	OxyHb								DeoxyHb							
	A vs DA		DA vs O		SC vs O		A vs O		DA vs A		DA vs O		A vs O		SC vs O	
Conditions	punc	pFDR	punc	pFDR	punc	pFDR	punc	pFDR	Punc	PFDR	Punc	PFDR	Punc	PFDR	Punc	PFDR
1									0,016	0,381						
3	0,024	0,588	0,023	0,316												
4			0,026	0,316												
5					0,011	0,261										
6										0,043	0,628					
7							0,032	0,254								
8													0,018	0,441		
9							0,006	0,138								
16							0,02	0,238								
24															0,027	0,642

3.1.5 Expérience 5 : Contraste de pitch chez des nouveau-nés monolingues français

3.1.5.1 Méthodes

3.1.5.1.1 Participants

Dix-huit nouveau-nés, nés à terme (12 filles, 6 garçons, moyenne d'âge: 2,86 jours, étendue [2-5 jours]) dans des familles parlant exclusivement français, ont participé à cette cinquième expérience. Les autres critères de sélection des enfants étaient les mêmes que pour l'Expérience 3. Dix-huit nouveau-nés supplémentaires ont été testés mais exclus des analyses pour cause de mouvement ou de réveil (9), cheveux noirs très épais (4), parce qu'ils ont pleuré avant la fin de l'expérience (2) ou à cause d'un problème technique (3).

3.1.5.1.2 Matériel

Nous avons utilisé différentes paires de tons purs (bip). Les deux tons d'une paire étaient contrastés uniquement en pitch (d'une octave), mais étaient identiques en durée (soit deux notes de 250ms ou de 500ms) et en intensité (70dB).

Les séquences de ces paires de tons étaient présentées selon 3 conditions (voir Figure 3.3). Dans la condition en accord avec l'ITL, les paires de tons suivaient le groupement (aigu-grave) prédit par l'ITL. Dans la condition en désaccord avec l'ITL, les paires de tons suivaient (grave- aigu) le type non prédit par l'ITL. Dans la condition sans contraste, il n'y avait aucune variation prosodique, tous les indices (durée, intensité et pitch) ayant des valeurs égales.

3.1.5.1.3 *Design expérimental, procédure, traitement des données et analyses*

Le design expérimental, la procédure, le traitement des données et les analyses étaient identiques à l'Expérience 3.

3.1.5.2 *Résultats et discussion*

Les résultats ont été moyennés par canal et par condition et sont présentés dans la Figure 3.8. Des tests de Student ont été réalisés canal par canal (illustré dans la Tableau 3-3 et la Figure 3.10). Pour la comparaison avec la baseline, nous avons trouvé des activations significatives ou marginales dans plusieurs canaux, pour les séquences en désaccord avec l'ITL [CH 22 ($p < .01$); CH 1, 2, 8, ($p < .05$) et CH 14 ($p < .1$)], pour les séquences en accord avec l'ITL [CH 8, 13, 19 ($p < .05$) et pour les séquences sans contraste [CH 4 ($p < .05$), CH 8,19 ($p < .1$)]. Lors de la comparaison directe entre les deux conditions principales, un effet marginal a été retrouvé en faveur des séquences en désaccord avec l'ITL dans le canal 22 ($p < .1$).

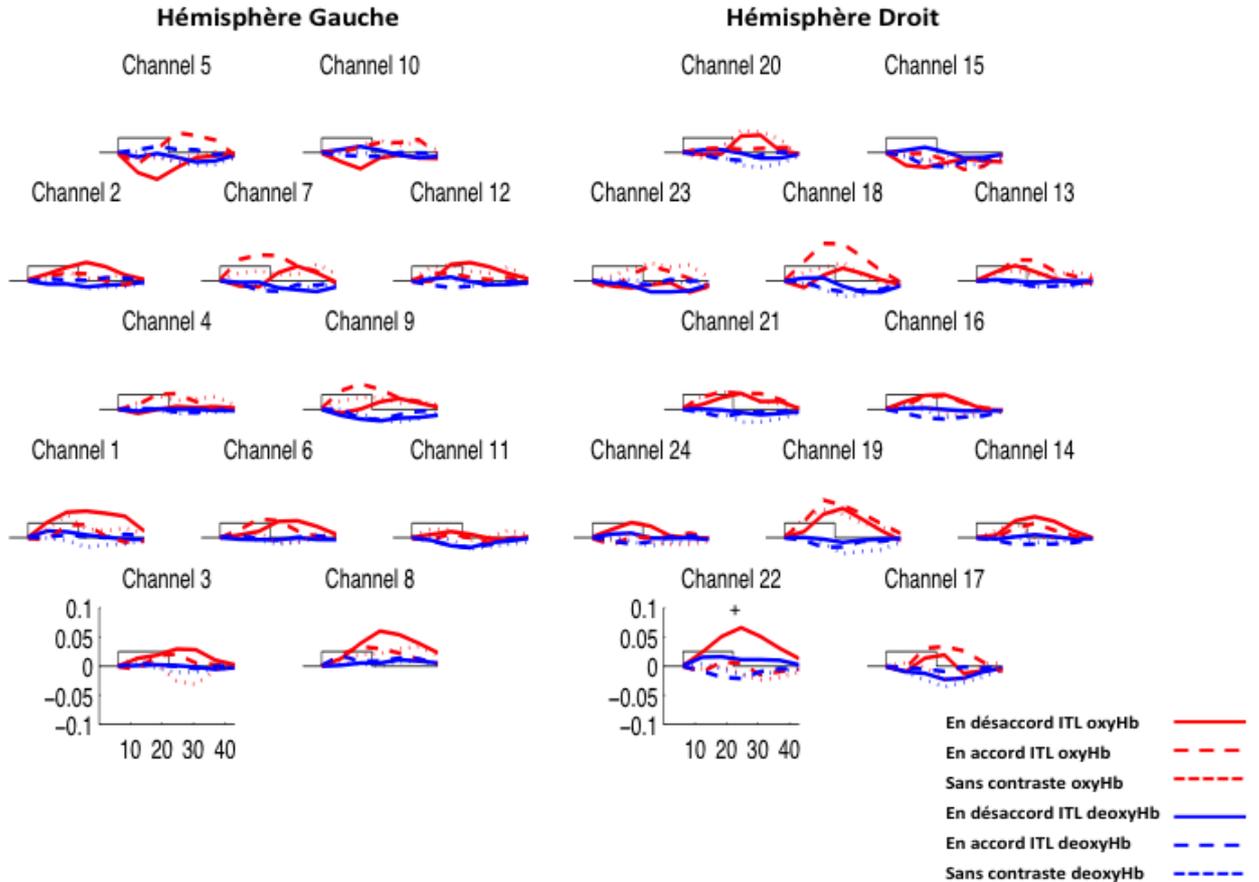


Figure 3.8 : Résultats des grandes moyennes de l'Expérience 5 (condition pitch, monolingues). Toutes les conventions du graphique sont identiques à la Figure 3.6.

Les résultats de cette expérience montrent globalement des effets plus faibles pour la perception du groupement rythmique basé sur le pitch que pour celui de la durée. Les profils d'activation sont très semblables à ceux retrouvés pour l'intensité dans l'Expérience 4, des réponses plus importantes étant présentes pour les séquences en désaccord avec l'ITL, mais atteignant un effet marginal dans seulement un canal. Comme pour l'intensité, ces faibles effets peuvent être expliqués par le rôle plus faible du pitch par rapport à la durée dans la prosodie du français. D'ailleurs, des études chez l'adulte montrent que les adultes français ont des difficultés à utiliser le pitch pour grouper les sons en accord avec l'ITL (Bhatara et al., sous presse, 2013).

Tableau 3-3 : Résultats des tests de Student pour l'oxyHb et deoxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FDR}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste]

Mesures	oxyHb						deoxyHb			
Conditions	A vs O		DA vs O		SC vs O		SC vs O		C vs O	
CH	p _{unc}	p _{FDR}								
1			0,02	0,201						
2			0,045	0,268						
4					0,037	0,601				
8	0,019	0,383	0,025	0,201						
9							0,023	0,178		
13	0,048	0,383								
16									0,036	0,587
17							0,038	0,178		
18							0,045	0,178		
19	0,035	0,383					0,006	0,143		
20							0,043	0,178		
21							0,042	0,178		
22			0,005	0,116						

3.1.6 Expérience 6 : Contraste de pitch chez les nouveau-nés bilingues français-autre langue

3.1.6.1 Méthode

3.1.6.1.1 Participants

Dix-huit nouveau-nés, nés à terme (12 filles, 6 garçons, moyenne d'âge: 2,17 jours, étendue [1-4 jours]) ayant une mère parlant français et une autre langue [arabe (8), mandarin (2), cantonais (1), amérindien (1), bulgare (1), kabyle (1), koyaka (1), anglais (1), polonais (1), turc (1)], ont participé à cette sixième expérience. Les mères ont déclaré avoir parlé entre 40 et 60% du temps français, (moyenne = 47%) et le reste du temps l'autre langue pendant la grossesse. Les autres critères de sélection des enfants étaient les mêmes que pour l'Expérience 3. Vingt-quatre nouveau-nés supplémentaires ont été testés mais exclus des analyses pour cause de mouvement ou de réveil (13), cheveux noirs très épais (8) ou à cause d'un problème technique (3).

3.1.6.1.2 Matériel

Le matériel était identique à celui utilisé dans l'Expérience 5.

3.1.6.1.3 Design expérimental, procédure, traitement des données et analyses

Le design expérimental, la procédure, le traitement des données et les analyses étaient identiques à l'Expérience 5.

3.1.6.2 Résultats et discussion

Les résultats ont été moyennés par canal et par condition et sont présentés dans la Figure 3.9. Des t-tests canal par canal ont été réalisés (illustrés dans la Tableau 3-4, et la Figure 3.10). Pour les comparaisons avec la baseline, nous avons trouvé des activations significativement plus fortes dans de nombreux canaux, pour les conditions en désaccord avec l'ITL [10 ($p < .01$), 7, 12 ($p < .05$), et 5, 6, 21, 23 ($p < .1$)]. Cet effet est retrouvé de façon prédominante dans l'hémisphère gauche. Pour les séquences sans contraste, une activation significative est aussi retrouvée dans le canal 17 ($p < .05$). Lors de la comparaison directe entre les deux conditions expérimentales, des activations significativement plus fortes pour la condition en désaccord avec l'ITL, par rapport à la condition en accord avec l'ITL, ont été retrouvées, surtout dans des canaux de l'hémisphère gauche [7, 10, 21 ($p < .05$) et 6 ($p < .1$)].

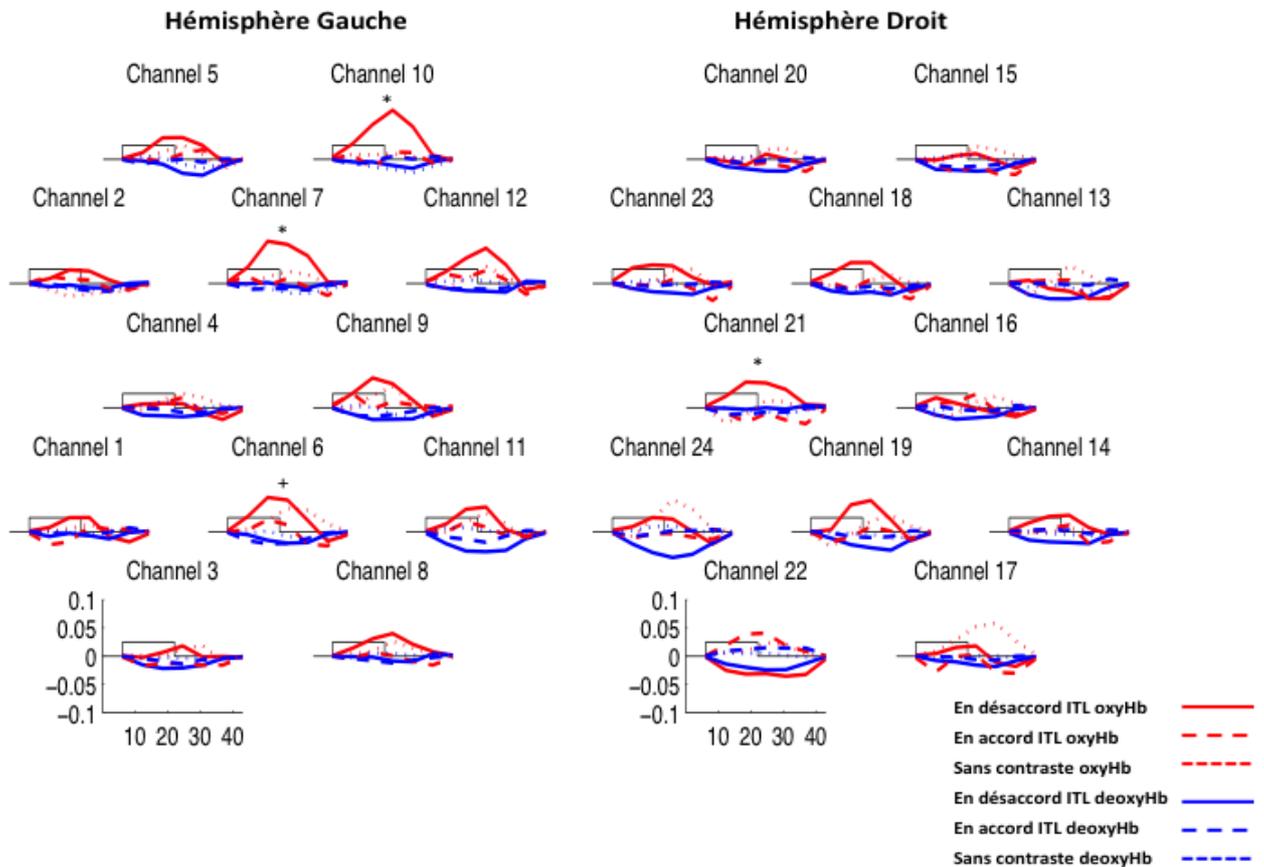


Figure 3.9 : Résultats des grandes moyennes de l'Expérience 6 (Condition pitch, bilingues). Toutes les conventions du graphique sont identiques à la Figure 6.

Les résultats de cette expérience montrent globalement des effets similaires à l'Expérience 3 pour les contrastes de durée trouvés chez les nouveau-nés exposés au français, suggérant que les nourrissons bilingues testés montrent un biais de groupement stable pour le pitch. De plus, ces activations différentielles sont trouvées également dans l'hémisphère gauche, dans des canaux relativement similaires à ceux de l'Expérience 3.

Tableau 3-4 : Résultats des tests de Student pour l'oxyHb uniquement pour $p < .05$ [valeur de p non corrigés (p_{unc}) ; valeur de p corrigée (p_{FDR}) ; CH : canaux ; Condition, A : en accord avec l'ITL DA : en désaccord avec ITL, baseline : 0, SC : sans contraste]

Mesures	OxyHb						DeoxyHb							
Conditions	A vs DA		DA vs O		SC vs O		DA vs A		DA vs O		A vs O		SC vs O	
CH	p _{unc}	p _{FDR}												
2													0,026	0,462
3									0,015	0,071	0,032	0,274		
4									0,026	0,077				
5									0,032	0,077				
6											0,039	0,274		
7	0,02	0,231	0,021	0,258										
10	0,029	0,231	0,006	0,148										
11									0,015	0,071				
12			0,041	0,325										
13									0,007	0,071				
15											0,021	0,274		
17									0,03	0,077				
18													0,039	0,462
19							0,006	0,151	0,003	0,062				
20									0,028	0,077				
21	0,02	0,231			0,041	0,893								
23									0,019	0,076				
24							0,02	0,241	0,01	0,071				

3.1.7 Discussion

Nous avons réalisé quatre expériences en NIRS montrant que les nouveau-nés, ayant eu différents types d'exposition linguistique in utero (une ou deux langues), présentaient des préférences de groupement prosodique à la naissance. Pour la première fois dans ce champ de recherche, nous avons montré que des nourrissons, provenant de milieu monolingue et bilingue, étaient sensibles au groupement prosodique dès la naissance. En effet, à notre connaissance, aucune des études précédentes n'a testé des nourrissons aussi jeunes, les études sur le développement de biais prosodiques commençant à partir de nourrissons âgés de 5-6 mois. De plus, les nourrissons dans les études précédentes provenaient d'environnement monolingue pour la majorité (Bion et al., 2011; Hay & Saffran, 2012; Yoshida et al., 2010), à une récente exception près (Molnar, Lallier, et al., 2014).

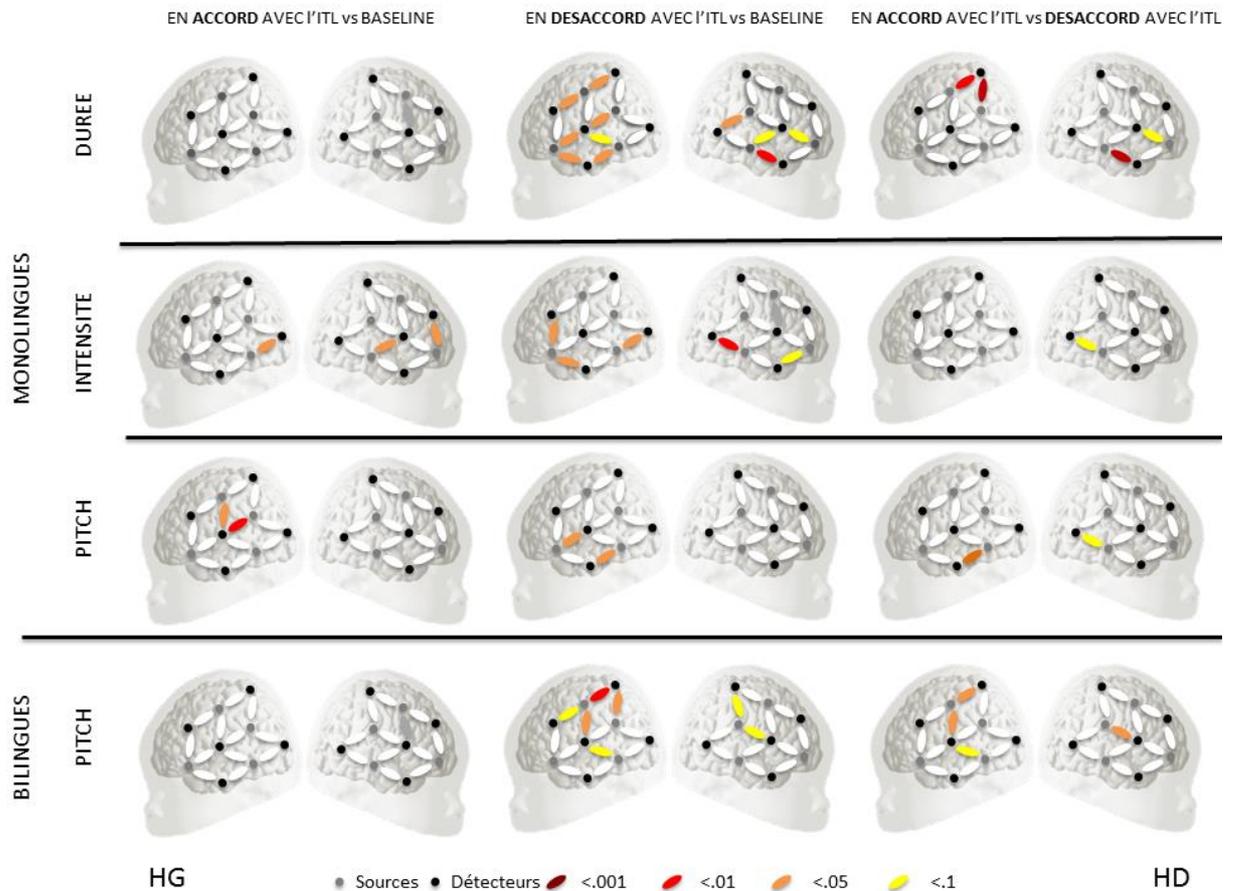


Figure 3.10 : Carte statistique (valeurs de p) comparant les réponses aux séquences : en accord avec ITL vs baseline, en désaccord avec ITL vs baseline et en accord avec ITL vs désaccord avec ITL. Pour chaque canal, les valeurs de p ont un code couleur : $p < .001$, bordeaux; $p < .01$, rouge ; $p < .05$, orange, $p < .1$, jaune, sans correction.

Chez les nouveau-nés exposés exclusivement au français (Exp. 3-5), un biais stable pour la durée a été trouvé, alors que les biais pour le pitch et l'intensité sont beaucoup plus faibles. Cette effet très fort pour la durée, un des indices les plus importants dans la prosodie du français, suggère un effet prénatal de la langue. Les plus faibles effets trouvés pour les deux autres indices pourraient indiquer que des biais perceptifs automatiques de groupement prosodique pour ces deux autres indices pourraient être présents à la naissance, mais seraient moins prononcés dans la population d'enfants que nous avons testés (issus de famille monolingues francophones). De plus, une activation plus forte a été relevée pour les séquences en désaccord avec l'ITL dans les trois expériences, indiquant que le traitement des patterns prosodiques, qui vont à l'inverse des biais perceptifs et/ou des patterns retrouvés dans la langue maternelle, demande plus d'effort que celui des patterns naturels et/ou familiers.

Chez les nouveau-nés exposés au français et à une autre langue, une réponse différentielle stable a été retrouvée pour le pitch (le seul indice testé pour ce groupe). Ce profil d'activation ressemble fortement à celui retrouvé dans l'Expérience 3 pour le contraste de durée chez les monolingues français, suggérant un biais stable pour le pitch chez les bilingues. De plus, cela confirme également que les différences de pitch utilisées dans cette étude sont assez grandes pour éliciter de fortes préférences de groupement. L'effet assez faible retrouvé chez les nouveau-nés exposés exclusivement au français dans l'Expérience 3 indique alors un probable manque de sensibilité aux variations de pitch dans ce groupe de monolingues.

Prises ensemble, ces données suggèrent donc qu'il est possible que des biais universels existent à la naissance, mais qu'ils seraient déjà fortement influencés par l'exposition linguistique prénatale. Cependant, plusieurs aspects de notre recherche méritent de plus amples discussions.

