



HAL
open science

La prosodie de l'enfant à l'interface de la musique et de la parole

Karine Martel, Chantal Caracci, Marie-Thérèse Le Normand

► **To cite this version:**

Karine Martel, Chantal Caracci, Marie-Thérèse Le Normand. La prosodie de l'enfant à l'interface de la musique et de la parole. *Enfance*, 2020, N° 4 (4), pp.451-473. 10.3917/enf2.204.0451 . hal-04313690

HAL Id: hal-04313690

<https://hal.science/hal-04313690>

Submitted on 29 Nov 2023

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

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

LA PROSODIE DE L'ENFANT À L'INTERFACE DE LA MUSIQUE ET DE LA PAROLE

[Karine Martel](#), [Chantal Caracci](#), [Marie-Thérèse Le Normand](#)

Presses Universitaires de France | « [Enfance](#) »

2020/4 N° 4 | pages 451 à 473

ISSN 0013-7545

Article disponible en ligne à l'adresse :

<https://www.cairn.info/revue-enfance-2020-4-page-451.htm>

Distribution électronique Cairn.info pour Presses Universitaires de France.

© Presses Universitaires de France. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

La prosodie de l'enfant à l'interface de la musique et de la parole

Karine Martel^a, Chantal Caracci^b
et Marie-Thérèse Le Normand^c

RÉSUMÉ

La musique et la parole sont des signaux sonores complexes, basés sur les mêmes configurations acoustiques que sont la durée, l'intensité et la hauteur, qui suivent plusieurs niveaux d'organisation : la morphologie, la phonologie, la sémantique, la syntaxe et la pragmatique pour la parole ; le rythme, la mélodie, et l'harmonie pour la musique. L'une des composantes les plus saillantes de la musique est sa dimension mélodique, résultant d'un ensemble de variations de « hauteur » sonore – corrélât perceptif de la fréquence – intervenant au fur et à mesure qu'un morceau se déroule. De même, pour la parole, l'une des composantes les plus saillantes est la mélodie qui, combinée au tempo et au timbre de la voix, forme une véritable partition musicale. En nous appuyant sur les données de la littérature, nous nous demanderons dans quelle mesure ces deux systèmes de communication, parole et musique, s'appuient sur des phénomènes prosodiques communs, partagés ou distincts que perçoit le bébé dans le milieu utérin et au cours de son développement. Dès le 3^e trimestre de grossesse, le fœtus est déjà capable de percevoir des rythmes qui reposent sur une organisation temporelle très régulière s'apparentant à ceux de la musique. Ensuite, le nouveau-né présente des capacités de perception de la parole relatives à des indices communs à la musique tels que l'accentuation, le rythme, le débit et les pauses. Parallèlement, le langage que les adultes adressent au bébé aide le nourrisson non seulement à parfaire ses connaissances sur les formes prosodiques du babillage, des mots et des phrases de sa langue maternelle mais aussi à exprimer ses émotions dans les aspects pragmatiques du langage.

MOTS-CLÉS : PROSODIE, MUSIQUE, PERCEPTION DE LA PAROLE, LANGAGE ADRESSÉ AU BÉBÉ (LAB OU MAMANAI), INTO-NOSYNTAXE.

^a Grhapes (EA 7287 - Groupe de recherche sur le handicap, l'accessibilité, les pratiques éducatives et scolaires), INSHEA, UPL, 58 avenue des Landes, 92150 Suresnes. *E-mail* : karine.martel@inshea.fr

^b CAP enfants, 1 allée des Barbanniers, 92230 Gennevilliers. *E-mail* : chcaracci@capenfants.com

^c (Auteur correspondant.) INSERM & Laboratoire de Psychopathologie et Processus de Santé, Université de Paris, 71 avenue Édouard Vaillant, Boulogne-Billancourt Cedex, 92774. *E-mail* : marie-therese.le-normand@u-paris.fr

ABSTRACT

Prosody at the interface of music and speech in infants and toddlers

Music and speech signals are complex sounds, based on similar acoustic patterns like duration, intensity and pitch, which follow several levels of organization: morphology, phonology, semantics, syntax and pragmatics for speech; rhythm, melody, and harmony for music. One of the most salient components of music is melody, resulting from a set of variations in pitch – correlate of frequency in perception – as a piece unfolds. Similarly, for speech, one of the most prominent components is melody which, combined with tempo and timbre of the voice, forms a real musical partition. In this literature overview, the main question is to know to what extent these two systems of communication, speech and music, are based on common prosodic phenomena, shared or distinct in infants and toddlers whereas the baby perceives in the uterine environment and during its development. From the third trimester of pregnancy, the fetus is already able to perceive rhythms based on a very regular temporal organization similar to those of music. Next, speech perception in newborn infants is related to music cues like accents, rhythm, speech-rate and pauses. At the same time, infant directed speech helps them not only to develop the prosodic forms of their native babbling, first words and early grammar, but also to refine the pragmatic aspects of language expressing their emotions.

KEYWORDS: PROSODY, MUSIC, SPEECH PERCEPTION, INFANT-DIRECTED SPEECH (IDS), INTONOSYNTAX.

INTRODUCTION

La musique comme la parole reposent sur une organisation séquentielle et hiérarchique d'événements, respectivement les notes et les mots, qui impliquent différents niveaux de représentations (phrase ou phrasé). Une façon d'examiner leurs points communs ou leurs différences consiste à les analyser à leur état initial, chez le fœtus, chez le nouveau-né puis au cours du développement chez le jeune enfant, du stade du babillage à l'apparition des premiers mots et des premières formes syntaxiques. Observer l'expérience linguistique et musicale de ces jeunes auditeurs exposés à ces deux systèmes de communication peut en effet permettre de mieux comprendre en quoi la perception de la musique et de la parole s'appuie sur des phénomènes prosodiques communs partagés ou distincts. Nous focaliserons notre réflexion sur l'organisation temporelle et mélodique de la musique et de la parole (rythme, intonation et harmonie) pendant la période fœtale et au cours du développement avec l'apparition des premiers formats syntaxiques. *In utero*, les circuits neuronaux spécialisés visant

à traiter l'information prosodique de la voix humaine sont déjà très actifs (Lecanuet & Granier-Deferre, 1993 ; Granier-Deferre, Ribeiro, Jacquet & Bassereau, 2011 ; Lecanuet, 1997, 2000). Après la naissance, le bébé est sensible à des variations fines de tonalité, de fréquence et d'intensité, ainsi qu'à des composantes musicales comme la mélodie et le rythme (Nazzi, Bertoncini & Mehler, 1998 ; Plantinga & Trainor, 2005). La prosodie règle d'ailleurs les premiers échanges conversationnels en jouant un rôle central dans la synchronie interactionnelle et la régulation des émotions (Crown, Feldstein, Jasnow & Jaffe, 2002 ; Trevarthen, 1999). Il suffit d'observer les parents parlant avec leur nouveau-né pour se rendre compte que le langage adressé au bébé (LAB) est caractéristique d'une prosodie particulière : le débit est ralenti par rapport à celui habituellement utilisé pour parler à un autre adulte, la fréquence fondamentale est élevée, il y a aussi de grandes variations de hauteurs (des pics de hauteur qui donnent la sensation de montées mélodiques extrêmes dans les plages fréquentielles aiguës), de longues pauses et des contours d'intonation répétés. Les parents exagèrent leurs intonations pour capter l'attention de leur bébé. Si en musique, la mélodie ou le « phrasé musical » est l'un des premiers paramètres que reconnaissent les nourrissons, le chant maternel et les berceuses partageraient aussi des caractéristiques transculturelles incluant des contours simples et redondants (Trehub & Russo, 2020). Puis, au moment de l'arrivée du babillage, le jeune locuteur utilise les indices prosodiques des phrases comme l'accentuation, l'intonation et le rythme pour s'approprier progressivement les mots grammaticaux en repérant les frontières syntaxiques des groupes de mots.

