

Université de Montréal

**L'apprentissage des tons lexicaux à partir du signal acoustique de la parole chez les
nourrissons âgés de 4 mois**

par Maude De Carufel-Tougas

Département de Psychologie
Faculté des arts et des sciences

Essai doctoral d'intégration présenté
en vue de l'obtention du grade de doctorat
en psychologie, option neuropsychologie clinique (D.Psy.)

16 mai 2019

© Maude De Carufel-Tougas, 2019

Résumé

Durant la première année de vie, le nourrisson apprend les catégories phonétiques de sa langue maternelle, incluant les tons lexicaux. Cependant, il n'est pas clair comment cette réorganisation perceptive s'effectue à partir d'un signal acoustique si variable. Dans une étude de simulations par réseaux neuronaux artificiels, la dynamique du signal acoustique permettait de traiter cette variabilité et catégoriser les tons lexicaux du chinois mandarin. La présente étude explore l'habileté du nourrisson à apprendre les tons du mandarin à partir des patrons de fréquence fondamentale (F0) du signal acoustique de la parole, mais aussi de leur mouvement, c'est-à-dire des patrons de vitesse de F0 (D1). Trente-deux nourrissons francophones âgés de 4 mois ont été familiarisés aux quatre tons mandarins selon un continuum de stimuli en 10 étapes qui variait à la fois sur F0 et D1. Dans la première condition, le continuum allait du ton montant au ton descendant de sorte que D1 était distribué de façon bimodale et F0, de façon unimodale. Dans la seconde condition, le continuum allait du ton bas au ton haut, avec une distribution bimodale des variations de F0 et unimodale des variations de D1. Les performances durant la phase test suggèrent que le nourrisson peut discriminer les tons dès l'âge de 4 mois, tel qu'observé dans les études antérieures, mais qu'il semble trop jeune pour apprendre la catégorisation des tons à partir des patrons de distribution des indices acoustiques. D'autres études sont nécessaires pour mieux comprendre les mécanismes sous-jacents à l'acquisition des catégories phonétiques.

Mots-clés : perception de la parole, réorganisation perceptive, apprentissage statistique, tons lexicaux, nourrissons.

Abstract

In the first year of life, the newborn learns the phonetic categories of his native language, including the lexical tones. However, it is unclear how this perceptive reorganisation takes place from an acoustic input that is so variable. In a study of simulations by artificial neural networks, the dynamic of the acoustic signal would be effective to process this variability and categorize the lexical tones of Chinese Mandarin. This study explored the newborn ability to learn the tones in Mandarin, from the fundamental frequency (F0), but also from their movement, that is to say the velocity patterns of F0 (D1). Thirty-two francophone newborns 4 months of age were initially familiarized with tones by a continuum of stimuli comprising 10 steps, that varied on both F0 and D1. In the first condition, the continuum was going from rising tone to descending tone so that D1 was distributed in a bimodal distribution and F0 was distributed in a unimodal distribution. In the second condition, the continuum was going from low tone to high tone, with a bimodal distribution of F0 variations and unimodal variations of D1. During the test phase, the performances suggest that the infant can discriminate tones as early 4 months of age, as seen in previous studies, but seems too young to learn the categorization of tones from the patterns of distribution of acoustic indices. Further studies are needed to better understand the underlying mechanisms of phonetic categories acquisition.

Keywords : speech perception, perceptual reorganization, distributional learning, lexical tone, infants.

Table des matières

Résumé.....	i
Abstract.....	ii
Liste des figures.....	iv
Liste des sigles.....	v
Liste des abréviations.....	vi
Remerciements.....	vii
Introduction.....	8
Contexte théorique.....	9
Perception de la parole chez le nourrisson.....	9
La réorganisation perceptive chez les nourrissons.....	10
La réorganisation perceptive et l'apprentissage statistique.....	18
Les tons lexicaux.....	22
Réorganisation perceptive des tons lexicaux.....	23
Les tons lexicaux et l'apprentissage statistique.....	28
Le mouvement de F0 comme indice acoustique.....	31
Objectifs de l'étude.....	32
Hypothèses.....	33
Méthodologie.....	33
Participants.....	33
Procédure.....	37
Phase d'apprentissage.....	37
Phase test.....	38
Analyses statistiques.....	40
Prédictions.....	40
Résultats.....	43
Phase d'apprentissage.....	43
Phase test.....	43
Discussion.....	46
Conclusion.....	53
Bibliographie.....	55

Liste des figures

Figure 1. Continuum de l'expérience de Maye, Werker et Gerken (2002)	20
Figure 2. Continuum de l'étude de Liu et Kager (2014).....	29
Figure 3. Patrons F0 des stimuli	34
Figure 4. Continuum F0.....	35
Figure 5. Continuum D1.	36

Liste des sigles

D1 : Le patron de vitesse de la fréquence fondamentale

F0 : Fréquence fondamentale

T : Ton

Liste des abréviations

c.-à-d. : C'est-à-dire

etc : Et cetera

É.-T. : Écart-type

M : Moyenne

Max : Maximum

Min : Minimum

ms : Millisecondes

n : Nombre d'individus

par ex. : Par exemple

p : seuil de signification

n^2 : eta-carré partiel

Remerciements

Tout d'abord, je remercie mon directeur de recherche principal, M. Bruno Gauthier, qui par sa confiance et son support, m'a permis d'évoluer à travers la grande aventure de la recherche. Merci d'avoir cru en moi et d'avoir accepté mon rythme de travail. De plus, je tiens à remercier ma codirectrice, Mme Rushen Shi (UQAM), qui m'a permis de développer ma rigueur scientifique et le dépassement de soi. Merci également à Patricia Laniel, membre du Laboratoire d'études en neuropsychologie de l'enfant et de l'adolescent (LÉNEA) ainsi qu'à Mme Sarah Lippé, fondatrice du laboratoire Neuroscience of Early Development (NED) de l'Hôpital Sainte-Justine de Montréal et à Inga Sophia Knoth, coordonnatrice de ce laboratoire, pour leur collaboration à mon projet.

Je tiens également à remercier ma famille, Thérèse, Johanne, Daniel, Julie, Jonathan, Patrick, Brigitte, qui ont su m'encourager et me soutenir à travers ces nombreuses années d'études.

Je tiens aussi à remercier des êtres chers, mes amies de la cohorte, Gabrielle, Marianne et Édith ainsi que mes amies de longue date, Claudine, Marion, Caroline, Geneviève et Lise. Dans les bons moments, elles ont souligné mes réussites, dans les mauvais, elles m'ont écouté, soutenu, encouragé et jamais elles ne m'ont abandonné. Elles m'ont poussé vers l'avant sans jamais douter de moi. Merci d'avoir partagé ces moments avec moi, j'ai été si bien accompagnée!

Introduction

La recherche en psycholinguistique et en neurolinguistique a apporté beaucoup à la compréhension des circuits cérébraux de la parole et de l'acquisition du langage. Les chercheurs tentent de préciser, le plus minutieusement possible, les étapes d'acquisition du langage des nourrissons. Selon Tomasello (2003), chaque enfant passerait à travers les mêmes étapes d'acquisition du langage et selon une chronologie semblable du développement des habiletés langagières. La perception de la parole est une étape préliminaire cruciale pour les étapes subséquentes de l'acquisition du langage (Werker et Yeung, 2005). Dès les premiers moments de vie, les bébés sont dotés de remarquables habiletés de perception de la parole. En effet, les recherches menées dans ce domaine montrent que, dès la naissance, certains circuits cérébraux sont déjà en place et en mesure de percevoir et de traiter le signal de la parole, notamment le circuit temporopariétofrontal (Dehaene-Lambertz, 2004). D'ailleurs, depuis les années 70, des chercheurs se sont intéressés au développement de la perception de la parole et ont découvert que les nourrissons, en tout début de vie, peuvent distinguer la majorité des sons présents dans les langues du monde (par ex. Eimas, Siqueland, Jusczyk et Vigorito, 1971). Toutefois, un nombre limité d'études se sont intéressées à la perception des tons lexicaux.

Cet essai débute avec un bilan des connaissances à propos de la perception de la parole des nourrissons, incluant les habiletés précoces du nouveau-né et la réorganisation perceptive durant la première année de vie. Par la suite, la réorganisation perceptive et l'apprentissage statistique des tons lexicaux seront abordés, pour ensuite exposer les buts de l'étude ainsi que les hypothèses. Enfin, la méthodologie de l'étude sera décrite en détail suivie des résultats, de la discussion et de la conclusion.

Contexte théorique

Perception de la parole chez le nourrisson

Tout d'abord, aux vues de la présente étude, la conceptualisation de l'aspect perceptif de la parole est primordiale. Ainsi, Bertoncini et Boysson-Bardies (2000) offrent une définition selon laquelle la perception de la parole correspond à un mécanisme de traitement perceptif qui s'organise graduellement afin d'en arriver à un système modulé par les propriétés phonologiques et prosodiques de la langue de l'enfant.

Dès la naissance, les mécanismes de perception auditive du nourrisson sont aptes à fonctionner (Graven et Browne, 2008). La parole est transmise via un signal acoustique continu variable. Ainsi, le nourrisson doit trouver une façon de structurer et d'organiser cette information pour éventuellement en faire du sens. Le jeune poupon arrive à effectuer une structuration du signal de la parole dès les premiers jours de vie. À 4 jours, les nourrissons discriminent leur langue maternelle d'une langue étrangère (Mehler, Lambertz, Jusczyk et Amiel-Tison, 1986), suggérant qu'ils sont en mesure de reconnaître et percevoir les propriétés de leur langue comme la prosodie par exemple. De plus, dès l'âge de 2 mois, le nourrisson utilise les propriétés mélodiques et rythmiques de la parole afin de reconnaître des phrases de sa langue maternelle (Dehaene-Lambertz et Houston, 1998).

Le nourrisson arrive à effectuer une structuration du signal de la parole également par perception des catégories phonétiques, même s'il est dépourvu d'expérience langagière et de la capacité de produire des sons de la parole (Bertoncini et Boysson-Bardies, 2000). Ainsi, le bébé âgé entre 1 et 4 mois est en mesure de percevoir des catégories phonétiques, notamment en discriminant les contrastes de voisement (par ex. [b] et [p]; Eimas, Siqueland, Jusczyk et Vigorito, 1971), de mode d'articulation (par ex. [b] et [m]; Eimas et Miller, 1980) et du lieu

d'articulation (par ex. [b] et [d]; Morse, 1972). Par ailleurs, jusqu'à l'âge d'environ 4 à 6 mois, le nourrisson perçoit les contrastes de la parole même dans une langue autre que la sienne (Werker, Gilbert, Humphrey et Tees, 1981), ce qui lui a valu le titre d'auditeur universel (*universal listener*).

La réorganisation perceptive chez les nourrissons

Plusieurs études montrent que la capacité des nourrissons à distinguer les contrastes de la parole des langues non familières diminue progressivement au cours de la première année de vie (Kuhl et al., 1992 ; Polka et Werker, 1994 ; Werker et Lalonde, 1988 ; Werker et Tees, 1984), mais qu'elle peut être retrouvée sous certaines conditions (par ex. l'exposition à une langue tonale (Liu et Kager, 2014; Tsao, 2017)). Ce phénomène de déclin de la capacité à discriminer les contrastes de la parole s'appelle la réorganisation perceptive (Werker et Tees, 1984). Bien que cela puisse sembler contre-intuitif, ce déclin de performance est une indication de développement cognitif (Best, McRoberts et Sithole, 1988). Les changements développementaux dans la perception des sons de la parole comme les voyelles (vers 4 à 6 mois) et les consonnes (vers 10 mois) comptent parmi les preuves qu'un enfant a commencé l'acquisition de sa langue maternelle (Kuhl et al., 1992 ; Polka et Werker, 1994 ; Werker et Lalonde, 1988 ; Werker et Tees, 1984). Ces changements reflètent la concentration efficace de l'attention des nourrissons aux seules dimensions acoustiques qui sont pertinentes aux sons de leur langue maternelle. D'ailleurs, la réorganisation perceptive est nécessaire pour l'acquisition ultérieure du système phonémique propre à la langue maternelle du nourrisson (Aslin, Jusczyk et Pisoni, 1998). Ainsi, durant la première année de vie du nourrisson, la perception universelle du langage se développe en une perception spécifique à la langue maternelle du bébé.

Werker et Tees (1984) ont étudié le processus de réorganisation perceptive avec des bébés anglophones de 6 à 8 mois, de 8 à 10 mois et de 10 à 12 mois. Les bébés devaient discriminer deux contrastes phonétiques, le premier en langue hindi et l'autre en langue thompson (langue amérindienne parlée en Colombie-Britannique). Les résultats montrent que presque tous les nourrissons de 6 à 8 mois étaient en mesure de discriminer les deux contrastes phonétiques, qu'une plus petite proportion d'enfants de 8 à 10 mois pouvait le faire, et que la plupart des enfants de 10 à 12 mois n'y arrivaient pas. De plus, les nourrissons âgés de 11-12 mois dont la langue maternelle était l'hindi ou le thompson discriminaient seulement les contrastes phonétiques propres à leur langue maternelle. De surcroît, des nourrissons anglophones de 11 à 13 mois ne sont plus aptes à discriminer le contraste phonémique appartenant à la langue hindi (da-Da) qui est non existante dans leur langue maternelle (Werker et Lalonde, 1988). Ceci suggère qu'au cours de la première année de vie, le nourrisson affine ses habiletés perceptives en portant davantage attention aux sons de sa langue maternelle. Cela expliquerait pourquoi les bébés perdent progressivement la capacité de discriminer les contrastes de la parole autres que ceux de leur langue maternelle et à mieux discriminer les sons de leur langue maternelle.