Le premier aspect concerne la question des caractéristiques prosodiques des langues entendues par les nouveau-nés bilingues et à quel point la modulation du biais serait liée à la présence de contrastes de pitch ou non dans ces langues. Ces nouveau-nés ont tous reçu une exposition équilibrée entre le français et une autre langue (entre 40-60 % d'exposition à chacune de leurs langues respectives). Les 2^{èmes} langues entendues par le groupe de nouveau-nés bilingues sont hétérogènes, mais elles ont toutes un point commun au niveau de leur structure prosodique ; elles présentent toutes par rapport au français de plus importantes variations de pitch. En effet, le pitch est utilisé pour marquer la proéminence au niveau de l'accentuation lexicale en kabyle, arabe, polonais et bulgare (Angoujard, 1986; Chaker, 1995; Dimitrova, Redeker, & Hoeks, 2009; Malisz, 2013), la proéminence au niveau du syntagme en turc (Nespor et al., 2008), et il est utilisé dans le système lexical tonal en koyaka, cantonais ou mandarin (Creissels & Gregoire, 1993; Lee, Chen, Luke, & Shen, 2002). Entendre donc des contrastes de pitch avant la naissance (*in utero*) pourrait avoir un effet pour l'utilisation de cet indice pour le groupement prosodique, entraînant la perception d'un biais stable à la naissance, similaire à celui retrouvé chez les nouveau-nés français monolingues pour la durée. Nos recherches confirment qu'une exposition répétée à certaines caractéristiques prosodiques de la parole avant la naissance (contrastes de durée ou de pitch) pourrait développer ou stabiliser des préférences de pattern prosodique. Pour cela, il serait intéressant d'avoir des groupes de bilingues homogènes (par exemple, franco-arabe ou franco-chinois), afin de relier leur profil d'activation au poids acoustique du pitch dans

chacune de ces langues ou encore les comparer à des groupes de nouveau-nés monolingues de différentes langues (par exemple, arabe ou chinois).

Deuxièmement, il est important de noter que nous avons seulement testé le pitch chez les nouveau-nés bilingues. Cependant, si les résultats obtenus dans les Expériences 1 et 4 reflètent une modulation très précoce du biais de groupement, on pourrait se demander si les bilingues ont besoin d'un seuil minimal d'exposition à leur deuxième langue maternelle pour développer ce biais. Une étude récente s'est justement intéressée à cette question chez des nourrissons de 9 mois bilingues, recevant une exposition dominante à l'espagnol ou au basque (Molnar, Lallier, et al., 2014). Ces auteurs ont comparé la préférence de ces deux groupes d'enfants pour des groupements de tons trochaïques versus iambiques contrastant en durée. Les résultats montrent une préférence claire pour le groupement de tons trochaïque (effet de nouveauté) chez les bilingues de dominance espagnole, mais pas chez les bilingues de dominance basque. Les auteurs expliquent ces résultats par le fait que ces langues diffèrent sur leur prosodie au niveau du syntagme, l'espagnol ayant un pattern rythmique iambique alors que le basque est trochaïque. Ces résultats suggèrent ainsi que la perception du biais prosodique basé sur la durée résulte de la qualité, mais également de la quantité du signal linguistique, puisque les bilingues de dominance basque ne montrent pas de préférence dans cette tâche, alors qu'ils étaient également exposés à de l'espagnol. De futures recherches seront nécessaires pour mieux comprendre l'émergence de ces biais en contrôlant le pourcentage d'exposition à la langue et ses effets.

Troisièmement, il existe une explication alternative à l'émergence du biais basé sur le pitch chez les nouveau-nés bilingues, pas nécessairement exclusive à la première. Nos participants bilingues sont peut-être plus sensibles au groupement de pitch par rapport aux monolingues, non seulement parce que leur autre langue utilise des contrastes de pitch, mais aussi parce que le fait d'être bilingue confère à ces nouveau-nés une flexibilité perceptuelle et cognitive. En effet, de récentes observations ont mis en exergue des effets cognitifs précoces de l'exposition simultanée à plusieurs langues et dès la naissance. Les bilingues montrent des avantages cognitifs et perceptifs en comparaison des monolingues, dont une meilleure attention aux détails perceptifs et contextuels permettant la discrimination des langues (Sebastián-Gallés, Albareda-Castellot, Weikum, & Werker, 2012; Werker, 2012), des phonèmes (Sundara, Polka, & Molnar, 2008) ou encore de l'accentuation lexicale (Abboub et al., 2015; Bijeljac-Babic et al., 2012). Ils sont aussi capables d'adapter rapidement leur attention et d'appliquer

de nouvelles règles contrairement aux monolingues (Kovács & Mehler, 2009a, 2009b). Évidemment, ces études ont porté sur des nourrissons plus âgés que les nouveau-nés testés ici, lesquels avaient entre 7 et 10 mois pour la majorité de ces études et avaient donc une plus grande quantité d'exposition linguistique. Cependant, nos résultats sont compatibles avec l'idée que l'exposition reçue lors du dernier trimestre de la grossesse pourrait être suffisante pour que ces nouveau-nés bilingues développent des capacités plus flexibles ou avancées dans le traitement du langage en général, ou prêtent plus d'attention à des indices prosodiques plus subtils, que les monolingues. Si de futures études devront dissocier les deux précédentes hypothèses, notamment en testant des groupes de bilingues exposés *in utero* à des langues dont la prosodie linguistique utilise peu de contrastes de pitch, nos résultats montrent toutefois clairement que des différences en termes d'exposition à la langue modulent les biais de groupement prosodique, confirmant un impact très précoce de la langue maternelle.

Le quatrième point, relié au premier, concerne les origines développementales des faibles biais trouvées chez les monolingues pour l'intensité et le pitch. Ces deux indices acoustiques sont peu utilisés dans la prosodie du français. Les nouveau-nés ont été testés à la naissance, donc bien avant la spécialisation neuronale et perceptive pour leur langue maternelle. Nous pouvons alors interpréter la faiblesse des réponses que nous avons observée selon deux hypothèses. D'une part, si la spécialisation pour la langue maternelle commence dès lors que le système auditif est opérationnel (dès le début du troisième trimestre de grossesse), alors les biais de groupement prosodique sur la base des trois indices prosodiques pourraient être à la fois déjà présents à la naissance et déjà modulés par les indices entendus majoritairement *in utero*. Une sensibilité initiale au pitch et à l'intensité pourrait ainsi être présente, mais aurait déjà diminué du fait de l'exposition au seul français. A l'inverse, il est également possible que les biais de groupement universaux soient initialement faibles et que la langue entendue les renforce sélectivement, en fonction des indices présents dans le signal linguistique entendu. Des études supplémentaires impliquant des populations n'ayant pas reçu la même quantité d'exposition à la langue maternelle (des fœtus ou des nourrissons prématurés) seront nécessaires pour préciser l'origine de ces biais et leur développement.

Cinquièmement, alors que des études chez l'adulte suggèrent qu'un effet de modulation de la perception du groupement spécifique à la langue maternelle ne se retrouve que si ce matériel est suffisamment complexe (Bhatara et al., sous presse,

2013 ; Hay & Diehl, 2007), nous avons observé à la naissance une modulation partielle des capacités de perception du groupement prosodique pour des tons simple purs. Et d'autres études, chez des nourrissons plus âgés, ont également observé la modulation du biais sur des tons non-linguistiques à partir de 7-8 mois (Molnar, Lallier, et al., 2014; Yoshida et al., 2010). Il est possible que la modulation de l'ITL par l'environnement linguistique n'implique pas les mêmes processus entre les adultes et les nourrissons. Les facteurs pouvant moduler l'ITL cross-linguistiquement doivent donc être explorés en détails : type de matériel : linguistique versus non linguistique; degré de complexité ; caractéristiques prosodiques des langues testées; type de tâche et de procédure; âge ; ...)

Sixièmement, dans les Expériences 3 et 6, nous avons principalement observé des activations bilatérales lors de la présentation des séquences en désaccord avec l'ITL par rapport à la baseline, dans des canaux correspondants aux zones auditives temporales (Gervain, Macagno, et al., 2008; May et al., 2011). Par contre, dans ces mêmes expériences, les différences les plus fortes entre les deux conditions expérimentales (en accord versus en désaccord avec l'ITL) sont beaucoup plus prononcées dans les canaux temporo-pariétaux de l'hémisphère gauche (impliquant peut-être le Gyrus Temporal Supérieur, le Gyrus Temporal Moyen et le cortex Pré-moteur). Ce résultat est au premier abord surprenant, étant donné que de nombreuses recherches ont montré que le traitement des mélodies musicales et de la prosodie linguistique se fait préférentiellement dans l'hémisphère droit, que ce soit chez les adultes ou les nourrissons (Friederici, 2012; Homae, Watanabe, Nakano, Asakawa, & Taga, 2006; Wartenburger et al., 2007). Cependant, ce qui distingue ici les deux types de groupement présentés, ce n'est pas la présence ou non de prosodie, mais la conformité de la structure prosodique présentée avec l'ITL. Or, des travaux ont justement montré que l'hémisphère gauche serait préférentiellement impliqué dans le traitement auditif lorsqu'il s'agit notamment de traiter la structure de sons complexes (Gervain, Macagno, et al., 2008; Peña et al., 2003) ou de faire un apprentissage séquentiel (Dehaene-Lambertz et al., 2002). Nos résultats pourraient aller dans le même sens. Mais de façon alternative, étant donné les limites de notre technique pour étudier les structures cérébrales profondes, il reste possible que les régions mises en évidence dans cette étude fassent partie d'un réseau plus large impliqué dans le traitement du groupement prosodique.

En résumé, nos résultats sont les premiers à démontrer que, dès la naissance, des nouveau-nés exposés à une ou deux langues *in utero* présentent des préférences de

groupement prosodique suivant les prédictions de l'ITL (Bion et al., 2011; Hayes, 1995). De plus, les biais que nous avons observés sont en cohérence avec les propriétés prosodiques des langues entendues, suggérant ainsi que l'expérience prénatale module déjà la perception et l'apprentissage du langage. Ces biais perceptifs pourraient être des outils très puissants guidant les nourrissons lors de l'acquisition de la structure grammaticale, en particulier l'ordre des mots de leur langue maternelle (Gervain, Nespor, et al., 2008; Gervain & Werker, 2013; Nespor et al., 2008) ou encore de son lexique, notamment en permettant de segmenter le flux continu de la parole en unités linguistiques comme les mots, grâce à la détection ou reconnaissance des patterns rythmiques spécifiques de leur langue maternelle. C'est justement sur ce dernier point que va porter le dernier chapitre.

3.1.8 Points clefs

- Les nouveau-nés sont sensibles à l'ITL pour la durée et le pitch.
- La sensibilité à l'ITL est trouvée essentiellement pour les indices acoustiques présents dans la prosodie de leur langue maternelle.
- L'ITL est modulée par l'exposition prénatale (monolingue et bilingue).

3.2 CHAPITRE 4 : Chez des nourrissons français monolingues ou bilingues de 10 mois

-- Cette partie est issue de la traduction libre de Abboub N., Bijeljac-Babic R., Serres J & Nazzi T. (2015). On the importance of being bilingual: Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, 132, 111-120 --

Dans les chapitres précédents, nous avons vu que non seulement les nourrissons de 7 mois, mais aussi les nouveau-nés, sont capables de traiter des patterns rythmiques variant en durée, intensité et en pitch. Nous avons également vu que, dès la naissance, la sensibilité aux indices acoustiques est différente chez les nouveau-nés ayant entendu exclusivement du français et ceux qui ont entendu en plus une autre langue. Ces résultats suggèrent que des mécanismes très précoces sont présents à la naissance pour détecter les indices de la prosodie linguistique et ses patterns rythmiques. Dans la parole naturelle, les indices comme la durée, le pitch ou l'intensité contribuent tous les trois à contraster les unités prosodiques de la parole. Par exemple, au niveau du mot, une augmentation d'intensité, de pitch et/ou de durée sur une syllabe signale la présence de l'accentuation et ces trois indices sont parfois utilisés en isolation, parfois en combinaison. De plus, en fonction des langues, la position de l'accentuation lexicale peut être soit variable, soit fixe et être présente par exemple sur la syllabe initiale, finale, ou sur l'avant-dernière syllabe des mots. Si elle est variable, l'accentuation lexicale peut changer le sens d'un mot dans certaines langues (accentuation lexicale), comme par exemple en anglais ou en allemand. Par contre, dans d'autres langues comme le français, l'accentuation lexicale n'existe pas.

Dans ce chapitre, nous allons nous intéresser à la perception de l'accentuation au niveau des mots, notamment de la sensibilité aux patterns rythmiques. Cette fois-ci, nous allons présenter les mots dans un contexte naturel où plusieurs indices sont présents pour marquer cette accentuation. La première question que pose notre étude concerne la perception de l'accentuation des mots iambiques et trochaïques : les nourrissons sont-ils sensibles aux différences de patterns rythmiques/accidentuels ? La deuxième question posée, au vu des précédentes études, est : comment ces compétences sont-elles modulées par l'environnement linguistique ? Pour répondre à ces questions,

nous allons présenter à deux groupes de nourrissons de 10 mois, exposés soit au français exclusivement, soit au français et à une autre langue, des listes de mots bisyllabiques accentués soit en position initiale (pattern trochaïque), soit en position finale (pattern iambique).

3.2.1 Introduction

Aujourd'hui, plus de la moitié de la population mondiale est bilingue (Grosjean, 2010). Des recherches récentes ont exploré les mécanismes précoces sous tendant l'acquisition du langage chez des enfants grandissant dans un environnement bilingue et notamment les mécanismes permettant l'apprentissage des régularités de chacune des langues (Byers-Heinlein et al., 2010; Weikum et al., 2007). Un des premiers pas importants dans l'acquisition du langage concerne l'acquisition de l'inventaire des phonèmes de la ou des langue(s) maternelle(s), mais aussi des propriétés prosodiques comme le rythme et l'intonation (au niveau de la phrase et aussi du mot : accentuation lexicale ou tons lexicaux). La présente étude s'intéresse au traitement de l'accentuation lexicale. Cette accentuation est marquée par des différences d'intensité, de pitch ou de durée, entre des syllabes accentuées et non-accentuées. Dans les langues avec une accentuation lexicale variable comme l'anglais ou encore l'allemand, la position de l'accentuation, ou de la syllabe accentuée, peut changer le sens du mot ; par exemple IMPort est un nom, alors que imPORT est un verbe (Cutler, 2005). Toutefois, des patterns d'accentuation peuvent prédominer dans certaines langues, c'est le cas des patterns trochaïques en allemand ou encore en anglais. Le français ne possède pas d'accentuation au niveau du mot, même s'il existe une accentuation au niveau du syntagme, se situant généralement sur la dernière syllabe de cette unité (Jun & Fougeron, 2002). Pour des nourrissons apprenant une langue comme l'anglais (avec une accentuation variable), il est donc important d'être capable de détecter la position de l'accentuation pour l'accès au lexique. Cependant, cette capacité n'est pas nécessaire pour le français. Notre étude a donc pour but de comprendre comment de très jeunes nourrissons, exposés à la fois au français et à une autre langue accentuelle (à accentuation variable), apprennent deux systèmes prosodiques différents. Pour ce faire, nous avons comparé les capacités de discrimination de patterns d'accentuation par des nourrissons de 10 mois apprenant le français (monolingues) et par des nourrissons apprenant le français et une autre langue accentuelle (bilingues).

Seules quelques études ont comparé la perception précoce de la prosodie du langage entre des nourrissons monolingues et bilingues. Chez les monolingues, la reconnaissance de la langue maternelle est présente dès la naissance (Mehler et al., 1988; Moon et al., 1993). De plus, des nouveau-nés et des nourrissons de 2 mois sont capables de discriminer des langues entre elles si, et seulement si, ces langues diffèrent rythmiquement (Byers-Heinlein et al., 2010; Christophe & Morton, 1998; Nazzi, Bertoncini, et al., 1998). Ce n'est qu'à partir de 3,5 à 5 mois que les nourrissons sont capables de discriminer des langues ayant un rythme similaire (Butler, Floccia, Goslin, & Panneton, 2011; Molnar, Gervain, et al., 2014; Nazzi et al., 2000). Les bilingues, quant à eux, préfèrent leurs deux langues maternelles, au même niveau, dès la naissance et sont capables de les discriminer entre elles, si elles présentent des rythmes différents (Byers-Heinlein et al., 2010). A partir de 3,5 à 5 mois, ils deviennent aussi capables de discriminer leurs langues maternelles même si elles sont rythmiquement similaires (Bosch & Sebastián-Gallés, 1997; Molnar, Gervain, et al., 2014). De subtiles différences entre nourrissons monolingues et bilingues ont été observées dans ces études, notamment des temps d'écoute plus longs, pouvant suggérer que les enfants bilingues traitent leurs langues maternelles d'une façon différente que les monolingues. Cependant, ces études concernant les capacités de discrimination des langues révèlent des performances assez similaires entre bilingues et monolingues, donc concernant le traitement des propriétés prosodiques globales.

D'autres études ont exploré la sensibilité des nourrissons à l'information prosodique au niveau du mot, et en particulier leurs capacités à discriminer des patterns d'accentuation et à préférer des caractéristiques spécifiques de leurs langues maternelles. Deux types de contexte expérimental ont été utilisés dans ces études. Dans le premier contexte, simple (i.e., sans variabilité segmentale), un seul type d'items (e.g. : /gaba/) possédant deux types d'accentuation (en position initiale : trochaïque versus en position finale : iambique) est répété plusieurs fois. Ce contexte permet ainsi de mesurer les capacités de discrimination de sons de parole sur la base de propriétés acoustiques de bas niveau. Dans le deuxième contexte, plus complexe (i.e., avec variabilité segmentale), des listes de différents items (e.g. : /datu/, /sapi/, ...) possédant un pattern trochaïque ou iambique sont présentées. Ce contexte permet d'évaluer les capacités d'abstraction et de généralisation des patterns prosodiques dans des listes d'items présentant de la variabilité segmentale.

En l'absence de variabilité segmentale, des capacités de discrimination très précoces sont observées chez : des nouveau-nés italiens (Sansavini et al., 1997), des nourrissons américains de 2 mois (Jusczyk & Thompson, 1978), des nourrissons allemands de 4 mois (données électro-physiologiques, ERP : Friederici, Friedrich, & Christophe, 2007 ; données comportementales, HPP : Höhle, Bijeljac-Babic, Herold, Weissenborn, & Nazzi, 2009) et des nourrissons espagnols de 6 mois (Skoruppa et al., 2013). Par ailleurs, ces capacités sont également observées chez des nourrissons français de 4 à 10 mois (Friederici et al., 2007; Höhle et al., 2009; Skoruppa et al., 2009, 2013), même si l'accentuation lexicale n'est pas utilisée en français. Ces résultats concernant la capacité de discrimination de patterns d'accentuation dans un contexte simple (sans variabilité segmentale) suggèrent que cette capacité est indépendante des caractéristiques de la langue maternelle. Cependant, dès 10 mois, la sensibilité des nourrissons français à discriminer les patterns accentuels décline. Bien que les nourrissons français soient capables de discriminer des patterns d'accentuation après une familiarisation de 2 minutes, ils ne le sont pas après seulement 1 minute de familiarisation, alors que des nourrissons français de 6 mois réussissent avec seulement 1 minute (Bijeljac-Babic et al., 2012). Ces études suggèrent que ces nourrissons français de 10 mois ont besoin de plus de temps pour caractériser les différents patterns rythmiques des stimuli entendus. Etant donné qu'en français, l'accentuation au niveau lexical n'existe pas, ces changements développementaux pourraient refléter une réorganisation perceptive liée aux propriétés prosodiques de la langue maternelle. Dans un contexte sans variabilité segmentale, des changements développementaux chez des nourrissons allemands entre 4 et 6 mois sont aussi observés, concernant l'émergence d'une préférence pour les items trochaïques par rapport aux items iambiques, alors qu'aucun biais de la sorte n'est observé à 6 mois chez des nourrissons français (Höhle et al., 2009).

Seulement deux études ont exploré les capacités de traitement de l'accentuation en l'absence de variations segmentales chez des enfants bilingues. Bijeljac-Babic et collègues (2012) ont trouvé que des nourrissons âgés de 10 mois, apprenant simultanément le français et une langue accentuelle, peuvent discriminer les contrastes d'accentuation, à l'inverse des nourrissons n'apprenant que le français. Cependant, cet effet n'est présent que dans un sous-groupe d'enfants bilingues, celui pour lequel la langue dominante est celle possédant une accentuation lexicale. Les bilingues équilibrés ne présentent qu'un effet marginal. Dans une seconde étude, les capacités de

nourrissons plus jeunes, de 6 mois, apprenant le français et l'allemand simultanément, ont été évaluées (Bijeljic-Babic, Serres, Höhle, & Nazzi, 2013). Dans ce groupe, une préférence pour les items trochaïques est observée au même âge que chez des nourrissons monolingues allemands (Höhle et al., 2009). De plus, le degré de préférence ne semble pas être corrélé au taux d'exposition à la langue. Ces deux études suggèrent donc qu'apprendre une langue avec une accentuation lexicale permet de maintenir la sensibilité à détecter différents patterns accentuels chez ces nourrissons bilingues. Cependant, ces études ont été réalisées dans un contexte expérimental où aucune variabilité segmentale n'est présente, amenant à penser que ces capacités pourraient résulter d'un traitement acoustique de bas niveau.

Est-ce que de telles capacités pourraient s'observer si le contexte expérimental contenait de la variabilité segmentale ? Les nourrissons sont-ils capables de détecter un pattern prosodique commun lors de la présentation d'une liste de mots différents ? Chez des monolingues, une telle capacité est assez limitée à la naissance. En effet, ils sont capables de détecter des patterns d'accentuation lors de la présentation de listes de mots variant uniquement par leurs consonnes (Sansavini et al., 1997), mais pas lorsque les listes de mots varient aussi par leurs voyelles (Sansavini, Bertoncini, & Giovanelli, 1994). Les nourrissons de 6 mois français et espagnols ont également des difficultés à discriminer des patterns d'accentuation lors de la présentation de listes de mots possédant de la variabilité segmentale (Skoruppa et al., 2013). Ce n'est que pour des nourrissons un peu plus âgés que des preuves claires de discrimination sont observées : chez des espagnols de 9 mois et des anglais de 8 et 12 mois (Skoruppa et al., 2009, 2011). Néanmoins, ces capacités semblent être modulées par la langue maternelle et pas seulement par l'âge. En effet, les nourrissons français de 9 mois ne sont pas capables de discriminer différents patterns d'accentuation dans ces conditions expérimentales, ce qui met en évidence des difficultés à traiter les patterns d'accentuation dans un contexte de variabilité segmentale (Skoruppa et al., 2009). En dernier lieu, une étude utilisant des listes de mots possédant de la variabilité segmentale montre l'émergence d'un biais trochaïque chez des nourrissons américains entre 6 et 9 mois (Jusczyk et al., 1993). Tous ces résultats suggèrent qu'entre 6 et 9 mois, des nourrissons apprenant une langue avec une accentuation lexicale variable traitent efficacement l'accentuation en présence de variabilité segmentale, contrairement aux nourrissons n'entendant que du français. Ces observations posent la question de savoir comment cette capacité se développe chez des nourrissons bilingues entendant le français et une langue possédant une accentuation

lexicale variable. En effet, comment ces enfants traitent-ils l'accentuation dans un contexte complexe où la variation segmentale est présente, contexte qui implique d'extraire et de généraliser les patterns d'accentuation ?