PERCEPTION DU RYTHME ET DE LA MÉLODIE DE LA MUSIQUE ET DE LA PAROLE DANS LE MILIEU UTÉRIN : L'EXPÉRIENCE PRÉ-NATALE

De nombreuses recherches ont montré, à l'aide de capteurs du rythme cardiaque, que les fœtus, durant le dernier trimestre de grossesse, sont capables de différencier la musique de la parole (Granier-Deferre, Bassereau, Ribeiro, Jacquet & Decasper, 2011 ; Kisilevsky, Hains, Jacquet, Granier-Deferre & Lecanuet, 2004), leur langue maternelle d'une langue inconnue (Kisilevsky *et al.*, 2009), la voix de leur mère de celle d'une locutrice étrangère (Kisilevsky *et al.*, 2003), et de reconnaître la voix du père, même si on observe un biais pour la voix maternelle (Lee & Kisilevsky, 2014). Ces études mettent en avant que, dès le début du troisième trimestre de grossesse, le système auditif est assez mature pour distinguer, notamment sur la base des indices prosodiques, deux types de sons avec une préférence pour les stimuli familiers (Lecanuet & Granier-Deferre, 1993 ; Lecanuet, 1997, 2000). Par ailleurs, plus les fœtus se rapprochent du terme, plus ils sont en mesure de discriminer de fines différences acoustiques, comme des différences de sonie, de hauteur et de spectre, qui sont nécessaires pour percevoir les contours mélodiques (Decasper, Lecanuet,

Busnel, Granier-Deferre & Maugeais, 1994 ; Granier-Deferre, Ribeiro, Jacquet & Bassereau, 2011). Ils peuvent aussi différencier des notes musicales de basse *versus* haute fréquence (Lecanuet, Granier-Deferre, Jacquet & DeCasper, 2000). L'ensemble de ces études attestent donc de l'existence de sensibilités très précoces aux propriétés prosodiques de la parole, avant même la naissance.

Ces préférences pour la voix maternelle, ou encore la langue maternelle, dans le milieu utérin, reflètent non seulement la capacité à discriminer et à traiter ces stimuli auditifs, mais également la mise en mémoire de certaines caractéristiques auxquelles les fœtus ont été exposés au cours de la grossesse. Il existe par conséquent un apprentissage pendant la vie fœtale. Le cerveau de l'enfant peut garder une trace d'un mot fréquemment entendu pendant les derniers mois de la grossesse, et il semble qu'un processus de mémorisation a déjà commencé à se mettre en place à ce stade. Ces capacités reflètent alors, probablement, la présence de puissants mécanismes d'apprentissage.

Afin de mettre ce processus en évidence, Partanen, Kujala, Naatanen *et al.* (2013) ont comparé deux groupes d'enfants, l'un exposé à l'« apprentissage » et un autre servant de groupe contrôle. Pour le premier groupe, 17 jeunes femmes enceintes ont participé à l'expérience. Chacune a reçu un CD de 4 minutes sur lequel était enregistrée une suite de trois syllabes (*tatata*), répétée plus de 400 fois. L'enregistrement comportait aussi des variantes de cette suite, dans lesquelles le *ta* du milieu était prononcé de façon plus aiguë ou plus grave, mais ces variantes étaient répétées beaucoup moins souvent ; entre les suites syllabiques était intercalée une musique non vocale. Les mères ont fait entendre le CD à leur bébé cinq à sept fois par semaine, à partir de la vingt-neuvième semaine de grossesse jusqu'à la naissance. Au moment de l'accouchement, les bébés avaient entendu la suite syllabique en moyenne 25 000 fois, et chacune des variantes environ 4 400 fois. Ils ont ensuite été testés dans les jours ou les premières semaines après leur naissance : on leur a fait entendre le *tatata* et ses différentes variantes, en enregistrant par électro-encéphalogramme (EEG) la réponse de leur cerveau. L'EEG montrait alors clairement que les enfants différençaient les suites syllabiques des variantes. Comme l'enregistrement était fait en moyenne cinq jours après la dernière fois que le bébé avait entendu le CD, l'expérience a donc mis en évidence une forme de mémorisation. Les bébés du groupe de contrôle qui n'avaient pas écouté le CD pendant la grossesse ne présentaient pas les mêmes réponses différenciées à l'EEG.

La même équipe finlandaise a aussi constaté, dans une autre étude, que des bébés, dont les mères avaient fait entendre une berceuse 5 fois par jour durant le dernier trimestre de la grossesse, testés à la naissance et à 4 mois, avaient des patterns de potentiels évoqués plus importants que d'autres bébés qui n'avaient pas été exposés à la même stimulation musicale. Les auteurs avaient en effet demandé à un groupe de mères enceintes de chanter « Brille, brille petite étoile » cinq fois par semaine au cours du dernier trimestre de leur grossesse pour les comparer à un groupe contrôle. Puis, immédiatement après la naissance et à nouveau à l'âge de 4 mois, ils avaient fait écouter aux bébés deux enregistrements de la chanson pour déterminer si un apprentissage s'était

produit. L'une des versions correspondait à la mélodie originale et l'autre comportait des notes modifiées. À l'issue des tests, les auteurs avaient constaté que les niveaux d'activité cérébrale étaient plus élevés pour la mélodie originale pour le groupe des enfants qui l'avaient entendue *in utero*. Ce résultat était constant lors des deux temps d'expérimentations. Pour ces auteurs, cette exposition intensive en période prénatale a induit l'élaboration de représentations neuronales actives pendant plusieurs mois après la naissance. Une telle stimulation au moment où les capacités auditives deviennent fonctionnelles serait déjà stockée en mémoire à long terme pour le devenir de l'enfant (Partanen *et al.*, 2013).

PERCEPTION DU RYTHME ET DE LA MÉLODIE DE LA MUSIQUE ET DE LA PAROLE DU NOUVEAU-NÉ

Une littérature abondante sur les capacités précoces des nourrissons a mis en lumière le rôle de la perception des corrélats rythmiques et mélodiques des sons de parole et de la musique. Avant même d'avoir conscience de l'existence de propositions, de groupes de mots tels que les syntagmes, les nouveau-nés accordent beaucoup d'importance aux contours intonatifs de la langue et à ses caractéristiques rythmiques.

Les études utilisant des méthodes comportementales, comme l'orientation du regard préférentiel, ont montré que les nouveau-nés sont sensibles aux composantes rythmiques du langage et peuvent discriminer entre des langues en fonction de leurs caractéristiques rythmiques, que le contraste contienne ou non leur langue maternelle (Nazzi, Bertoncini & Mehler, 1998). Les nouveau-nés ont également une préférence pour leur langue maternelle mais, cela n'a été exploré qu'en utilisant des langues appartenant à deux classes rythmiques différentes. Par exemple, la capacité de distinguer entre deux langues de la même classe rythmique comme l'anglais et l'allemand n'apparaît pas avant l'âge de 4 mois (Gervain & Mehler, 2010 ; Nazzi, Bertoncini & Mehler, 1998). Les nouveaux nés peuvent montrer une préférence pour toute langue appartenant à la même classe rythmique que leur langue maternelle. Ainsi, les bébés sont capables de distinguer deux langues l'une de l'autre, à condition que celles-ci appartiennent à des groupes rythmiques différents (Moon, Cooper & Fifer, 1993). Des recherches ont toutefois dans certains cas aussi démontré la capacité des nouveau-nés à discriminer deux langues appartenant à la même famille, à condition que l'une d'entre elles soit la langue maternelle de l'enfant (Bosch & Sebastian-Gallés, 1997 ; Nazzi, Jusczyk & Johnson, 2000).

Les études utilisant les techniques électrophysiologiques ou de neuroimagerie cérébrale ont montré que la musique et la parole partagent des ressources neuronales communes pour le traitement de la prosodie. Cette spécialisation précoce des systèmes neuronaux reposerait sur un mode de traitement spécifique. Dès la naissance, des réseaux communs seraient utilisés pour traiter la

structure, qu'elle soit musicale ou langagière (Patel, 2010 ; Fava *et al.*, 2014 ; Kotilahti *et al.*, 2014).