En 1994, Polka et Werker ont examiné la capacité de bébés anglophones âgés de 6-8 mois et de 10-12 mois à distinguer deux contrastes de voyelles en allemand ([Y] - [U] et [y]- [u]). Une variante de la procédure *Conditioned Headturn Procedure* (CHP) a été utilisée. Les résultats ont confirmé qu'une plus grande proportion de bébés âgés de 6-8 pouvaient discriminer les contrastes allemands en comparaison aux bébés de 10-12 mois. Les auteurs ont alors introduit un groupe de bébés anglophones âgés de 4 mois exposés aux mêmes stimuli lors d'une deuxième expérience. Afin de pouvoir mesurer les habiletés de discrimination des bébés de 4

mois, la méthodologie de la deuxième expérience a été modifiée au niveau de la procédure (procédure par habituation visuelle). Les bébés âgés de 6-8 mois ont également participé à la deuxième expérience. Les résultats ont révélé que seulement les plus jeunes bébés discriminaient les deux contrastes allemands et qu'un déclin de la discrimination des voyelles est présent dès 6 mois. Dès lors, en comparaison aux consonnes, ces résultats suggèrent qu'il y a présence d'une réorganisation perceptive des voyelles, et ce, à un plus jeune âge.

Kuhl et ses collègues (1992) ont montré que l'exposition à une langue spécifique pendant la première moitié de vie d'un nourrisson modifie sa perception phonétique. Cette étude était basée sur les travaux de Grieser et Kuhl (1989) selon lesquels l'expérience linguistique influence un prototype propre à langue maternelle du bébé. Du point de vue linguistique, le terme prototype fait référence à un stimulus d'une catégorie occupant une position prééminente tandis que le non-prototype occupe la position inverse (Taylor, 2008). En ce sens, le bébé détient l'habileté de se représenter des catégories phonétiques similaires sous forme d'un prototype qui a pour fonction d'agir comme un aimant, ce que les auteurs appellent l'effet magnétique (*magnet effect*), en diminuant les différences perceptives entre le centre et les extrémités de la catégorie (Kuhl, Williams, Lacerda, Stevens et Lindblom, 1992). Les auteurs ont exposé des bébés américains anglais ($n=16$) et suédois ($n=16$) âgés de 6 mois à deux voyelles de langue maternelle et de langue étrangère (voyelle en anglais /i/ et voyelle en suédois /y/). L'hypothèse consistait à ce que les bébés qui apprennent l'anglais traitent la voyelle en anglais /i/ comme un prototype (propre à leur langue maternelle) et la voyelle en suédois /y/ (langue étrangère) comme un non-prototype et la réponse inverse pour les bébés apprenant le suédois. Les résultats ont révélé que les deux groupes de nourrissons présentent un effet magnétique significatif pour le prototype propre à leur langue maternelle. Les auteurs confirment que des changements

développementaux dans la perception des sons de la parole apparaissent à cet âge et que les nourrissons réagissent particulièrement aux voyelles de leur langue maternelle.

Jusqu'ici il a été question de traiter d'études comportementales qui abordaient la réorganisation perceptive des voyelles et des consonnes. Maintenant, la réorganisation perceptive sera abordée sous l'angle de la prosodie qui est décrite par le ton, la tonalité, l'accentuation et l'intonation (Di Cristo, 2013). Une étude par potentiel évoqué révèle que les nourrissons de 4 mois répondent spécifiquement à l'accentuation propre à leur langue maternelle (Friederici, Friedrich, et Christophe, 2007). Dans cette étude, les nourrissons français et allemands âgés de 4 mois devaient écouter une série de stimuli (bàba et babà) qui différaient selon l'accent sur la première ou la deuxième syllabe. Certains stimuli étaient fréquemment répétés (standards) et les autres stimuli (déviants) étaient peu fréquents et intercalés. Les résultats montrent que les nourrissons allemands de 4 mois évoquent une réponse électrophysiologique, la *mismatch response* (MMR), aux stimuli dont l'accent est sur la deuxième syllabe et les nourrissons français de 4 mois évoquent une MMR également, mais aux stimuli dont l'accent est sur la première syllabe. Chacun des groupes présente une réponse électrophysiologique (MMR) aux stimuli déviants relatifs à l'accent dominant propre à sa langue maternelle. Ainsi, ces différentes activations cérébrales de chacun des groupes de nourrissons suggèrent que ceux-ci auraient la capacité de se représenter en mémoire ce qui appartient à leur langue maternelle.

D'autres recherches ont étudié la réorganisation perceptive quant à la prosodie. Notamment, Höhle, Bijeljac-Babic, Herold, Weissenborn et Nazzi (2009) se sont intéressés aux langues accentuelles (*stress-timed language*). Parmi ce type de langues, il y a l'anglais, le néerlandais et l'allemand. Contrairement aux langues accentuelles, le français, l'espagnol et le

cantonais sont des langues syllabiques dont les phrases sont rythmées par les syllabes. Höhle et ses collègues (2009) se sont donc intéressés à savoir si des nourrissons apprenant l'allemand ou le français avaient une préférence pour des stimuli trochaïques (première syllabe accentuée et deuxième syllabe non accentuée) ou bien iambiques (première syllabe non accentuée et deuxième syllabe accentuée). Plus précisément, les auteurs voulaient explorer les habiletés de discrimination de deux pseudo-mots : un trochaïque (gába) et un iambique (gabá). Les stimuli étaient produits par un locuteur allemand. Les résultats de cette étude comportementale (*Headturn Preference Paradigm*) suggèrent que des nourrissons français âgés de 6 mois n'ont pas de préférence trochaïque ou iambique, mais qu'ils peuvent tout de même discriminer les stimuli. En ce qui concerne les bébés allemands, ceux-ci montrent une préférence trochaïque à 6 mois. À 4 mois, aucune préférence pour les mots trochaïques ou iambiques n'est présente chez le groupe de bébés allemands et le groupe de bébés français. Ainsi, le biais trochaïque est spécifique à la langue accentuelle et émerge chez l'enfant allemand entre 4 et 6 mois. Les auteurs concluent que le biais trochaïque semble être prédit par la structure et les propriétés rythmiques spécifiques de la langue maternelle de l'enfant, et ce, tôt dès l'apprentissage de ladite langue accentuelle. Par ailleurs, un nourrisson tant d'une langue accentuelle que syllabique utilisera les accents comme source d'indices dans le traitement du langage lorsque celui-ci est exposé à des variations systémiques interlinguistiques. Autrement dit, le nourrisson a les capacités de s'adapter aux indices phonologiques présents dans sa langue maternelle.

En 2009, Skoruppa et ses collègues ont réalisé une expérience avec la langue française (langue où l'accentuation lexicale est non contrastive) et l'espagnol (langue où l'accentuation lexicale est contrastive) afin d'investiguer le rôle de l'accentuation dans la perception de la parole chez les bébés de 9 mois. Pendant la phase de familiarisation, un groupe de nourrissons

étaient exposés à 8 stimuli (pseudo-mots de deux syllabes) avec l'accent sur la première syllabe (p.ex. 'kiba, 'nuki) et l'autre groupe à 8 stimuli dont l'accent était sur la deuxième syllabe (p.ex. ki'ba, nu'ki). Les stimuli étaient prononcés par un locuteur espagnol. Pour la phase test, les auteurs ont comparé les habiletés de perception des bébés français et espagnols avec 8 nouveaux stimuli dont 4 avaient l'accent sur la syllabe initiale (p.ex. 'lapi, 'naku, 'nila, 'tuli) et 4 avaient l'accent sur la syllabe finale (p.ex. ki'bu, lu'ta, pi'ma, pu'ki). Afin de savoir si les bébés discriminaient les nouveaux stimuli, les moyennes des temps de regard étaient enregistrées quant aux différents types d'essais de la phase test (essais similaires ou essais nouveaux). Chacun des bébés entendait deux types d'essais : des essais similaires (accentuation familiarisée) et des essais nouveaux (accentuation non familiarisée). Les résultats ont révélé que les bébés espagnols discriminaient avec succès les deux patrons d'accentuation puisque les temps de regard enregistrés étaient significativement plus longs pour les nouveaux essais que pour les essais similaires. En ce qui concerne les bébés français, les résultats ne révélaient aucune différence quant aux temps de regard enregistrés pour les deux types d'essais de la phase test, indiquant qu'ils n'ont pas discriminé les patrons d'accentuation. Les résultats suggèrent que les capacités de discrimination des patrons d'accentuation des nourrissons espagnols relèvent donc du contraste entre ceux-ci puisque l'accentuation contrastive est utilisée pour la signification des mots dans la langue espagnole. Comme les accents ne sont pas contrastifs en français, cela explique pourquoi les bébés francophones n'ont pas discriminé les contrastes de patrons d'accentuation. Les auteurs se sont donc questionnés à savoir si les bébés français avaient des difficultés quant à la perception phonétique de l'accent ou bien s'ils avaient des difficultés à discriminer les stimuli en raison d'une variabilité segmentale importante de ces stimuli. Par conséquent, les chercheurs ont réalisé une deuxième expérience avec des stimuli

présentant moins de variabilité. Pour ce faire, ils ont utilisé 24 exemplaires d'un même pseudo-mot dont 12 avaient l'accent sur la syllabe initiale et 12 avaient l'accent sur la syllabe finale (réalisés par un locuteur espagnol) ('pima et pi'ma). Les stimuli étaient semblables à l'étude d'Hohle et al. (2009), deux pseudo-mots de deux syllabes, un stimulus iambique et un stimulus trochaïque. Bien que les bébés français n'aient pas réussi à discriminer les deux patrons d'accentuation à l'expérience 1, ils ont réussi à discriminer les stimuli avec succès à l'expérience 2 puisque les exemplaires du même pseudo-mot présentaient des segments ayant moins de variabilité. Les auteurs concluent qu'à l'âge de 9 mois les habiletés de perception de l'accent sont déjà influencées par la langue maternelle. C'est pourquoi les nourrissons français dont la langue maternelle, langue où l'accentuation lexicale est non-contrastive, ignorent la dimension prosodique lorsqu'ils écoutent les mêmes pseudo-mots contrairement aux bébés espagnols dont la langue maternelle a un contraste d'accent au niveau lexical. Les résultats de cette étude suggèrent que les bébés espagnols, exposés à une langue à accentuation lexicale contrastive, développeraient donc des habiletés robustes de perception de la variabilité segmentale à l'intérieur des patrons d'accentuation alors que les bébés français n'ont pas ces capacités.

En 2013, Skoruppa et ses collègues ont réalisé une étude en utilisant la même méthode et les mêmes stimuli que dans leur étude de 2009 (Skoruppa et al.) avec des bébés français et espagnols, mais cette fois-ci âgés de 6 mois plutôt que 9 mois. La première expérience consistait à étudier les habiletés de discrimination des bébés exposés à des stimuli (24 pseudo-mots différents) de patrons d'accentuation variés (initial et final). La deuxième expérience, moins exigeante en termes de variabilité phonologique (absence de variation segmentale et exposition à 24 exemplaires du même pseudo-mot), a permis de tester les habiletés de discrimination de patrons d'accentuation variés (initial et final) chez les nourrissons. Les résultats de l'expérience

1 révèlent qu'aucun des groupes de bébés n'a discriminé les différences entre les patrons d'accentuation des pseudo-mots, alors que dans la deuxième expérience, les deux groupes ont réussi à discriminer les patrons d'accentuation de plusieurs exemplaires d'un même pseudo-mot. En tenant compte des études antérieures (Höhle, Bijeljac-Babic, Herold, Weissenborn et Nazzi., 2009 ; Skoruppa et al., 2009), les auteurs concluent que les nourrissons, jusqu'à l'âge de 9 mois, peuvent discriminer des patrons d'accentuation lorsqu'il n'y a pas de variabilité phonologique, et ce, indépendamment du type de langage appris.

Les études décrites précédemment confirment que les bébés sont influencés par leur langue maternelle et donc qu'une réorganisation perceptivo-motrice semble s'opérer sur les voyelles et les consonnes ainsi que sur la prosodie durant la première année de vie du nourrisson (Friederici, Friedrich, et Christophe, 2007 ; Höhle, Bijeljac-Babic, Herold, Weissenborn et Nazzi, 2009 ; Kuhl, Williams, Lacerda, Stevens et Lindbloom, 1992 ; Polka et Werker, 1994 ; Skoruppa et al., 2009 ; Skoruppa et al., 2013 ; Werker et Tees, 1984). Plus précisément, les changements développementaux dans la perception des sons de la parole s'opèrent vers 4 à 6 mois pour les voyelles et vers 10 mois pour les consonnes. En ce qui concerne l'accentuation, dès 4 mois, les nourrissons commencent à ajuster leurs habiletés à percevoir l'accent propre à leur langue maternelle. Vers 6 mois, les bébés qui apprennent une langue dont l'accentuation est non contrastive (par ex. le français) portent moins attention aux contrastes d'accent en comparaison aux bébés qui apprennent une langue dont l'accentuation est contrastive, comme l'espagnol. Enfin, à 9 mois, la perception de l'accent au niveau phonologique est complètement altérée par la langue maternelle du nourrisson. Par contre, même si les bébés français dont leur langage natif est une langue qui n'utilise pas l'accentuation et la prosodie contrastive, ceux-ci

peuvent discriminer différents patrons d'accentuation dans une tâche impliquant un traitement phonétique plutôt qu'un traitement phonologique.