Le but de la présente étude était donc d'examiner la perception de l'accentuation lexicale dans un contexte complexe, chez des nourrissons âgés de 10 mois, entendant soit exclusivement du français, soit du français combiné à une autre langue possédant une accentuation lexicale variable. La même procédure que celle de Bijeljac-Babic et al. (2012) a été utilisée, ainsi que des stimuli identiques à l'étude de Skoruppa et al. (2009), afin d'étudier la capacité de discrimination des nourrissons dans une condition difficile (i.e., une seule minute de familiarisation). Chaque nourrisson a été familiarisé avec une liste de mots prononcés avec un type d'accentuation donné (i.e., trochaïque ou iambique), puis de nouvelles listes de mots ont été présentées, soit ayant le même type d'accentuation que pendant la familiarisation (pattern familial), soit correspondant à l'autre type (pattern nouveau). Si les nourrissons discriminent les deux types de patterns accentuels, alors des temps de regard différents entre les patterns nouveaux et familiers devraient être observés dans la phase de test. Ceci suggérerait que les nourrissons ne sont pas seulement capables de percevoir les indices acoustiques des contrastes prosodiques, mais sont également capables d'extraire et de généraliser ces patterns rythmiques en présence de variabilité segmentale. Étant donné les résultats obtenus par Skoruppa et al. (2009), ce pattern de résultats n'est attendu que chez les nourrissons bilingues de 10 mois et non chez les nourrissons français. De plus, la dominance de la langue (i.e., la balance d'exposition au français versus à la langue avec accentuation lexicale variable) sera également considérée, ce facteur pouvant moduler la performance des bilingues.

3.2.2 Expérience 7

3.2.2.1 *Méthode*

3.2.2.1.1 Participants

Aucun participant n'avait de problème de santé connu et tous étaient nés à terme. Les participants ont été recrutés à l'aide des listes d'état civil obtenues par les Archives de France ou par la base de données de notre laboratoire (des familles ayant déjà participé à des études menées dans notre laboratoire). Un consentement écrit des parents a été obtenu pour chaque enfant participant. 48 nourrissons, âgés en moyenne de

10 mois, ont participé à cette expérience. Vingt-quatre d'entre eux ont été exposés exclusivement au français dans leur environnement familial (11 filles, 13 garçons, $M_{age} = 10;02$ [mois; jours], rang = 9;10–10;27). Six nourrissons monolingues supplémentaires ont été testés, mais leurs données ont été exclues de l'analyse, 4 du fait de problèmes techniques ou erreurs expérimentales et 2 pour cause de pleurs et/ou fatigue. Les 24 autres nourrissons, sont issus de famille bilingue et étaient exposés au français et à une autre langue (15 filles, 9 garçons, $M_{age} = 10;08$, rang = 9;08–11;06). Les données de 11 nourrissons bilingues supplémentaires ont été exclues de l'analyse, 10 pour cause de pleurs et/ou fatigue, et un en raison d'un problème technique.

Tous les nourrissons bilingues ont été exposés à deux langues simultanément depuis la naissance, mais le taux d'exposition à celles-ci varie entre les enfants. L'exposition aux langues a été mesurée avec un questionnaire d'exposition (Bosch & Sebastián-Gallés, 1997), comptabilisant le temps d'interaction entre les parents et le nourrisson en fonction de la langue utilisée, durant la semaine et le week-end. Avec cette mesure, deux sous-groupes ont été distingués au sein du groupe de nourrissons bilingues, en fonction de la dominance du français dans leur environnement linguistique : un groupe « français dominant », composés de 12 nourrissons exposés majoritairement au français (entre 60-70%, avec une moyenne de 67,9%), et un groupe « français non-dominant », composés de 12 nourrissons exposés minoritairement au français (entre 30 à 50 % du temps avec une moyenne de 42,1%). Dans cette étude, le groupe de bilingues est hétérogène, c'est-à-dire que les nourrissons sont exposés à 11 langues différentes reflétant la diversité culturelle présente à Paris et dans sa banlieue. Ainsi, différentes langues sont parlées dans les familles de nos participants bilingues : arabe, anglais, allemand, hébreux, italien, japonais, portugais, russe, serbe turc et wolof (voir Tableau 3-5). La stratégie d'explorer les effets du bilinguisme sur des groupes de bilingues hétérogènes (où les langues partagent toutes une propriété commune, qui est la présence de l'accentuation lexicale à des positions variables) a été adoptée dans plusieurs études récentes (Bijeljac-Babic et al., 2012; Fennell, Byers-Heinlein, & Werker, 2007; Poulin-Dubois, Blaye, Coutya, & Bialystok, 2011; Singh et al., 2015). Il est important de noter que les nourrissons entendant du français et de l'espagnol n'ont pas été inclus dans la présente étude, étant donné que les stimuli ont été produits par une locutrice espagnole. Les stimuli sont donc prononcés dans une langue étrangère pour tous les enfants.

Pour les 11 langues entendues par le groupe d'enfants bilingues, nous avons vérifié qu'un accent lexical est réalisé au niveau du mot (nous n'avons par contre pas pu évaluer précisément les facteurs acoustiques utilisés pour marquer l'accentuation dans toutes ces langues, voir aussi la discussion à propos de ce sujet), et que la position de l'accentuation est variable, même s'il existe des patterns d'accentuations prédominants dans certaines langues. La présence d'accentuation variable est rapportée pour l'anglais, l'allemand, l'italien, le russe, le turc, le portugais et le serbe (Goedemans & van der Hulst, 2010), l'hébreu (Segal & Kishon-Rabin, 2012), le wolof et l'arabe parlé au Maghreb (qui est le type d'arabe entendu par nos participants ; (Downing, 2004). Enfin, le japonais possède un accent lexical de pitch (hauteur) variable (Cutler & Otake, 1999).

3.2.2.1.2 Stimuli

Les stimuli utilisés sont les mêmes que ceux de l'étude de Skoruppa et al. (2009), comprenant un total de 16 séquences CVCV : 8 utilisées pour la familiarisation, 8 pour la phase de test. Les stimuli ont été prononcés dans un registre de parole adressée à l'enfant (*Infant-directed speech*) par une locutrice de langue maternelle espagnole. Les 8 pseudo-mots CVCV pour la familiarisation ont été enregistrés soit avec l'accentuation sur la syllabe initiale (trochaïque : /'datu/, /'sapi/, /'kiba/, /'nuki/, /'latu/, /'buki/, /'luma/, /'tiku/) pour présentation à la moitié des nourrissons, soit avec l'accentuation sur la syllabe finale (iambique: /da'tu/, /sa'pi/, /ki'ba/, /nu'ki/, /la'tu/, /bu'ki/, /lu'ma/, /ti'ku/) pour présentation à l'autre moitié des nourrissons. Huit stimuli ont été utilisés pour la phase test (utilisés pour tous les nourrissons), 4 avec accentuation sur la syllabe initiale (trochaïque: /'lapi/, /'naku/, /'nila/, /'tuli/) et 4 autres avec accentuation sur la syllabe finale (iambique: /ki'bu/, /lu'ta/, /pi'ma/, /pu'ki/). Les pseudo-mots utilisés dans notre expérience sont pour la plupart des non-mots dans toutes les langues auxquelles les nourrissons étaient exposés³.

³ A l'aide de locuteurs natifs, nous avons identifié seulement 8 séquences pouvant être de vrais mots dans ces langues. Toutefois, ces mots sont très peu fréquents et donc probablement peu entendus par des nourrissons de 10 mois. De plus, 6 de ces mots correspondent à des stimuli utilisés dans la phase de familiarisation (ce qui ne peut donc pas influencer les performances de discrimination durant la phase test). Exemples de mots : pour le japonais, /kiba/ [croc], /nuki/ [enlever], et /buki/ [arme]; pour l'italien, /buki/ [trous]; pour le portugais, /latu/ [large] et /naku/ [morceau]; pour le wolof, /sapi/ [sans saveur] and /nila/ [c'est ça].

Tableau 3-5 : Caractéristiques des nourrissons bilingues (% d'exposition au français, deuxième langue) et performances individuelles (temps d'orientation du regard en secondes pour les patterns d'accentuations familiers ou nouveaux à la phase de test et les scores de différence), en fonction du sous-groupe (français dominant versus français non-dominant).

Langues	% d'exposition au français	Familier	Nouveau	Différence
arabe	70	6,12	7,13	1,02
arabe	70	5,37	7,6	2,23
anglais	70	7,94	11,06	3,12
allemand	70	5,45	7,71	2,26
hébreu	70	8,61	9,48	0,87
Italien	70	6,97	9,33	2,37
Italien	70	6,15	7,61	1,46
Russe	70	6,84	5,73	-1,11
Serbe	70	5,98	6,98	0,99
allemand	65	4,4	5,12	0,72
Italien	60	7,12	7,61	0,49
Wolof	60	8,6	13,05	4,45
arabe	50	7,88	6,67	-1,21
arabe	50	6,42	7,49	1,08
anglais	50	5,61	6,44	0,83
allemand	50	7,41	7,74	0,32
Italien	50	4,65	7,56	2,91
Russe	50	5,5	5,63	0,13
Turque	50	5,56	7,19	1,63
Portugais	35	4,75	6,11	1,35
arabe	30	6,29	9,92	3,63
Japonaise	30	5,93	7,18	1,25
Portugais	30	6,74	9,4	2,66
Russe	30	4,45	7,12	2,68

Des mesures acoustiques ont révélé que l'accentuation est marquée par des différences de durée, d'intensité et de pitch entre les syllabes accentuées et non accentuées. Cependant, les différences de durée étaient minimales pour les items trochaïques (pour un résumé et des détails statistiques voir Tableau 3-6). Toutes les différences rapportées montrent des valeurs plus grandes (plus aiguë, plus forte et plus longue) pour les syllabes accentuées que les syllabes non-accentuées, excepté pour la durée des items trochaïques.

Tableau 3-6 : Analyses acoustiques (durée, intensité et pitch) des voyelles contenues dans les stimuli utilisés dans notre étude (issue de Skoruppa et al. 2009) et celle de Bijeljac-Babic et al. (2012).

	Item trochaïque		Item Iambique	
	Syllabe initiale	Syllabe finale	Syllabe initiale	Syllabe finale
Skoruppa et al. (2009); présente étude				
Durée (ms)	209,9	220,2	90,6	342,9***
Intensité (dB)	85,4	74,9***	77,9	81,3***
Pitch (Hz)	291,7	191,2***	216,1	244,1***
Höhle et al. (2009); Bijeljac-Babic et al. (2012)				
Durée (ms)	208,8	248,4*	93,6	349,8***
Intensité (dB)	79,5	73,4***	77,5	76,7*
Pitch (Hz)	174,6	134,6***	170,6	150,5*

*** : $p < ,001$; * : $p < ,05$

3.2.2.1.3 Procédure, équipement/dispositif et dessin expérimental

L'expérience a été conduite dans une pièce insonorisée. La cabine de test est disposée de la manière suivante : une lumière rouge se situe de chaque côté de la cabine et une lumière verte se situe au centre. Ces lumières sont fixées à des panneaux, à la hauteur du regard des nourrissons. Sous les lumières rouges, des hauts parleurs (SONY xs-F1722) sont cachés derrière des panneaux. En dessous de la lumière verte, une caméra est placée et permet de mesurer et d'enregistrer le comportement des nourrissons. Un ordinateur (Dell Optiplex), une télévision connectée à la caméra et un « boîtier réponse » se trouvent à l'extérieur de la cabine insonorisée. Le boîtier réponse, équipé de plusieurs boutons, est connecté à l'ordinateur. L'observateur mesurait le temps d'orientation du regard des nourrissons lors de la présentation des stimuli en regardant la vidéo du nourrisson sur l'écran de télévision. En pressant les boutons de réponse en fonction de la direction du regard du nourrisson (i.e., gauche, droite, milieu), l'observateur contrôlait le début et la fin du fonctionnement des lumières et la présentation des stimuli. Afin d'éviter que l'observateur et les parents n'entendent les stimuli présentés au nourrisson, tous les deux portaient des bouchons d'oreilles et écoutaient de la musique avec un casque fermé.

La même version de la procédure d'orientation préférentielle de la tête (head preferential procedure, HPP) que Bijeljac-Babic et al. (2012) a été utilisée. Chaque

nourrisson était assis sur les genoux d'un des parents au centre de la cabine de test. Chaque test commençait avec la lumière verte clignotant au centre jusqu'à ce que le nourrisson s'oriente vers cette source lumineuse. Dès que le nourrisson avait regardé cette lumière verte, celle-ci s'éteignait et une des lumières rouges sur le côté se mettait à clignoter. Quand l'enfant se tournait vers la bonne direction (lumière clignotante), l'essai était initié. Le stimulus sonore était délivré par les haut-parleurs via un amplificateur audio (Marantz PM4000). Chaque stimulus était joué jusqu'à sa fin ou arrêté après que l'enfant ait détourné son regard de plus de 2 s de la lumière rouge clignotante. Quand l'enfant détournait le regard de la lumière moins de 2 s et revenait sur la cible, alors le son continuait d'être joué. Cependant, le temps passé à « regarder ailleurs » était automatiquement soustrait des temps d'orientation pour cet essai. Donc, le temps maximum d'orientation du regard pour un essai correspondait à la durée de chaque séquence de son. Si un essai avait un temps d'orientation de moins de 1,5 s, l'essai était répété une seconde fois et le temps d'orientation initial était remplacé par le temps obtenu lors de ce deuxième essai. Toutes les informations à propos de la durée des temps d'orientation pour chaque essai étaient enregistrées par l'ordinateur.

Notre procédure avait pour but de tester la discrimination et pour ce faire, elle comprenait une phase de familiarisation suivie d'une phase de test. Les nourrissons étaient familiarisés soit aux items de type trochaïque, soit aux items de type iambique, jusqu'à ce qu'ils atteignent le critère de 1 min de temps d'orientation. La moitié des 2 sous-groupes de bilingues et la moitié du groupe de monolingues ont été familiarisés avec le pattern trochaïque et les autres enfants ont été familiarisés avec le pattern iambique. Une fois le critère de 1 min de temps d'orientation atteint, tous les nourrissons étaient exposés aux 4 essais de la phase de test: 2 essais correspondant au même type de pattern d'accentuation (« familier ») et 2 essais correspondant au pattern opposé (« nouveau »). Ce bloc de 4 essais était répété 3 fois avec un ordre de présentation aléatoire. Un total de 12 essais tests a donc été présenté, avec 6 essais correspondant au pattern d'accentuation familier et 6 essais au pattern d'accentuation nouveau.

3.2.2.2 Résultats

Les résultats sont illustrés dans la Figure 3.11. Dans un premier temps, une analyse de variance (ANOVA) sur les temps d'orientation du regard a été réalisée avec, comme facteur intra-sujet, la Familiarité (pattern d'accentuation familier vs nouveau en

test) et, comme facteurs inter-sujets, le Pattern d'accentuation de la familiarisation (pattern iambique vs trochaïque) et la Langue maternelle (monolingue vs bilingue). Cette analyse a révélé un effet significatif de la Familiarité [$F(1, 44) = 17.16, p < .001, \eta^2_p = .281$], et une interaction entre la Familiarité et la Langue maternelle [$F(1, 44) = 6.56, p = .014, \eta^2_p = .130$]. Des comparaisons planifiées ont montré que cette interaction était due au fait que le groupe d'enfants bilingues avait des temps d'orientation plus longs pour les listes d'items avec le nouveau pattern d'accentuation ($M_{\text{nouveau}} = 7.78 \text{ s}, SE = 0.37$) que pour les listes d'items avec le pattern d'accentuation familier ($M_{\text{familier}} = 6.28 \text{ s}, SE = 0.25$) [$F(1, 44) = 22.47, p < .001, \eta^2_p = .338$]. Par contre, le groupe de monolingues ne présentait pas de différence de temps d'orientation entre les deux types de listes ($M_{\text{nouveau}} = 7.21 \text{ s}, SE = 0.28$ vs $M_{\text{familier}} = 6.85 \text{ s}, SE = 0.28$) [$F(1, 44) = 1.25, p = .27$]. Sur les 24 nourrissons bilingues, 22 ont eu des temps d'orientation plus longs pour les listes d'items nouveaux (test binomial, $p = .84$). Tous les autres effets ou interactions n'étaient pas significatifs [$F_s(1, 44) < 1$].

Dans un second temps, l'influence du taux d'exposition à la langue sur les temps d'orientation du regard des nourrissons bilingues a été testée. Une ANOVA a été réalisée sur les résultats des enfants bilingues uniquement, avec, comme facteur intra-sujet, la Familiarité (pattern d'accentuation familier vs nouveau en test) et, comme facteurs inter-sujets, le pattern d'accentuation de la familiarisation (pattern iambique vs trochaïque) et l'exposition à la langue (français dominant vs français non dominant). Cette analyse a révélé un effet principal de la Familiarité [$F(1, 22) = 27.80, p < .001, \eta^2_p = .998$], mais aucune interaction entre la Familiarité et l'Exposition à la langue [$F(1, 22) < 1$]. Afin d'évaluer si l'effet de la familiarité était influencé par la balance des deux langues, des scores de préférence différentielle ont été calculés pour chaque enfant (i.e., différence dans les temps d'orientation pour les patterns d'accentuation nouveaux et familiers). La corrélation entre ces scores et le taux d'exposition au français a été calculée, et aucune corrélation significative n'a été observée [$r = -.17, p = .427$].

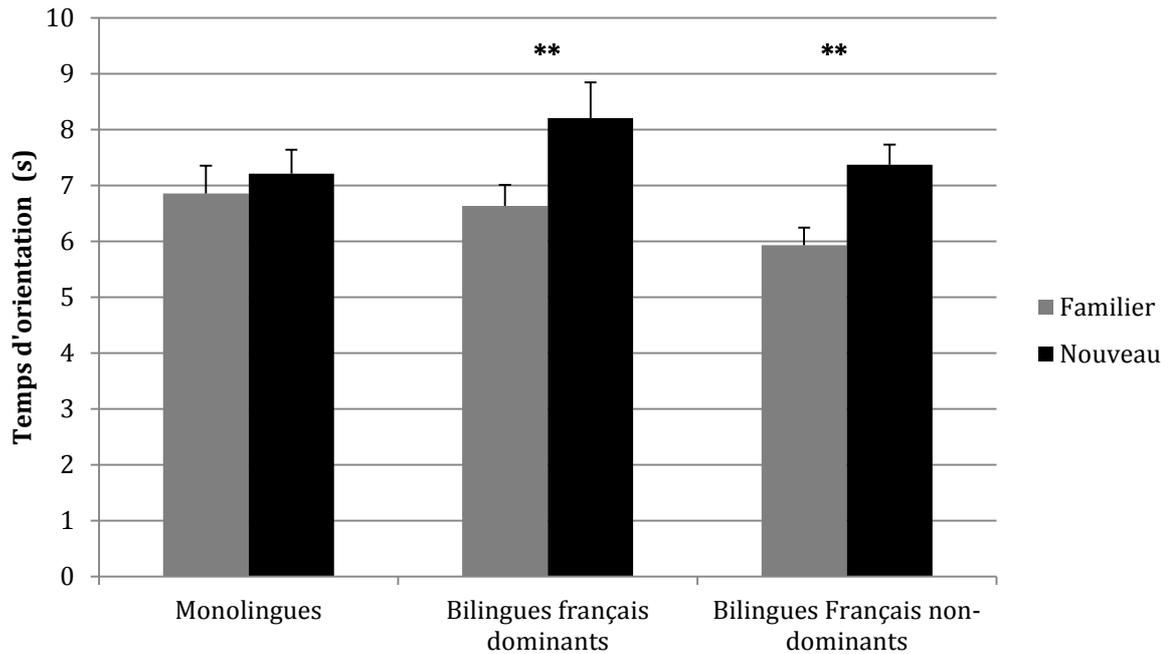


Figure 3.11: Moyenne des temps d'orientation (et les erreurs standard) pour les patterns d'accentuations familiers et nouveaux, séparés par groupe linguistique : français monolingues, français dominant bilingues (60 à 70% de français) et les français non dominant bilingues (30 à 50% de français), ** $p < .01$

3.2.3 Discussion

Cette étude a exploré la sensibilité à un contraste d'accentuation lexicale dans des listes d'items présentant de la variabilité segmentale chez des nourrissons de 10 mois, monolingues français et bilingues apprenant le français et une autre langue à accentuation lexicale variable. Dans ce contexte expérimental complexe, les nourrissons monolingues n'ont pas réussi à discriminer les patterns d'accentuation alors que les bilingues ont réussi.

L'absence de préférence significative chez les monolingues français réplique de nombreux travaux antérieurs, notamment entre 6 et 9 mois de vie (Skoruppa et al., 2009, 2013). Ces résultats confirment ainsi que ces nourrissons monolingues ne discriminent pas les patterns d'accentuation dans ce contexte complexe avec variabilité segmentale. Ces résultats étaient attendus, étant donné que les conditions de test étaient plus difficiles dans notre tâche, utilisant une familiarisation de seulement 1 min (contre 2 minutes pour les études de Skoruppa et al., 2009, 2011, 2013). De plus, plusieurs études (Bijeljac-Babic et al., 2012; Skoruppa et al., 2009) ayant utilisé la même tâche, ont montré que les nourrissons monolingues français ont déjà des difficultés à

discriminer les patterns d'accentuation dans un contexte sans variabilité segmentale avec 1 seule min de familiarisation (alors qu'ils réussissent avec 2 min). Ces résultats diffèrent de ceux des nourrissons monolingues espagnols chez lesquels ces capacités émergent entre 6 et 9 mois. Il apparaît donc que les capacités à traiter les patterns d'accentuation en présence de variabilité segmentale apparaissent sur cette période développementale. Par ailleurs, ces changements développementaux semblent également être spécifiques à la langue maternelle puisqu'ils ne sont retrouvés que chez les nourrissons exposés à des langues utilisant des contrastes d'accentuation.

Les nourrissons bilingues, contrairement aux nourrissons monolingues, apparaissent capables de discriminer les contrastes accentuels en présence de variabilité segmentale. Nos résultats montrent encore une fois une « supériorité » des bilingues français-autre langues à accentuation, à traiter l'information d'accentuation lexicale par rapport aux monolingues français. Cette capacité de discrimination est ici mise en évidence en présentant des stimuli avec de la variabilité segmentale, indiquant donc que les nourrissons bilingues ne perçoivent pas uniquement les corrélats acoustiques des contrastes prosodiques, mais peuvent aussi extraire et généraliser ces patterns d'accentuation. De plus, les nourrissons bilingues montrent ce pattern de discrimination avec seulement 1 min de familiarisation, considérée comme étant une condition expérimentale difficile. Ces résultats s'ajoutent donc aux travaux précédents montrant que des nourrissons bilingues grandissant dans le même type d'environnement linguistique (groupe de bilingues hétérogène : français et autre langue accentuelle) sont capables de discriminer des patterns d'accentuation lorsqu'un seul item répété est présenté (Bijeljac-Babic et al., 2012). Ainsi, notre recherche étend nos connaissances concernant l'impact de l'environnement linguistique, ici des langues à accentuation variable, sur la capacité du traitement de l'information des patterns d'accentuation à un niveau plus abstrait.