À l'aide d'une méthode de neuro-imagerie, la spectroscopie proche infra-rouge (NIRS), Perani *et al.* (2010) ont cherché à savoir dans quelle mesure la spécialisation des systèmes neuronaux pour le traitement de la hauteur, de la mélodie et de l'harmonie et la structure émergent des séquences musicales que l'on observe chez l'adulte est due à une exposition à long terme ou à des contraintes neurobiologiques. Pour cela, ils ont examiné la réaction de ces systèmes à la naissance, chez 18 nouveau-nés âgés de 1 à 3 jours à l'écoute d'extraits de musique tonale occidentale originaux ou modifiés (avec changements de tonalité ou dissonants). Ils ont constaté une activation au niveau du cortex auditif primaire supérieur dans l'hémisphère droit dans la condition des extraits originaux. Dans la condition où les extraits étaient modifiés, ces réponses hémodynamiques étaient significativement réduites et des activations avaient émergé dans le cortex frontal inférieur gauche ainsi que les structures limbiques. Les auteurs avaient conclu qu'il existe déjà chez le nourrisson une spécialisation hémisphérique dans le traitement de la musique à la naissance.

En 2016, Abboub, Nazzi & Gervain ont réalisé une expérience dans laquelle ils ont également employé la NIRS pour enregistrer les réponses hémodynamiques lors de la présentation à des nouveau-nés de séquences de tons groupés en paires, suivant soit le pattern iambique (contraste faible-fort ou bas-haut ou court-long), soit le pattern trochaïque (fort-faible ou haut-bas ou long-court), soit sans contraste prosodique. Les nouveau-nés exposés *in utero*, exclusivement au français ont traité différemment les patterns iambiques des patterns trochaïques en présence de variations de durée, mais pas en présence de variations d'intensité ou de proéminence de hauteur. Prises ensemble, ces données démontrent que l'expérience prénatale a un rôle important dans les sensibilités au groupement rythmique observées dès la naissance. Cette prédisposition s'expliquerait par les conditions de survie de l'enfant qui, ainsi, serait apte à s'affilier au groupe social des locuteurs natifs (Kinzler, Dupoux & Spelke, 2007). Ce serait aussi sur les caractéristiques prosodiques de la langue que les bébés s'appuieraient pour segmenter des phrases musicales (De Carvalho, Dautriche, Lin & Christophe, 2017).

Dès 6 mois, les nourrissons seraient capables de traiter la structure temporelle d'extraits musicaux, autrement dit de classer des schémas rythmiques et de les distinguer de schémas arythmiques. Pour ce faire, il semble qu'ils regroupent les événements en unités distinctes en fonction de leur placement temporel et qu'ils utilisent l'allongement final comme indice pour segmenter un flux musical en unités perceptuelles (Hannon & Johnson, 2005). Ils seraient aussi capables de distinguer des modifications de tempo (Trehub, Schneider & Henderson, 1995) et de hauteur (Plantinga & Trainor, 2005).

LE LANGAGE ADRESSÉ AU BÉBÉ (LAB) : UN RYTHME ET UNE MÉLODIE INTERACTIVE

Tous les auteurs qui se sont intéressés au langage adressé au bébé (LAB) ont observé que les adultes utilisent un registre singulier pour communiquer avec le jeune enfant en l'occurrence un langage empreint de musicalité (Kempe, Schaeffer & Thoresen, 2010 ; Kitamura & Burnham, 2003 ; Kitamura, Thanavithuth, Luksaneeyanawin & Burnham, 2002 ; Liu, Tsao & Kuhl, 2009). Les adultes, mais aussi les enfants plus âgés, utilisent en effet des formats de phrases syntaxiquement simplifiés en exagérant la prosodie de la langue notamment les caractéristiques mélodiques et rythmiques (Papoušek, Papoušek & Symmes, 1991 ; Smith & Trainor, 2008). Sur le plan segmental d'une part, à travers une articulation marquée, l'emploi d'un vocabulaire désignant la plupart du temps des entités concrètes et faisant partie du contexte de communication immédiat, incluant des diminutifs, des onomatopées. D'autre part, le LAB se caractérise par des propriétés suprasegmentales spécifiques telles qu'une élévation de la hauteur de la voix, un tempo ralenti, des pauses vides plus fréquentes et plus longues, et des allongements de la syllabe finale des derniers mots des syntagmes et des énoncés produits (Fernald & Mazzie, 1991 ; Kemler, Hirsh-Pasek, Jusczyk & Cassidy, 1989 ; Masataka, 1992 ; McRoberts & Best, 1997 ; Vosoughi & Roy, 2012).

Le fait de parler plus lentement et de simplifier la syntaxe des énoncés est un moyen de s'adapter aux capacités cognitives et linguistiques des bébés qui ne sont pas encore totalement efficaces dans le traitement du discours, mais déjà de remarquables partenaires de communication. Selon Trainor & Desjardins (2002), les contours intonatifs exagérés du LAB aident les nourrissons à reconnaître et acquérir les catégories de voyelles de la langue. Les prééminences, comme par exemple les pics de hauteur ou les allongements de syllabes, contenues dans le LAB, servent aussi à attirer l'attention et à transmettre les affects. Plusieurs expériences ont mis en évidence la préférence des bébés pour écouter du LAB par rapport au langage adressé aux adultes (Fernald, 1985 ; Fernald & Kuhl, 1987 ; Panneton Cooper, Abraham, Berman & Staska, 1997) en montrant par exemple que les nourrissons de 4 mois, contrairement à ceux âgés de 1 mois, préfèrent la voix de leur propre mère produisant du LAB à tout autre locuteur inconnu. Le ralentissement du débit de parole semble, quant à lui, améliorer la reconnaissance des mots par les bébés de 19 mois (Graf Estes & Hurley, 2013 ; Song, Demuth & Morgan, 2010).

La qualité et la quantité du LAB fluctuent au cours de la première année de vie du bébé, sans doute en réponse à l'évolution de ses progrès comportementaux (Henning, Striano & Lieven, 2005). Le LAB évolue parallèlement aux progrès cognitifs et socio-langagiers de l'enfant et ses fonctions changent au cours du temps, relativement aux exigences des adultes vis-à-vis de leur enfant. L'enfant, lui, enrichit ses vocalises et se spécialise progressivement pour la

prosodie de sa ou de ses langues selon qu'il grandit dans un environnement monolingue ou plurilingue (Spinelli, Fasolo & Mesman, 2017).

LA SYNCHRONIE DE LA PAROLE COORDONNÉE ET PARTAGÉE : UNE HARMONIE MÉLODIQUE ADULTES-BÉBÉS

Cet intérêt précoce pour les caractéristiques prosodiques, notamment les contours intonatifs, transparaît dans la façon dont ils sont mis en scène lors des interactions précoces. On a longtemps considéré le nouveau-né comme étant passif dans la relation à autrui, et envisagé que seule la mère guidait la conversation. Or, un certain nombre d'auteurs ont démontré l'existence d'une synchronisation coordonnée et harmonieuse entre adultes et bébés lors de leurs interactions même les plus précoces (Feldman, 2007). Chaque partenaire adapte ses actions en fonction du comportement et des sollicitations de l'autre. Ils régulent mutuellement les intérêts et les émotions de l'un et de l'autre en réalisant des imitations et des signaux multimodaux notamment vocaux (Bateson, 1975, 1979 ; Fogel, 1993a ; Fogel & Hannan, 1985 ; Stern, Jaffe, Beebe & Bennett, 1975 ; Stern, Beebe, Jaffe & Bennett, 1977 ; Weinberg & Tronick, 1994). C'est ce que Trevarthen (1977) appelle l'*intersubjectivité primaire*. Elle se met en place pendant le premier semestre de vie de l'enfant. Au début de sa vie, le bébé est très attiré par la communication avec d'autres êtres humains, il ne s'intéresse que peu aux objets. Elle conduit ensuite à l'*intersubjectivité secondaire* vers 9-12 mois, au cours de laquelle le bébé est capable d'intégrer un objet dans la relation avec une personne. Cette théorie fait écho aux travaux sur l'attachement, auquel le parent et l'enfant participent de façon synchrone. Cette synchronie est un élément très important de la relation mère-bébé ; c'est la cohérence temporelle que l'on retrouve entre les actes, les affects et les rythmes biologiques de la mère et du bébé (Fogel, 1993b ; Tronick, 1989). Cette capacité est présente depuis la naissance. La synchronie ne va pas sans le concept de réciprocité, qui est le fait que les stimulations du parent entraînent des réactions chez le bébé, et que celles-ci modifient et guident les attitudes du parent (Gratier, 2003 ; Gratier & Trevarthen, 2007 ; Gratier & Bertau, 2012 ; Rochat, Querido & Striano, 1999).