La réorganisation perceptive et l'apprentissage statistique

D'après Saffran, Aslin et Newport (1996), une analyse des régularités statistiques présentes dans le langage permettrait de percevoir et donc de traiter le langage. L'étude de Saffran, Aslin et Newport (1996) consistait à faire entendre à de jeunes bébés anglais américains âgés de 8 mois quatre pseudo-mots de trois syllabes dont aucune information prosodique ni acoustique n'était fournie telle que l'accent ou des pauses entre les mots, qui habituellement, limitent les frontières des mots. Pendant la phase de familiarisation, les bébés entendaient un flux de parole continu (enchaînements aléatoires répétés de quatre pseudo-mots trisyllabiques différents) (p.ex. bidakupadotigolabubidaku) pendant 2 minutes et le seul indice pour segmenter les pseudo-mots était les probabilités de transition entre les paires de syllabes dont certaines étaient plus élevées. En phase test, les bébés ont entendu quatre répétitions de trois enchaînements de syllabes dont deux étaient des pseudo-mots cibles du langage artificiel préalablement présenté en phase de familiarisation et deux étaient des inter-mots du flux continu (où la probabilité transitionnelle était nulle). Les nourrissons ont réussi à effectuer une discrimination tel qu'indiqué par des temps de fixations plus longs accordés aux inter-mots. Cela indique que les bébés détiennent l'habileté à utiliser les probabilités de transitions entre des syllabes consécutives afin de reconnaître les frontières lexicales, et ce, après seulement deux minutes d'exposition aux stimuli. Cette étude est la première à avoir confirmé que les nourrissons peuvent utiliser les propriétés statistiques du langage dans la segmentation des mots.

Maye, Werker et Gerken (2002) se sont également intéressés à l'apprentissage statistique en tant que mécanisme à la base de la réorganisation perceptive. Pour ce faire, les auteurs ont manipulé les propriétés distributionnelles de stimuli langagiers dans le but d'observer si les nourrissons peuvent utiliser cette information pour apprendre à discriminer les sons. L'étude a été effectuée auprès de nourrissons anglophones âgés de 6 mois et de 8 mois. L'hypothèse était qu'une propriété acoustique (par ex. le moment du début du voisement), lorsqu'elle est distribuée de façon bimodale et reflète un contraste phonétique (par ex. voisé – non-voisé), permettrait au nourrisson de former deux catégories phonétiques lui permettant de discriminer les deux sons, alors que si la propriété acoustique est distribuée de façon unimodale, l'enfant ne formerait qu'une seule catégorie.

Afin de vérifier leur hypothèse, les chercheurs ont manipulé les propriétés distributionnelles des syllabes [da] et [ta] dans les stimuli présentés aux enfants. La différence phonétique entre ces syllabes relève du moment du début du voisement (*voice-onset time*, VOT). Les auteurs ont créé huit stimuli situés sur le continuum [da-ta] en faisant varier le VOT (les chiffres 1 à 8 sur la Figure 1 représentent différentes valeurs du VOT), pour ensuite créer deux ensembles de stimuli, l'un selon une distribution unimodale et l'autre selon une distribution bimodale de ces stimuli (Figure 1).

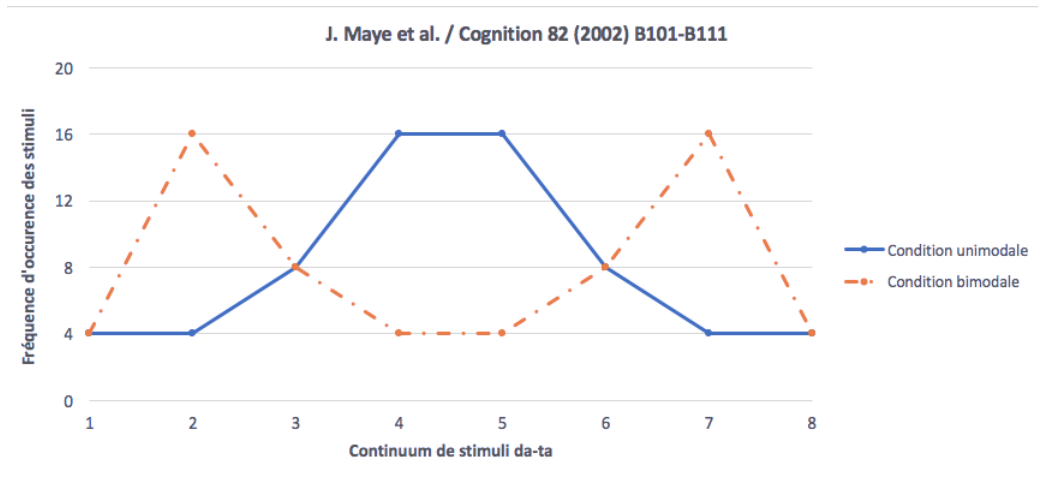


Figure 1. Continuum de l'expérience de Maye, Werker et Gerken (2002).

L'étude comprenait deux phases, une phase d'apprentissage et une phase test. Durant la phase d'apprentissage, qui durait 2,3 minutes, la moitié des poupons des deux groupes d'âge était exposée à la condition bimodale, c'est-à-dire que les nourrissons entendaient davantage les stimuli situés aux extrémités du continuum que ceux au centre du continuum. L'autre moitié des nourrissons était exposée à la condition unimodale où les stimuli présentés étaient en majorité situés au centre du continuum. Lors de la phase test, tous les nourrissons étaient exposés à huit essais de 8 stimuli, dont la moitié étaient alternants (alternance entre deux stimuli) et l'autre moitié non-alternants (répétition du même stimulus). Les quatre essais alternants étaient composés de l'alternance entre les stimuli 1 et 8 se situant aux extrémités du continuum. La moitié des essais non-alternants était composée de la répétition du stimulus 3 et les deux autres essais non-alternants étaient composés de la répétition du stimulus 6 (voir Figure 1). La procédure de regards préférentiels était utilisée pour cette expérience. Un temps de regard plus long pour un type d'essai comparativement à l'autre type signifie que le bébé aurait une

préférence pour ce type d'essai, mais aussi qu'il peut discriminer les deux types d'essais, et par le fait même les stimuli aux extrémités du continuum

Les résultats ont été analysés à partir d'une ANOVA mixte (2 conditions d'apprentissage X 2 groupes d'âge X 2 types d'essais test). Le facteur intrasujet est le type d'essai à deux niveaux (alternants et non-alternants), les facteurs intersujets sont l'âge à deux niveaux (6 et 8 mois) et le type de condition d'apprentissage à deux niveaux (condition unimodale et condition bimodale). La prédiction voulait que seuls les bébés de la condition bimodale présentent des temps de regard distincts aux différents types d'essais, indiquant qu'ils discriminaient ceux-ci et donc qu'ils auraient formé deux catégories phonétiques. Les résultats ont révélé une interaction non-significative, mais très près du seuil de signification entre le type d'essai (essais alternants et essais non-alternants) et la condition d'apprentissage (bimodale et unimodale). L'analyse des effets principaux indique un effet significatif de la condition d'apprentissage, où les temps de regard des nourrissons de la condition bimodale étaient plus longs que ceux des nourrissons de la condition unimodale, mais aucun effet significatif de l'âge des groupes et aucun effet significatif des types d'essais. En raison de l'interaction presque significative, les auteurs ont effectué des analyses par tests *t* afin de comparer les moyennes des types d'essais dans chacune des conditions. Les résultats ont confirmé l'hypothèse d'une différence dans les temps de regard quant aux types d'essais (alternants et non-alternants) chez les nourrissons de la condition bimodale, mais non de la condition unimodale. Plus précisément, les nourrissons de la condition bimodale regardaient plus longtemps les essais non-alternants que les essais alternants tandis que les nourrissons de la condition unimodale ne montraient pas de préférence entre les deux types d'essais. La phase d'apprentissage semble donc avoir influencé les habiletés de discrimination des nourrissons de la condition bimodale, résultant en une meilleure capacité

à discriminer le contraste de stimuli. Ainsi, ces résultats suggèrent que le nourrisson peut apprendre à modifier sa structure perceptive de la parole selon sa sensibilité à la distribution statistique des sons, soit par apprentissage statistique.

Dans un même ordre d'idée, Wanrooij, Boersma et Van Zuijen (2014) ont montré que l'exposition à une distribution bimodale influence et facilite la perception de stimuli langagiers chez des nourrissons allemands âgés de 2 et 3 mois. Les nourrissons de la condition bimodale, exposés à des distributions bimodale et unimodale d'un contraste de voyelle en anglais, montrent une large réponse électrophysiologique (MMR), indiquant une discrimination de ce contraste.

Les tons lexicaux

Les tons lexicaux, comme les voyelles, les consonnes et l'accentuation contrastive, sont des catégories phonétiques utilisées dans les langues tonales afin de distinguer la signification des mots. Parmi les langues tonales, il y a le mandarin, le vietnamien, le thaï, le cantonais et plusieurs autres. Le chinois mandarin est une langue tonale qui comprend quatre tons (Yip, 2002), deux tons statiques : le ton haut (T1) et le ton bas (T3), ainsi que deux tons dynamiques : le ton montant (T2) et le ton descendant (T4). Au niveau acoustique, les tons lexicaux se caractérisent principalement par des variations de la fréquence fondamentale (F0), mais aussi de l'amplitude et de la durée (Abramson, 1962 ; Chao, 1933 ; Fry, 1958). La fréquence fondamentale se définit comme le nombre de pulsations par seconde que le signal contient, chaque pulsation étant produite par une simple vibration des cordes vocales, et la fréquence de ces pulsations est mesurée en Hertz où un Hertz est un cycle par seconde (Yip, 2002). Au niveau perceptif, F0 détermine la hauteur du son et est habituellement considérée comme l'indice acoustique

principal permettant de distinguer les tons lexicaux (Abramson, 1978; Gandour et Harshman, 1978).

Réorganisation perceptive des tons lexicaux

Comme pour les voyelles et les consonnes, le phénomène de réorganisation perceptive est observé avec les tons lexicaux. En effet, un déclin de la discrimination des tons est présent chez les nourrissons apprenant des langues non-tonales. Dès l'âge de 2 mois, un nourrisson exposé à l'anglais peut distinguer les tons montants et descendants des contours de fréquence fondamentale dans le signal acoustique (Karzon et Nicholas, 1989). Kuhl et Miller (1982) ont montré que les nourrissons plus jeunes que 4 mois pouvaient discriminer les variations de F0. S'ensuivrait une sensibilité accrue aux contrastes de tons chez les nourrissons apprenant une langue tonale comparativement aux nourrissons apprenant une langue non-tonale.

L'étude de Harrison (2000) montre que les bébés apprenant le yoruba (langue tonale africaine) âgés de 6 à 8 mois discriminent mieux les tons yoruba (tons hauts, moyens et bas) comparativement aux bébés anglophones (langue non-tonale) du même âge. D'autres études suggèrent que vers 6 mois, la capacité à discriminer les tons commence à diminuer chez les bébés apprenant une langue non-tonale. Mattock et Burnham (2006) ont étudié la réorganisation perceptive chez des bébés anglophones et chinois de 6 et 9 mois en utilisant des tons thaïs. Les nourrissons étaient exposés à deux conditions de discrimination. La première condition consistait à une exposition à un contraste de tons montant versus descendant et l'autre condition exposait les nourrissons à un autre contraste de tons montant versus bas. De plus, les bébés étaient exposés à deux contextes : l'un avec la parole (syllabe [ba]) et l'autre sans parole (sons d'un violon). Les résultats indiquent que les nourrissons chinois de 6 et 9 mois présentent une

habileté à discriminer les tons aussi bien dans les deux contextes. À l'inverse, les habiletés de discrimination des tons lexicaux des bébés anglophones de 6 et 9 mois déclinent alors que les habiletés de discrimination ne déclinent pas dans un contexte avec des sons de violons.

De même, Mattock, Molnar, Polka et Burnham (2008) ont montré que des nourrissons âgés de 9 mois apprenant une langue non-tonale (français et anglais) échouaient quant à la discrimination d'un contraste de tons thaïs (contraste de tons montant versus bas) alors que des nourrissons de 4 à 6 mois discriminaient le contraste. Liu et Kager (2014) ont également observé une diminution de l'habileté à discriminer le contraste de tons T1 et T4 en chinois mandarin (contraste d'un ton haut versus descendant) vers l'âge de 9 mois chez les nourrissons néerlandais (langue non-tonale). D'autres auteurs ont observé ce même déclin perceptuel dans la discrimination des tons cantonais haut-montant versus moyen chez des nourrissons anglophones âgés de 9 mois en comparaison aux nourrissons âgés de 4 mois qui discriminaient le contraste de tons (Yeung, Chen et Werker, 2013). En effet, les bébés anglais de 4 mois réussissent à discriminer les contrastes de tons en cantonais puisqu'ils sont sensibles aux indices acoustiques (p.ex. les variations de fréquence fondamentale) (Yeung, Cheng et Werker (2013). Dans cette étude, les bébés anglais préféraient le ton 33 (ton bas montant) au ton 25 (ton haut montant) puisqu'il a une hauteur de fréquence fondamentale constante, donc plus facile à identifier en comparaison au ton 25. Un déclin perceptuel est également observé dans la discrimination des tons cantonais (ton 25 et ton 33) entre 6 et 9 mois chez des nourrissons allemands (Götz, Yeung, Krasotkina, Schwarzer et Höhle, 2018). En effet, les nourrissons âgés de 6 mois discriminent avec succès le contraste de ton tandis que les nourrissons âgés de 9 mois ne le discriminent pas. À cet égard, il semble que les nourrissons d'une langue non-tonale perdent progressivement la capacité à discriminer les contrastes de tons des langues tonales au cours de la première année

de vie, mais qu'en plus cette réorganisation perceptivo-motrice des tons lexicaux semble débiter plus tôt (4 mois) que prévu.