Par ailleurs, cette capacité à traiter les patterns d'accentuation est observée dans la quasi-totalité du groupe de bilingues (i.e., la préférence pour les patterns « nouveaux » dans la phase de test est présente chez 22 des 24 enfants testés), alors même qu'aucun de ces nourrissons n'est familier avec l'espagnol (langue utilisée pour enregistrer les stimuli). Ces résultats suggèrent donc que ces nourrissons peuvent percevoir et traiter précisément les indices d'accentuation lexicale, même s'ils ne sont pas prononcés avec les mêmes caractéristiques que dans leur langue maternelle. Ce point renforce notre conclusion précédente concernant la perception de l'accentuation

qui serait traitée à un niveau abstrait (en détectant et généralisant cette information à travers la liste variée de non-mots). Cependant, notre étude ne nous permet pas de savoir comment la distance acoustique entre la façon dont est réalisée l'accentuation lexicale entre l'espagnol et la langue maternelle de chaque enfant affecte leurs performances/préférences. Pour cela, des études avec des groupes homogènes de nourrissons bilingues seront nécessaires.

L'effet de la balance d'exposition entre le français et une langue accentuelle a aussi été exploré dans cette recherche. Les résultats établissent que même 30 % d'exposition à une langue avec une accentuation lexicale variable suffisent pour que les nourrissons bilingues soient capables de discriminer les patterns accentuels de listes constituées de pseudo-mots variés. Ceci suggère que même une faible exposition à une telle langue pendant l'enfance peut avoir un effet significatif sur le développement du traitement du langage. D'autres études corroborent cette idée en montrant que l'exposition à une langue pendant l'enfance impacte fortement les capacités de réapprentissage de cette langue à l'âge adulte (Au, Knightly, Jun, & Oh, 2002). Par ailleurs, aucune différence n'est observée entre les performances des deux sous-groupes de bilingues. Ces résultats diffèrent de ceux obtenus par Bijeljac-Babic et collègues lors de la présentation d'un seul type d'items (/gaba/). En effet, dans cette précédente étude, les bilingues français non-dominants de 10 mois (30 à 40% d'exposition au français) ont montré des capacités de discrimination, alors que le groupe de bilingues plus équilibrés (40 à 60 % d'exposition au français) n'ont montré qu'un effet marginal.

Comment pouvons-nous expliquer les différences d'effet entre ces études ? Une première possibilité pourrait être liée à des répartitions différentes des sous-groupes. Cependant, la différence de moyenne d'exposition dans les 2 sous-groupes de chaque étude est similaire (27% vs. 52% dans Bijeljac-Babic et al., 2012, avec 25% de différence; 42% vs. 68% dans la présente étude avec 26% de différence). Dans l'étude de Bijeljac-Babic (2012), les nourrissons ont eu donc moins d'exposition au français et auraient donc dû être plus avantagés. Une deuxième possibilité pourrait être liée à des différences de salience dans les contrastes d'accentuation utilisés. Les analyses acoustiques des différences de durée, d'intensité et de pitch des stimuli utilisés dans la présente étude (identiques à ceux de Skoruppa et al., 2009) et ceux de Bijeljac-Babic et al., (2012) sont présentés dans le Tableau 3-6. Les syllabes accentuées sont réalisées de manière consistante dans notre étude. En effet, les syllabes accentuées dans les items trochaïques sont marquées par une plus forte intensité et un pitch plus élevé sur la

première syllabe (sans différence de durée entre les deux syllabes). Les items iambiques sont marqués par une intensité plus forte et un pitch plus élevé sur la dernière syllabe, mais également par une durée plus longue. Dans les stimuli de Bijeljac-Babic et al. (2012), les syllabes accentuées apparaissent comme étant moins marquées. En effet, si les syllabes accentuées dans les items trochaïques sont marquées par une plus forte intensité et un pitch plus élevé sur la première syllabe, il y a également une différence de durée mais en faveur de la syllabe finale (non accentuée). De plus, si les syllabes accentuées dans les items iambiques sont marquées par une augmentation de la durée sur la dernière syllabe (accentuée), à l'inverse, l'intensité est plus forte et le pitch est plus élevé sur la syllabe initiale (non accentuée). Donc, les patterns accentuels des stimuli de la présente étude sont clairement plus contrastés. Ceci pourrait alors expliquer que des capacités de discrimination soient observées, même dans le groupe entendant plus de français. Il est à noter cependant que ces différences de salience des patterns accentuels ne sont pas suffisantes pour permettre leur discrimination par des nourrissons monolingues français de 6 et 9 mois ou des nourrissons espagnols de 6 mois (Skoruppa et al., 2009, présente étude).

Une troisième possibilité non exclusive pouvant expliquer la différence entre les résultats de la présente étude et celle de Bijeljac-Babic et collègues est liée à des différences de traitement global entre les enfants bilingues et monolingues. Un grand nombre d'études ont souligné les avantages cognitifs des bilingues par rapport aux monolingues, en particulier concernant les fonctions exécutives (voir pour revues, (Bialystok, Craik, & Luk, 2012; Kroll & Bialystok, 2013). De tels effets peuvent être détectés chez des nourrissons très jeunes, dès 7 ou 12 mois (Kovács & Mehler, 2009a, 2009b). Les différences observées ici entre les enfants monolingues et bilingues pourraient être dues au fait que les nourrissons bilingues soient meilleurs que les monolingues pour discriminer des patterns accentuels dans une situation de variabilité segmentale (i.e., situation complexe requérant un coût cognitif plus important) que dans une situation non variée (i.e, situation moins complexe avec des types d'items répétés). Ces effets potentiels de « facilitation » dans des situations de variabilité sont encore débattus dans le domaine de l'acquisition du langage, mais également dans d'autres domaines comme le traitement des visages ou le traitement audiovisuel (pour une récente revue de la littérature, voir (Watson, Robbins, & Best, 2014). Ce mécanisme de facilitation pourrait jouer un rôle dans le développement cognitif à la fois des nourrissons monolingues et bilingues. Cependant, comme la présence de deux langues

dans l'environnement d'enfants bilingues augmente la variabilité des sons auxquels les enfants sont exposés, cette plus grande variabilité pourrait accentuer l'effet de facilitation en comparaison d'enfants monolingues. Cet effet de facilitation pourrait donc influencer les capacités des nourrissons bilingues à discriminer des patterns d'accentuation en présence de variabilité segmentale. Des études supplémentaires sont nécessaires afin de mieux comprendre ces mécanismes.

En conclusion, les résultats de la présente étude contribuent à mieux comprendre l'acquisition précoce et simultanée de deux langues. Ces nouvelles données démontrent une meilleure sensibilité aux patterns d'accentuation lexicale chez les bilingues apprenant le français et une autre langue à accentuation lexicale variable que chez les monolingues apprenant seulement le français. Celles-ci soulignent donc un effet important des caractéristiques prosodiques de la langue maternelle sur le développement des compétences linguistiques. Les implications de cette augmentation de sensibilité aux patterns d'accentuation à de plus hauts niveaux de traitement (e.g., segmentation de la forme des mots dans un flux continu de parole, acquisition de paires de mots qui diffèrent par leur seul pattern d'accentuation) demandent à être explorées dans de futures recherches.

3.2.4 Points clefs

- L'exposition à une langue avec une accentuation lexicale variable augmente la sensibilité aux patterns d'accentuation.
- Les nourrissons bilingues extraient et généralisent les patterns d'accentuation en présence de variabilité segmentale.
- Ces effets sont trouvés même si le taux d'exposition à la langue accentuelle est de 30%.
- Ces effets ne sont pas présents chez des nourrissons monolingues français.

4 DISCUSSION GENERALE

Ce qui apparaît comme évident pour un adulte, à savoir percevoir et comprendre le signal de parole, est loin de l'être pour un nourrisson et encore moins pour un nouveau-né. Dans le chapitre introductif, nous avons souligné les limites des théories classiques d'apprentissage (Skinner, 1967) ne pouvant pas expliquer à elles seules un apprentissage aussi rapide et aussi spécifique de la langue maternelle. L'alternative proposée par la théorie du *bootstrapping*, dans le cadre des théories innéistes de l'acquisition du langage, et étudiée dans cette thèse, souligne l'importance des indices prosodiques dans l'apprentissage de la langue maternelle (Gleitman & Wanner, 1982). La revue de la littérature sur les régularités présentes dans les langues et les mécanismes de traitement précoce de la parole a révélé l'importance de ces indices dans l'acquisition par les nourrissons du lexique et des règles de grammaire rudimentaires de leur langue maternelle.

Dans le cadre théorique du *prosodic bootstrapping*, nous avons entrepris d'étudier un mécanisme de groupement rythmique qui pourrait guider les nourrissons pour acquérir les propriétés abstraites de leur langue maternelle : la Loi Iambo-Trochaïque (ITL). Cette hypothèse dérive de recherches qui ont montré que l'ITL corrèle avec les proéminences présentes dans les langues au niveau des syntagmes et des mots. Il a été proposé que l'ITL serait un mécanisme universel reposant sur des compétences générales (Bion et al., 2011; Cooper & Meyer, 1960; Hay & Diehl, 2007; Hayes, 1995; Woodrow, 1909) et partagé avec d'autres espèces (de la Mora et al., 2013). Or, des données ont mis en évidence des différences dans la perception du groupement, en fonction de la langue maternelle des locuteurs (Iversen et al., 2008; Yoshida et al., 2010). L'émergence du groupement rythmique fait alors débat. Il a été proposé que l'ITL pourrait aussi reposer sur des mécanismes purement statistiques, étant la conséquence de l'analyse des régularités prosodiques présentes dans la langue en acquisition (Iversen et al., 2008; Yoshida et al., 2010). Très récemment, une alternative à ces hypothèses a émergé, postulant que l'ITL pourrait reposer sur des mécanismes universaux et être modulée par l'environnement linguistique (Bhatara et al., 2013; Yoshida et al., 2010).

Afin d'étudier la perception de l'ITL, nous avons fait le choix d'étudier plusieurs populations : des locuteurs monolingues adultes, experts de leur langue maternelle ; des nourrissons ayant eu seulement 7,5 mois d'exposition à leur langue maternelle ; et enfin, des nouveau-nés ayant eu très peu d'exposition à la langue maternelle (quasi-exclusivement prénatale). Pour comprendre de quelle manière

l'environnement linguistique module l'ITL, nous avons observé ce phénomène chez trois types de groupes linguistiques : des groupes de monolingues exposés soit à l'allemand, soit au français ; des groupes de bilingues exposés au français et à une autre langue. Nos hypothèses principales étaient :

- Les adultes et les nourrissons sont sensibles à l'ITL.
- L'expérience linguistique module la perception de l'ITL.

4.1 Résumé des résultats

4.1.1 L'ITL guide le processus de segmentation chez l'adulte et le nourrisson

Dans l'Expérience 1, nous avons présenté à des adultes une tâche de groupement basée sur la mémorisation de séquences linguistiques. Puis, dans l'Expérience 2, nous avons adapté cette tâche à des nourrissons de 7,5 mois, en utilisant la méthode comportementale de regard préférentiel (HPP, Headturn Preference Procedure), suivant en cela la démarche de Bion et al. (2011) sur l'italien. Premièrement, nous avons présenté à des locuteurs adultes allemands et français des séquences variant en durée, intensité, pitch ou sans aucune modulation prosodique. Les deux groupes linguistiques ont groupé les syllabes par paires iambiques, quand la séquence variait en durée, et ont groupé les syllabes par paires trochaïques quand la séquence variait en intensité. Le groupement basé sur le pitch est apparu comme moins stable, l'effet semblant dépendre de la position de la rampe appliquée en début de séquence. De plus, nous avons mis en évidence de faibles résultats cross-linguistiques, et ce uniquement pour le groupement basé sur l'intensité où les français ont une perception moins stable. L'Expérience 2 chez les nourrissons de 7,5 mois montre une sensibilité au groupement rythmique basée sur la durée et le pitch dans les deux groupes linguistiques. En revanche, ces nourrissons ne semblent pas utiliser l'intensité pour grouper les syllabes de façon stable et les effets dans cette condition sont faiblement modulés par la langue maternelle, comme chez les adultes.

L'ensemble de ces données confirme que les variations de durée, d'intensité ou de pitch semblent guider le groupement par paires d'une séquence de syllabes chez l'adulte et le nourrisson. De façon importante, des différences existent entre les adultes et les nourrissons, notamment pour le pitch et l'intensité, soulignant des changements

développementaux dans la perception de ces indices, dont nous allons discuter dans les parties suivantes.

4.1.2 L'ITL et préférences de pattern rythmique chez le nourrisson

Dans la première partie de cette thèse, nous avons donc montré que les adultes et nourrissons de 7,5 mois possèdent des procédures ou mécanismes similaires et se servent du groupement rythmique pour organiser et mémoriser une séquence continue de sons linguistiques. Dans les Expériences 3-6, nous nous sommes intéressés à l'origine de ces capacités. À l'aide d'une méthode de neuro-imagerie (NIRS), nous avons enregistré les réponses hémodynamiques lors de la présentation à des nouveau-nés de séquences de tons cette fois-ci déjà groupés en paires, suivant soit le pattern iambique, soit le pattern trochaïque, soit sans contraste prosodique. Les nouveau-nés exposés *in utero* exclusivement au français ont traité différemment les patterns iambiques des patterns trochaïques en présence de variations de durée, mais pas en présence de variations d'intensité ou de pitch. Afin de comprendre le rôle de l'expérience linguistique prénatale dans ces résultats, nous avons présenté les mêmes stimuli de la condition pitch à des nourrissons exposés *in utero* au français et à une autre langue possédant des variations de pitch plus importantes au niveau phonologique que le français. Les résultats obtenus montrent que ces nouveau-nés bilingues traitent différemment les patterns iambiques des patterns trochaïques. Prises ensemble, ces données démontrent que l'expérience prénatale a un rôle important dans les sensibilités au groupement rythmique observées dès la naissance.

Finalement, dans l'Expérience 7, nous avons mesuré l'impact de l'expérience linguistique sur le maintien de la sensibilité aux patterns accentuels, notamment l'expérience avec des langues qui utilisent ces contrastes prosodiques au niveau lexical. Nous savons des études précédentes, qu'à la fin de la première année de vie, ces sensibilités déclinent chez les nourrissons exposés au français, du fait de l'absence d'accentuation lexicale dans cette langue. Nous avons donc présenté, à des nourrissons de 10 mois français monolingues ou bilingues français / autre langue à accentuation lexicale variable, des non-mots dissyllabiques accentués sur la première syllabe (trochaïque) ou la deuxième syllabe (iambique) dans un contexte de variabilité segmentale. Nous avons observé que les nourrissons bilingues discriminaient les mots sur la base de leurs patterns accentuels. Ces résultats suggèrent qu'ils ont donc extrait

les patterns d'accentuations présents dans les mots, en détectant et en encodant les variations d'indices acoustiques relatifs à l'accentuation lexicale à un niveau abstrait, par-delà la variabilité phonétique des mots. Ces résultats n'étaient pas modulés par la dominance relative des deux langues entendues par ces nourrissons bilingues. Par contre, les enfants monolingues français échouaient.

Tableau 4.1: Résumé des résultats

Chapitre	Age	Tâche	Durée de la familiarisation	Stimuli	Langue		Effet		
							Durée	Intensité	Pitch
1	Adultes	Segmentation	1mn	Syllabes	Français	monolingues	oui	oui	non
					Allemand	monolingues		non	
2	7-8 mois	Segmentation	3mn	Syllabes	Français	monolingues	oui	non	oui
					Allemand	monolingues			
3	Nouveau-nés	Discrimination		Tons	Français	monolingues	oui	non	non
					Français- autres langues	bilingues			oui
4	10 mois	Discrimination	1mn	Non-mots	Français	monolingues	non		
					Français- autres langues	bilingues français dominant	oui		
						bilingues français non-dominant	oui		

En résumé, les expériences présentées dans nos quatre chapitres expérimentaux ont montré l'impact des propriétés de la langue maternelle sur les compétences de perception des indices acoustiques relatif à la prosodie linguistique (voir Tableau 4.1). De plus, nous avons mis en évidence des réorganisations de ces compétences lors de la première année de vie, en fonction de la nature de l'expérience linguistique reçue.

Si nos résultats apportent à la littérature de nouvelles données sur l'étude des biais prosodiques et de la Loi Iambo-Trochaïque (ITL), des différences sont retrouvées avec les études antérieures. Nous n'avons mis en évidence que de faibles différences cross-linguistiques entre les allemands et les français. De plus, les adultes et les nourrissons ne présentent pas les mêmes compétences pour traiter le groupement basé sur le pitch et l'intensité. A quoi ces effets sont-ils dus? Plusieurs explications sont possibles.

4.2 Les limites dans la comparaison de nos études

Si nos résultats apportent à la littérature de nouvelles données sur l'étude du biais prosodique, des différences sont retrouvées entre nos expériences et les études antérieures. La comparaison de la perception de l'ITL chez l'adulte et le nourrisson laisse de nombreuses questions ouvertes : Avons-nous utilisé une bonne méthode et un matériel adapté ? De plus, de faibles différences cross-linguistiques ont été retrouvées entre les allemands et les français. A quoi sont dus ces effets ? Par ailleurs, les méthodes et la nature des stimuli utilisées entre les Expériences 1 et 2 étaient différentes des Expériences 3 à 6. Quel pourrait être l'impact de ces facteurs sur les résultats obtenus ? Et enfin plus globalement, la perception du groupement dépend à la fois des capacités auditives et de la nature du signal reçu, qui sont loin d'être identiques pour les adultes et les nourrissons ; comment ces deux éléments ont-ils pu jouer sur nos résultats ?

4.2.1 Une bonne méthode pour comparer adultes et nourrissons ?

Dans les Expériences 1 et 2, nous avons voulu utiliser le même matériel afin de pouvoir comparer au mieux les nourrissons et les adultes, mais l'une des limites de cette méthodologie est que les adultes et nourrissons n'ont pas eu la même difficulté de traitement, au vu de la nature des stimuli utilisés. Nous avons probablement rendu la tâche difficile pour les nourrissons et trop facile pour les adultes. Comme nous l'avons rapporté dans le chapitre 2, nous avons réalisé des études pilotes utilisant le même nombre de syllabes pour les adultes et les nourrissons (8 syllabes). Ces études préliminaires ont montré que les nourrissons de 7,5 mois étaient incapables de traiter la séquence linguistique, quel que soit l'indice acoustique variant. Cette absence de résultat suggérait que le matériel devait être trop complexe pour des nourrissons aussi jeunes, notre tâche nécessitant l'encodage des syllabes par paires et la mise en mémoire de cette information afin de se souvenir de la séquence. Bien que des capacités basiques de mémorisation des sons aient été mises en évidence à la naissance (Benavides-varela, Hochmann, Macagno, Nespor, & Mehler, 2012; Decasper & Fifer, 1980; Mampe, Friederici, Christophe, & Wermke, 2009; Mehler et al., 1988; Nazzi, Bertoncini, et al., 1998), la mémoire à court terme reste encore limitée à 7,5 mois (Kuhl, 2004; Nelson, 1995). De plus, nous savons que la complexité acoustique des stimuli impacte la

capacité des nourrissons à traiter les patterns accentuels (voir chapitre expérimental 4), et la plupart des études sur le biais prosodique chez les nourrissons ont utilisé des stimuli avec une faible variabilité segmentale. Donc l'absence de résultats dans les expériences pilotes pourrait être due au trop grand nombre de syllabes à mémoriser. Dans le chapitre 2, nous avons obtenu de meilleurs résultats (en durée et pitch), en présentant des séquences de 6 au lieu de 8 syllabes. Il est possible qu'en réduisant encore plus le nombre de syllabes et, a fortiori, la variabilité segmentale, nous aurions pu obtenir des effets encore plus significatifs.

4.2.2 De faibles différences cross-linguistiques ?

D'autre part, concernant les faibles différences cross-linguistiques, il apparaît qu'en ayant utilisé une tâche assez simple chez l'adulte (du fait d'une faible variabilité segmentale liée à la présentation de seulement 8 syllabes répétées, contre 32 dans Bhatara et al., 2013), les participants français ont pu faire appel à des représentations plus acoustiques (universelles) que linguistiques, obtenant alors des performances similaires aux participants allemands. Cette interprétation s'accorde avec d'autres données de la littérature. Par exemple, une étude a montré que, dans une tâche de répétition immédiate, des participants peuvent imiter des détails phonétiques qui ne sont pas utilisés phonologiquement dans leur langue maternelle (Goldinger et al. 1998). Ceci suggère que des adultes peuvent, dans une tâche simple, faire appel à des compétences générales pour traiter des propriétés/contrastes n'entrant pas dans le répertoire phonologique de leur langue maternelle. De plus, comme nous l'avons souligné dans l'introduction générale, des études chez l'adulte ont révélé que l'augmentation de la complexité acoustique (de stimuli linguistiques comme non-linguistiques) fait émerger des différences cross-linguistiques dans la perception de l'ITL (Bhatara et al, 2013, sous presse). Ces études soulignent ainsi la nécessité d'utiliser une tâche suffisamment complexe lors de l'investigation cross-linguistique de la perception de l'ITL.

4.2.3 Quel niveau prosodique module l'ITL ?

Si des différences cross-linguistiques faibles ont été retrouvées chez l'adulte et le nourrisson pour l'intensité, les nouveau-nés exposés au français (monolingues) et ceux exposés au français et à une autre langue (bilingues) ont eu des réponses très différentes lors de la présentation des séquences variant en pitch. Il apparaît donc que

l'environnement linguistique, ici prénatal, a un impact très fort à la naissance. Ces résultats nous permettent-ils de préciser quelles caractéristiques de cet environnement impactent la perception de l'ITL ? Deux hypothèses ont été proposées : la modulation du biais prosodique pourrait être liée à la forme rythmique, accentuelle des mots (Jakobson, Fant, & Halle, 1952; Kusumoto & Moreton, 1997). Cette première hypothèse soutient le fait que le placement des syllabes accentuées dans les mots ou le poids relatif des syllabes (accentuées et non-accentuées) pourrait avoir un impact sur le groupement rythmique. Nous avons vu que les nourrissons exposés à des langues à accentuation lexicale variable, comme l'anglais ou l'allemand, développent des préférences pour les patterns rythmiques fort-faible vers 6/9 mois (Höhle et al., 2009; Jusczyk et al., 1993), alors que ceux exposés au français, langue n'ayant pas d'accentuation lexicale, ne montrent pas de préférence (Höhle et al., 2009) et perdent de leur sensibilité à l'accentuation lexicale (Abboub et al., 2015; Bijeljic-Babic et al., 2012; Skoruppa et al., 2009). Toutefois, comme ces acquisitions au niveau lexical n'ont jusqu'ici été observées qu'après 6 mois, il est peu probable qu'elles puissent expliquer les effets cross-linguistiques que nous avons obtenu à la naissance.