Les interactions précoces sont essentielles au développement du lien entre une mère et son bébé, et ce dernier y participe pleinement. Cette participation s'observe de façon importante au niveau vocal. Trevarthen (1999) a montré qu'en interaction face-à-face, les mères et les nourrissons donnent de la cohérence à leur production vocale de façon temporelle et mélodique et qu'ils utilisent des règles de proto-conversation, comme le *turn-taking* (tour de parole). Les interactions précoces sont constituées de vocalisations entrecoupées de pauses courtes et comprises entre deux silences plus longs. Elles ressemblent ainsi aux dialogues adultes, en termes d'organisation. Mais même s'ils utilisent

aussi bien l'un que l'autre ces règles, il reste des différences dans leur façon de participer à l'interaction. Les interactions semblent plus initiées par les parents que par les bébés de 6 mois (Brundin, Rijdhholm & Larson, 1988). La fréquence des vocalisations serait plus importante chez les parents. Gratier & Devouche ont montré, en 2011, que, de manière générale, les mères vocalisent trois fois plus que leur bébé âgé de 3 mois. En revanche, les tours de parole ne sont pas différents en longueur suivant qu'ils ont été initiés par les enfants ou les parents. En ce qui concerne le répertoire vocal des bébés, il est constitué de contours unitonaux et unidirectionnels, soit montants, soit descendants (Snow, 1998, 2002, 2006). Les mères, quant à elles, possèdent un répertoire plus complexe avec des contours montant, descendant, en cloche, en U, sinusoïdal, complexe, même si elles utilisent le plus souvent des contours unidirectionnels. L'objectif est d'obtenir l'attention du bébé et de la garder le plus longtemps possible pendant l'interaction sociale (Garnica, 1977). Les mères utilisent des contours prosodiques spécifiques pour exprimer différentes intentions de communication. Les contours intonatifs montants servent plutôt à obtenir l'attention du bébé et à encourager une réponse de sa part (Ferrier, 1985 ; Stern, Spieker, Barnett & MacKain, 1983). Les contours en cloche sont surtout utilisés par les mères, pour maintenir l'attention du bébé, une fois qu'elle a été obtenue (Stern, Spieker & MacKain, 1982). Cette exagération des caractéristiques vocales se retrouverait également dans d'autres modalités sensorielles, au niveau des gestes et des regards adressés à l'enfant (Lickliter, Bahrick & Vaillant-Mekras, 2017). Cela peut servir de soutien aux vocalisations maternelles. Dans le cadre des proto-conversations, l'adulte est incité à faire des vocalisations musicales, accompagnées d'expressions faciales animées et des gestes adaptés (Stern, 1974 ; Trehub, 1990). Papoušek *et al.* (1991) ont montré que les vocalisations maternelles s'harmonisent avec la tonalité des vocalisations du bébé. Cette synchronie des vocalisations du parent en réponse à celles du bébé participe à la maturation des vocalisations de l'enfant et à une coordination vocale « suffisamment bonne », elle-même corrélée à la qualité de l'attachement (Beebe, Jaffe, Lachmann, Feldstein, Crown & Jasnow, 2000). Cette synchronisation ne se fait pas seulement dans le sens du parent vers l'enfant, elle est bidirectionnelle. Déjà à 6 semaines, un bébé accorde le rythme de ses regards au rythme des vocalisations de l'adulte, tandis que celui-ci accorde ses paroles aux regards de son enfant (Crown, Feldstein, Jasnow & Jaffe, 2002). Le regard du bébé fonctionnerait comme un signal de sa disponibilité à l'interaction. Les nourrissons qui regardent plus leur mère reçoivent plus de réponses tactiles et de plus longues vocalisations que ceux qui évitent de la regarder (Noble, Shafaie & Self, 1982). À 4 mois, le bébé module déjà le rythme de ses vocalisations, en termes de prise de parole et de pauses, à celui des vocalisations et verbalisations de l'adulte (Feldstein, Jaffe, Beebe, Crown, Jasnow, Fox & Gordon, 1993 ; Papoušek & Papoušek, 1977 ; Trevarthen & Aitken, 2003). Plusieurs études ont mis en évidence que les comportements du bébé ont une influence sur la parole maternelle. En effet, dans un développement typique, les réactions positives du bébé aux vocalisations de sa mère améliorent l'amplitude des courbes de la

prosodie maternelle (Burnham, Kitamura & Vollmer-Conna, 2002). La présence et la qualité des réponses de l'enfant augmentent la qualité du LAB (Smith & Trainor, 2008). Werker et McLeod (1989) parlent ainsi d'une boucle interactive complète, car le bébé s'engage affectivement et vocalement pour répondre à sa mère et cela le rend plus réactif, et donc intéressant pour l'adulte, afin de continuer cette interaction. L'échange de vocalisations entre la mère et son bébé est donc très important pendant une interaction.

Les interactions précoces ne se limitent pas à des échanges vocaux/verbaux. Les adultes, souvent amusent, apaisent ou endorment les enfants à l'aide de berceuses et de comptines chantées. Plusieurs études ont analysé l'intérêt des bébés pour le chant maternel (Corbeil, Trehub & Peretz, 2013 ; Nakata & Trehub, 2004 ; Trehub, Unyk & Trainor, 1993 ; Tsang, Falk & Hessel, 2017). Si la parole maternelle est plus efficace que le chant maternel pour capter l'attention du nourrisson, elle l'est moins pour la maintenir dans le temps, tandis que les répétitions dans le chant réguleraient le niveau d'éveil et d'attention des nourrissons (Trehub, 2001). Cette propriété a aussi été observée auprès de bébés nés prématurément, avec un très faible poids de naissance (Shenfield, Trehub & Nakata, 2003). Scola, Bourjade & Jover (2015) notent quant à elles une réduction des mouvements des membres inférieurs chez les bébés entre 5 et 9 mois, lors d'épisodes interactifs chantés avec la mère. Ce phénomène pourrait refléter soit un intérêt prononcé pour ce chant, soit les effets relaxants du chant maternel. Il pourrait aussi s'expliquer par l'existence d'un biais pour les émotions à valence positive et leur perception à travers les émotions transmises par la musique adressée aux bébés (Delavenne, Gratier & Devouche, 2013, Gratier & Devouche, 2020). Les chants maternels tels que les berceuses, souvent porteurs d'expressions intimes, joyeuses, retiendraient alors davantage l'attention des bébés (Nawrot, 2003). Les premiers échanges entre le bébé et les adultes de son entourage proche se présentent donc comme de véritables partitions musicales, dans lesquelles chacun s'ajuste, avec une extrême finesse, au rythme, aux prises de parole et aux sollicitations de l'autre (Gratier *et al.*, 2015 ; Keller, 2008 ; Trevarthen, 1999). Le LAB ainsi que le chant adressé au bébé posséderaient des propriétés communes pour donner du sens aux situations partagées entre les adultes et les nourrissons avant que ce dernier ne maîtrise le langage et se distingueraient au fil du temps en deux types d'oralité ; avec un usage spécifique et évolutif de la prosodie de la parole pour intégrer le modèle de la langue de l'environnement.