Selon les études citées jusqu'à présent, les bébés apprenant une langue tonale maintiendraient une sensibilité aux tons lexicaux tandis que les nourrissons n'apprenant pas une langue tonale présenteraient un déclin de la sensibilité aux tons (Liu et Kager, 2014; Mattock et Burnham, 2006 ; Mattock, Molnar, Polka, et Burnham, 2008 ; Yeung, Chen et Werker, 2013). Or, d'autres recherches ont fait différentes découvertes quant à la perception des tons chez les nourrissons apprenant une langue non-tonale. Ces derniers pourraient récupérer une sensibilité aux tons lexicaux après la première année de vie malgré la présence d'un déclin perceptif de 6 à 9 mois. C'est ce que suggère l'étude de Liu et Kager (2014) auprès de nourrissons néerlandais (langue non-tonale) de 5 à 18 mois. Dans cette étude, tous les nourrissons avaient réussi à discriminer le contraste de ton T1 et T4 en chinois mandarin. Cependant, lorsque les différences acoustiques entre ces deux tons étaient réduites artificiellement, seulement les nourrissons de 5-6 mois et de 17-18 mois ont montré une discrimination de ce contraste alors que les nourrissons de 8 à 15 mois présentaient un déclin de la discrimination de ce contraste. Cela suggère que les nourrissons de 5-6 mois montrent une sensibilité aux tons, par la suite cette sensibilité est perdue et ensuite regagnée vers 17-18 mois. De façon similaire, des nourrissons allemands (langue non-tonale) âgés de 6 mois et de 18 mois sont en mesure de discriminer des tons cantonais (ton 25 et ton 33) alors que les nourrissons âgés de 9 mois ne réussissent pas à les discriminer (Götz, Yeung, Krasotkina, Schwarzer et Höhle, 2018). Par ailleurs, l'étude de Tsao (2017) n'a pas montré de déclin dans la discrimination du contraste de tons mandarins T1 et T3 par des nourrissons anglophones (langue non-tonale) âgés de 6 à 8 mois et de 10 à 12 mois. De plus, les bébés anglophones plus âgés (10-12 mois) présentent une meilleure performance que les plus

jeunes (6-8 mois) dans la discrimination du contraste de tons T1 et T3 (Tsao, 2017). Ces recherches soutiennent que les nourrissons apprenant une langue non-tonale récupéreraient une sensibilité aux tons lexicaux après la première année de vie. De plus, ces études suggèrent que différents contrastes seraient plus faciles à discriminer que d'autres et par conséquent, selon les contrastes étudiés, les résultats diffèrent entre les études.

Ceci a amené Shi, Santos, Gao et Li (2017) à tester la perception de deux contrastes de tons en chinois mandarin, soit T2 et T3 (contraste de tons acoustiquement similaires) et T1 et T4 (contraste de tons acoustiquement différents) chez des nourrissons francophones (langue non-tonale) âgés de 4, 8 et 11 mois. Un groupe était exposé au contraste T2 et T3 et l'autre groupe au contraste T1 et T4 (chaque groupe comprenait les trois groupes d'âge). Les nourrissons étaient exposés à une phase d'habituation et une phase test. Comme les auteurs s'y attendaient, les résultats indiquent une tendance de la discrimination des tons similaires (T2 et T3) qui décline avec l'âge, tandis que le contraste T1 et T4 était discriminé chez les trois groupes d'âge. D'ailleurs, dans l'étude de Liu et Kager (2014), les auteurs avaient également utilisé le contraste T1 et T4 en chinois mandarin pour leur expérience avec des nourrissons néerlandais de 5 à 18 mois. Les résultats de Liu et Kager (2014) pointent dans la même direction que ceux de Shi, Santos, Gao et Li (2017) où tous les nourrissons avaient réussi à discriminer le contraste de tons T1 et T4. D'après Shi et ses collègues (2017), il est possible que les contrastes de tons des études énoncées dans la section précédente (ayant observées un déclin de la discrimination des tons) soient moins acoustiquement saillants que ceux de leur étude, ce qui expliquerait pourquoi leurs résultats diffèrent des autres études où des nourrissons de plus de 9 mois apprenant une langue non-tonale présentaient un déclin de la discrimination des tons, notamment des tons thaïs (contraste de tons montant versus descendant et contraste de tons

montant versus bas) (Mattock et Burnham, 2006; Mattock, Molnar, Polka et Burnham, 2008) et des tons cantonais (contraste de tons haut montant versus moyen) Yeung, Chen et Werker, 2013). De plus, Shi, Santos, Gao et Li (2017) rapportent que les procédures d'expérimentation varient d'une étude à l'autre, pouvant en partie expliquer les différents résultats.

Les travaux de Hay, Graf Estes, Wang et Saffran (2015) corroborent ceux de Shi, Santos, Gao et Li (2017). Hay et ses collègues (2015) ont réalisé différentes expériences auprès de bébés anglophones âgés de 14, 17 et 19 mois dans le but d'examiner la trajectoire développementale de leur interprétation lexicale quant aux variations de la hauteur de F0 à partir du chinois mandarin. Dans la première expérience, les bébés anglophones âgés de 14 mois étaient exposés à deux exemplaires de la même syllabe (sans signification) en chinois mandarin, soit une produite avec un ton montant (T2) et l'autre produite avec un ton descendant (T4) dans une tâche d'association syllabes-objets. Chacune des syllabes était associée à un objet à l'écran. Par exemple, la syllabe du ton montant était combinée à l'objet A. Les résultats indiquent que les nourrissons ont appris à identifier les associations syllabe-objet, et ce, malgré que les tons lexicaux ne soient pas utilisés en anglais. Ainsi, l'expérience a permis de démontrer que les nourrissons de 14 mois sont capables de distinguer le contraste de tons T2 et T4, et ce, en utilisant les contours de hauteur de F0 (indices lexicaux pertinents pour les bébés de 14 mois).

Les auteurs ont ensuite refait la même expérience, une fois avec des nourrissons âgés de 17 et une autre fois avec des nourrissons âgés de 19 mois. Les résultats révèlent que les bébés semblent ignorer les contrastes de tons dans la tâche d'association syllabes-objets. De fait, à cet âge, ils n'utiliseraient pas les variations de la hauteur de F0, ce qui voudrait dire que ce n'est pas un indice lexical pertinent pour eux. Ainsi, entre 14 et 19 mois, l'interprétation de la pertinence lexicale des variations de la hauteur de F0 des nourrissons semble changer.

Cependant, il en demeure que même après 19 mois, les nourrissons conservent la sensibilité aux contrastes de tons lorsque ceux-ci sont exposés à une tâche de discrimination seulement. C'est ce que Hay et ses collègues (2015) ont confirmé dans la dernière expérience de cette étude. Les bébés anglophones âgés de 19 mois exposés aux mêmes stimuli que les expériences précédentes, mais sans tâche d'association syllabes-objets présentaient une discrimination significative des tons montants et descendants. Ainsi, le déclin perceptif des nourrissons varie en raison de nombreux facteurs, notamment la saillance acoustique et le traitement phonologique, fortement présents dans le contraste de tons T2 et T4 de cette étude. Les résultats de l'étude de Hay et ses collègues (2015) corroborent les travaux de Skoruppa et al. (2009) où les nourrissons francophones possèdent les capacités de discriminer les différents patrons d'accentuation quand une tâche ne requiert pas de traitement phonologique.

En considérant ce qui précède, il est possible de constater que certaines études observent un déclin de la performance des nourrissons qui apprennent une langue non-tonale à discriminer les tons lexicaux tandis que d'autres études n'en observent pas. Ces différences pourraient possiblement relever de la saillance acoustique et du traitement phonologique des stimuli utilisés dans les études. Par ailleurs, la variabilité des résultats à travers les différentes études concerne les nourrissons âgés de plus de 6 mois. En revanche, toutes les études publiées jusqu'à maintenant montrent de façon consistante que les bébés plus jeunes (de moins de 6 mois) peuvent discriminer tous les tons lexicaux, peu importe s'ils apprennent une langue tonale ou non-tonale.

Les tons lexicaux et l'apprentissage statistique

L'étude de Liu et Kager (2014) cherchait à savoir comment l'apprentissage statistique influence la discrimination des tons chez les nourrissons apprenant une langue non-tonale après

le déclin perceptuel survenant vers l'âge de 9 mois. Les stimuli utilisés étaient deux tons du chinois mandarin (T1 et T4), situés sur un continuum allant du ton haut au ton descendant. Le continuum était divisé en huit stimuli qui variaient selon la fréquence fondamentale allant du stimulus 1 [ta1] au stimulus 8 [ta4] (Figure 2).

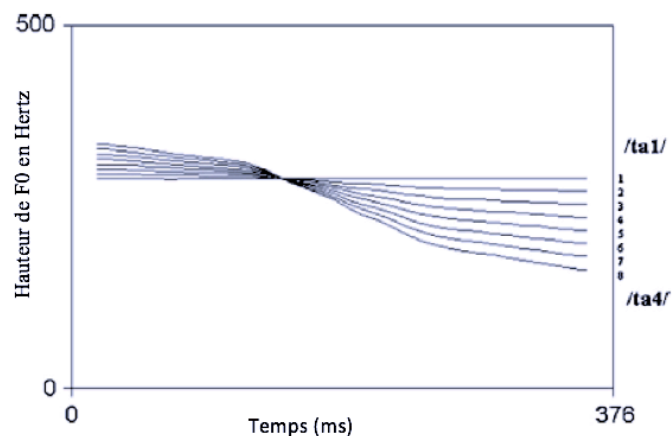


Figure 2. Continuum de l'étude de Liu et Kager (2014).

Durant la phase d'apprentissage, un groupe de nourrissons était exposé à la distribution unimodale et un autre groupe à la distribution bimodale. Les deux distributions différaient selon la fréquence d'exposition aux stimuli tonaux du continuum. La distribution bimodale présentait davantage de stimuli aux extrémités du continuum alors que la condition unimodale présentait davantage de stimuli au centre du continuum. Un contraste de tons était alors présent dans la condition bimodale, mais non dans la condition unimodale. À la suite de la phase d'apprentissage, les nourrissons étaient exposés à une phase d'habituation où des exemplaires du stimulus 6 étaient présentés à répétition. Une fois que les nourrissons étaient habitués au stimulus, la phase test les exposait à deux essais de différents exemplaires du stimulus 3. Afin de savoir si les nourrissons discriminaient les tons, les temps de regard étaient enregistrés. Les

résultats indiquent que les bébés de la condition bimodale et unimodale s'habituent à la présentation du nouveau stimulus en phase d'habituation. En comparant les moyennes de temps de regard des nourrissons entre les deux derniers essais de la phase d'habituation et les deux essais de la phase test, une interaction de la condition d'apprentissage et du changement de phase est significative. En analysant chacune des conditions d'apprentissage, les résultats révèlent que seulement les bébés de la condition bimodale discriminent le contraste de tons au changement de phase comparativement aux bébés de la condition unimodale. Ces résultats suggèrent que les nourrissons néerlandais peuvent regagner l'habileté à discriminer les tons hauts et descendants après avoir été exposés à une distribution bimodale de ces tons. Ces résultats sont compatibles avec la conclusion qu'une exposition à une distribution bimodale peut modifier significativement les habiletés perceptives des nourrissons (Maye, Werker et Gerken, 2002). D'ailleurs, en 2017, les résultats de l'étude de Liu et Kager auprès de nourrissons néerlandais suggèrent que seulement ceux âgés de 11 mois exposés à une distribution bimodale discriminaient le contraste de ton T1 et T4. À 5 mois, peu importe la condition (bimodale ou unimodale) à laquelle les nourrissons étaient exposés, ces derniers discriminaient tous le contraste T1 et T4. À 14 mois, les nourrissons d'aucune des conditions ne discriminaient le contraste (Liu et Kager, 2017). Cela dit, il n'existe pas encore de données précises et systématiques concernant l'âge de l'apprentissage statistique des tons en chinois mandarin chez les nourrissons. Malgré certaines évidences suggérant que la réorganisation perceptive débute dès l'âge de 4 mois (voir Yeung, Chen et Werker, 2013), aucune étude n'a vérifié la capacité des nourrissons âgés de 4 mois à apprendre les tons dynamiques (montant et descendant) du chinois mandarin à partir de la fréquence d'occurrence de ces tons. De plus, la façon dont

l'enfant peut traiter la variabilité du signal acoustique pour en dégager les catégories tonales demeure mal comprise.

Le mouvement de F0 comme indice acoustique

Le signal acoustique de la parole affiche une grande variabilité qui relève de différents facteurs, comme le contexte (Lieberman, Delattre, Cooper et Gerstman, 1954) et le locuteur (Johnson et Mullennix, 1997; Nearey, 1989). C'est en raison de cette variabilité que des auteurs ont cherché à trouver un indice acoustique plus stable dans la perception et l'acquisition des tons lexicaux. Gauthier, Shi et Xu (2007) ont fait l'hypothèse que les patrons de vitesse de F0 pourraient réduire cette variabilité, car ceux-ci permettent d'éliminer, entre autres, la variabilité induite par les différences entre les locuteurs. Afin de mieux comprendre comment le nourrisson peut faire face à la variabilité du signal acoustique, les auteurs ont utilisé un réseau de neurones artificiels non-supervisé afin de vérifier le type d'information que celui-ci pouvait utiliser pour apprendre à catégoriser les quatre tons du chinois mandarin (Gauthier, Shi et Xu, 2007). La syllabe [ma] était utilisée avec les quatre tons du chinois mandarin. Au total, 900 stimuli produits par différents locuteurs (donc très variables) étaient présentés au réseau lors d'une phase d'apprentissage. Dans une condition, le réseau était entraîné avec les patrons de F0 alors que dans l'autre condition, avec les patrons de vitesse de F0 (D1), obtenus à partir du calcul de la première dérivée de F0. Les résultats de la phase test, où de nouveaux stimuli étaient présentés au réseau afin d'en évaluer la capacité de généralisation à de nouveaux tons, indiquent que le réseau ayant appris à partir de D1 était meilleur pour classer les tons en comparaison au réseau exposé à F0. Ceci suggère que le patron de vitesse de F0 (D1) serait un indice acoustique important dans la perception et l'acquisition des catégories tonales.