Une seconde possibilité est que le pattern rythmique des syntagmes pourrait affecter le groupement rythmique (Iversen, 2008 ; Yoshida, 2010). Dans l'Expérience 6, les nouveau-nés bilingues ont été exposés à une deuxième langue, dans laquelle les variations de pitch sont plus systématiques d'un point de vue phonologique qu'en français. Ces propriétés ont pu développer ou maintenir la sensibilité de ces nouveau-nés à des variations de pitch. D'autre part, comme nous l'avons vu dans le chapitre introductif, le français est une langue à tête initiale (VO), avec une proéminence marquée par un allongement de la durée en fin de syntagme (Nespor et al., 2008). Cette expérience avec les variations de durée pour les nouveau-nés français pourrait maintenir ou développer leur sensibilité à des variations de durée et leur permettre d'utiliser ces variations pour détecter les deux types de groupement basés sur la durée (mais pas sur l'intensité ou le pitch). Donc l'expérience avec ces corrélats acoustiques de la structure grammaticale de la langue maternelle pourrait influencer le développement du groupement rythmique. Ceci s'accorde avec les données de la littérature sur la discrimination des langues et la reconnaissance de la langue maternelle à la naissance (Mehler et al., 1988; Moon et al., 1993; Nazzi, Bertoncini, et al., 1998; Ramus et al., 2000), qui semblent reposer sur les propriétés prosodiques globales au niveau des syntagmes/phrases, plutôt que sur le niveau « lexical ».

4.2.4 L'impact du matériel et de la méthode sur les préférences rythmiques

Les différences de matériels et de procédures entre nos études chez les nouveau-nés (Exp 3-6) et celles chez les adultes et nourrissons (Exp 1-2) sont également à prendre en compte dans l'interprétation des données. Premièrement, alors que nous avons utilisé des séquences de syllabes dans les Expériences 1 et 2, les nouveau-nés ont entendu des paires de tons non-linguistiques dans les Expériences 3-6. Afin de comprendre si l'ITL s'applique différemment chez les nourrissons en fonction de la nature des stimuli (linguistiques/non-linguistiques), une seule étude a utilisé ces deux types de matériels (Hay & Saffran, 2012). Aucune différence n'a été trouvée. Cependant, un seul indice (l'intensité) a été testé, à un seul âge (6 mois), et dans un seul groupe linguistique (nourrissons exposés à l'anglais). Cette absence de différence est donc difficile à généraliser. Des études supplémentaires chez le nourrisson et l'adulte avec des séquences non-linguistiques et chez les nouveau-nés avec des stimuli linguistiques seront nécessaires. Deuxièmement, la structure des stimuli était différente entre les Expériences 1-2 et les Expériences 3-6. En effet, les expériences sur les nouveau-nés ont directement étudié l'effort de traitement pour des séquences de tons purs déjà groupées en patterns trochaïques versus iambiques, alors que nos expériences sur les adultes et les nourrissons de 7,5 mois ont étudié indirectement comment ils groupaient une séquence linguistique continue. Des études complémentaires seront nécessaires pour comprendre le lien entre capacités de segmentation et de groupement, et préférences pour différents types d'unités déjà segmentées (paires trochaïques versus iambiques).

4.2.5 Capacité auditive et seuil de discrimination

Pour qu'un groupement soit perçu, il faut que les différences de pitch, d'intensité et de durée que nous avons utilisées soient perceptibles et discriminables. Les seuils de discrimination entre les nourrissons et les adultes ne sont pas les mêmes sur chaque dimension prosodique (pour une revue de la littérature, voir (Saffran, Werker, & Werner, 2006). En effet, comme nous l'avons souligné dans le chapitre introductif, le système auditif est fonctionnel à la naissance, mais il est loin d'être similaire à l'adulte. Concernant les capacités de discrimination fréquentielle (donc de pitch), les nourrissons de 3 mois requièrent une différence de fréquence entre deux sons trois fois plus grande

que les adultes (Aslin, 1989). Pour le traitement de changement d'intensité, des études ont montré que les nouveau-nés/nourrissons détectent des changements d'intensité, mais que leur sensibilité est encore relativement immature, de l'ordre de 6-9dB, comparé à 1-2dB chez l'adulte (Sinnott & Aslin, 1985; Tarquinio, Zelazo, & Weiss, 1990). Enfin, concernant le traitement de la durée, des mesures électrophysiologiques montrent que les nourrissons de 3 mois détectent de courtes pauses temporelles dans des séquences de sons de manière similaire à l'adulte (Trainor et al., 2001).

Il est donc généralement admis que les capacités de discrimination de ces trois indices acoustiques s'améliorent à partir de 6 mois (Mattock, Amitay, & Moore, 2010), avec peut-être moins de changements pour la durée. Donc il est possible que les différences utilisées pour la durée soient perçues de manière équivalente entre le nouveau-né, le nourrisson et l'adulte, mais différemment pour l'intensité et le pitch. Ce qui pourrait ainsi expliquer pourquoi la durée semble être la plus stable à travers nos trois populations, par rapport à l'intensité et au pitch.

4.2.6 Saillance acoustique

Il est également possible que les adultes et nourrissons ne traitent pas de la même manière ces trois indices acoustiques. Les nourrissons pourraient traiter préférentiellement les indices les plus saillants perceptivement et ignorer les autres. Les nourrissons sont souvent confrontés à du parler-bébé (IDS, Infant Directed Speech), dans lequel les différences acoustiques, notamment de pitch, sont plus marquées (Fernald & Kuhl, 1987; Fernald et al., 1989; Shroads, Johnson, Andrews, & Parramore Wilbourn, 2015). D'autre part, des études ont montré que les nourrissons ne s'appuient pas sur les composantes fréquentielles d'un son complexe, comme le font les adultes, dans une tâche psychophysique de discrimination (Bargones, Werner, & Marean, 1995; Bargones & Werner, 1994; Werner & Boike, 2001), suggérant ainsi que leur approche pour traiter ces sons diffère de celle des adultes. Ces différents facteurs pourraient expliquer que, dans nos expériences, les nourrissons de 7,5 mois sont plus sensibles au pitch que les adultes qui, eux, sont plus sensibles à l'intensité. (..) De futures études comparant les nourrissons et adultes devront prendre en compte ces différences de capacités auditives, mnésiques et attentionnelles.

4.3 L'ITL : vers une théorie unifiée

Les mécanismes précis du développement du groupement rythmique sont encore peu compris. Comme nous l'avons vu dans le chapitre introductif, l'origine de l'ITL fait débat : est-elle innée ? Provient-elle de mécanismes d'apprentissage purement statistiques des patterns prosodiques présents dans les langues naturelles ? Ou résulte-t-elle de mécanismes déjà fonctionnels in utero et ensuite modulés par l'expérience linguistique ?

Une des difficultés dans la littérature de l'ITL est la diversité des méthodes utilisées, de la nature des stimuli présentés et des langues maternelles des participants. De plus, aucune étude, jusqu'à celles effectuées dans cette thèse, n'a testé les trois indices en même temps, avec la même méthode et dans des populations exposées à différentes langues. Dans ce qui suit, nous allons donc interpréter nos données indice par indice, en référence aux résultats obtenus dans les études antérieures.

4.3.1 Durée

Nos études ont mis en évidence, chez les adultes comme chez les nourrissons, chez les allemands comme chez les français, la capacité d'exploiter des variations de durée pour grouper les syllabes lors de l'écoute d'une séquence continue. De plus, nous avons trouvé chez les nouveau-nés de familles monolingues francophones la capacité de traiter différemment les groupements iambiques des groupements trochaïques. Nos résultats révèlent donc un groupement ITL basé sur la durée assez stable dans les populations que nous avons étudiées (3 âges x 2 langues maternelles). Ces données rejoignent les résultats trouvés dans des études précédentes chez des adultes (anglais : Iversen et al., 2008 ; Trainor & Adams, 2000 ; Hay & Diehl ; espagnols : de la Mora et al., 2013 ; italiens : Bion, Benavides-Varela, & Nespor, 2011) et chez des nourrissons anglais de 7-8 mois (Trainor & Adams, 2000 ; Yoshida et al., 2010). Par ailleurs, d'autres études ont observé un groupement basé sur la durée chez des nourrissons du même âge (8 mois), mais avec un paradigme de grammaire artificielle couplant des indices fréquentiels avec des indices prosodiques (monolingues français : Bernard & Gervain, 2012 ; bilingues anglais et farsi/coréen/hindi/japonais : Gervain & Werker, 2013). En revanche, des adultes japonais ne montrent pas de préférence claire de groupement basé sur la durée (Iversen et al., 2008), de même que des nourrissons

monolingues italiens et japonais de 7-8 mois (Bion et al., 2011 ; Yoshida et al., 2010) et des rats (de la Mora et al., 2013).

Cette hétérogénéité de résultats a été jusqu'à présent justifiée par l'hypothèse que le groupement basé sur la durée émergerait de l'expérience linguistique, contrairement à l'hypothèse selon laquelle il serait relié à un biais inné, universel. Cependant, au vu de nos résultats chez les nourrissons et les adultes concernant l'effet de la position de la rampe, nous pouvons remettre en question les résultats de Bion et al. (2011) et de de la Mora et al. (2013). Le fait que ces deux études ne retrouvent pas d'effet de groupement, chez les nourrissons de 7,5 mois et les rats, n'est peut-être pas lié à un manque de sensibilité, surtout que l'indice acoustique majeur de l'accentuation lexicale en italien est la durée (Bertinetto, 1980). Il est possible que ces deux populations soient sensibles au groupement de la durée, mais que la position de la rampe de début de séquence, utilisée dans ces deux dernières expériences, était telle, qu'elle pouvait générer un groupement (« par défaut ») en désaccord avec l'ITL. Ainsi, l'effet du groupement de l'ITL et du groupement par défaut ont pu s'annuler, d'où les résultats négatifs obtenus. Il serait intéressant de re-tester les nourrissons italiens et les rats avec l'autre type de rampe que nous avons utilisée, qui serait en accord avec le groupement ITL.

Nos données montrent qu'il est possible que des prédispositions pour la perception du biais basé sur la durée existent dès la naissance, qui se stabilisent dans l'enfance et sont présentes chez l'adulte, sauf si elles vont contre les propriétés de la langue maternelle (comme c'est le cas pour le japonais).

4.3.2 Intensité

Nos résultats ont mis en évidence que les adultes français et allemands peuvent utiliser des variations d'intensité pour grouper et mémoriser des séquences de syllabes. Ce pattern de résultat est en accord avec ce qui a été observé précédemment chez des locuteurs français, allemands, anglais et japonais (Bhatara, Boll-Avetisyan, Agus, Höhle, & Nazzi, sous presse ; Bhatara et al., 2013; Hay & Diehl, 2007; Iversen et al., 2008). En revanche, si certaines études rapportent des effets de groupement basé sur l'intensité à 6 mois, chez des nourrissons monolingues exposés à l'anglais, et à 9 mois, chez des nourrissons bilingues exposés à l'espagnol et au basque (Hay & Saffran, 2012 ; Molnar et al. 2014), nous n'avons retrouvé aucun effet chez nos nourrissons de 7,5

mois, ni chez nos nouveau-nés. Nos résultats vont dans le sens de l'étude de Trainor & Adams (2000) qui ont aussi échoué à obtenir des résultats significatifs de groupement lié à l'intensité. Nos travaux suggèrent que les nourrissons ne sont pas sensibles au groupement basé sur l'intensité et que ces compétences émergeraient plus tardivement (puisqu'on la retrouve chez les adultes). Or, comme nous en avons discuté plus haut, il est également possible que des différences méthodologiques aient pu moduler nos résultats et que ces compétences soit présentes, mais initialement faibles.

D'autre part, nous retrouvons des différences cross-linguistiques pour la perception de cet indice chez l'adulte et le nourrisson de 7,5 mois. Ces résultats sont en accord avec des travaux antérieurs chez l'adulte qui observaient aussi des patterns de résultats similaires dans le jugement des séquences variant en intensité, mais avec des différences de stabilité dans la perception du groupement entre allemands et français, à l'avantage des allemands (Bhatara et al., 2013). De plus, bien que l'intensité ne semble pas être un indice qui leur permette de segmenter la séquence, les nourrissons allemands sont sensibles à la position de la rampe. Elle pourrait agir comme un indice supplémentaire qui pourrait les aider à grouper les syllabes. Ces différences cross-linguistiques retrouvées chez l'adulte et le nourrisson peuvent être expliquées par des différences liées à l'accentuation au niveau du mot, entre l'allemand et le français (Cutler, 2005), et la prédominance de pattern trochaïque dans le lexique allemand. Il a d'ailleurs été montré que ces différences de patterns rythmiques, notamment au niveau du mot, peuvent dès 6 mois moduler cross-linguistiquement les préférences de patterns rythmiques. Les nourrissons allemands préfèrent écouter des listes de mots qui suivent un pattern trochaïque, alors que ce n'est pas le cas des nourrissons français (Höhle et al., 2009). Ces travaux peuvent donc expliquer pourquoi nous retrouvons peut-être ici une indication d'une plus grande sensibilité à l'intensité chez les nourrissons allemands que français.

En résumé, nos études ne mettent pas en évidence de préférence claire de groupement basé sur l'intensité, ni à la naissance, ni à 7,5 mois, mais des sensibilités liées à la position de la rampe semblent être présentes chez les nourrissons allemands. En effet, un indice supplémentaire (le groupement par défaut) semble être requis pour que les nourrissons allemands de 7,5 mois puissent grouper les syllabes. Cette sensibilité pourrait être liée à leur exposition à des variations d'intensité, très présentes dans la prosodie de l'allemand, contrairement au français qui utilise très peu l'intensité comme marqueur prosodique. Des prédispositions pour traiter le groupement basé sur

l'intensité ne sont donc pas à exclure. Ainsi, la perception du groupement rythmique à 7,5 mois n'est peut-être pas encore très stable, mais suite à l'exposition linguistique, elle pourrait se stabiliser et se renforcer ensuite à l'âge adulte, parallèlement à l'apprentissage de la langue maternelle. Au vu des différences du rôle de l'intensité entre la prosodie de l'allemand et du français, il est possible que ces stratégies de groupement et ces sensibilités émergeraient d'abord chez les allemands, puis plus tardivement chez les français.

4.3.3 Pitch

Nos expériences mettent en évidence des patterns de résultats assez hétérogènes pour le pitch dans nos trois populations. Premièrement, chez les adultes, des réponses inconsistantes ont été retrouvées, le groupement semblant dépendre de la position de la rampe. Des études précédentes chez l'adulte ont trouvé des patterns de réponses similaires aux nôtres chez des locuteurs anglais et japonais, ayant conduit leurs auteurs à postuler que le pitch ne serait pas un indice aussi pertinent pour le groupement trochaïque que l'intensité (Kusumoto & Moreton, 1997; Rice, 1992; Woodrow, 1911). Or, plus récemment, d'autres études ont montré, qu'à l'inverse, le pitch est aussi pertinent que l'intensité pour percevoir le groupement trochaïque (Bhatara et al., 2013; Bion et al., 2011). Deuxièmement, chez les nourrissons allemands et français de 7,5 mois, nous avons trouvé une utilisation du pitch pour grouper les syllabes, confirmant des travaux antérieurs chez les nourrissons monolingues italiens (Bion et al., 2011). Par ailleurs, une étude utilisant un paradigme de grammaire artificielle, couplant les indices prosodiques aux indices fréquentiels, a également mis en évidence un effet chez les bilingues exposés à l'anglais et au farsi/coréen/hindi/japonais (Gervain & Werker, 2013). Troisièmement, à la naissance, les nouveau-nés bilingues exposés au français et à une autre langue traitent différemment les patterns iambiques des patterns trochaïques si le pitch varie, ce qui n'est pas le cas des nouveau-nés exposés exclusivement au français. Ces résultats montrent donc que la sensibilité ITL basée sur le pitch semble dépendre de l'exposition linguistique entendue dans le dernier trimestre de grossesse, puisque les langues entendues par les nouveau-nés bilingues étaient des langues où le pitch possédait un rôle de marquage de la proéminence plus important au niveau phonologique que dans la prosodie du français.

En résumé, l'ensemble de nos données met en évidence, pour la première fois, certains phénomènes de perception de l'ITL à la naissance et chez des nourrissons de 7,5 mois français et allemands. Nous retrouvons également certaines différences cross-linguistiques en fonction de la nature du signal linguistique entendu. De plus, à travers les trois populations, certains indices acoustiques semblent être plus stables que d'autres dans la perception du groupement. La durée semble être très stable alors que l'intensité révèle des résultats plus faibles. Ceci est consistant avec Hayes (1985), qui observait déjà dans ces travaux que les effets du groupement trochaïque étaient plus faibles que les effets de groupement iambique, suggérant un rôle plus important du groupement iambique dans les stratégies de segmentation. D'ailleurs, de nombreuses recherches portant sur les processus de segmentation ont montré que la durée serait le corrélât acoustique le plus universel de l'accentuation lexicale dans les positions non-accentuées (accentuation lexicale sans proéminence de syntagme ; (Langus et al., 2012; Tyler & Cutler, 2009). Les voyelles qui portent l'accentuation lexicale sont souvent plus longues que les voyelles non-accentuées dans beaucoup de langues, notamment l'allemand (Dogil & Williams, 1999) et le français (Delattre, 1966). Ceci pourrait en partie expliquer pourquoi la durée est l'indice le plus stable dans nos trois études. Mais de nombreuses questions restent en suspens, ce dont nous discutons dans ce qui suit.

4.4 Mécanismes sous tendant l'ITL

4.4.1 Un héritage évolutif : contraintes perceptives

Le traitement du langage repose sur des processus évolutifs (voir Werker & Gervain, 2013). Des recherches ont montré que le système perceptif humain partage de nombreux points communs avec celui d'autres espèces. Plusieurs études ont montré des sensibilités prosodiques chez le tamarin (Ramus et al., 2000) ou encore le rat (Toro et al., 2003). D'autres études comportementales ont montré que les singes (Koda and Masataka, 2002, Kojima et al., 2003 ; Coye, Ouattara, Kuberbuhler & Lemasson, 2015), les grenouilles (Capranica, 1966) et les oiseaux (Nelson, 1989) utilisent aussi des indices rythmiques pour interpréter les vocalisations spécifiques de leur espèce. Reconnaître des variations de durée, pitch et intensité semble donc être une caractéristique fondamentale dans la perception d'une large variété de vocalisations animales. Or, bien que ces études chez des animaux révèlent des similarités cross-

espèces de fonctionnement des systèmes perceptifs, ici prosodiques, ces animaux ont eu besoin d'un entraînement leur permettant de discriminer ces patterns auditifs, où un temps plus long est requis par rapport à leurs homologues humains. Ces études soulignent des différences possibles dans la manière de traiter les sons entre les humains et d'autres espèces. Toutefois, elles permettent de mettre en évidence des prédispositions certaines pour apprendre à traiter ces variations prosodiques contenues dans la parole. D'ailleurs, certaines recherches ont révélé des capacités communes pour des compétences qui ne sont pas dans le répertoire naturel de ces animaux. En effet, des études ont rapporté des sensibilités aux relations d'identité (ABB versus AAB) chez les tamarins (Hauser, Marcus, & Weiss, 2002), des capacités d'apprentissage de catégories de consonnes chez les cailles (Kluender, Diehl, & Killeen, 1987), ou encore la capacité de calculer les probabilités de transition entre syllabes contenues dans des séquences de parole (Pons et al., 2006). Ces travaux mettent donc en évidence des compétences profondément ancrées d'un point de vue évolutif, soulignant de nombreux points communs entre les espèces.

A notre connaissance, une seule recherche a étudié la perception du groupement rythmique chez l'animal. Elle montre que les rats sont sensibles au groupement rythmique (de la Mora et al., 2013). Ce mécanisme auditif de groupement pourrait faire partie des contraintes perceptives dont disposerait les humains pour organiser leur environnement auditif. Des études comparatives supplémentaires seront nécessaires pour comprendre les origines de ce phénomène.

Ainsi, les mécanismes sous tendant l'apprentissage du langage, dont l'ITL, pourraient provenir de contraintes du système perceptif. Ce mécanisme de groupement rythmique, qui traite la structure du signal acoustique et l'organise selon un certain ordre, semble être modulé par la langue maternelle, mais comment ?

4.4.2 Des régularités prosodiques

L'apprentissage statistique des régularités prosodiques présentes dans la (les) langue(s) apprise(s) pourrait ainsi moduler ce mécanisme de groupement. D'ailleurs, des études ont montré que des nourrissons sont capables de reproduire certains sons possédant les propriétés prosodiques spécifiques de leur langue maternelle et ce dès la naissance. Des travaux ont notamment mis en évidence qu'une montée des contours de pitch et un allongement sont présents au niveau de la dernière syllabe dans les

productions bi-syllabiques des nourrissons français de 7 à 18 mois (Hallé, de Boysson-Bardies, & Vihman, 1991, Levitt & Wang, 1991). Ceci pourrait refléter le fait que, comme nous l'avons vu dans le chapitre introductif, l'intonation en français est caractérisée par une augmentation de pitch au niveau de la fin de certaines unités prosodiques (mots, syntagme), sauf dans la dernière unité de la phrase qui présente un contour descendant. Par contre, en allemand, langue trochaïque, l'intonation est caractérisée par un contour de pitch descendant et un allongement sur la dernière unité (Wiese, 1996) ; et chez les nourrissons allemands de 6 à 12 mois, on a observé dans leur production une tendance à répliquer un pattern trochaïque (Fischer, 2009). Ces mêmes patterns de spécificité précoce ont aussi été retrouvés dans les cris de nouveau-nés allemands et français (Mampe et al., 2009). De plus, nous avons discuté à plusieurs reprises du fait que les nourrissons ont des capacités perceptives précoces reflétant aussi les caractéristiques de la prosodie de la langue maternelle (Bijeljac-Babic et al., 2012; Höhle et al., 2009; Skoruppa et al., 2009). Ces études montrent que la production précoce reflète les régularités prosodiques retrouvées dans la langue maternelle des nourrissons, probablement du fait d'un apprentissage pré- et post-natal. Sur quels mécanismes reposent ces compétences ?

Des études ont montré que les nourrissons et les rongeurs sont sensibles aux indices statistiques présents dans les stimuli qui leur sont présentés, comme des probabilités simples de distribution et des probabilités plus complexes comme les probabilités de transitions (Maye, Werker, & Gerken, 2002; Pons, 2006; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Toro & Trobalón, 2005). Des études en électrophysiologie chez l'animal suggèrent que le cortex auditif en développement peut utiliser ces indices statistiques pour construire des représentations acoustiques, en montrant par exemple que des neurones se synchronisent à des sons présentés de manière répétitive (probabilité simple de distribution ; (Bao, Chang, Teng, Heiser, & Merzenich, 2013; Chang & Merzenich, 2003; de Villers-Sidani, Simpson, Lu, Lin, & Merzenich, 2008; Zhou, Nagarajan, Mossop, & Merzenich, 2008). Une récente étude indique que le cortex auditif en développement peut aussi encoder des probabilités de transition de haut niveau présentes dans des séquences (Kover, Gill, Tseng, & Bao, 2013). Certes, ces dernières recherches ont été effectuées chez l'animal, mais il est très probable que les propriétés de leur système perceptif soient partagées avec l'espèce humaine, et que les nourrissons disposent de mécanismes spécialisés qui

leur permettraient d'exploiter ces régularités prosodiques pour commencer à construire des représentations pré-linguistiques.