Gerry, Unrau et Trainor (2012) ont cherché à savoir si la formation musicale entre parents et enfants améliorerait le développement socio-émotionnel et communicationnel de bébés âgés de 6 à 12 mois. Pour ce faire, ils ont observé 34 enfants répartis en trois groupes : un premier groupe pratiquait le chant à la crèche et avec les parents (chansons et berceuses), un deuxième groupe était invité à écouter des morceaux de musique de façon passive (musique sans paroles) et un troisième ne pratiquait ni n'écoutait de musique ou de chants. Il est ressorti qu'une expérience musicale active favorise le développement de gestes communicatifs prélinguistiques et de comportements sociaux. Les résul-

tats laissent supposer non seulement que les nourrissons peuvent participer à une formation musicale significative lorsque des approches pédagogiques appropriées sont mises en place, mais aussi qu'une initiation musicale au cours de la petite enfance a un réel impact sur le développement socio-communicatif.

Zhao et Kuhl (2016) ont testé les effets d'une intervention musicale contrôlée en laboratoire sur le traitement neuronal de la musique et de la parole chez 47 nourrissons de 9 mois. Les sujets étaient répartis en deux groupes, l'un participant à des activités de musique (groupe expérimental), l'autre à des activités de jeux avec un adulte (groupe contrôle) pendant 12 séances de 15 minutes (étalées sur environ 4 semaines). L'intervention consistait à faire écouter des rythmes musicaux à trois temps, en l'occurrence des extraits de valses. Les bébés du groupe contrôle étaient invités à partager des jeux sociaux sans musique. À la fin des séances, le traitement neuronal de la structure temporelle par les nourrissons a été mesuré grâce à la technique de magnétoencéphalographie à la fois pour la parole et pour la musique. Les sujets du groupe expérimental présentaient une activation des régions corticales auditives temporales et préfrontales significativement plus importante en réponse aux violations de la structure temporelle musicale à laquelle ils s'étaient adaptés. Des résultats similaires étaient constatés pour la parole. Les auteurs en ont alors conclu que l'intervention avait amélioré la capacité des bébés à extraire des informations sur la structure temporelle des stimuli et à anticiper les événements dans le temps aussi bien pour la musique que la parole. En outre, ils ont suggéré la possibilité que les mécanismes qui sous-tendent l'acquisition du langage pendant cette période sensible ne sont sans doute pas exclusifs aux entrées vocales de la parole, mais aussi dédiés aux stimuli musicaux.

L'INTERFACE PROSODIE/BABILLAGE : UN PHRASÉ MUSICAL EN HARMONIE AVEC LA LANGUE MATERNELLE

La structure syllabique et rythmique de la langue parlée dans l'environnement va moduler les différents babillages des enfants comme si l'on observait des partitions musicales en harmonie avec la langue maternelle. Les enfants manifestent une préférence perceptuelle pour des suites syllabées qui respectent les propriétés prosodiques de la langue parlée de leur environnement (l'accentuation, le rythme, le débit, les pauses, l'intonation). Chacune de ces propriétés prosodiques se manifeste par des variations au niveau de la hauteur, de l'intensité et/ou de la durée.

Du point de vue de la hauteur tonale, Whalen, Levitt et Wang (1991) en comparant les contours prosodiques du babillage d'enfants français et anglais, suivis entre 5 et 13 mois, ont trouvé que le pattern d'intonation dominant chez les enfants anglais est un contour tombant, confirmant ainsi des résultats de Kent et Murray (1982). Chez les enfants français, en revanche, ils ont trouvé

autant de contours montants que de contours tombants. Un résultat en accord avec les différences prosodiques constatées entre les deux langues (Delattre, 1961).

Du point de vue de la durée, Levitt et Wang (1991) montrent que les enfants français produisent davantage d'allongements finals, entre 4 et 13 mois que les enfants américains du même âge. Les enfants français diffèrent des anglophones en ce qu'ils produisent des « énoncés » plus longs (en nombre de syllabes) et que l'allongement final y est encore plus marqué. Snow (1994) et Konopczynski (1986) retrouvent aussi ce phénomène d'allongement final respectivement chez des enfants anglais et français, juste au début de la période lexicale. L'étude interlangue de DePaolis et collaborateurs (2008) pointe également cette tendance, avec des différences entre le français et l'anglais d'une part, le finnois et le gallois d'autre part.

Les études interlangues montrent que les suites syllabiques du babillage de 8-9 mois se conforment, souvent, aux patterns intonatifs propres à la langue. Ainsi, l'enfant francophone préférera un rythme iambique au rythme trochaïque, qui aura à l'inverse la préférence de l'enfant anglophone. Vers 9-12 mois, l'enfant contrôle les modulations mélodiques de sa voix à des fins linguistiques. Le babillage s'enrichit encore en suites syllabées, longues et intonées, adaptées à la situation communicationnelle (babillages de surprise, de joie, de charme). Vers 12-16 mois, il y a une stabilisation de la voix au même moment qu'il commence à entrer dans le lexique. Le système accentuel se stabilise (l'accent final est définitivement installé à 16 mois chez l'enfant français, réalisé d'abord par un allongement syllabique). L'enfant sait utiliser la prosodie pour questionner, déclarer, s'exclamer, ordonner.

Dans une autre étude, Hallé, de Boysson-Bardies et Vihman (1991) ont analysé certains aspects prosodiques des productions dissyllabiques (babillage et mots) de cinq enfants français et cinq enfants japonais d'environ 18 mois. Hallé et ses collègues ont étudié les durées des voyelles. D'après les résultats, il est ressorti que ces durées sont clairement spécifiques à la langue maternelle de l'enfant. Les contours de la fréquence fondamentale (Fo) ont été aussi analysés. Les auteurs ont retrouvé pour les enfants français des contours de Fo montants et un allongement final, et pour les enfants japonais des contours de Fo descendants et une absence d'allongement final. D'après les données, il apparaît donc que les productions reflètent les patterns de Fo des adultes. Les productions dissyllabiques produites par les enfants japonais sont semblables aux formes adultes non seulement en termes d'intonation globale, mais aussi en termes de tons et de durée.

Davis *et al.* (2000) ont conduit une étude consistant en la comparaison des séquences dissyllabiques de quatre enfants normaux entre 7 et 12 mois et cinq adultes américains, portant sur la Fo, l'intensité et la durée des syllabes accentuées et non accentuées. Le but était d'étudier les propriétés acoustiques des mots dissyllabiques perçus comme différemment accentués dans le babillage et d'explorer les contraintes physiologiques dues à la position des syllabes. Les résultats n'avancent aucune différence significative entre les trois corrélats

acoustiques (Fo, durée, intensité) pour les items perçus uniformément accentués. À l'opposé, des différences significatives sont relevées pour les items asymétriquement accentués, entre syllabes accentuées et non accentuées pour les trois corrélats. Pour les items adultes, les différences sont significatives. Chez les enfants, la durée plus longue de la deuxième voyelle varie en fonction de l'accent, mais pas de façon significative. La Fo des voyelles fluctue aussi en fonction de la position et de l'accent : la Fo est plus haute dans les voyelles initiales et dans les syllabes accentuées, significativement pour l'accent, mais pas pour la position. On observe également ces mêmes variations pour l'intensité, et les deux effets (position et accent) sont significatifs.

D'autres auteurs (Nathani, Oller & Cobo-Lewis, 2003) encore ont voulu vérifier le pattern d'allongement final, pour voir si c'est un effet biologique (comportement inné) ou acquis par l'enfant. Ils ont enregistré huit enfants normaux de 1 an et huit enfants sourds de 4 ans à plusieurs étapes du développement vocal prélinguistique : précanonique, canonique, post-canonique. Bien qu'ils aient constaté une variabilité interindividuelle considérable, les auteurs ont conclu que la surdité peut affecter de façon significative les aspects temporels de la production des enfants, ce qui laisserait supposer que l'allongement final est bien acquis par l'enfant.