Objectifs de l'étude

À ce jour, aucune étude comportementale n'a montré la présence d'habiletés de discrimination de contrastes de tons par apprentissage statistique chez les bébés de moins de 6 mois. Par conséquent, il est possible que les nourrissons âgés de 4 mois n'aient pas encore tout à fait développé la capacité d'utiliser l'apprentissage statistique par l'exposition à diverses distributions pour l'apprentissage de catégories phonétiques. Dès lors, il est possible que les nourrissons ne tiennent pas compte des conditions d'apprentissage formées à partir de distributions statistiques d'indices acoustiques reflétant le contraste. Enfin, d'après la littérature sur les capacités perceptives des nourrissons, il est possible de prédire que les bébés pourront à tout le moins discriminer les tons à 4 mois (Yeung, Chen et Werker, 2013) même si l'apprentissage statistique n'a pas lieu à cet âge.

La présente étude vise à explorer l'habileté des nourrissons à apprendre les tons mandarins dynamiques (montant et descendant) à partir du signal de la parole. Plus précisément, elle vise à examiner si l'enfant de 4 mois est en mesure de former une représentation des tons dynamiques du mandarin selon la prédiction du réseau de neurones artificiels proposé par les travaux de Gauthier, Shi et Xu (2007), c'est-à-dire à partir de l'utilisation des patrons de vitesse de la fréquence fondamentale (D1). Pour ce faire, une méthodologie similaire à celle de Maye, Werker et Gerken (2002) et Liu et Kager (2014) sera utilisée. Il sera question d'évaluer l'habileté du nourrisson à discriminer un contraste tonal après une exposition à des distributions d'indices acoustiques reflétant ce contraste (condition d'apprentissage). En d'autres termes, il s'agit d'évaluer si des nourrissons français âgés de 4 mois sont en mesure d'apprendre les tons lexicaux à partir d'une distribution statistique. Cette étude permettra une meilleure compréhension des mécanismes sous-jacents à la réorganisation perceptive des tons lexicaux.

Hypothèses

De manière générale, il est attendu que les bébés de 4 mois soient en mesure d'apprendre à discriminer le contraste de tons dynamiques du mandarin à partir de la distribution des indices de ce contraste présent dans le signal de la parole, étant donné les données récentes qui suggèrent que la réorganisation perceptive des tons débute à cet âge (Yeung, Chen et Werker, 2013). De plus, en se basant sur les résultats obtenus avec les réseaux de neurones artificiels, il est attendu que les bébés apprennent le contraste de tons en portant attention à D1 (Gauthier, Shi et Xu (2007).

Méthodologie

Participants

Les participants étaient 32 nourrissons (16 filles et 16 garçons) âgés de 4 mois plus ou moins une semaine ($M = 121,63$ jours, $\acute{E}.-T. = 5,17$ jours, min = 113 jours, max = 132 jours) exposés en majorité au français (96,9%). Au total, 35 bébés ont été testés. Parmi ceux-ci, trois bébés ont été retirés de l'étude, puisqu'ils pleuraient lors de l'expérience ($n = 2$) ou parce qu'il avait été exposé à une langue tonale ($n = 1$). En effet, pour participer à l'étude, les bébés ne devaient pas avoir été exposés à une langue tonale. Ce critère d'exclusion était important à respecter pour éviter de biaiser les résultats puisque les habiletés perceptives des nourrissons diffèrent selon l'exposition à une langue tonale et non-tonale. Le recrutement des participants a été réalisé par le biais des réseaux sociaux et de l'unité des naissances de l'Hôpital Sainte-Justine de Montréal. Les mères ayant donné naissance étaient rencontrées la journée du congé de l'hôpital. Après avoir accepté de donner leurs coordonnées aux chercheurs par le biais du formulaire de sollicitation qui a été remis par l'étudiante responsable de l'étude, les parents ont reçu un formulaire de consentement. Les parents ont pris connaissance de ce formulaire et ont

ensuite été contactés par téléphone par un membre de l'équipe de recherche afin de confirmer leur intérêt à participer à la recherche et convenir d'un rendez-vous. Les enfants ont été assignés aléatoirement par un logiciel de randomisation de données en ligne à l'une des deux conditions d'apprentissage (condition 1 et condition 2 – voir section suivante). L'étude a été conduite selon les normes éthiques de l'Hôpital Ste-Justine et de l'Université de Montréal.

Stimuli

Un continuum de contraste de tons qui varient à la fois sur F0 et D1 a été créé, ceci afin d'examiner la capacité des nourrissons à utiliser les patrons de vitesse de F0 pour apprendre à catégoriser les tons dynamiques du chinois mandarin. Pour la création des stimuli, la syllabe [lao] a été produite avec un ton montant par une locutrice dont la langue maternelle est le chinois mandarin. Le logiciel informatique Praat (Boersma, 2001) a été utilisé pour manipuler la fréquence fondamentale (F0) de la syllabe pour créer 10 stimuli (Figure 3).

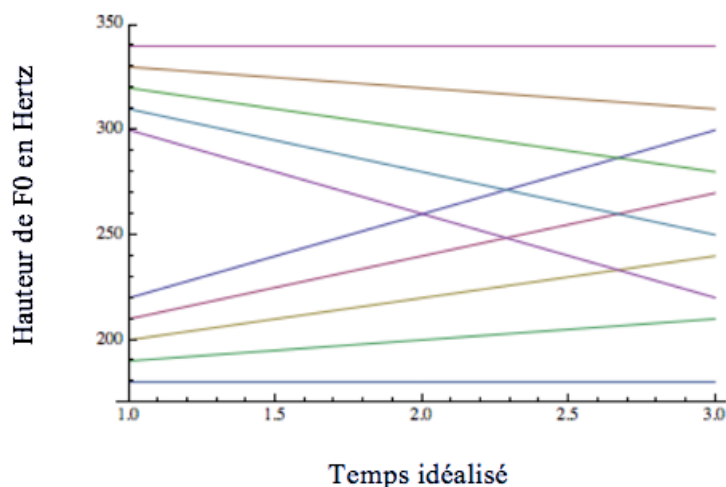


Figure 3. Patrons F0 des stimuli.

La figure 3 illustre les patrons de F0 des 10 stimuli utilisés pour l'expérience. Les valeurs en abscisse représentent le temps idéalisé et celles en ordonnée représentent la hauteur de la F0 en Hertz. Les patrons F0 à partir du bas jusqu'en haut du graphique correspondent aux stimuli 1

(ton bas) à 10 (ton haut) en fonction du continuum selon F0 (voir Figure 4). Lorsque les deux moitiés du continuum sont inversées pour représenter le continuum selon D1 (les stimuli 1 à 5 deviennent les stimuli 6 à 10 et vice versa) (voir Figure 5), les tons montants et descendants se retrouvent aux extrémités du continuum.

Les stimuli du continuum selon F0 varient en fonction de la hauteur moyenne de F0 (Figure 4), par étape de 20 Hertz, du ton bas au ton haut. Sur cette représentation du continuum, les stimuli 1 et 10 ont une F0 moyenne de 180 et 340 Hertz, respectivement, et les stimuli centraux (5 et 6) ont une F0 moyenne de 260 Hertz. Lorsque les deux moitiés du continuum sont inversées pour représenter le continuum selon D1, le continuum varie en fonction de la direction de F0. Les stimuli de ce continuum selon D1 vont du ton montant (stimulus 1) au ton descendant (stimulus 10) (Figure 5). Les fréquences initiales et finales sont respectivement de 200 et 340 Hertz pour le ton montant et de 340 et 200 Hertz pour le ton descendant. Les stimuli aux extrémités du continuum ont donc une excursion de 140 Hertz et le continuum varie en fonction de l'excursion de F0 par étape de 30 Hertz, de telle sorte que les stimuli du centre du continuum (5 et 6) ont une excursion de 0 Hertz (voir Figure 3).

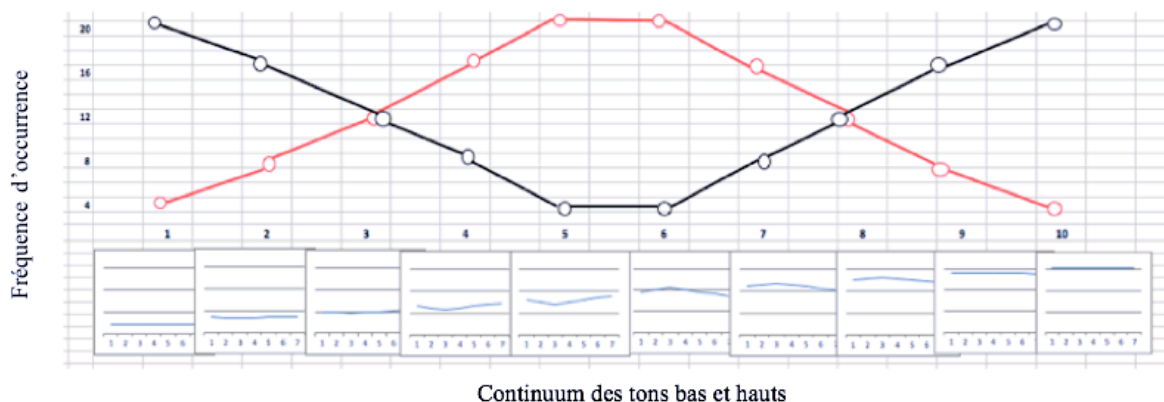


Figure 4. Continuum F0. Variation des tons en fonction de la moyenne de la hauteur de F0.

La ligne rouge de la figure 4 représente la condition 1 (distribution unimodale en F0 et bimodale en D1) et la ligne noire de la même figure représente la condition 2 (distribution bimodale en F0 et unimodale en D1) des tons hauts et bas durant la phase d'apprentissage. Le continuum des tons se trouve sur l'abscisse et la fréquence d'occurrence des stimuli durant la phase d'apprentissage se trouve en ordonnée. Le stimulus 1 correspond au ton bas et le stimulus 10 au ton haut. La F0 des stimuli est illustrée dans la série de graphiques du bas de la figure.

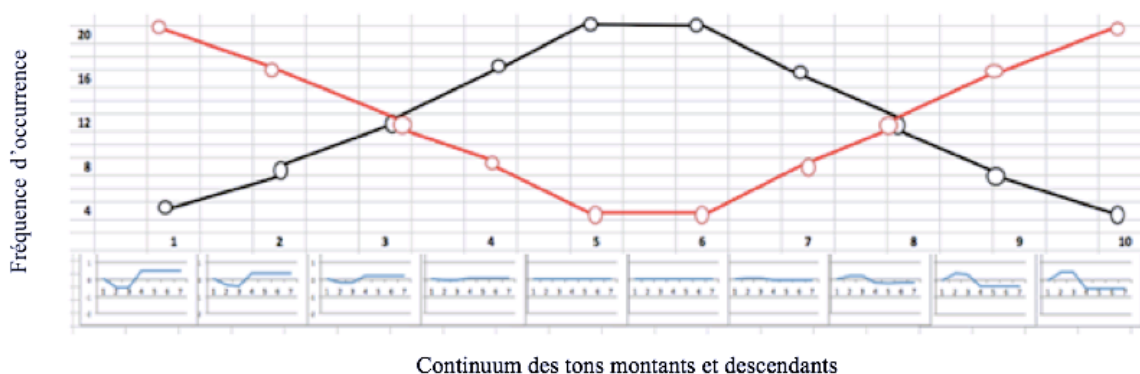


Figure 5. Continuum D1. Variation des tons en fonction des patrons de vitesse de F0.

La ligne rouge de la figure 5 représente la condition 1 (distribution bimodale en D1 et unimodale en F0) et la ligne noire de la même figure représente la condition 2 (distribution unimodale en D1 et bimodale en F0) des tons montants et descendants durant la phase d'apprentissage. Le continuum des tons se trouve sur l'abscisse et la fréquence d'occurrence des stimuli durant la phase de familiarisation se trouve en ordonnée. Le stimulus 1 correspond au ton montant et le stimulus 10 au ton descendant. Le patron de vitesse de F0 (D1) idéalisé des stimuli est illustré dans la série de graphiques du bas de la figure.

Procédure

L'expérience se déroulait dans une chambre Faraday et insonorisée du laboratoire *Neuroscience of Early Development* de l'Hôpital Sainte-Justine de Montréal. Un impédancemètre était utilisé afin d'effectuer une tympanométrie qui est une mesure de l'état de l'oreille moyenne, dont les résultats indiquent si les tympons du bébé bougent bien ou non. Quand un tympan présentait un mouvement, cela indiquait que le conduit de l'oreille n'était pas obstrué (par ex. absence de sécrétions). À l'inverse, si aucun mouvement du tympan n'était présent, il était alors possible que le conduit soit obstrué. Par conséquent, il était recommandé de consulter un médecin afin d'effectuer des tests plus approfondis, notamment détecter la présence d'une otite chez le bébé. Sans pouvoir se prononcer sur l'audition, la tympanométrie permettait tout de même d'obtenir un indice afin d'éliminer la possibilité que les bébés ne puissent entendre les stimuli de l'expérience (en raison d'otites par exemple).