Conclusions :

Il est possible que le groupement perceptif rythmique (formalisé sous la notion d'ITL) fasse partie des contraintes computationnelles de notre cerveau, dont l'humain a hérité à travers l'évolution. Les régularités prosodiques présentes dans le signal acoustique, qui pourraient être extraites grâce aux mécanismes qui traitent les indices statistiques, moduleraient l'ITL. Mais sur quels réseaux cérébraux ces processus reposent-ils ?

4.5 Bases cérébrales de l'ITL

4.5.1 Partage de compétence

Pour la première fois, dans cette thèse, nous nous sommes intéressés aux bases cérébrales de l'ITL, en utilisant la NIRS pour tester des nouveau-nés. Comme nous l'avons vu dans le chapitre expérimental 3, à la naissance, le cerveau est capable de traiter un signal aussi complexe que la parole. L'organisation anatomique qui soutient le traitement du langage est en place dans le cerveau humain dès 29 semaines de gestation, avant même d'avoir une expérience d'écoute spécifique (Dehaene-Lambertz et al., 2008). Par ailleurs, ces zones du langage dans la scissure sylvienne semblent être actives et fonctionnelles dès 28-32 semaines de gestation, comme le suggère une récente étude rapportant l'activation des aires classiques du langage chez l'enfant prématuré lors de la discrimination de phonèmes (Mahmoudzadeh et al., 2013). De plus, des projections ventrales, partant des aires temporales vers les aires préfrontales, sont présentes à la naissance, mais également des connexions dorsales, partant du cortex auditif et se terminant dans le cortex pré-moteur (Brauer et al. 2013). Dans nos études NIRS, nous avons présenté des tons purs non-linguistiques et nous avons retrouvé des activations dans les aires temporales, avec des activations différentielles pour les séquences trochaïques versus iambiques, préférentiellement retrouvées dans l'hémisphère gauche. De nombreuses études en imagerie fonctionnelle utilisant la NIRS chez des nouveau-nés ont aussi retrouvé dans l'hémisphère gauche des activations différentielles, mais pour de la parole continue versus inversée (Pena et al. 2003), ou

encore pour des mots suivant un pattern structurel ABB versus ABC, c'est-à-dire avec ou sans répétition (Gervain et al., 2012; Gervain, Macagno, et al., 2008). Or, le traitement de la prosodie linguistique chez les nourrissons et l'adulte a été mis en évidence dans l'hémisphère droit (Friederici, 2012; Homae et al., 2006; Wartenburger et al., 2007). Toutefois, comme nous l'avons mentionné dans le chapitre expérimental 3, les nouveau-nés dans nos expériences devaient détecter la conformité de la structure prosodique en accord avec l'ITL. Le traitement de ces stimuli pouvait donc impliquer l'hémisphère gauche, préférentiellement recruté lors de traitement structurel (Gervain et al., 2008), ce que semblent confirmer nos données.

De nombreuses théories de « partage de compétences » entre la perception musicale et langagière ont été émises. Des études en comportement (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009) soulignent notamment des similarités dans le traitement structurel de stimuli musicaux et langagiers. De plus, des études en neuro-imagerie trouvent que ce type de traitement repose sur des réseaux communs (Fedorenko & Thompson-Schill, 2014). Un modèle a été proposé par Patel (2003) : SSIRH pour *Shared Syntactic Integration Resource Hypothesis* (voir Figure 4.1). Son hypothèse repose sur des études montrant que des réseaux complexes dans les aires inféro-frontales et temporales sont recrutés pour traiter la structure des sons langagiers et musicaux. Des régions intègreraient donc ces informations acoustiques, qu'elles soient présentes dans des stimuli de nature linguistique ou non linguistique. Ces études ont d'ailleurs montré, plus précisément, que les aires frontales (IFG orbitalis) seraient plus impliquées lors du traitement rythmique chez les sujets musiciens (Vuust et al. 2006, 2011). Toutefois, des études neuropsychologiques observent que des patients amusicaux ne sont pas nécessairement aphasiques ; il est donc probable que ces réseaux codent la structure, mais qu'ils ne soient pas spécifiquement responsables des compétences langagières ou musicales (Peretz, 1993; Peretz et al., 1994).

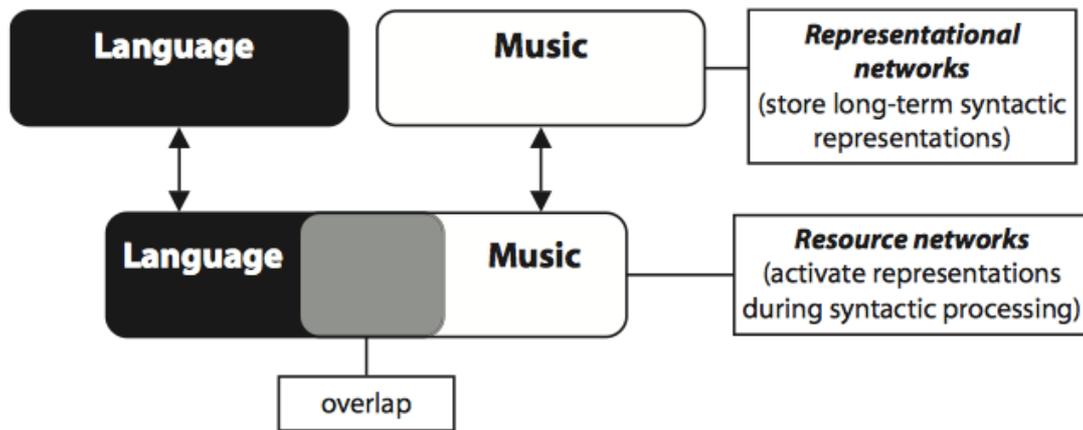


Figure 4.1 : Diagramme schématisant la relation fonctionnelle entre le traitement linguistique et le traitement syntaxique musical.

Selon Patel (2003), il serait possible que le traitement du groupement prosodique repose à la naissance sur des réseaux communs utilisés pour traiter la structure, qu'elle soit musicale ou langagière. Mais comment se développent ces compétences dans la première année de vie ? Dans nos expériences, les nouveau-nés, testés 2 jours en moyenne après la naissance, ont eu une exposition à leur langue maternelle quasi-exclusivement prénatale ; il est possible, voire probable, que les réponses préférentielles que nous avons observées dans l'hémisphère gauche se renforcent dans la première année de vie, suite à l'exposition à leur langue maternelle. En effet, de nombreuses études indiquent des réorganisations cérébrales à la suite de l'expérience langagière (Perani et al. 2011, Sato, et al. 2010). Des études sur les bases cérébrales de l'ITL chez des nourrissons plus âgés seront nécessaires pour comprendre le développement du groupement prosodique.

4.5.2 Les compétences des nouveau-nés bilingues : le rôle de l'expérience prénatale

Nous avons retrouvé de très fortes activations chez les nouveau-nés bilingues pour les séquences variant en pitch, alors que les monolingues ne semblaient pas traiter différemment les séquences. Nous avons expliqué ces différences par la nature de l'exposition linguistique. Mais est-ce que ces différences entre les deux groupes linguistiques reflètent une exposition prénatale spécifique au pitch, ou résultent-elles du fait que des différences puissent déjà exister à la naissance dans la manière dont bilingues et monolingues traitent les sons ? Des études en neuro-imagerie ont indiqué

que les bases cérébrales associées à la discrimination phonétique sont moins matures chez des nourrissons bilingues de 6-9 mois, par rapport à des nourrissons monolingues du même âge (Garcia-Sierra et al., 2011). D'autres études sur cette question rapportent une maturité équivalente dans les deux populations, mais suggèrent que le traitement phonétique reposerait sur un réseau différent, avec une connectivité plus importante dans les aires frontales chez les bilingues (Jasinska & Petitto, 2013; Petitto et al., 2012). Par ailleurs, des études en comportement suggèrent que la discrimination phonétique est plus efficace chez les nourrissons bilingues (Conboy, Sommerville, & Kuhl, 2008), peut-être en lien avec des capacités attentionnelles et exécutives plus importantes chez les nourrissons bilingues (Kovács & Mehler, 2009a). Évidemment, nous avons testé des nouveau-nés qui sont loin d'avoir reçu une exposition à leurs langues maternelles aussi riche que des nourrissons de 6-9 mois, mais nous pouvons nous demander si, à un niveau cérébral, la maturation des réseaux langagiers serait équivalente.

Nous savons, sur la base d'études antérieures, que l'expérience langagière prénatale a un impact important sur les préférences retrouvées à la naissance: pour la voix maternelle (Decasper & Fifer, 1980), et pour des histoires ou encore des chansons entendues pendant la grossesse (Decasper & Spence, 1986). Une récente étude en EEG chez des nouveau-nés a établi que des traces neurales de cet apprentissage prénatal sont présentes à la naissance, et leurs résultats montrent également une corrélation positive entre le pourcentage d'exposition prénatale et l'activité cérébrale enregistrée (Partanen, Kujala, Naatanen, et al., 2013). Par ailleurs, cet effet d'apprentissage pouvait être généralisé à d'autres types de stimuli linguistiques auxquels les nouveau-nés n'avaient pas été exposés pendant la grossesse, suggérant donc que l'exposition prénatale pourrait avoir des effets positifs sur les compétences langagières en général. D'ailleurs, une récente étude sur de très grands prématurés, ayant reçu un enrichissement auditif sous la forme d'enregistrement audio de sons maternels (battements cardiaques, voix), a mis en évidence des différences en termes d'épaisseur dans le cortex auditif, comparé à ceux n'ayant pas reçu cet enrichissement (Webb, Heller, Benson, & Lahav, 2015). Cette étude suggère que ce signal acoustique (physiologique et linguistique) a des effets bénéfiques sur le développement des structures cérébrales. Malheureusement, étant donné que les auteurs ont présenté ces sons de parole mélangés à des bruits physiologiques, il est difficile de dé-corréler les deux effets possibles, mais cela reste une étude importante montrant l'impact de l'environnement sonore sur la plasticité cérébrale chez de très jeunes nourrissons. Notons que le groupe contrôle n'a pas reçu

d'enrichissement et a été exposé aux bruits quotidiens des salles de soins à l'hôpital. Ce contexte est à prendre en compte, étant donné des études montrant des effets négatifs de cet environnement sonore sur le développement auditif des nouveau-nés (Lahav & Skoe, 2014; Rand & Lahav, 2014).

Nous pouvons ainsi spéculer que cette exposition plus complexe et plus riche en termes de variabilité acoustique pourrait jouer sur la plasticité cérébrale des nouveau-nés bilingues et donc avoir un impact sur leur capacité à discriminer les patterns iambiques des patterns trochaïques pour la condition pitch. Des études testant les deux autres indices sont également requises pour comprendre plus précisément l'impact de l'exposition à deux langues maternelles *in utero*.

4.6 Perspectives

4.6.1 Interaction des indices

La durée, le pitch et l'intensité sont des corrélats de l'accentuation dans les langues. Dans cette thèse, nous les avons étudiés séparément alors que, dans les langues naturelles, les indices peuvent marquer l'accentuation ensemble ou séparément, ce qui pourrait avoir un impact sur la perception. Des études ont souligné ce point en montrant que la perception du groupement était d'autant plus stable que le signal contient des indices en co-occurrence (Crowhurst & Teodocio Olivares, 2014). De plus, nous avons également vu que, lors de la présentation de séquences variant en durée mélangées à des séquences variant en pitch, la perception de la durée était moins stable que lors de la présentation de séquences variant en durée mélangées à des séquences variant en intensité (Bhatara et al. 2013). De futures recherches seront nécessaires pour tester les combinaisons de ces trois indices, en prenant en compte la nature acoustique de la réalisation de la proéminence dans les mots et les syntagmes, en fonction de la langue maternelle du locuteur.

4.6.2 Les valeurs de différences

Dans cette thèse, nous nous sommes basés sur des valeurs de différences de durée, intensité et pitch communément utilisées lors de l'étude de la perception de l'ITL, reposant sur les valeurs qui élicitaient chez l'adulte des performances élevées (Bhatara et al., 2013; Hay & Diehl, 2007). Or, comme nous l'avons souligné plus haut,

les seuils de discrimination ne sont pas les mêmes entre les adultes et les nourrissons. D'autre part, nous avons vu que le développement de l'audition semble se stabiliser à partir de 6 mois. Il est donc possible que des différences existent, même entre les nouveau-nés et les nourrissons. Auquel cas, les valeurs de différences utilisées ne sont peut-être pas aussi bien adaptées pour faire émerger la sensibilité au groupement à tous les âges testés. De plus, il est difficile de déterminer, au travers des trois indices, des équivalences sur la magnitude des différences utilisées, compliquant le travail de comparaison des résultats obtenus pour les différents indices. Des études complémentaires testant différentes valeurs seront nécessaires.

4.6.3 Le choix des langues dans les études cross-linguistiques

Nous avons fait le choix de comparer les allemands aux français pour de nombreuses raisons que nous avons évoquées dans la partie introductive. Cependant, bien que le français et l'allemand diffèrent au niveau du mot, ils partagent un système prosodique commun, étant donné que l'allemand est une langue mixte au niveau des syntagmes, qui peut suivre l'ordre OV ou VO, mais avec une prépondérance VO, le pattern utilisé en français (Nespor et al., 2008). Ce pattern commun d'accentuation pourrait limiter l'apparition de différences cross-linguistiques dans notre étude et ne nous permet pas totalement de comprendre à quel niveau prosodique s'applique la modulation de l'ITL. Ainsi, d'autres études cross-linguistiques, incluant des langues plus éloignées comme le japonais, des langues tonales (ne possédant pas d'accentuation lexicale, mais des tons lexicaux) ou des langues comme le farsi ou le turc (ayant au niveau du syntagme un pattern complètement opposé OV) pourraient nous éclairer un peu plus sur le niveau prosodique (mot ou syntagme) auquel s'applique la modulation de la loi ITL.

4.6.4 Privation sensorielle précoce : surdit  et implant cochl aire

Nous avons  tudi  la perception de l'ITL dans un cadre typique chez des enfants recevant un ou deux types d'input linguistique. Or, qu'en serait-il du cas inverse, o  le signal linguistique est appauvri ou quasi-inexistant ? Afin de d terminer le r el impact de l'environnement linguistique sur le d veloppement de l'ITL, tester ce type de population serait d'un grand support. Les enfants n s avec une surdit  cong nitale n'ont

pas eu accès au signal acoustique du langage et pourrait nous aider à trancher sur la question de l'émergence de l'ITL. Nécessite-t-elle un input linguistique pour émerger ?

De plus, on observe une grande variabilité entre les individus sourds ayant une aide auditive et/ou des implants cochléaires ; certains d'entre eux étant capables de développer une perception du langage parlé et des capacités de production proche des normo-entendant, particulièrement si ces aides ont été appliquées très tôt dans le développement (pour une revue, voir Lyness et al., 2013), alors que d'autres font preuve de grandes difficultés. Des études chez des enfants n'ayant pas été appareillés au même âge et capables de percevoir le signal de parole pourraient nous renseigner sur les fenêtres temporelles critiques de l'émergence de l'ITL. Ces questions restent ouvertes.

4.7 Conclusions

D'énormes progrès ont été faits pour comprendre les fondations de l'acquisition du langage chez l'humain. Nous avons vu à travers cette thèse que l'acquisition du langage commence bien avant la production des premiers mots. Nous avons montré que des biais perceptifs prosodiques sont présents dès la naissance et sont modulés par l'environnement linguistique. Nous avons également montré que des sensibilités perceptives prosodiques sont communes entre des nourrissons grandissant dans des milieux linguistiques différents, soulignant donc la présence de capacités perceptives générales. En revanche, nous avons relevé de nombreuses controverses concernant la manière dont ces biais émergent et interagissent avec la nature du signal linguistique. Ce travail soutient l'idée que le développement de ces biais reflèterait une sorte de processus « épigénétique », avec des facteurs endogènes (perception initiale du groupement) et exogènes (régularités présentes dans le signal acoustique) qui interagissent ensemble lors du développement de ces capacités.

REFERENCES

- Abboub, N., Bijeljic-Babic, R., Serres, J., & Nazzi, T. (2015). On the importance of being bilingual : Word stress processing in a context of segmental variability. *Journal of Experimental Child Psychology*, *132*, 111–120.
- Allen, G. D., & Hawkins, S. (1978). The development of phonological rhythm. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 173–185). Amsterdam: North-Holland.
- Angoujard, J. P. (1986). Les hiérarchies prosodiques en arabe. *Revue Québécoise de Linguistique*, *16*(1), 11.
- Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 117–134).
- Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of “segmental anchoring” of F0: evidence from German. *Journal of Phonetics*, *32*(2), 177–197.
- Au, T. K., Knightly, L. M., Jun, S. ., & Oh, J. S. (2002). Overhearing a Language During Childhood. *Psychological Science*, *13*(3), 238–243. doi:10.1111/1467-9280.00444
- Bao, S., Chang, E. F., Teng, C.-L., Heiser, M. a, & Merzenich, M. M. (2013). Emergent categorical representation of natural, complex sounds resulting from the early post-natal sound environment. *Neuroscience*, *248C*, 30–42. doi:10.1016/j.neuroscience.2013.05.056
- Bargones, J. Y., & Werner, L. A. (1994). Adults listen selectively: Infants do not. *Psychological Science*, *5*(3), 170–174.
- Bargones, J. Y., Werner, L. A., & Marean, G. C. (1995). Infant psychometric functions for detection: mechanisms of immature sensitivity. *The Journal of the Acoustical Society of America*, *98*(1), 99–111. doi:10.1121/1.414446
- Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonation structure in Japanese and English. *Phonology*, *3*(01), 255–309.
- Benavides-varela, S., Hochmann, J., Macagno, F., Nespors, M., & Mehler, J. (2012). Newborn ’s brain activity signals the origin of word memories. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), 17908–17913. doi:10.1073/pnas.1205413109/
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.
- Bergelson, E., & Swingley, D. (2012). At 6-9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. doi:10.1073/pnas.1113380109
- Bernard, C., & Gervain, J. (2012). Prosodic cues to word order: what level of representation? *Frontiers in Psychology*, *3*(451), 1–6. doi:10.3389/fpsyg.2012.00451
- Bertinetto, P. M. (1980). The perception of stress by Italian speakers. *Journal of Phonetics*, *8*, 385–395.
- Bhatara, A., Boll-Avetisyan, N., Agus, T., Höhle, B., & Nazzi, T. (sous presse). Language experience affects grouping of musical instrument sounds. *Psychological Science*.
- Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T., & Höhle, B. (2013). Native language affects rhythmic grouping of speech. *The Journal of the Acoustical Society of America*, *134*(5), 3828–3843.

- Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240–50.
doi:10.1016/j.tics.2012.03.001
- Bijeljac-Babic, R., Serres, J., Höhle, B., & Nazzi, T. (2012). Effect of bilingualism on lexical stress pattern discrimination in French-learning infants. *PloS One*, 7(2), e30843.
- Bion, R. A. H., Benavides-varela, S., & Nespor, M. (2011). Acoustic markers of prominence influence infants' and adults' segmentation of speech sequences. *Language and Speech*, 54(1), 123–140.
- Bion, R. A. H., Höhle, B., & Schmitz, M. (2007). The role of prosody on the perception of word-order differences by 14-month-old German infants. In J. Barry & W. J. Trouvain (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1537–1540). Saarbrücken: Germany: International Congress of Phonetic Sciences.
- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Boersma, P., & Weenink, D. (2010). PRAAT: Doing phonetics by computer. <http://www.praat.org>.
- Bolton, T. L. (1894). Rhythm. *The American Journal of Psychology*, 6(2), 145–238.
- Bortfeld, H., Fava, E., & Boas, D. (2009). Identifying cortical lateralization of speech processing in infants using near-infrared spectroscopy. *Developmental Neuropsychology*, 34(1), 52–65. doi:10.1080/87565640802564481
- Bosch, L., Figueras, M., Teixidó, M., & Ramon-Casas, M. (2013). Rapid gains in segmenting fluent speech when words match the rhythmic unit: evidence from infants acquiring syllable-timed languages. *Frontiers in Psychology*, 4(106), 1–12. doi:10.3389/fpsyg.2013.00106
- 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.
- Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81(2), 33–44. doi:10.1016/S0010-0277(01)00122-6
- Butler, J., Floccia, C., Goslin, J., & Panneton, R. (2011). Infants' discrimination of familiar and unfamiliar accents in speech. *Infancy*, 16(4), 392–417. doi:10.1111/j.1532-7078.2010.00050.x
- Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*, 21(3), 343–348.
- Chaker, S. (1995). Données exploratoires en prosodie berbère: I. L'accent en kabyle. *Comptes Rendus Du GLECS*, 31(I), 27–54 & p. 55–82.
- Chance, B., Zhuang, Z., UnAh, C., Alter, C., & Lipton, L. (1993). Cognition-activated low-frequency modulation of light absorption in human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 90(8), 3770–4.
- Chang, E. F., & Merzenich, M. M. (2003). Environmental noise retards auditory cortical development. *Science*, 300(5618), 498–502. doi:10.1126/science.1082163
- Cheour-Luhtanen, M., Alho, K., Kujala, T., Sainio, K., Reinikainen, K., Renlund, M., ... Näätänen, R. (1995). Mismatch negativity indicates vowel discrimination in newborns. *Hearing Research*, 82(1), 53–58.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT.
- Chomsky, N. (1975). *Reflections on Language*. New York, 3.