Vihman, Nakai et DePaolis (2006) ont réalisé une étude sur plusieurs langues et montré que les enfants utilisent leurs connaissances des patterns segmentaux et de la prosodie de la langue de leur environnement pour produire du babillage ou des mots conformes à ces patterns. Ils ont enregistré du babillage dissyllabique et des mots identifiables produits par cinq enfants américains, français et écossais, de 10 à 18 mois. Ils ont aussi enregistré des productions dissyllabiques d'adultes (mots et non-mots), productions choisies pour leurs ressemblances segmentales et phonotactiques avec celles des enfants. Le but de ce travail était d'étudier les patterns rythmiques (durées) dans ces trois langues, qui ont différents systèmes d'accentuation. Les consonnes étudiées étaient des occlusives. Les résultats ont mis en avant des différences significatives dans les durées des productions voyelle-consonne-voyelle (VCV) chez les adultes : en anglais, la première voyelle (V1) est plus longue, en français la seconde voyelle (V2) est plus longue et en écossais, la consonne (C) est plus longue. Un seul enfant anglais, deux Écossais et trois Français produisaient des éléments avec des durées correspondant à la langue adulte. Ces différences n'étaient cependant pas significatives. Vihman et ses collègues ont par ailleurs retrouvé un changement entre les deux périodes repères (10 *versus* 18 mois). À 18 mois, les patterns sont plus proches de ceux de la langue adulte. En anglais, V1 devient plus longue et C plus courte ; en français, V1 devient plus courte et V2 plus longue ; et enfin en écossais, V1 devient plus courte et C ainsi que V2 deviennent plus longues. Les enfants français et écossais sont plus proches du modèle adulte plus tôt que les enfants anglais et la différence est significative. En conclusion, les auteurs ont montré que les enfants peuvent être assignés à la bonne langue sur la seule base du pattern rythmique. Vers 10/14 mois, ils ont l'allongement final. Les Français semblent plus homogènes dans

leurs productions, mais la différence n'est pas significative. Au fil du temps, les trois groupes se rapprochent du modèle adulte.

L'INTERFACE PROSODIE/MOT

L'interface prosodie/mot fonctionnerait comme un mécanisme de filtrage entre l'input et l'acquisition aidant l'enfant à détecter les unités ou chaînes acoustiques dotées d'informations structurelles pertinentes. De nombreuses études, comme celle de Cooper et Paccia-Cooper (1980), ont commencé à montrer le rôle de la prosodie dans l'émergence des premiers mots (Morgan & Newport, 1981 ; Gleitman & Wanner, 1982). Le premier aspect remarquable de la forme sonore des premiers mots est la continuité avec les formes du babillage. Les comparaisons interlangues montrent cette continuité de façon éloquente. Toutes les tendances relevées dans les formes du babillage qui sont spécifiques à la langue de l'environnement se maintiennent voire s'accroissent. Les différences interlangues se renforcent et se stabilisent.

La relation entre le prélinguistique et le linguistique existe dans les capacités d'apprentissage des premiers mots, notamment au niveau métrique et rythmique. Les patterns principaux se retrouvent, justement, dans les deux stades et témoignent d'une évolution progressive. Cela induirait que l'enfant apprend à maîtriser ce système rythmique de l'input avant les autres éléments, ou qu'il s'agit d'un des premiers éléments qu'il est capable de relier entre perception et production (Konopczynski, 1990, 1991 ; Vihman, 1996). Ces patterns rythmiques vont servir de cadre à l'insertion et à la restitution des configurations de sons les plus fréquents. A ces productions s'ajoute un critère important, celui du contour prosodique qui semble correspondre, depuis plusieurs mois déjà, en interaction, à des modalités de phrase. Ces productions prélinguistiques ont, comme les productions linguistiques, les premiers mots, des contours intonatifs associés correspondant aux mêmes modalités que durant la période linguistique (Vihman & Miller, 1988).

L'INTERFACE PROSODIE/SYNTAXE : LA PROSODIE DES PHRASES ET LES MOTS GRAMMATICaux

Selon cette conception, les enfants pourraient construire la structure syntaxique des phrases qu'ils entendent en utilisant à la fois les frontières prosodiques et les mots grammaticaux : les frontières prosodiques indiqueraient à l'enfant l'emplacement des frontières syntaxiques, tandis que les mots grammaticaux permettraient de trouver la nature de ces unités syntaxiques. L'hypothèse d'initialisation prosodique postule que les enfants pourraient utiliser les caractéristiques prosodiques des phrases comme l'accentuation, l'intonation et le rythme pour apprendre la syntaxe (Christophe, Guasti, Nespors & van Ooyen, 2003 ;

Gleitman & Wanner, 1982 ; Morgan, 1986). Les frontières de phrase ont des effets acoustiques sur la parole comme la durée, les pauses, les chutes tonales. Au début de l'apparition de la morphosyntaxe, les enfants ont une grammaire prosodique limitée qui peut expliquer l'omission des mots grammaticaux ou de syllabes inaccentuées (Carter & Gerken, 2004 ; Demuth, 2006 ; Demuth, 2007, 2019 ; Morgan & Demuth, 1996).

CONCLUSION

La prosodie est à l'interface de la musique et de la parole chez le jeune enfant. Elle est impliquée dans la différenciation progressive entre les structures linguistiques et les structures musicales en fonction de la nature des unités qui vont progressivement s'organiser et s'affiner au cours du développement. Les études comportementales montrent une sensibilité au contour intonatif général, c'est-à-dire à la prosodie de la langue cible, dans le milieu intra-utérin alors que la sensibilité aux frontières intonatives entre constituants marquées par une pause, un changement de hauteur vocale et l'allongement de l'avant-dernière syllabe se développe au cours de la première année. Quand un locuteur s'adresse à un enfant, la hauteur tonale est plus élevée, la gamme de hauteur est plus large, le débit de parole est ralenti, les mots sont plus courts et les pauses sont plus longues que dans le cas où un locuteur s'adresse à des adultes. Il est intéressant de noter que ces quatre paramètres acoustiques : la fréquence fondamentale (Fo) ou hauteur tonale, la durée (débit de parole, tempo et pauses), l'intensité, et les caractéristiques spectrales qui définissent la prosodie au niveau phonétique sont les mêmes que ceux qui définissent la mélodie d'une phrase musicale. Les variations de la fréquence fondamentale (Fo) de la voix décrite en termes de contours intonatifs sont prioritairement traitées par les bébés. Ce qui est moins le cas au cours de la période linguistique : la prosodie est en interaction avec toutes les autres composantes du langage : sur le plan phonologique, la prosodie correspond aux mécanismes d'intonation et d'accentuation des langues, et aux contraintes rythmiques qui régissent les groupements prosodiques. Enfin, la prosodie participe à la structuration syntaxique des énoncés à travers les accents démarcatifs comme la pause ou l'allongement de la dernière syllabe, placés en fin de groupes de mots ou de phrase. Le rythme n'explique pas tout : il existe dans le message parlé des contraintes beaucoup plus fondamentales auxquelles viennent également se modeler les constructions rythmiques, il s'agit des contraintes sémantico-pragmatiques. Les indices prosodiques sont par conséquent une source d'information cruciale, non seulement pour distinguer les propriétés grammaticales de la langue maternelle mais aussi pour rendre compte de la dimension expressive du message en augmentant les contours de la Fo, l'intensité et la durée pour exprimer des émotions, des sentiments, des attitudes et des intentions de communication.