Durant l'expérience, les nourrissons étaient assis sur les genoux du parent. Le parent portait des écouteurs diffusant une musique classique pour masquer les stimuli expérimentaux de façon à ce qu'il ne puisse pas influencer les comportements de l'enfant. De plus, le parent portait une paire de lunettes (lentilles teintées) afin que l'appareil d'oculométrie capte seulement le regard de l'enfant. Afin de tester la capacité des nourrissons à apprendre les tons dynamiques à partir du signal de la parole, ceux-ci ont été exposés aux différentes distributions présentées par le logiciel E-Prime.

Phase d'apprentissage. L'expérimentation débutait avec une phase d'apprentissage de 96 secondes durant laquelle une séquence de 120 stimuli auditifs était présentée à l'enfant selon la condition 1 ou la condition 2 (460 ms par stimulus et un intervalle inter-stimuli de 300 ms). Les 120 stimuli étaient répartis en 8 blocs comprenant 15 stimuli distribués aléatoirement à

l'intérieur de chaque bloc. Les stimuli ont été randomisés à l'aide d'un logiciel de randomisation de données en ligne. Les sons étaient présentés à un volume confortable par deux haut-parleurs situés de chaque côté de l'écran de télévision. Un capteur d'attention audiovisuel (sons de bulles et image colorée abstraite en mouvement circulaire sur un fond noir) était présenté au début de la phase d'apprentissage. Lorsque le bébé regardait l'écran, la phase d'apprentissage débutait et une image statique colorée et claire apparaissait (damier multicolore). La procédure de regards préférentiels était utilisée pour l'expérience et les temps de fixations visuelles étaient enregistrés par un appareil d'oculométrie *eye-tracking* par le logiciel Tobi. Afin de s'assurer de ne perdre aucune donnée, une caméra enregistrait également les réponses de fixations visuelles de l'enfant.

Durant la phase d'apprentissage, les bébés de la condition 1 étaient exposés à la distribution de fréquence bimodale selon le continuum D1 (ligne rouge de la Figure 5) (mais unimodale sur F0; ligne rouge de la Figure 4), signifiant que les stimuli des extrémités du continuum D1 (ton montant et descendant) apparaissaient plus souvent que ceux du centre (ton bas et haut). Les bébés de la condition 2 étaient exposés à une distribution de fréquence unimodale sur D1 (mais bimodale sur F0), où les stimuli des extrémités du continuum F0 (ton bas et haut) apparaissaient plus souvent que ceux du centre (ton montant et descendant). À titre de rappel, les stimuli variaient selon deux dimensions (D1 et F0). Ainsi, les nourrissons de la condition 1 étaient exposés à davantage de stimuli qui variaient sur D1 et moins sur F0 et les nourrissons de la condition 2 étaient exposés à davantage de stimuli qui variaient sur F0 que sur D1.

Phase test. Après la phase d'apprentissage suivait la phase test. Celle-ci s'appuyait sur le paradigme développé par Best et Jones (1998), dans lequel il y a deux types d'essais. Chaque

essai est composé d'une séquence de stimuli. La moitié des essais sont alternants (alternance entre deux stimuli) et l'autre moitié non-alternants (répétition du même stimulus). Ces auteurs ont montré que les nourrissons préféraient de façon significative l'alternance de deux stimuli qu'une répétition constante d'un même stimulus. Toutefois, cette préférence dépend de l'habileté des nourrissons à discriminer les sons présentés en alternance. Quand le nourrisson arrive à discriminer les types d'essais (alternants versus non-alternants) (tel qu'indiqué par un temps de regard significativement plus long pour un type d'essai), cela indique qu'il arrive à discriminer les stimuli compris dans l'essai alternant.

La phase test durait 118 secondes et comprenait 12 essais, comprenant chacun 12 stimuli de 460 ms chacun et séparés d'un silence de 300 ms. Chaque essai avait une durée de 10 secondes. D'une part, six essais étaient alternants, c'est-à-dire que les stimuli 3 et 8 étaient présentés en alternance à l'intérieur de chaque essai. D'autre part, six autres essais étaient non-alternants, c'est-à-dire que les essais présentaient la répétition du même stimulus (trois essais pour le stimulus 3 et trois essais pour le stimulus 8). Le capteur d'attention apparaissait au début de la phase test et entre les essais. Le début des essais était initié par l'enfant lorsqu'il regardait l'écran de télévision. À ce moment, une image statique représentée par un damier de couleurs vives apparaissait et l'essai débutait. Les essais étaient présentés jusqu'à la fin indépendamment si le bébé regardait ou non. À la fin de chaque essai, le damier disparaissait et l'image dynamique colorée du capteur d'attention réapparaissait. L'ordre des essais présentés variait entre un essai alternant et un essai non-alternant jusqu'au douzième essai (par ex. un essai alternant (alternance entre le stimulus 3 et 8) suivi d'un essai non-alternant (répétition du stimulus 3), ensuite suivi d'un essai alternant (alternance entre le stimulus 3 et 8) et suivi d'un essai non-alternant (répétition du stimulus 8), etc. La moitié des bébés de chaque condition entendait un essai

alternant en premier et l'autre moitié un essai non-alternant en premier. Les temps de regard vers l'écran étaient enregistrés par le logiciel Tobi et transférés dans une base de données confidentielle dédiée à l'équipe de recherche. Les données ont par la suite été analysées.

Analyses statistiques

La moyenne de temps de regard par essai pour chacun des deux types d'essais (alternants et non-alternants) a été calculée pour chaque nourrisson. En ce qui concerne l'analyse des résultats, les temps de regard des nourrissons de chacune des conditions (condition 1 et condition 2) ont été analysés par des tests t pour échantillons indépendants pour la phase d'apprentissage. Les moyennes de temps de regard pour les deux types d'essais (alternants et non-alternants) dans les deux conditions ont été analysées par des ANOVAs à mesures répétées pour la phase test. Ces analyses avaient pour but de vérifier lequel des groupes (condition 1 ou 2) a appris à catégoriser les tons lexicaux en comparant les moyennes des temps de regard entre les deux types d'essais (alternants et non-alternants) sur les 12 essais pendant la phase test. Pour l'ensemble des analyses, le seuil de signification est de 0,05 (α critique = 0,05).

Prédictions

Quatre issues sont possibles à la suite de l'expérimentation. Tout d'abord, si seulement les bébés de la condition 1, mais non ceux de la condition 2, discriminent les deux types d'essais durant la phase test, c'est qu'ils ont formé deux catégories phonétiques à partir de la distribution bimodale basée sur D1, alors que les bébés de la condition 2 n'ont formé qu'une seule catégorie à partir de la distribution unimodale basée sur D1, suggérant que les bébés sont sensibles aux variations de D1 plus qu'à celles de F0 et qu'ils peuvent utiliser l'information statistique relative aux variations de D1 pour apprendre les tons lexicaux dynamiques (montant et descendant) du

chinois mandarin. D'après cette prédiction, il serait attendu de retrouver une interaction significative entre les facteurs types de conditions et types d'essais dans une ANOVA à mesures répétées. Ensuite, des analyses par tests *t* confirmeraient l'hypothèse spécifique qu'une différence significative est présente quant au facteur type d'essais dans la condition 1.

Deuxièmement, si au contraire seulement les bébés de la condition 2, mais non ceux de la condition 1, discriminent les deux types d'essais, c'est qu'ils ont formé deux catégories phonétiques à partir de la distribution bimodale basée sur F0, alors que les bébés de la condition 1 n'ont formé qu'une seule catégorie à partir de la distribution unimodale basée sur F0, suggérant que les bébés sont sensibles aux variations de F0 plus qu'à celles de D1 et qu'ils peuvent utiliser l'information statistique relative aux variations de F0 pour apprendre les tons lexicaux statiques (haut et bas) du chinois mandarin. D'après cette prédiction, il serait attendu de retrouver une interaction significative entre les facteurs types de conditions et types d'essais dans une ANOVA à mesures répétées, suivie de tests *t* qui confirmeraient l'hypothèse spécifique qu'une différence significative est présente quant au facteur type d'essais dans la condition 2.

En ce qui concerne les prédictions 1 et 2, une interaction significative est prévue (dans chacune des prédictions) puisque les deux groupes (condition 1 et condition 2) se comportent différemment. Des analyses par tests *t* sont pertinentes pour indiquer de quelle façon les deux groupes discriminent. L'effet principal du facteur type d'essai n'est donc pas important pour ces prédictions.

Troisièmement, si les bébés de la condition 1 et 2 discriminent les deux types d'essais, c'est qu'ils ont formé deux catégories phonétiques à partir de la distribution bimodale basée sur D1 pour les nourrissons de la condition 1, alors que les bébés de la condition 2 ont formé deux catégories phonétiques à partir de la distribution bimodale basée sur F0, suggérant que les bébés

sont sensibles aux variations de D1 et de F0 en fonction de leur fréquence d'apparition dans le signal et qu'ils peuvent autant utiliser l'information statistique relative à D1 qu'à F0 pour apprendre les tons lexicaux statiques et dynamiques du chinois mandarin. En ce qui concerne les analyses, il serait attendu de retrouver une interaction non-significative entre les facteurs types de conditions et types d'essais dans une ANOVA à mesures répétées. Toutefois, il y aurait un effet principal significatif du facteur type d'essais. Il est également possible de retrouver une interaction significative entre les facteurs types de conditions et types d'essais dans la même ANOVA sans effet principal du facteur type d'essai. Dans ce cas, des analyses par tests *t* montreraient une discrimination pour les deux groupes, mais dans des directions différentes. Par exemple, il serait possible de retrouver un scénario où les bébés exposés à D1 montrent une préférence pour les essais alternants et les bébés exposés à F0 montrent une préférence pour les essais non-alternants. À l'inverse, il serait possible de retrouver un scénario où les bébés exposés à D1 montrent une préférence pour les essais non-alternants et les bébés exposés à F0 montrent une préférence pour les essais alternants. Les deux scénarios suggéreraient que D1 et F0 contribuent tous les deux aux habiletés perceptives de l'enfant, mais impliquent différents processus de traitements. Une seconde explication est possible si les bébés des deux conditions discriminent les deux types d'essais : il se peut qu'ils ne soient pas capables d'utiliser les informations statistiques des distributions bimodales ou unimodales afin d'apprendre les tons lexicaux dynamiques du chinois mandarin, mais que le système auditif du nourrisson soit tout de même en mesure de discriminer les tons par une sensibilité à un indice acoustique comme F0 par exemple.

Il est important de mentionner pour les trois prédictions précédentes que peu importe si le facteur condition est significatif ou pas cela n'a pas d'incidence sur les hypothèses.

Quatrièmement, si les bébés de la condition 1 et 2 ne discriminent pas les deux types d'essais, il se peut qu'ils ne soient pas sensibles aux variations de D1 et de F0 et cela signifie qu'ils n'utilisent pas l'information statistique relative à D1 et F0 pour apprendre les tons lexicaux du chinois mandarin et que le système auditif du nourrisson n'est pas en mesure de discriminer les tons. D'après cette prédiction, il serait attendu de retrouver aucune interaction entre les facteurs types de conditions et types d'essais dans une ANOVA à mesures répétées ainsi que l'absence d'effet principal significatif du facteur type d'essais.

Résultats

Phase d'apprentissage

Afin de vérifier si les deux groupes de bébés diffèrent entre la condition 1 bimodale D1 (équivalente à la condition unimodale F0) et la 2 condition unimodale D1 (ou bimodale F0), pendant la phase d'apprentissage, une analyse par test t pour échantillons indépendants a été effectuée pour comparer les moyennes de temps de regard des deux groupes. Les résultats révèlent qu'il n'y a pas de différence statistiquement significative des temps de regard entre les groupes, $t(30) = 0,274$, $p = 0,786$. En effet, les nourrissons de la condition 1 ($M = 38\ 917$ ms, $\acute{E}.-T. = 14\ 494$ ms) regardent pratiquement autant que les nourrissons de la condition 2 ($M = 40\ 453$ ms, $\acute{E}.-T. = 17\ 083$ ms) pendant la phase d'apprentissage.

Phase test

La moyenne de temps de regard pour les deux types d'essais (alternants et non-alternants) a été calculée pour chaque nourrisson. Les postulats de base de l'ANOVA ont été vérifiés, concernant par exemple la normalité des données. Afin de vérifier l'hypothèse suggérant qu'une exposition à D1 est plus utile pour l'apprentissage des tons chez les

nourrissons qu'une exposition à F0, différentes analyses ont été effectuées. Tout d'abord, il s'agissait de vérifier lequel des groupes (condition 1 ou 2) a appris à catégoriser les tons lexicaux en comparant les moyennes des temps de regard sur les 12 essais (6 essais alternants et 6 essais non-alternants) pendant la phase test. Pour ce faire, une ANOVA à mesures répétées à deux facteurs a été effectuée, soit un facteur intrasujet qui est le type d'essai à deux niveaux (alternants et non-alternants) et un facteur intersujet qui est le type de condition d'apprentissage à deux niveaux (condition 1 et condition 2). L'analyse révèle qu'il n'y a pas d'interaction entre la condition d'apprentissage et le type d'essai ($F(1, 30) = 0,139, p = 0,712, \eta^2 = 0,005$). Les résultats révèlent également qu'il n'y a pas d'effet entre les essais alternants et non-alternants ($F(1, 30) = 0,103, p = 0,751, \eta^2 = 0,003$) indépendamment de la condition d'apprentissage. De plus, il n'y a pas d'effet entre les conditions 1 et 2 ($F(1, 30) = 0,511, p = 0,480, \eta^2 = 0,017$) indépendamment des types d'essais.