- Christophe, A., & Dupoux, E. (1996). Bootstrapping lexical acquisition: The role of prosodic structure. *The Linguistic Review*, *13*, 383–412.
- 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 the Acoustical Society of America*, *(95)*, 1570–1580.
- Christophe, A., Gout, A., Peperkamp, S., & Morgan, J. L. (2003). Discovering words in the continuous speech stream: the role of prosody. *Journal of Phonetics*, *31*(3-4), 585–598.
- Christophe, A., Guasti, T., Nespors, M., Dupoux, E., & Van Ooyen, B. (1997). Reflections on phonological bootstrapping: Its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, *12*(5-6), 585–612. doi:10.1080/016909697386637
- Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, *2*(3), 385–394. doi:10.1207/S15327078IN0203_6
- Christophe, A., & Morton, J. (1998). Is Dutch native English? Linguistic analysis by 2-month-olds. *Developmental Science*, *1*(2), 215–219. doi:10.1111/1467-7687.00033
- Christophe, A., Nespors, M., Teresa Guasti, M., & Van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: The case of the head-direction parameter. *Developmental Science*, *6*(2), 211–220. doi:10.1111/1467-7687.00273
- Conboy, B. T., Sommerville, J. A., & Kuhl, P. K. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, *44*(5), 1505–1512.
- Cooper, Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: evidence from native and non-native listeners. *Language and Speech*, *45*(3), 207–228. doi:10.1177/00238309020450030101
- Cooper, G., & Meyer, L. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.
- Creissels, D., & Gregoire, C. (1993). La notion de ton marqué dans l'analyse d'une opposition tonale binaire: le cas du mandingue. *Journal of African languages and linguistics*, *14*(2), 107–154.
- Cristia, A., Minagawa-Kawai, Y., Egorova, N., Gervain, J., Filippin, L., Cabrol, D., & Dupoux, E. (2014). Neural correlates of infant accent discrimination: an fNIRS study. *Developmental Science*, 1–8. doi:10.1111/desc.12160
- Crowhurst, M. J., & Teodocio Olivares, A. (2014). Beyond the Iambic-Trochaic Law: the joint influence of duration and intensity on the perception of rhythmic speech. *Phonology*, *31*(01), 51–94.
- Crystal, D. (1997). *The Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.
- Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233–262. doi:10.1016/j.cognition.2004.08.005
- Cutler, A. (1975). Prosody and sentence perception in English. In *Perspectives on mental representation* (pp. 201–216).
- Cutler, A. (2005). Lexical Stress. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264–289). Blackwell Publishing.
- Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, *2*(3-4), 133–142. doi:10.1016/0885-2308(87)90004-0

- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25(4), 385–400.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113–121.
- Cutler, A., & Otake, T. (1999). Pitch accent in spoken-word recognition in Japanese. *The Journal of the Acoustical Society of America*, 105(3), 1877. doi:10.1121/1.426724
- De la Mora, D. M., Nespor, M., & Toro, J. M. (2013). Do humans and nonhuman animals share the grouping principles of the iambic-trochaic law? *Attention, Perception & Psychophysics*, 75, 92–100.
- De Villers-Sidani, E., Simpson, K. L., Lu, Y.-F., Lin, R. C. S., & Merzenich, M. M. (2008). Manipulating critical period closure across different sectors of the primary auditory cortex. *Nature Neuroscience*, 11(8), 957–965. doi:10.1038/nn.2144
- DeCasper, a. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, 17(2), 159–164.
- Decasper, A. J., & Fifer, W. P. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, (208), 1174–1176.
- Decasper, A. J., Lecanuet, J., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior & Development*, 17, 159–164.
- Decasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant and Child Development*, 9, 133–150.
- Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298(5600), 2013–2015.
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., & Dehaene, S. (2008). How does early brain organization promote language acquisition in humans? *European Review*, 16(04), 399. doi:10.1017/S1062798708000513
- Dehaene-Lambertz, G., Hertz-Pannier, L., Dubois, J., Mériaux, S., Roche, A., Sigman, M., & Dehaene, S. (2006). Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proceedings of the National Academy of Sciences of the United States of America*, 103(38), 14240–5. doi:10.1073/pnas.0606302103
- Dehaene-Lambertz, G., & Pena, M. (2001). Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport*, 12(14), 3155–8.
- Delattre, D. (1966). A comparison of syllable length conditioning among languages. *IRAL*, 4, 183–198.
- Dell, F., Hirst, D., & Vergnaud, J.-R. (1984). *Forme sonore du langage: Structure des représentations en phonologie*. (Paris: Her., pp. 65–122).
- Dimitrova, D. V., Redeker, G., & Hoeks, J. C. J. (2009). Did you say a BLUE banana? The prosody of contrast and abnormality in Bulgarian and Dutch. *INTERSPEECH*, 999–1002.
- Ding, X. P., Fu, G., & Lee, K. (2014). Neural correlates of own- and other-race face recognition in children: a functional near-infrared spectroscopy study. *NeuroImage*, (85), 335–44.
- Dogil, G., & Williams, B. (1999). The phonetic manifestation of word stress. In Van der Hulst (Ed.), *Word prosodic systems in the languages of Europe* (pp. 273–311). Berlin: Mouton de Gruyter.

- Downing, L. J. (2004). What african languages tell us about accent typology. *ZAS Papers in Linguistics*, 37, 101–136.
- Dryer, M. S. (1992). The greenbergian word order correlations. *Language*, 62(1), 81–138.
- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., & Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensori-motor networks. *Cerebral Cortex*, 19(2), 414–23.
- Dubois, J., Poupon, C., Thirion, B., Simonnet, H., Kulikova, S., Leroy, F., ... Dehaene-Lambertz, G. (2015). Exploring the early organization and maturation of linguistic pathways in the human infant brain. *Cerebral Cortex*, 1–16. doi:10.1093/cercor/bhv082
- Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness.” *The Journal of the Acoustical Society of America*, 110(3), 1606. doi:10.1121/1.1380437
- Dupoux, E., Sebastian, N., Pallier, C., & Mehler, J. (1997). A destressing “ deafness ” in French ? *Journal of Memory and Language*, 36, 406–421.
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & van der Vrecken, O. (1996). The MBROLA project: towards a set of high quality speech synthesizers free of use for non commercial purposes. In Alfred I. duPont Institute (Ed.), *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96* (Vol. 3, pp. 1393–1396). Wilmington: DE. doi:10.1109/ICSLP.1996.607874
- Echols, C., Crowhurst, M. J., & Childers, J. B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language*, 225(36), 202–225.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171, 303–306.
- Fallgatter, a J., & Strik, W. K. (1998). Frontal brain activation during the Wisconsin Card Sorting Test assessed with two-channel near-infrared spectroscopy. *European Archives of Psychiatry and Clinical Neuroscience*, 248(5), 245–9.
- Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Memory & Cognition*, 37(1), 1–9. doi:10.3758/MC.37.1.1
- Fedorenko, E., & Thompson-Schill, S. L. (2014). Reworking the language network. *Trends in Cognitive Sciences*, 18(3), 120–6. doi:10.1016/j.tics.2013.12.006
- Fennell, C. T., Byers-Heinlein, K., & Werker, J. F. (2007). Using speech sounds to guide word learning: the case of bilingual infants. *Child Development*, 78(5), 1510–25. doi:10.1111/j.1467-8624.2007.01080.x
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, 21(01), 2000. doi:10.1017/S0142716400001053
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16(3), 477–501. doi:10.1017/S0305000900010679
- Féry, C. (1998). German word stress in optimality theory. *The Journal of Comparative Germanic Linguistics*, 2(2), 101–142.

- Féry, C., Hörnig, R., & Pahaut, S. (2011). Correlates of phrasing in French and German from an experiment with semi-spontaneous speech. *Intonational Phrasing in Romance and Germanic*, 11–41.
- Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 36(4), 680–703. doi:10.1016/j.wocn.2008.05.001
- Fischer, A. (2009). *Prosodic organization in the babbling of German-learning infants between the age of six and twelve months*.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development*, 67(6), 3192–3218. doi:10.1111/j.1467-8624.1996.tb01909.x
- Fraisse, P. (1974). *Psychologie du rythme*. Paris:Presses Universitaires de France.
- Fraisse, P. (1982). Rhythm and timing. In D. Deutsch (Ed.), *The psychology of music* (p. 149–180). New York: Academic Press.
- Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends in Cognitive Sciences*, 16(5), 262–268.
- Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology : CB*, 17(14), 1208–11.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, 27(4), 765–768.
- Gandour, J., Tong, Y., Wong, D., Talavage, T., Dziedzic, M., Xu, Y., ... Lowe, M. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage*, 23(1), 344–357.
- Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., ... Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. *Journal of Phonetics*, 39(4), 546–557. doi:10.1016/j.wocn.2011.07.002
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15, 870–878. doi:10.1006/nimg.2001.1037
- Gerhardt, K. J., Otto, R., Abrams, R. M., Colle, J. J., Burchfield, D. J., & Peters, a J. (1992). Cochlear microphonics recorded from fetal and newborn sheep. *American Journal of Otolaryngology*, 13(4), 226–233.
- Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, 51(3), 237–265. doi:10.1016/0010-0277(94)90055-8
- Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: the newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3), 564–74.
- Gervain, J., Macagno, F., Cogoi, S., Peña, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America*, 105(37), 14222–14227.
- Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, 61, 191–218.
- Gervain, J., Mehler, J., Werker, J. F., Nelson, C., Csibra, G., Lloyd-Fox, S., ... Aslin, R. N. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience*, 1(1), 22–46. doi:10.1016/j.dcn.2010.07.004

- Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study. *Cognitive Psychology*, *57*(1), 56–74.
- Gervain, J., & Werker, J. F. (2013). Prosody cues word order in 7-month-old bilingual infants. *Nature Communications*, *4*, 1490.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.
- Gleitman, L. R., Gleitman, H., Landau, B., & Wanner, E. (1988). Where the learning begins: Initial representations for language learning. In F. Newmeyer (Ed.), *The Cambridge Linguistic Survey* (pp. 150–193). Cambridge: Cambridge University Press.
- Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the art. In L. R. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art* (pp. 3–48). Cambridge, MA: Cambridge University Press.
- Golinkoff, R. M., Jacquet, R. C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, *67*(6), 3101–3119. doi:10.2307/1131769
- Gomez, D. M., Berent, I., Benavides-Varela, S., Bion, R. a. H., Cattarossi, L., Nespor, M., & Mehler, J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences*, (21), 1–5.
- Gonzalez-gomez, N., & Nazzi, T. (2013). Effects of prior phonotactic knowledge on infant word segmentation: The case of non-adjacent dependencies. *Journal of Speech, Language, and Hearing Research*, *56*(3), 840–850. doi:10.1044/1092-4388(2012/12-0138)
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., & Decasper, A. J. (2011). A melodic contour repeatedly experienced by human near-term fetuses elicits a profound cardiac reaction one month after birth. *PloS One*, *6*(2), e17304.
- Granier-Deferre, C., & Busnel, M.-C. (2011). L'audition prénatale, quoi de neuf ? *Spirale*, *59*(3), 17. doi:10.3917/spi.059.0017
- Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., & Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental Science*, *14*(2), 336–352.
- Grosjean, F. (2010). *Bilingual : life and reality*. Cambridge, MA: Harvard University Pres.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–79. doi:10.1126/science.298.5598.1569
- Hauser, M. D., Marcus, G. F., & Weiss, D. J. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*, B15–B22. doi:10.1016/S0010-0277(02)00139-7
- Hay, J. F., & Diehl, R. L. (2007). Perception of rhythmic grouping: testing the iambic/trochaic law. *Perception & Psychophysics*, *69*(1), 113–22.
- Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(6), 610–641.
- Hayashi, A., & Mazuka, R. (2002). Developmental change in infants' preferences for continuousness in Japanese speech. *Paper Presented at the the 13th Biennial International Conference on Infant Studies, Toronto*.
- Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and meter* (pp. 201–260). New York: Academic Press.

- Hayes, B. (1995). Metrical stress theory : principles and case studies. In *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics* (pp. 429–446). Chicago, IL: University of Chicago Press.
- Hepper, P. G., & Shahidullah, B. S. (1994). Development of fetal hearing. *Archives of Disease in Childhood*, *71*, 81–87.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.
- Hirsh-Pasek, K., Nelson, D. G. K., Jusczyk, P. W., Druss, K. wright C. B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269–286.
- Höhle, B. (2002). *Der Einstieg in die Grammatik: Die Rolle der Phonologie/Syntax-Schnittstelle für Sprachverarbeitung und Spracherwerb. Habilitationsschrift.*
- Höhle, B., Bijeljic-Babic, R., Herold, B., Weissenborn, J., & Nazzi, T. (2009). Language specific prosodic preferences during the first half year of life: evidence from German and French infants. *Infant Behavior & Development*, *32*(3), 262–274.
- 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.
- Homae, F., Watanabe, H., Nakano, T., & Taga, G. (2007). Prosodic processing in the developing brain. *Neuroscience Research*, *59*(1), 29–39.
- Hoshi, Y., & Tamura, M. (1993). Detection of dynamic changes in cerebral oxygenation coupled to neuronal function during mental work in man. *Neuroscience Letters*, *150*(1), 5–8.
- Houston, D. M., Jusczyk, P. W., Kuijpers, C., Coolen, R., & Cutler, A. (2000). Cross-language word segmentation by 9-month-olds. *Psychonomic Bulletin & Review*, *7*(3), 504–509.
- Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *The Journal of the Acoustical Society of America*, *124*(4), 2263–2271.
- Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385.
- Jakobson, R., Fant, G., & Halle, M. (1952). Preliminaries to speech analysis: The distinctive features and their correlates. In *Acoustics laboratory N°13*. MIT, Technical Report.
- Jasinska, K. K., & Petitto, L.-A. (2013). How age of bilingual exposure can change the neural systems for language in the developing brain: A functional near infrared spectroscopy investigation of syntactic processing in monolingual and bilingual children. *Developmental Cognitive Neuroscience*, *6C*, 87–101. doi:10.1016/j.dcn.2013.06.005
- Jessen, M., Marasek, K., Schneider, K., & Classen, K. (1995). Acoustic correlates of word stress and the tense/lax opposition in the vowel system of German. *Proceedings of the International Congress of Phonetic Sciences*, *13*(4), 428–431.
- Jöbsis, F. F. (1977). Noninvasive, Infrared Monitoring of Cerebral and Myocardial Oxygen Sufficiency and Circulatory Parameters. *Science*, *198*, 1264–1267.
- Jöbsis-vanderVliet, F. F. (1999). Discovery of the Near-Infrared Window into the body and the early development of Near-Infrared Spectroscopy. *Journal of Biomedical Optics*, *4*(4), 392–396.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: when speech cues count more than statistics. *Journal of Memory and Language*, *44*(4), 548–567.

- Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy*, *13*(5), 440–455. doi:10.1080/15250000802329321
- Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, *14*(1), 147–172. doi:10.1515/prbs.2002.002
- Jusczyk, P. W. (1989). Perception of cues to clausal units in native and non-native languages. In *Biennial meeting of the society for research of child development*. Kansas City.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of words' sound patterns in fluent speech. *Cognitive Psychology*, *29*, 1–23.
- 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., Hirsh-Pasek, K., Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*(2), 252–293. doi:10.1016/0010-0285(92)90009-Q
- Jusczyk, P. W., Hohne, E., & Bauman, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics*, *61*(8), 1465–1476. doi:10.3758/BF03213111
- Jusczyk, P. W., Hohne, E., & Mandel, D. (1995). Picking up regularities in the sound structure of the native language. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and methodological issues in cross-language speech research* (pp. 91–119). Timonium, MD: York Press.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*, 159–207.
- Jusczyk, P. W., & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old infants. *Perception & Psychophysics*, *23*(2), 105–9.
- Kato, T., Kamei, a, Takashima, S., & Ozaki, T. (1993). Human visual cortical function during photic stimulation monitoring by means of near-infrared spectroscopy. *Journal of Cerebral Blood Flow and Metabolism : Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, *13*(3), 516–20. doi:10.1038/jcbfm.1993.66
- Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*(1), 55–68. doi:10.1017/S030500090001343X
- Kisilevsky, B. S., Hains, S. M. J., Brown, C., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., ... Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior & Development*, *32*(1), 59–71. doi:10.1016/j.infbeh.2008.10.002
- Kisilevsky, B. S., Hains, S. M. J., Jacquet, a Y., Granier-Deferre, C., & Lecanuet, J. P. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*(5), 550–9.
- Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., ... Wang, Z. (2003). Effect of experience on fetal voice recognition. *Psychological Science*, *14*(3), 220–224.
- Kluender, K. R., Diehl, R. L., & Killeen, P. R. (1987). Japanese quail can learn phonetic categories. *Science*, *237*(4819), 1195–1197. doi:10.1126/science.3629235
- Kooijman, V., Hagoort, P., & Cutler, A. (2009). Prosodic structure in early word segmentation: ERP evidence from dutch ten-month-olds. *Infancy*, *14*(6), 591–612.

- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., ... Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping, 31*(4), 595–603. doi:10.1002/hbm.20890
- Kovács, A. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences of the United States of America, 106*(16), 6556–6560.
- Kovács, A. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. *Science, 325*(5940), 611–2.
- Kover, H., Gill, K., Tseng, Y.-T. L., & Bao, S. (2013). Perceptual and neuronal boundary learned from higher-order stimulus probabilities. *Journal of Neuroscience, 33*(8), 3699–3705. doi:10.1523/JNEUROSCI.3166-12.2013
- Kreitewolf, J., Friederici, A. D., & von Kriegstein, K. (2014). Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. *NeuroImage, 102*, 332–344.
- Kroll, J. F., & Bialystok, E. (2013). Understanding the Consequences of Bilingualism for Language Processing and Cognition. *Journal of Cognitive Psychology, 25*(5), 37–41.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience, 5*(11), 831–843. doi:10.1038/nrn1533
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science, 9*(2). doi:10.1111/j.1467-7687.2006.00468.x
- Kurby, C. a., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences, 12*(2), 72–9. doi:10.1016/j.tics.2007.11.004
- Kusumoto, K., & Moreton, E. (1997). Native language determines parsing of nonlinguistic rhythmic stimuli. *Journal of the Acoustical Society of America, 102*(2), 3204–3204.
- Lahav, A., & Skoe, E. (2014). An acoustic gap between the NICU and womb: a potential risk for compromised neuroplasticity of the auditory system in preterm infants. *Frontiers in Neuroscience, 8*(381), 1–8. doi:10.3389/fnins.2014.00381
- Langus, A., Marchetto, E., Bion, R. A. H., & Nespors, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language, 66*(1), 285–306. doi:10.1016/j.jml.2011.09.004
- Lecanuet, Granier-Deferre, C., Jacquet, A.-Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology, 36*(1), 29–39.
- Lecanuet, & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics, Gynecology, and Reproductive Biology, 68*, 1–23.
- Lee, & Kisilevsky, B. S. (2014). Fetuses respond to father's voice but prefer mother's voice after birth. *Developmental Psychobiology, 56*(1), 1–11. doi:10.1002/dev.21084
- Lee, W., Chen, F., Luke, K. K., & Shen, L. (2002). The prosody of bisyllabic and polysyllabic words in Hong Kong Cantonese. In *Proc. SP-2002* (pp. 451–454). Aix-en-Provence.
- Lehiste, L. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.
- Lenneberg, E. (1967). *Biological Foundations of Language*. New York, NY: Wiley.
- Locke, J. (1690). *An Essay concerning Human Understanding*.
- Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., ... Wallois, F. (2013). Syllabic discrimination in premature human

- infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(12), 4846–51.
doi:10.1073/pnas.1212220110
- Malisz, Z. (2013). *Speech rhythm variability in Polish and English : A study of interaction between rhythmic levels*.
- Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology : CB*, *19*(23), 1994–1997.
- Mandel, D. R., Jusczyk, P. W., & Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53*(2), 155–180.
doi:10.1016/0010-0277(94)90069-8
- Mandel, Kemler Nelson, D. G., & Jusczyk, P. W. (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development*, *11*(2), 181–196.
doi:10.1016/S0885-2014(96)90002-7
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*, 940–967.
- Matsuda, G., & Hiraki, K. (2006). Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: a NIRS study of children. *NeuroImage*, *29*(3), 706–11.
- Mattock, K., Amitay, S., & Moore, D. R. (2010). Auditory development and learning. In *Oxford Handbook of Auditory Science*, *3* (pp. 297–324).
- Mattys, S. L., & Jusczyk, P. W. (2001a). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology. Human Perception and Performance*, *27*(3), 644–655. doi:10.1037/0096-1523.27.3.644
- Mattys, S. L., & Jusczyk, P. W. (2001b). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, *78*(2), 91–121. doi:10.1016/S0010-0277(00)00109-8
- Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive Psychology*, *38*, 465–494.
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, *56*(2), 154–178.
doi:10.1002/dev.21177
- May, L., Byers-Heinlein, K., Gervain, J., & Werker, J. F. (2011). Language and the newborn brain: does prenatal language experience shape the neonate neural response to speech? *Frontiers in Psychology*, *2*, 222.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), 101–111.
doi:10.1016/S0010-0277(01)00157-3
- Mazuka, R. (1996). Can a grammatical parameter be set before the first word ? Prosodic contributions to early setting of a grammatical. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 313–330). Mahwah, NJ: Lawrence Erlbaum Associates.
- Meek, J. H. (2002). Basic principles of optical imaging and application to the study. *Developmental Science*, *3*, 371–380.
- Meek, J. H., Firbank, M., Elwell, C. E., Atkinson, J., Braddick, O., & Wyatt, J. S. (1998). Regional hemodynamic responses to visual stimulation in awake infants. *Pediatric Research*, *43*, 840–843.

- Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld. (1978). Infant recognition of mother's voice. *Perception*, 7, 491 – 497.
- 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., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143–178.
- Meisel, J. (1992). *The acquisition of verb placement: functional categories and V2 phenomena in language acquisition*. Dordrecht: Kluwer Academic Press.
- Michelas, A., & D'Imperio, M. (2010). Accentual phrase boundaries and lexical access in french. In *Proceedings of the speech prosody conference* (pp. 1–4).
- Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: A developmental scenario. *Developmental Cognitive Neuroscience*, 1(3), 217–232.
- Minagawa-Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language-specific processing in early infant development. *Cerebral Cortex*, 21, 254–61.
- Molnar, M., Gervain, J., & Carreiras, M. (2014). Within-rhythm class native language discrimination abilities of basque-spanish monolingual and bilingual infants at 3.5 months of age. *Infancy*, 19(3), 326–337.
- Molnar, M., Lallier, M., & Carreiras, M. (2014). The amount of language exposure determines nonlinguistic tone grouping biases in infants from a bilingual environment. *Language Learning*, 64(s2), 45–64.
- Moon, C., Panneton Cooper, R., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior & Development*, 16, 495–500.
- Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. (J. Morgan & K. Demuth, Eds.).
- Morgan, J. L., Edwards, S., & Wheeldon, L. R. (2014). The relationship between language production and verbal short-term memory: the role of stress grouping. *Quarterly Journal of Experimental Psychology (2006)*, 67(2), 220–46. doi:10.1080/17470218.2013.799216
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19(4), 498–550. doi:10.1016/0010-0285(87)90017-X
- Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Developmental Science*, 13(3), 407–420. doi:10.1111/j.1467-7687.2009.00898.x
- 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., Floccia, C., & Bertoncini, J. (1998). Discrimination of pitch contours by neonates. *Infant Behavior & Development*, 21(4), 779–784.
- Nazzi, T., Iakimova, G., Bertoncini, J., Fredonie, S., & Alcantara, C. (2006). Early segmentation of fluent speech by infants acquiring French: Emerging evidence for crosslinguistic differences. *Journal of Memory and Language*, 54(3), 283–299.