De surcroît, afin de vérifier s'il y avait une différence entre les regroupements d'essais en blocs quant au temps de regard des nourrissons de chaque condition et le type d'essais, une autre analyse a été réalisée par une ANOVA à mesures répétées à trois facteurs, soit deux facteurs intrasujets et un facteur intersujet. Les facteurs intrasujets sont le type d'essai à deux niveaux (alternants et non-alternants) et les changements de blocs à 3 niveaux (bloc 1, bloc 2 et bloc 3). Le bloc 1 comprenait les essais 1 à 4, le bloc 2 les essais 5 à 8 et le bloc 3 les essais 9 à 12. Le facteur intersujet est le type de condition d'apprentissage à deux niveaux (condition 1 et condition 2). Tout d'abord, les résultats indiquent une interaction non-significative entre le type de condition, le type d'essai et les différents blocs ($F(2, 60) = 2,094, p = 0,132, \eta^2 = 0,065$). En outre, il n'y a pas d'interaction entre le type de condition et les différents blocs ($F(2, 60) = 0,527, p = 0,593, \eta^2 = 0,017$). Les résultats indiquent également qu'il n'y a pas

d'interaction entre le type de condition et le type d'essai ($F(2, 60) = 0,139, p = 0,712, \eta^2 = 0,005$). De plus, il n'y a pas d'interaction entre les trois blocs et le type d'essai ($F(2, 60) = 2,043, p = 0,138, \eta^2 = 0,064$). Cependant, on retrouve un effet significatif du facteur bloc ($F(2, 60) = 9,384, p < 0,001, \eta^2 = 0,238$) indépendamment du type de condition et du type d'essai. Ces résultats indiquent donc que les temps de regard diffèrent entre les 3 blocs d'essais, peu importe la condition d'apprentissage et le type d'essai présenté. Les résultats de la comparaison des moyennes de temps de regard pour chaque bloc indiquent une différence statistiquement significative au bloc 1 et au bloc 2 ($p = 0,009$) ainsi qu'une différence statistiquement significative au bloc 1 et au bloc 3 ($p = 0,003$). Une diminution des temps de regard est observée à travers la phase test : bloc 1 ($M = 7\ 170$ ms, $\acute{E}.-T. = 411$ ms), bloc 2 ($M = 6\ 306$ ms, $\acute{E}.-T. = 381$ ms) et bloc 3 ($M = 5\ 695$ ms, $\acute{E}.-T. = 420$ ms).

Malgré qu'il n'y ait pas d'interaction avec le facteur bloc, mais qu'un effet significatif de celui-ci est présent indépendamment du type de condition et du type d'essai, nous avons tout de même analysé les types d'essais dans chaque bloc à titre exploratoire seulement. À cet égard, trois ANOVAs à mesures répétées ont été effectuées pour chaque bloc en comparant les moyennes de temps de regard entre les essais alternants et non-alternants dans le but de documenter les tailles d'effet. Le facteur intrasujet est le type d'essai à deux niveaux (alternants et non-alternants) et le facteur intersujet est le type de condition d'apprentissage à deux niveaux (condition 1 et condition 2). Tout d'abord, l'interaction entre les deux facteurs (types d'essai et condition) est non-significative pour chaque bloc (bloc 1 : $F(1,30) = 0,031, p = 0,861, \eta^2 = 0,001$; bloc 2 : $F(1, 30) = 2,882, p = 0,100, \eta^2 = 0,088$; bloc 3 : $F(1, 30) = 1,406, p = 0,245, \eta^2 = 0,045$). Le facteur intersujet, le type de condition, est non-significatif pour chaque bloc (bloc 1 : $F(1, 30) = 0,021, p = 0,885, \eta^2 = 0,001$; bloc 2 : $F(1, 30) = 1,123, p = 0,298, \eta^2 = 0,036$; bloc 3 :

$F(1,30) = 0,485, p = 0,491, \eta^2 = 0,016$). Les résultats révèlent qu'il n'y a pas de différence entre les essais alternants et non-alternants pour chaque bloc (bloc 1 : $F(1, 30) = 1,787, p = 0,191, \eta^2 = 0,056$); bloc 2 : $F(1, 30) = 0,001, p = 0,976, \eta^2 < 0,001$); bloc 3 : $F(1, 30) = 3,344, p = 0,077, \eta^2 = 0,100$). En ce qui concerne le troisième bloc (essais 9 à 12), bien que les résultats ne soient pas significatifs, c'est à cet endroit qu'il y a la plus grande taille d'effet (taille d'effet moyenne). Celle-ci expliquerait 10% de la variabilité des temps de regard. Il a été possible d'observer des temps de regard plus long pour les essais alternants ($M = 5\,960$ ms, $\acute{E}.-T. = 2\,485$ ms) en comparaison aux essais non-alternants ($M = 5\,431$ ms, $\acute{E}.-T. = 2\,507$ ms) au troisième bloc. En d'autres mots, une taille d'effet moyenne a été détectée aux essais 9 à 12, une taille d'effet qui est non négligeable. De ce fait, d'un point de vue qualitatif, une possible tendance est observée chez les nourrissons, vers la fin de la tâche, à regarder vers un type d'essai peu importe la condition.

Discussion

La présente étude visait à explorer les mécanismes sous-jacents à la réorganisation perceptive des tons lexicaux chez le nourrisson. Il s'agissait d'examiner si les bébés de 4 mois étaient en mesure de former une représentation des tons statiques (haut et bas) et dynamiques (montant et descendant) du chinois mandarin à partir du signal de la parole via l'apprentissage statistique, ceci à partir de l'information relative aux patrons de vitesse (D1) de la fréquence fondamentale (F0) du signal acoustique.

Les résultats de l'étude indiquent d'abord que lors de la phase d'apprentissage les temps de regards sont équivalents pour les bébés de la condition 1 et de la condition 2. Ainsi, l'absence de différence des temps de regards entre les deux groupes lors de la phase d'apprentissage

suggère que les deux groupes ont autant porté attention aux stimuli durant la tâche l'un que l'autre. Il en découle que les résultats subséquents relatifs à la phase test ne devraient pas être affectés par les réponses des deux groupes au matériel d'apprentissage, puisque ces réponses sont similaires.

Ensuite, l'absence d'un effet principal des conditions et du type d'essai sur les temps de regard ainsi que l'absence d'un effet d'interaction entre ces variables pendant la phase test (essais 1 à 12) indiquent que les prédictions 1 (c'est-à-dire catégorisation à partir de D1) et 2 (c'est-à-dire catégorisation à partir de F0) sont inexactes. Ceci va à l'encontre de l'hypothèse générale de l'étude puisque les bébés ne semblent pas avoir porté plus attention à un indice qu'à l'autre alors que l'hypothèse générale suggérerait que les bébés auraient dû apprendre à discriminer le contraste de tons dynamique du mandarin à partir de la distribution des indices de ce contraste (D1). Les résultats soutiennent plutôt la prédiction 3, selon laquelle les deux groupes de nourrissons auraient dû apprendre le contraste de tons à partir de la distribution et de l'indice propre à leur condition. Les résultats supportent également la seconde explication de la prédiction 3, selon laquelle le système auditif du nourrisson serait possiblement en mesure de discriminer les tons par une sensibilité à un des indices acoustiques (D1 ou F0). Selon cette explication, il est possible que la durée de la phase d'apprentissage de notre étude (96 secondes) n'ait pas été assez longue afin de permettre aux bébés un temps suffisant d'exposition aux distributions des indices D1 et F0. Par exemple, la phase d'apprentissage de l'étude de Maye, Werken et Gerken (2002) durait 2,3 minutes, celle de Liu et Kager (2014) durait 3 minutes et celle de Saffran, Aslin et Newport (1996) durait 2 minutes.

Ensuite, les résultats ne montrent aucune différence entre les temps de regard des essais alternants et non-alternants dans les deux conditions, mis à part vers la fin des essais (essais 9 à

12), où les temps de regards étaient plus longs envers les essais alternants pour les bébés des deux conditions. D'ailleurs, c'est aux essais 9 à 12 que la plus grande taille d'effet a été observée. L'analyse qualitative des résultats révèle une possible tendance à pouvoir discriminer les types d'essais, en conséquence les tons qui composent ces essais, et ceci vers la fin de la tâche. Les nourrissons de 4 mois discriminaient possiblement le contraste de tons (T1-T3 ou T2-T4), mais de façon tardive durant la tâche. Cela indique la possibilité qu'un processus de catégorisation était en train de s'installer. Comme la prédiction 3 le mentionnait, si les bébés de la condition 1 et 2 discriminaient les deux types d'essais, c'est qu'ils avaient formé deux catégories phonétiques à partir de la distribution bimodale basée sur D1 pour les nourrissons de la condition 1, alors que les bébés de la condition 2 avaient formé deux catégories phonétiques à partir de la distribution bimodale basée sur F0, suggérant que les bébés seraient sensibles aux variations de D1 et de F0 en fonction de leur fréquence d'apparition dans le signal et qu'ils pourraient autant utiliser l'information statistique relative à D1 qu'à F0 pour apprendre les tons lexicaux statiques et dynamiques du chinois mandarin. Il aurait été intéressant de vérifier si le processus de catégorisation était en train de s'installer en poursuivant l'expérience avec davantage d'essais dans la phase test.

Une autre possibilité qui pourrait expliquer la discrimination réfère à l'ampleur du travail demandé aux bébés pendant la phase d'apprentissage. En effet, les jeunes bébés de 4 mois ont possiblement travaillé avec vigueur pendant l'expérience à analyser les stimuli et tenter de former des catégories phonétiques. En plus de demander aux nourrissons de découvrir la structure statistique, les stimuli de la phase d'apprentissage variaient sur deux dimensions acoustiques (F0 et D1). Ainsi, en raison de l'exigence de la tâche, il est probable qu'ils soient revenus à leurs capacités habituelles et naturelles de perception universelle pour tenter de

discriminer les stimuli (par ex. Mattock et Burnham, 2006) plutôt que les catégoriser (Maye, Werker et Gerken, 2002; Liu et Kager, 2014). La discrimination de stimuli semble être un traitement plus facile que la catégorisation de stimuli. La possible discrimination tardive des deux tons dans notre étude suggère que les nourrissons ont probablement essayé de travailler sur la distribution statistique des exemplaires même s'ils n'y sont pas complètement parvenus.

Il est aussi possible que les nourrissons de notre étude aient déjà été influencés par leur langue maternelle et que l'expérience langagière aurait affecté leur représentation neuronale, diminuant ainsi leur performance de discrimination des tons lexicaux en chinois mandarin en début de la phase test. D'après Friederici, Friedrich, et Christophe (2007), les habiletés perceptives de la parole des nourrissons conformément à leur langue maternelle se développent plus tôt qu'initialement anticipé. L'expérience langagière affecterait la représentation neuronale de l'accent des mots dès l'âge de 4 mois, que cette capacité à discriminer les accents est aussi liée aux capacités de discrimination des tons lexicaux (Friederici, Friedrich, et Christophe, 2007).

Ainsi, bien que les nourrissons de 4 mois présentent une habileté auditive à discriminer des stimuli comme les tons lexicaux (Shi, Santos, Gao et Li, 2017; Yeung, Chen et Werker, 2013), il n'est pas clair s'ils sont en mesure à un si jeune âge d'utiliser une distribution unimodale ou bimodale d'indices acoustiques pour former des catégories phonétiques grâce à l'apprentissage statistique. Des études comportementales ont montré que des nourrissons plus âgés que 4 mois sont en mesure d'utiliser l'apprentissage statistique. En ce qui a trait à l'apprentissage de catégories phonétiques, l'étude de Liu et Kager (2014) comprenait des nourrissons de 11-12 mois, celle de Liu et Kager (2017), des nourrissons de 5 mois et celle de Maye, Werker et Gerken (2002), des nourrissons de 6 et 8 mois, et les groupes de bébés

pouvaient apprendre des tons et des consonnes, respectivement, selon une distribution bimodale et unimodale. Il en est de même pour les études sur la segmentation des mots à partir de l'apprentissage statistique. Les nourrissons étaient âgés de 8 mois dans l'étude de Saffran, Aslin, et Newport (1996) ainsi que dans celle de Saffran (2001). En ce sens, la prédiction du réseau de neurones artificiels proposé par les travaux de Gauthier, Shi et Xu (2007) suggérant que le réseau D1 est meilleur pour classer les tons en comparaison au réseau exposé à F0 et plus utile à l'apprentissage des tons qu'une exposition à F0 serait possiblement envisageable avec des nourrissons plus âgés. Il serait donc possible que des nourrissons plus âgés soient en mesure d'apprendre à partir de la distribution et de l'indice propre à leur condition, soit F0 ou D1.

Par ailleurs, les nourrissons apparaissent aptes à utiliser une information acoustique puisqu'une discrimination semble présente en fin de tâche. De ce fait, les nourrissons de la condition 1, exposés à la distribution bimodale de D1 et les nourrissons de la condition 2, exposés à la distribution bimodale de F0, utiliseraient possiblement les informations acoustiques relatives à D1 ou F0 afin de discriminer le contraste de tons (T2-T4). En effet, les nourrissons sont sensibles aux indices acoustiques comme F0 et D1 et ces indices sont utiles dans la perception des tons. Par exemple, un nourrisson âgé de 2 mois exposé à l'anglais peut distinguer les tons montants et descendants des contours de fréquence fondamentale (Karzon et Nicholas, 1989). Kuhl et Miller (1982) ont montré que les nourrissons plus jeunes que 4 mois pouvaient discriminer les variations de F0. Des bébés anglais de 4 mois réussissent à discriminer les contrastes de tons en cantonais en utilisant les variations de fréquence fondamentale (Yeung, Cheng et Werker, 2013). Dans cette étude, les bébés anglais préféraient le ton 33 (ton bas montant) au ton 25 (ton haut montant) puisqu'il a une hauteur de fréquence fondamentale constante, donc plus facile à identifier en comparaison au ton 25. En manipulant artificiellement

les distributions du contraste T1 et T4, des nourrissons néerlandais de 5-6 mois et de 17-18 mois ont montré une discrimination de ce contraste de tons (Liu et Kager, 2014). Les résultats quant à la possible discrimination tardive des tons lexicaux de notre étude sont cohérents avec ceux rapportés dans les études précédentes.

Une autre explication possible de nos résultats se base sur l'étude de Tsao (2008). À titre de rappel, le T1 est le ton haut, le T2 est le ton montant, le T3 est le ton bas et le T4 est le ton descendant. Dans cette étude, des nourrissons de 10 à 12 mois apprenant le chinois mandarin discriminaient plus facilement le contraste de tons T1 et T3 que le contraste de tons T2 et T3 et le contraste T2 et T4. D'après Tsao (2008), le contraste de tons T1 et T3 est dissimilaire en comparaison aux contrastes de tons T2-T3 et T2-T4, qui sont plutôt similaires. Le contraste de tons T1 et T3 serait plus facile à discriminer puisque les tons présentent une grande différence de hauteur de F0 ainsi qu'une différence de contour de F0. À l'inverse, le contraste de tons T2 et T4 serait plus difficile à discriminer puisqu'il serait plutôt similaire en raison d'une petite différence quant à la moyenne de la hauteur de F0, malgré une grande différence quant au contour de F0 des tons (montant versus descendant) (Liu, Tsao et Kuhl, 2007). Ainsi, les nourrissons mandarins présenteraient une meilleure discrimination des tons lexicaux lorsque le contraste de tons est dissimilaire sur le plan acoustique. Possiblement que les bébés de 10-12 mois de l'étude de Tsao (2008) portaient plus attention à F0 qu'à D1 (autrement le contraste T2-T4 leur aurait paru dissimilaire et ils l'auraient discriminé). Ceci pourrait alors expliquer pourquoi les bébés avaient plus de difficultés à discriminer le contraste T2-T4 dans l'étude de Tsao (2008). De plus, les résultats de l'étude de Cheng et ses collègues (2013) révèlent qu'à la naissance une réponse électrophysiologique MMR est présente pour le contraste de tons T1 et T3 ainsi qu'une réponse P-MMR à 6 mois. En ce qui concerne le contraste T2 et T3, aucune

réponse électrophysiologique MMR n'est présente à la naissance, mais une réponse P-MMR est présente à 6 mois. Une réponse électrophysiologique MMN est présente à 12, 18 et 24 mois pour le contraste T1 et T3 chez les nourrissons mandarins tandis qu'aucune réponse MMN n'est présente pour le contraste T2 et T3 à ces mêmes âges (Cheng et al., 2013; Chen et Lee, 2018). Les résultats de ces études (de la naissance à 24 mois) suggèrent que les mécanismes sous-jacents à la discrimination des contrastes de tons T1 et T2 et T1 et T3 diffèrent possiblement en fonction des variations sur le plan acoustique de ces contrastes de tons (Cheng et al., 2013; Chen et Lee, 2018). Ceci expliquerait aussi pourquoi les bébés de notre étude, où le contraste de tons présenté en phase test est possiblement similaire, ne parvenaient pas à discriminer les essais durant les deux premiers tiers de la tâche. Toutefois, comme les deux tons composant les essais alternants ne variaient pas seulement qu'en termes de contour (montant et descendant), mais aussi de hauteur moyenne de F0 et de D1, cela expliquerait pourquoi ils y sont possiblement parvenus après une certaine exposition aux stimuli.

Par ailleurs, il est important d'aborder le côté positif des résultats où les bébés de 4 mois apprenant une langue non-tonale ont montré une tendance à discriminer les tons malgré que la tâche fût difficile et que les stimuli étaient très variables en phase d'apprentissage. D'ailleurs, plusieurs études rapportent de bonnes habiletés de discrimination des tons lexicaux chez les nourrissons d'une langue non-tonale, par exemple les travaux de Mattock, Molnar, Polka et Burnham (2008). Ces chercheurs ont montré que des nourrissons français et anglais âgés de 4 et 6 mois discriminent un contraste de tons thaïs (ton montant versus bas) en comparaison aux nourrissons de 9 mois qui échouaient la discrimination de ce contraste. Les travaux de Shi, Santos, Gao et Li (2017) appuient également nos résultats où des bébés français de 4 mois discriminaient les contrastes de tons du chinois mandarin. L'étude de Yeung et ses collègues

(2013) a montré que dès 4 mois, les nourrissons anglophones discriminent un contraste de tons cantonais. La tendance des bébés francophones de 4 mois à pouvoir discriminer les tons corrobore les habiletés de discrimination précédemment observées chez le nourrisson. Ainsi, les jeunes nourrissons apprenant une langue non-tonale montrent sensiblement de bonnes habiletés de discrimination des tons.

Conclusion

La présente étude visait à explorer l'habileté des nourrissons à apprendre les tons mandarins à partir de la distribution statistique d'indices acoustiques. Plus précisément, elle visait à examiner si l'enfant de 4 mois est en mesure de former une représentation des tons dynamiques du mandarin à partir de l'utilisation des patrons de vitesse de la fréquence fondamentale (D1). Nos résultats suggèrent qu'il n'est pas clair si les bébés de 4 mois sont capables à cet âge d'utiliser les informations statistiques des distributions bimodales ou unimodales des indices F0 ou D1 afin de catégoriser les tons lexicaux dynamiques du chinois mandarin. Afin de vérifier à quelle variation (D1 ou F0) les nourrissons sont sensibles, une autre étude où seulement un indice acoustique est manipulé sur le continuum plutôt que les deux en même temps serait nécessaire. Néanmoins, les nourrissons semblent tout de même capables de discriminer les tons dès cet âge. Le développement des mécanismes de discrimination des nourrissons semble s'opérer sous différentes façons selon l'âge, la langue maternelle, les stimuli présentés et la procédure. Il apparaît nécessaire que des recherches futures s'intéressent aux mécanismes sous-jacents à l'acquisition des catégories phonétiques avec l'électroencéphalographie par exemple. Au-delà des mesures comportementales, cette technique permettrait de mesurer l'activité cérébrale du nourrisson en obtenant des mesures plus précises. De plus, il serait intéressant d'effectuer à nouveau cette expérience, mais avec des nourrissons

plus âgés puisqu'ils seraient possiblement aptes à catégoriser les tons du chinois mandarin en utilisant les informations statistiques de distributions bimodales ou unimodales.

Bibliographie

- Abramson, A. S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments, Vol. 28(2)*. Bloomington, IN: Indiana University Research Center in Anthropology, Folklore, and Linguistics.
- Abramson, A. S. (1978). Static and dynamic acoustic cues in distinctive tones. *Language and speech, 21(4)*, 319-325.
- Aslin, R. N., Jusczyk, P. W., & Pisoni, D. B. (1998). Speech and auditory processing during infancy: constraints on and precursors to language. Dans D. Kuhn & R. Siegler (Éds), *Handbook of child psychology: cognition, perception, and language* (vol.2, pp. 147–254). New York: Wiley.
- Bertoncini, J., & De Boysson-Bardies, B. (2000). La perception et la production de la parole avant deux ans. Dans M. Kail, M. Fayol (Éds), *L'acquisition du langage: Le langage en émergence. De la naissance à trois ans* (vol. 1, pp. 95-136). Paris, France : Presses universitaires de France.
- Best, C., & Jones, C. (1998). Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behavior and Development, (21)*, 295.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-

- speaking adults and infants. *Journal of Experimental Psychology: Human perception and performance*, 14(3), 345.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International* 5:9/10, 341-345.
- Chao, Y. R. (1933). Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology*, 4, 121-134.
- Cheng, Y. Y., & Lee, C. Y. (2018). The Development of Mismatch Responses to Mandarin Lexical Tone in 12-to 24-Month-Old Infants. *Frontiers in psychology*, 9, 448.
- Cheng, Y. Y., Wu, H. C., Tzeng, Y. L., Yang, M. T., Zhao, L. L., & Lee, C. Y. (2013). The development of mismatch responses to Mandarin lexical tones in early infancy. *Developmental neuropsychology*, 38(5), 281-300.
- Dehaene-Lambertz, G. (2004). Bases cérébrales de l'acquisition du langage: apport de la neuro-imagerie. *Neuropsychiatrie de l'enfance et de l'adolescence*, 52(7), 452-459.
- Dehaene-Lambertz, G., & Houston, D. (1998). Faster orientation latencies toward native language in two-month-old infants. *Language and Speech*, 41(1), 21-43.
- Di Cristo, A. (2013). *La prosodie de la parole*. De Boeck Supérieur.

- Eimas, P. D., & Miller, J. L. (1980). Discrimination of the information for manner of articulation by young infants. *Infant Behavior and Development, 3*, 367-375.
- Eimas P.D., Siqueland E.R., Jusczyk P., & Vigorito J. (1971). Speech perception in infants. *Science, 171*, 303-306.
- Friederici, A. D., Friedrich, M., & Christophe, A. (2007). Brain responses in 4-month-old infants are already language specific. *Current Biology, 17*(14), 1208-1211.
- Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech, 1*(2), 126-152.
- Gandour, J. T., & Harshman, R. A. (1978). Crosslanguage differences in tone perception: A multidimensional scaling investigation. *Language and speech, 21*(1), 1-33.
- Gauthier, B., Shi, R., & Xu, Y. (2007). Simulating the acquisition of lexical tones from continuous dynamic input. *Journal of the Acoustical Society of America, 121*(5), EL190-EL195.
- Götz, A., Yeung, H. H., Krasotkina, A., Schwarzer, G., and Höhle, B. (2018). Perceptual reorganization of lexical tones: effects of age and experimental procedure. *Front. psychol. 9*:477. doi: 10.3389/fpsyg.2018.00477

- Graven, S. N., & Browne, J. V. (2008). Auditory development in the fetus and infant. *Newborn and infant nursing reviews*, 8(4), 187-193.
- Grieser, D., & Kuhl, P. K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, 25(4), 577.
- Harrison, P. (2000). Acquiring the phonology of lexical tone in infancy. *Lingua*, 110, 581-616.
- Hay, J. F., Graf Estes, K., Wang, T., & Saffran, J. R. (2015). From flexibility to constraint: The contrastive use of lexical tone in early word learning. *Child development*, 86(1), 10-22.
- 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 and Development*, 32(3), 262-274.
- Johnson, K., & Mullennix, J. W. (1997). *Talker variability in speech processing*. New York: Academic Press.
- Karzon, R. G., & Nicholas, J. G. (1989). Syllabic pitch perception in 2-to 3-month-old infants. *Perception and Psychophysics*, 45(1), 10-14.

- Kuhl P.K., & Miller J.D. (1982). Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants. *Perception and Psychophysics* 31, 279-292.
- Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255(5044), 606-608.
- Liberman, A. M., Delattre, P. D., & Cooper, F. S. (1958). Some cues for the distinction between voiced and unvoiced stops in initial position. *Language and Speech*, 1,153-167.
- Liberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. 1. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, 68(8), 13.
- Liu, L., & Kager, R. W. J. (2014). Perception of tones by infants learning a non-tone language. *Cognition*, 133(2), 385–394.
- Liu, L., & Kager, R. (2017). Statistical learning of speech sounds is most robust during the period of perceptual attunement. *Journal of experimental child psychology*, 164, 192-208.

- Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2007). Acoustic analysis of lexical tone in Mandarin infant-directed speech. *Developmental Psychology, 43*(4), 912.
- Mattock, K., & Burnham, D. (2006). Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy, 10*(3), 241-265.
- Mattock, K., Molnar, M., Polka, L., & Burnham, D. (2008). The developmental course of lexical tone perception in the first year of life. *Cognition, 106*(3), 1367-1381.
- Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101-B111.
- Mehler, J., Lambertz, G., Jusczyk, P., & Amiel-Tison, C. (1986). Discrimination of the mother tongue by newborn infants. *Comptes rendus de l'Academie des sciences. Serie III, Sciences de la vie, 303*(15), 637-640.
- Morse, P. A. (1972). The discrimination of speech and nonspeech stimuli in early infancy. *Journal of Experimental Child Psychology, 14*(3), 477-492.
- Nearey, T. M. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America, 85*(5), 2088-2113.

- Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human perception and performance*, 20(2), 421.
- Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, 81(2), 149-169.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Shi, R., Santos, E., Gao, J., & Li, A. (2017) Perception of similar and dissimilar lexical tones by non-tone-learning infants. *Infancy*. doi: 10.1111/infa.1219110.
- 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), 914-919.
- 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
- Tomasello, M. (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.

- Taylor, J. R. (2008). Prototypes in cognitive linguistics. In *Handbook of cognitive linguistics and second language acquisition* (pp. 49-75). Routledge.
- Tsao, F.-M. (2008). The effect of acoustical similarity on lexical-tone perception of one-year-old Mandarin learning infants. *Chinese Journal of Psychology*, 50, 111–124.
- Tsao, F. M. (2017). Perceptual Improvement of Lexical Tones in Infants: Effects of Tone Language Experience. *Frontiers in Psychology*, 8.
- Wanrooij, K., Boersma, P., & Van Zuijen, T. (2014). Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study. *Frontiers in psychology*, 5, 77.
- Werker, J. F., Gilbert, J. H. V., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, 52, 349-355.
- Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental psychology*, 24(5), 672.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, 7(1), 49-63.

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning. *Trends in cognitive sciences*, 9(11), 519-527.

Yeung, H. H., Chen, K. H., & Werker, J. F. (2013). When does native language input reorganize phonetic perception? The precocious case of lexical tone. *Journal of Memory and Language*, 68(2), 123-139.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.