- 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. doi:10.1006/jmla.2000.2698
- Nazzi, T., Mersad, K., Sundara, M., Iakimova, G., & Polka, L. (2014). Early word segmentation in infants acquiring Parisian French: task-dependent and dialect-specific aspects. *Journal of Child Language*, 41(3), 600–33. doi:10.1017/S0305000913000111
- Nazzi, T., & Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. *Speech Communication*, 41(1), 233–243.
- Nelson, C. (1995). The ontogeny of human memory: A cognitive neuroscience perspective. *Developmental Psychology*, 31(5), 723–738. doi:10.1037/0012-1649.31.5.723
- Nelson, D. G. K., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development*, 18, 111–116.
- Nespor, M., Guasti, M. T., & Christophe, A. (1996). Selecting word order: The rhythmic activation principle. In U. Kleinhenz (Ed.), *Interfaces in phonology* (pp. 1–26). Berlin: Akademie.
- Nespor, M., Shukla, M., Vijver Van de, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realizations in VO et OV languages. *Lingue E Linguaggio*, 2, 1–29.
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Nishibayashi, L.-L., Goyet, L., & Nazzi, T. (2015). Early speech segmentation in French-learning infants: Monosyllabic words versus embedded syllables. *Language and Speech*, 1–17.
- Obrig, H., & Villringer, A. (2003). Beyond the visible--imaging the human brain with light. *Journal of Cerebral Blood Flow and Metabolism*, 23(1), 1–18.
- Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language*, 32, 258–278.
- Otsuka, Y., Nakato, E., Kanazawa, S., Yamaguchi, M. K., Watanabe, S., & Kakigi, R. (2007). Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *NeuroImage*, 34, 399–406.
- Partanen, E., Kujala, T., Naatanen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences of the United States of America*, 110(37), 15145–15150. doi:10.1073/pnas.1302159110
- Partanen, E., Kujala, T., Tervaniemi, M., & Huotilainen, M. (2013). Prenatal music exposure induces long-term neural effects. *PloS One*, 8(10), e78946. doi:10.1371/journal.pone.0078946
- Pascalis, O., De Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant Behavior and Development*, 18, 79–85.
- Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., & Mehler, J. (2003). Sounds and silence: an optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences of the United States of America*, 100(20), 11702–5.
- Petitto, L. a., Berens, M. S., Kovelman, I., Dubins, M. H., Jasinska, K., & Shalinsky, M. (2012). The “Perceptual Wedge Hypothesis” as the basis for bilingual babies’ phonetic processing advantage: New insights from fNIRS brain imaging. *Brain and Language*, 121(2), 130–143. doi:10.1016/j.bandl.2011.05.003

- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: University Press.
- Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 399–441). Hillsdale: Erlbaum.
- Pinker, S. (1999). Out of the minds of babes. *Science*, 283(5398), 40–41. doi:10.1126/science.283.5398.40
- Pons, F. (2006). The effects of distributional learning on rats' sensitivity to phonetic information. *Journal of Experimental Psychology. Animal Behavior Processes*, 32(1), 97–101. doi:10.1037/0097-7403.32.1.97
- Pons, F., & Bosch, L. (2010). Stress pattern preference in Spanish-learning infants: The role of syllable weight. *Infancy*, 15(3), 223–245. doi:10.1111/j.1532-7078.2009.00016.x
- Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology*, 108, 567–579. doi:10.1016/j.jecp.2010.10.009
- Pujol, R., Lavigne-Rebillard, M., & Lenoir, M. (1998). Development of sensory and neural structures in the mammalian cochlea. In *Development of the auditory System* (Springer N., pp. 146–192).
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Crèpin, G. (1988). Fetal hearing. *European Journal of Obstetrics and Gynecology and Reproductive Biology*, 29(1), 191–212.
- Querleu, D., Renard, X., Versyp, F., Paris-Delrue, L., & Vervoot, P. (1988). La transmission intra-amniotique des voix humaines. *Revue Française de Gynécologie et D'obstétrique*, 83(1), 43–50.
- Ramus, F. (2002). Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2, 85–115.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288(5464), 349–51.
- Rand, K., & Lahav, A. (2014). Impact of the NICU environment on language deprivation in preterm infants. *Acta Paediatrica (Oslo, Norway : 1992)*, 103(3), 243–8. doi:10.1111/apa.12481
- Rice, C. C. (1992). *Binarity and ternarity in metrical theory: parametric extensions*. University of Texas, Austin.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8 month-old infants. *Science*, 274, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27–52. doi:10.1016/S0010-0277(98)00075-4
- Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35(4), 606–621.
- Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In *Handbook of Child Development* (Vol. 6, pp. 58–108). doi:10.1002/9780470147658
- Sakatani, K., Chen, S., Lichty, W., Zuo, H., & Wang, Y. P. (1999). Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy. *Early Human Development*, 55(3), 229–36.

- Sakatani, K., Xie, Y., Lichty, W., Li, S., & Zuo, H. (1998). Language-activated cerebral blood oxygenation and hemodynamic changes of the left prefrontal cortex in poststroke aphasic patients : A near-infrared spectroscopy study. *Stroke*, *29*(7), 1299–1304. doi:10.1161/01.STR.29.7.1299
- Sansavini, A., Bertocini, J., & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology*, *33*(1), 3–11.
- Sato, Hirabayashi, Y., Tsubokura, H., Kanai, M., Ashida, T., Konishi, I., ... Maki, A. (2012). Cerebral hemodynamics in newborn infants exposed to speech sounds: a whole-head optical topography study. *Human Brain Mapping*, *33*(9), 2092–103. doi:10.1002/hbm.21350
- Sato, Y., Sogabe, Y., & Mazuka, R. (2007). Brain responses in the processing of lexical pitch-accent by Japanese speakers. *Neuroreport*, *18*(18), 2001–4.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010a). Development of hemispheric specialization for lexical pitch-accent in Japanese infants. *Journal of Cognitive Neuroscience*, *22*(11), 2503–13.
- Sato, Y., Sogabe, Y., & Mazuka, R. (2010b). Discrimination of phonemic vowel length by Japanese infants. *Developmental Psychology*, *46*(1), 106–119. doi:10.1037/a0016718
- Schmitz, M. (2008). *The Perception of Clauses in 6- and 8- Month-Old German-Learning Infants : Influence of Pause Duration and the Natural Pause Hierarchy*. University Potsdam.
- Sebastián-Gallés, N., Albareda-Castellot, B., Weikum, W. M., & Werker, J. F. (2012). A bilingual advantage in visual language discrimination in infancy. *Psychological Science*, *23*(9), 994–999.
- Segal, O., & Kishon-Rabin, L. (2012). Evidence for language-specific influence on the preference of stress patterns in infants learning an iambic language (Hebrew). *Journal of Speech, Language, and Hearing Research*, *55*, 1329–1341.
- Selkirk, E. O. (1984). *Phonology and Syntax: The Relation Between Sound and Structure*. Cambridge: Press, MIT.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21.
- Shroads, E., Johnson, K. M., Andrews, E., & Parramore Wilbourn, M. (2015). Listen here !: Parents ' use of prosodic highlighting during interactions with young infants. *Poster Presented at SRCD Biennial Meeting, Philadelphia*.
- Singh, & Dan, I. (2006). Exploring the false discovery rate in multichannel NIRS. *NeuroImage*, *33*, 542–549. doi:10.1016/j.neuroimage.2006.06.047
- Singh, L., Fu, C. S. L., Rahman, A. a, Hameed, W. B., Sanmugam, S., Agarwal, P., ... Rifkin-Graboi, A. (2015). Back to basics: A bilingual advantage in infant visual habituation. *Child Development*, *86*(1), 294–302. doi:10.1111/cdev.12271
- Singh, White, K. S., Morgan, J. L., & White, K. (2008). Building a lexicon in the face of phonetic variation: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, *4*, 157–178.
- Sinnott, J. M., & Aslin, R. N. (1985). Frequency and intensity discrimination in human infants and adults. *The Journal of the Acoustical Society of America*, *78*(6), 1986–1992. doi:10.1121/1.392654
- Skinner, B. . (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Skoruppa, K., Cristià, A., Peperkamp, S., & Seidl, A. (2011). English-learning infants' perception of word stress patterns. *The Journal of the Acoustical Society of America*, *130*(1), EL50–5. doi:10.1121/1.3590169

- Skoruppa, K., Pons, F., Bosch, L., Christophe, A., Cabrol, D., & Peperkamp, S. (2013). The development of word stress processing in French and Spanish infants. *Language Learning and Development*, 9(1), 88–104. doi:10.1080/15475441.2012.693881
- Skoruppa, K., Pons, F., Christophe, A., Bosch, L., Dupoux, E., Sebastián-Gallés, N., ... Peperkamp, S. (2009). Language-specific stress perception by 9-month-old French and Spanish infants. *Developmental Science*, 12(6), 914–9. doi:10.1111/j.1467-7687.2009.00835.x
- Small, L. H., Simon, S. D., & Goldberg, J. S. (1988). Lexical stress and lexical access: homographs versus nonhomographs. *Perception & Psychophysics*, 44(3), 272–280. doi:10.3758/BF03206295
- Soderstrom, M., Nelson, D. G. K., & Jusczyk, P. W. (2005). Six-month-olds recognize clauses embedded in different passages of fluent speech. *Infant Behavior and Development*, 28(1), 87–94. doi:10.1016/j.infbeh.2004.07.001
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language*, 49(2), 249–267. doi:10.1016/S0749-596X(03)00024-X
- Soto Faraco, S., Sebastián, N., & Cutler, A. (2001). Segmental and suprasegmental mismatch in lexical access. *Journal of Memory and Language*, 45(3), 412–432. doi:10.1006/jmla.2000.2783
- Spence, J., & Decasper, J. (1987). Prenatal experience with low-frequency maternal-voice sounds influence neonatal perception of maternal voice samples. *Infant Behavior & Development*, 10, 133–142.
- Sundara, M., Polka, L., & Molnar, M. (2008). Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition*, 108(1), 232–242.
- Swallow, K. M., & Zacks, J. M. (2008). Sequences learned without awareness can orient attention during the perception of human activity. *Psychonomic Bulletin & Review*, 15(1), 116–122. doi:10.3758/PBR.15.1.116
- Taga, G., & Asakawa, K. (2007). Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage*, 36(4), 1246–52. doi:10.1016/j.neuroimage.2007.04.037
- Taga, G., Asakawa, K., Maki, A., Konishi, Y., & Koizumi, H. (2003). Brain imaging in awake infants by near-infrared optical topography. *Proceedings of the National Academy of Sciences of the United States of America*, 100(19), 10722–27.
- Tarquinio, N., Zelazo, P. R., & Weiss, M. J. (1990). Recovery of neonatal head turning to decreased sound pressure level. *Developmental Psychology*, 26(5), 752–758. doi:10.1037/0012-1649.26.5.752
- 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. *The Journal of Neuroscience*, 29(47), 14726–33.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, 39(4), 706–716. doi:10.1037/0012-1649.39.4.706
- Thiessen, E. D., & Saffran, J. R. (2007). Language learning and development learning to learn : Infants ' acquisition of stress-based strategies for word segmentation. *Language Learning and Development*, 3(1), 73–100.
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, 10(2), 172–175. doi:10.1111/1467-9280.00127

- Tincoff, R., & Jusczyk, P. W. (2012). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432–444. doi:10.1111/j.1532-7078.2011.00084.x
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*(5), 867–875. doi:10.3758/BF03193539
- Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition*, *6*(2), 131–6. doi:10.1007/s10071-003-0172-0
- Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception And Psychophysics*, *62*(2), 333–340.
- Turk, A., Jusczyk, P. W., & Gerken, L. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech*, *38*(2), 143–158. doi:10.1177/002383099503800202
- Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *The Journal of the Acoustical Society of America*, *126*(1), 367–76.
- Vaissière, J., & Michaud, A. (2006). Prosodic constituents in French : a data-driven approach. In I. Fonagy, Y. Kawaguchi, & T. Moriguchi (Eds.), *Prosody and syntax* (pp. 47–64). Amsterdam: John Benjamins.
- Vanderwert, R. E., & Nelson, C. (2014). The use of near-infrared spectroscopy in the study of typical and atypical development. *NeuroImage*, *85*, 264–71. doi:10.1016/j.neuroimage.2013.10.009
- Villringer, a, Planck, J., Hock, C., Schleinkofer, L., & Dirnagl, U. (1993). Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults. *Neuroscience Letters*, *154*(1-2), 101–4.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science*, *10*(2), 159–164. doi:10.1111/j.1467-7687.2007.00549.x
- Warren, P. (1999). Prosody and language processing. In S. Garrod & M. J. Pickering (Eds.), *Language processing* (pp. 155–188). Hove: Psychology Press.
- Wartenburger, I., Steinbrink, J., Telkemeyer, S., Friedrich, M., Friederici, A. D., & Obrig, H. (2007). The processing of prosody: Evidence of interhemispheric specialization at the age of four. *NeuroImage*, *34*(1), 416–425.
- Watanabe, E., Maki, A., Kawaguchi, F., Takashiro, K., Yamashita, Y., Koizumi, H., & Mayanagi, Y. (1998). Non-invasive assessment of language dominance with near-infrared spectroscopic mapping. *Neuroscience Letters*, *256*(1), 49–52.
- Watson, T. L., Robbins, R. a, & Best, C. T. (2014). Infant perceptual development for faces and spoken words: an integrated approach. *Developmental Psychobiology*, *56*(7), 1454–81. doi:10.1002/dev.21243
- Webb, A. R., Heller, H. T., Benson, C. B., & Lahav, A. (2015). Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proceedings of the National Academy of Sciences*, *112*(10), 3152–3157. doi:10.1073/pnas.1414924112
- Weikum, W. M., Vouloumanos, A., Navarra, J., Soto Faraco, S., Sebastián-Gallés, N., & Werker, J. F. (2007). Visual language discrimination in infancy. *Science*, *316*, 1159. doi:10.1126/science.1137686
- Welby, P., & Loevenbruck, H. (2006). Anchored down in anchorage : Syllable structure and segmental anchoring in French. *Italian Journal of Linguistics*, *18*, 74–124.
- Werker, J. F. (2012). Perceptual foundations of bilingual acquisition in infancy. *Annals of the New York Academy of Sciences*, *1251*, 50–61.

- Werker, J. F., & Gervain, J. (2013). Speech perception in infancy : A fondation for language acquisition. In P. Zelazo (Ed.), *The Oxford Handbook of Developmental Psychology* (pp. 909–925). Oxford University Press.
- Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspect of cross-language speech perception. *Child Development*, *52*, 349–355.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7*, 49–63.
- Werner, L. A., & Boike, K. (2001). Infants' sensitivity to broadband noise. *The Journal of the Acoustical Society of America*, *109*(5), 2103–2111. doi:10.1121/1.417731
- Wiese, R. (1996). The Phonology of German. In *The Phonology of the World's Languages*. Oxford: Oxford University Press.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, *91*(3), 1707–1717. doi:10.1121/1.402450
- Woodrow, H. (1909). A quantitative study of rhythm: The effect of variations in intensity, rate, and duration. *Archives of Psychology*, (14), 1–66.
- Woodrow, H. (1911). The role of pitch in rhythm. *Psychological Review*, *18*(1), 54–77.
- Woodrow, H. (1951). Time perception. In S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224–1236). New York: Wiley.
- Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., & Werker, J. F. (2010). The development of perceptual grouping biases in infancy: a Japanese-English cross-linguistic study. *Cognition*, *115*(2), 356–61.
- Zhou, X., Nagarajan, N., Mossop, B. J., & Merzenich, M. M. (2008). Influences of unmodulated acoustic inputs on functional maturation and critical-period plasticity of the primary auditory cortex. *Neuroscience*, *154*(1), 390–396.

ANNEXES

Annexe 1 : Étude Pilote Expérience 2

1. Méthode

- Participants

Nous avons testé 92 nourrissons monolingues de 7.5 mois nés à terme, apprenant le français à Paris (France) et l'allemand à Potsdam (Allemagne). Voir Tableau 1, Figure 1 pour la répartition des effectifs par conditions.

- Stimuli

La même séquence de familiarisation et les mêmes items tests que dans l'Expérience 1 (adultes) ont été utilisés.

- Procédure

Nous avons mesuré les réponses des nourrissons à l'aide de la méthode de regard préférentiel (HPP). La procédure est décrite dans la section 2.2.3.2.

2. Résultats

Nous avons réalisé une ANOVA à mesures répétées sur les temps d'orientations de regard avec le facteur intra-sujets type items (iambique vs trochaïque) et les facteurs inter-sujets conditions (durée, intensité, pitch et contrôle) et langue maternelle (allemand vs français). Les résultats sont rapportés dans le Tableau 1, Figure 1. Aucun effet n'est significatif. Les nourrissons de 7,5 mois allemands et français n'ont pas été capables de grouper les syllabes quel que soit la condition, suggérant que la tâche était trop difficile. Ceci nous a conduits à répliquer ces études avec des stimuli plus simples, ce qui est présenté en détails dans l'expérience 2 du chapitre 2.

Tableau 1 : Analyse de l'Expérience pilote Expérience 2

Iambique vs Trochaïque

Contrastes	ddl (x)	F (x, 80)	p
Langage	1	0,03	0,87
Condition	2	1,19	0,31
Type d'item	1	1,07	0,35
Langage × Condition	2	1,80	0,18
Type d'item × Langage	1	0,19	0,66
Type d'item × Condition	2	0,17	0,84
Type d'item × Langage × Condition	2	0,26	0,77

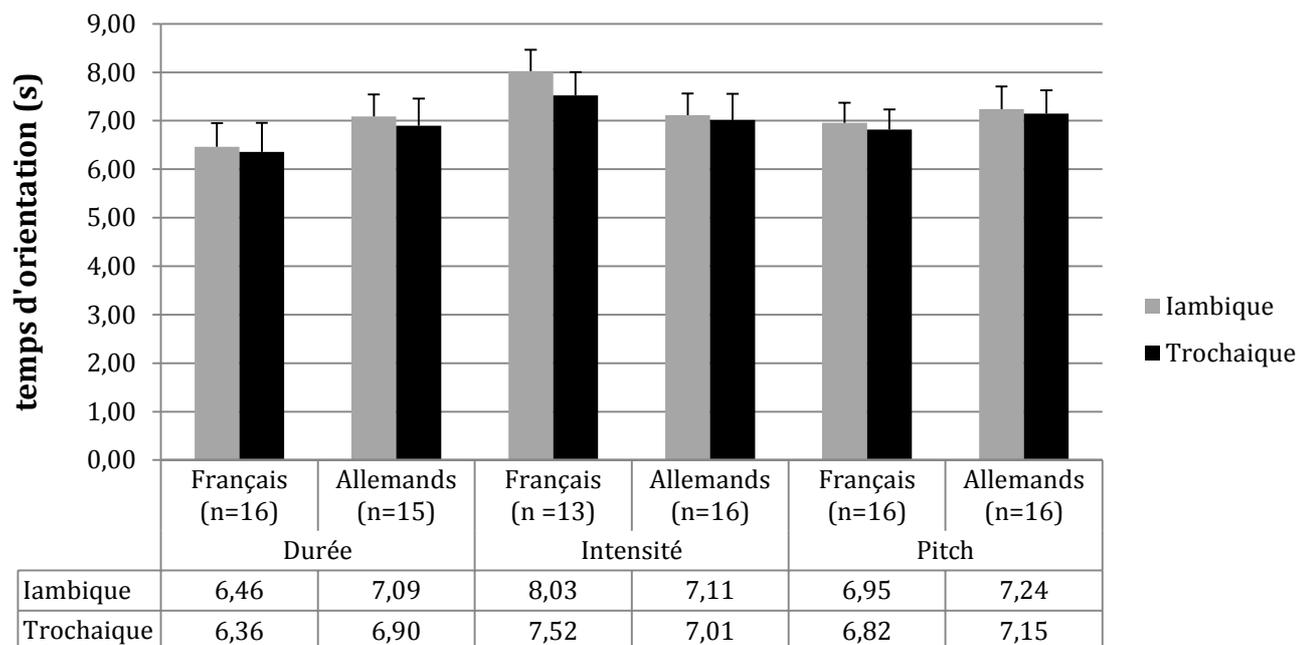


Figure 1 : Temps d'orientation de regard globaux ainsi que les effectifs dans les trois conditions expérimentales (durée, intensité et pitch), séparés par langues (français et allemands) et par types d'items (trochaïque et iambiques).

Annexe 2 : Pré-analyse Expérience 2

Étant donné que nous avons présenté nos 12 items tests (6 items trochaïque et 6 items iambique) dans un ordre aléatoire, nous avons séparé les temps de regard globaux en deux « blocs »: les trois premiers essais (essais 1-3) et les trois derniers essais de chaque type d'item (essais 4-6), voir Figure 1. En d'autre terme le bloc 1 (essais 1-3) contenait les trois premiers essais de chaque type (trochaïque et iambique) et le bloc 2 (essais 4-6) contenait les trois derniers essais de chaque type (trochaïques et iambiques).

Nous avons réalisé une ANOVA à mesures répétées sur les temps d'orientation de regard avec le facteur intra-sujets (bloc) et le facteur inter-sujets condition (durée, intensité, pitch et contrôle). Les résultats indiquent un effet principal du bloc $F(1, 155)=82.75$, $p < .00001$, mais pas d'effet significatif de condition $F(3, 155)=2.63$, $p = .052$, ni d'interaction avec la condition et le bloc $F(3, 155) = 1.63$, $p = .18$. Ces résultats montrent ainsi que les temps de regards baissent significativement entre les trois premiers essais et les trois derniers. De plus cet effet ne diffère pas en fonction de la condition, voir Figure

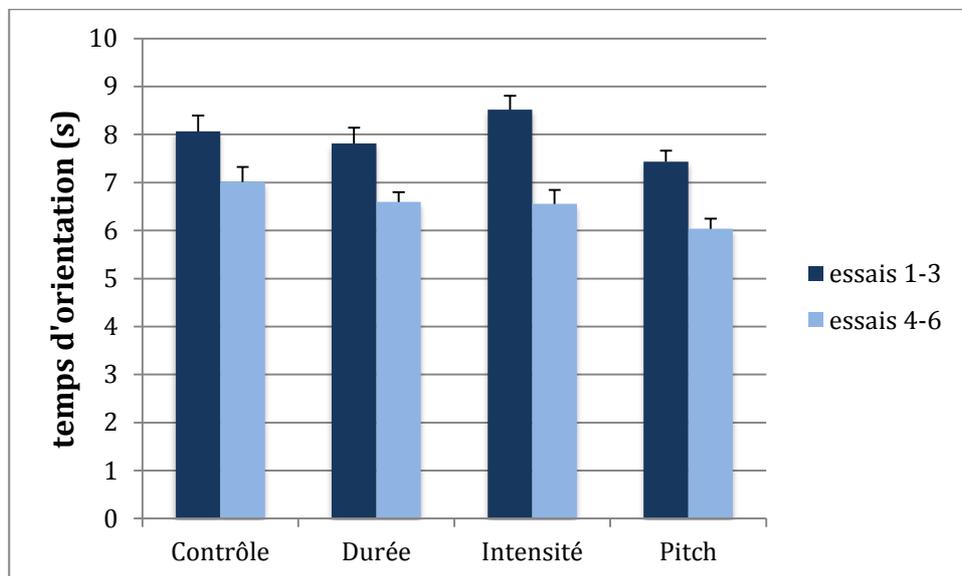


Figure 2 : Temps d'orientation de regard globaux dans les quatre conditions expérimentales (durée, intensité, pitch et contrôle), séparés par types d'items et des trois premiers et trois derniers essais

**Annexe 3 : Abboub, Bijeljac-Babic, Serres & Nazzi
(2015)**