RÉFÉRENCES

- Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, 162, 46-59.
- Bateson, M. C. (1975). Mother-infant exchanges: The epigenesis of conversational interaction. *Annals of the New York Academy of Sciences*, 263, 101-113.
- Bateson, M. C. (1979). The epigenesis of conversational interaction: A personal account of research development. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 63-77). London, England: Cambridge University Press.
- Beebe, B., Jaffe, J., Lachmann, F., Feldstein, S., Crown, C., & Jasnow, M. (2000). Systems models in development and psychoanalysis: the case of vocal rhythm coordination and attachment. *Infant Mental Health Journal*, 21(1-2), 99-122.
- Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new Pussycat? On talking to babies and animals. *Science*, 296(5572), 1435.
- Bosch, L., & Sebastián-Gallés, N. (1997). Native-language recognition abilities in 4-month-old infants from monolingual and bilingual environments. *Cognition*, 65(1), 33-69.
- Brundin, K., Rijdhalm, M., & Larson, K. (1988). Vocal communication between parents and infants. *Early Human Development*, 16, 35-53.
- Carter, A., & Gerken, L. (2004). Do children's omissions leave traces? *Journal of Child Language*, 31(3), 561-586. doi:10.1017/S030500090400621X
- Carvalho (de), A., Dautriche, I., Lin, I., & Christophe, A. (2017). Phrasal prosody constrains syntactic analysis in toddlers. *Cognition*, 163, 67-79. doi:10.1016/j.cognition.2017.02.018
- Christophe, A., Guasti, M. T., Nespore, M., & van Ooyen, B. (2003). Prosodic structure and syntactic acquisition: The case of the head-complement parameter. *Developmental Science*, 213-222.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and speech*. Harvard University Press
- Corbeil, M., Trehub, S. E., & Peretz, I. (2013). Speech vs. singing: Infants choose happier sounds. *Frontiers in Psychology*, 4. doi:10.3389/fpsyg.2013.00372
- Crown, C. L., Feldstein, S., Jasnow, M. D., & Jaffe, J. (2002). The cross-modal coordination of interpersonal timing: Six-week-olds infants' gaze with adults' vocal behavior. *Journal of Psycholinguistic Research*, 31(1), 1-23.
- Davis, B., MacNeilage, P., Matyear, C., & Powell, J. (2000). Prosodic Correlates of Stress in Babbling: An Acoustical Study. *Child Development*, 71(5), 1258-1270.
- DeCasper, A. J., Lecanuet, J.-P., Busnel, M.-C., Granier-Deferre, C., & Maugeais, R. (1994). Fetal reactions to recurrent maternal speech. *Infant Behavior and Development*, 17(2), 159-164. doi:10.1016/0163-6383(94)90051-5
- Delattre, P. C. (1962). Le jeu des transitions des formants et la perception des consonnes. *Proceedings of the Fourth International Congress of Phonetic Sciences*, Helsinki, 1961. Gravenhage: Mouton, pp. 407-417.
- Delavenne, A., Gratier, M., & Devouche, E. (2013). Expressive timing in infant-directed singing between 3 and 6 months. *Infant Behavior and Development*, 36, 1-13.
- Demuth, K. (2006). Crosslinguistic perspectives on the development of prosodic words. Guest Editor, Special Issue, *Language and Speech*, 49(2), 129-297.

- Demuth, K. (2007). Acquisition at the prosody-morphology interface. In A. Belikova, L. Meroni & M. Umeda (Eds.), *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 84-91). Somerville, MA: Cascadilla Proceedings Project.
- Demuth, K. (2019). Prosodic constraints on children's use of grammatical morphemes. *First Language*, 39(1), 80-95. doi:10.1177/0142723717751984
- Fava, E., Hull, R., Baumbauer, K., & Bortfeld, H. (2014). Hemodynamic responses to speech and music in preverbal infants. *Child Neuropsychology*, 20(4), 430-448. doi:10.1080/09297049.2013.803524
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181-195.
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-293.
- Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry*, 48(3-4), 329-354.
- Feldstein, S., Jaffe, J., Beebe, B., Crown, C. L., Jasnow, M., Fox, H., & Gordon, S. (1993). Coordinated interpersonal timing in adult-infant vocal interactions: a cross-site replication. *Infant Behavior and Development*, 16(4), 455-470.
- Ferrier, L. J. (1985). Intonation in discourse: Talk between 12-month-olds and their mothers. In K. Nelson (Ed.), *Children's language* (Vol. 5, pp. 35-60) Hillsdale, NJ: Erlbaum.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27(2), 209-221. doi:10.1037/0012-1649.27.2.209
- Fogel, A. (1993a). *Developing through relationships: Communication, self, and culture in early infancy*. Cambridge: Harvester Press.
- Fogel, A. (1993b). Two principles of communication: co-regulation and framing. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicative development* (pp. 9-22). London, Routledge.
- Fogel, A., & Hannan, T. E. (1985). Manual actions of nine- to fifteen-week-old human infant during face-to-face interaction with their mothers. *Child Development*, 56, 1271-1279.
- Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In G. E. Snow & G. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63-88). Cambridge: Cambridge University Press.
- Gerry, D., Unrau, A., & Trainor, L. J. (2012). Active music classes in infancy enhance musical, communicative and social development: Active music classes enhance development. *Developmental Science*, 15(3), 398-407. doi:10.1111/j.1467-7687.2012.01142.x
- Gervain, J., & Mehler, J. (2010). Speech Perception and Language Acquisition in the First Year of Life. *Annual Review of Psychology*, 61(1), 91-218.
- Gleitman, L. R., & Wanner, E. (1982). Language acquisition: The state of the art. In L. R. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge, MA: Cambridge University Press.
- Graf Estes, K., & Hurley, K. (2013). Infant-directed prosody helps infants map sounds to meanings. *Infancy*, 18(5), 797-824. doi:10.1111/Infa.12006
- Granier-Deferre, C., Bassereau, S., Ribeiro, A., Jacquet, A.-Y., & DeCasper, A. J. (2011). A Melodic Contour Repeatedly Experienced by Human Near-Term

- Fetuses Elicits a Profound Cardiac Reaction One Month after Birth. *PLoS ONE*, 6(2): e17304.
- Granier-Deferre, C., Ribeiro, A., Jacquet, A.-Y., & Bassereau, S. (2011). Near-term fetuses process temporal features of speech. *Developmental Science*, 14(2), 336-352.
- Gratier, M. (2003). Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development*, 18, 533-554.
- Gratier, M., & Trevarthen, C. (2007). Voice vitality and meaning: On the shaping of the infant's utterances in willing engagement with culture. Comment on Bertau's "On the notion of voice". *International Journal for Dialogical Science Fall*, 1(2), 169-181.
- Gratier, M., & Devouche, E. (2011). Imitation and Repetition of Prosodic Contour in Vocal Interaction at 3 Months. *Developmental Psychology*, 47(1), 67-76.
- Gratier, M., & Bertau, M. C. (2012). Polyphony: a vivid source of self and symbol. In M.-C. Bertau, M. M. Gonçalves, & P. T. F. Raggatt (Eds.), *Dialogic formations: Investigations into the origins and development of the dialogical self*. Charlotte, N. C.: Information Age Publishers.
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, 6, 1167. doi:10.3389/fpsyg.2015.01167
- Gratier, M., & Devouche, E. (2020). Participation vocale du bébé au chant des parents en milieu naturel. *Enfance*, 5(1), 67. doi:10.3917/enf2.201.0067
- Hallé, P. A., De Boysson-Bardies, B., & Vihman, M. M. (1991). Beginnings of Prosodic Organization: Intonation and Duration Patterns of Disyllables Produced by Japanese and French Infants. *Language and Speech*, 34(4), 299-318. doi:10.1177/002383099103400401
- Henning, A., Striano, T., & Lieven, E. V. M. (2005). Maternal speech to infants at 1 and 3 months of age. *Infant Behavior and Development*, 28, 519-536. doi:10.1016/j.infbeh.2005.06.001
- Jusczyk, P. W., & Krumhansl, C. L. (1993). Pitch and rhythmic patterns affecting infants' sensitivity to musical phrase structure. *Journal of Experimental Psychology: Human Perception and Performance*, 19(3), 627-640. doi:10.1037/0096-1523.19.3.627
- Karzon, R. G., & Nicholas, J. G. (1989). Syllabic pitch perception in 2- to 3-month-old infants. *Perception & Psychophysics*, 45, 10-14.
- Keller, P. E. (2008). Joint action in music performance. In F. Morganti, F. A. Carassa, and G. Riva (Eds.), *Enacting Intersubjectivity: A cognitive and social perspective on the study of interactions* (pp. 205-221). Amsterdam: IOS Press.
- Kemler, N., Hirsh-Pasek, K., Jusczyk, P., & Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16(1) 55-68.
- Kempe, V., Schaeffler, S., & Thoresen, J. C. (2010). Prosodic disambiguation in child-directed speech. *Journal of Memory and Language*, 62, 204-225.
- Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, 72(2), 353-365. doi:10.1121/1.388089

- Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences*, *104*(30), 12577-12580. doi:10.1073/pnas.0705345104
- Kisilevsky, B. S., Hains, S. M. J., Jacquet, A.-Y., Granier-Deferre, C., & Lecanuet, J.-P. (2004). Maturation of fetal responses to music. *Developmental Science*, *7*(5), 550-559. doi:10.1111/j.1467-7687.2004.00379.x
- Kisilevsky, B. S., Hains, S. M. J., Lee, K., Xie, X., Huang, H., Ye, H. H., Zhang, K., & Wang, Z. (2003). Effects of Experience on Fetal Voice Recognition. *Psychological Science*, *14*(3), 220-224. doi:10.1111/1467-9280.02435
- Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., Swansburg, M. L., Lee, K., Xie, X., Huang, H., Ye, H.-H., Zhang, K., & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, *32*(1), 59-71. doi:10.1016/j.infbeh.2008.10.002
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2001). Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development*, *24*, 372-392.
- Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mothers' speech: adjustments for age and sex in the first year. *Infancy*, *4*, 85-110.
- Konopczynski, G. (1986). Vers un modèle développemental du rythme français : problèmes d'isochronie reconsidérés à la lumière des données de l'acquisition du langage. *Bulletin de l'Institut de Phonétique de Grenoble*, *XV*, 157-190.
- Konopczynski, G. (1990). *Le langage émergent I : caractéristiques rythmiques*. Hambourg, Helmut Buske Verlag.
- Konopczynski, G. (1991). *Le langage émergent II. Aspects vocaux et mélodiques*. Hambourg Buske Verlag.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., Huotilainen, M., & Fellman, V. (2009). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping*, NA-NA. doi:10.1002/hbm.20890
- Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: Spotlight on auditory working memory. *Annals of the New York Academy of Sciences*, *1252*(1), 100-107. doi:10.1111/j.1749-6632.2012.06463.x
- Lecanuet, J.-P. (2000). L'audition fœtale. *Journal de Pédiatrie et de Puériculture*, *13*(6), 349-358. doi:10.1016/S0987-7983(00)80071-7
- Lecanuet, J.-P., Granier-Deferre, C., Jacquet, A. Y., & DeCasper, A. J. (2000). Fetal discrimination of low-pitched musical notes. *Developmental Psychobiology*, *36*(1), 29-39.
- Lecanuet, J.-P., & Granier-Deferre, C. (1993). Speech Stimuli in the Fetal Environment. In B. Boysson-Bardies, S. Schonen, P. Juszyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 237-248). Springer Netherlands. doi:10.1007/978-94-015-8234-6_20
- Lee, G. Y., & Kisilevsky, B. S. (2014). Fetuses respond to father's voice but prefer mother's voice after birth: Fetuses Respond to Father's Voice. *Developmental Psychobiology*, *56*(1), 1-11. doi:10.1002/dev.21084

- Levitt, A. G., & Wang, Q. (1991). Evidence for Language-Specific Rhythmic Influences in the Reduplicative Babbling of French-and English-Learning Infants. *Language and Speech*, 34(3), 235-249. doi:10.1177/002383099103400302
- Levitt, A. G., & Aydelott Utman, J. G. (1992). From babbling towards the sound systems of English and French: A longitudinal two-case study. *Journal of Child Language*, 19(1), 19-49. doi:10.1017/S0305000900013611
- Lickliter, R., Bahrick, L. E., & Vaillant-Mekras, J. (2017). The intersensory redundancy hypothesis: Extending the principle of unimodal facilitation to prenatal development. *Developmental Psychobiology*, 59(7), 910-915. doi:10.1002/dev.21551
- Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2009). Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language*, 36(4), 909-922. doi:10.1017/S030500090800929X
- Malloch, S., & Trevarthen, C. (2008). *Communicative musicality: Exploring the basis of human companionship*. Oxford: Oxford University Press.
- Masataka, N. (1992). Pitch characteristics of Japanese maternal speech to infants. *Journal of Child Language*, 19(2), 213-223. doi:10.1017/S0305000900011399
- May, L., Byers-Heinlein, K., Gervain, J., & Werker, J. F. (2011). Language and the newborn brain: Does prenatal language experience shape the neonate neural response to speech? *Frontiers in Psychology*, 2, 222.
- McMullen, E., & Saffran, J. R. (2004). Music and Language: A Developmental Comparison. *Music Perception*, 21(3), 289-311. doi:10.1525/mp.2004.21.3.289
- McRoberts, G. W., & Best, C. T. (1997). Accommodation in mean f0 during mother-infant and father-infant vocal interactions: a longitudinal case study. *Journal of Child Language*, 24, 719-736. doi:10.1017/s030500099700322x
- Moon, C., Cooper, R., & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16, 495-500.
- Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning & Verbal Behavior*, 20(1), 67-85. doi:10.1016/S0022-5371(81)90312-1
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge Mass: MIT Press.
- Morgan, J. L., & Demuth, K. (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 313-330). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455-464. doi:10.1016/j.infbeh.2004.03.002
- Nathani, S., Oller, D. K., Cobo-Lewis, A. B. (2003). Final syllable lengthening (FSL) in infant vocalizations. *Journal of Child Language*, 30, 3-25.
- Nawrot, E. S. (2003). The perception of emotional expression in music: Evidence from infants, children and adults. *Psychology of Music*, 1, 75-92.
- Nazzi, T., Bertocini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.
- Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, 43(1), 1-19. doi:10.1006/jmla.2000.2698

- Noble, C., Shafaie, M. L., & Self, P. (1982). Individual differences in early social: neonatal gaze aversion. *International Conference on Infant Studies*, Austin, Texas.
- Panneton Cooper, R., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development*, 20(4), 477-488.
- Papoušek, H., & Papoušek, M. (1977). Mothering and cognitive head-start: Psychobiological considerations. In H. R. Schaffer (Eds.), *Studies in mother-infant interaction* (pp. 63-85). London: Academic Press.
- Papoušek, M., Papoušek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behaviour and Development*, 14(4), 415-440.
- Partanen, E., Kujala, T., Tervaniemi, M., & Huotilainen, M. (2013). Prenatal Music Exposure Induces Long-Term Neural Effects. *PLoS ONE*, 8(10), e78946. doi:10.1371/journal.pone.0078946.
- Partanen, E., Kujala, T., Naatanen, R., Liitola, A., Sambeth, A., & Huotilainen, M. (2013). Learning-induced neural plasticity of speech processing before birth. *Proceedings of the National Academy of Sciences*, 110(37), 15145-15150. doi:10.1073/pnas.1302159110
- Patel, A. D. (2010). *Music, language, and the brain*. Oxford University press.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758-4763. doi:10.1073/pnas.0909074107
- Plantinga, J., & Trainor, L. J. (2005). Memory for melody: infants use a relative pitch code. *Cognition*, 98(1), 1-11. doi:10.1016/j.cognition.2004.09.008
- Ramus, F. (2002). Language discrimination by newborns. Teasing apart phonotactic, rhythmic, and intonational cues. *Annual Review of Language Acquisition*, 2, 85-115.
- Rochat, P., Querido, J., & Striano, T. (1999). Emerging sensitivity to the timing and structure of protoconversation in early infancy. *Developmental Psychology*, 35, 950-957.
- Scola, C., Bourjade, M., & Jover, M. (2015). Social interaction is associated with changes in infants' motor activity. *Socioaffective Neuroscience & Psychology*, 5, 28256.
- Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal Singing Modulates Infant Arousal. *Psychology of Music*, 31(4), 365-375. doi:10.1177/03057356030314002
- Shuka, M., Nespors, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, 54(1), 1-32.
- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy*, 13(4), 410-420.
- Snow, C. E. (1994). Beginning from Baby Talk: Twenty years of research on input in interaction. In C. Gallaway & B. J. Richards (Eds.), *Input and Interaction in Language Acquisition* (1st ed., pp. 3-12). Cambridge University Press. doi:10.1017/CBO9780511620690.002
- Snow, D. (1998). Children's imitations of intonation contours: are rising tones more difficult than falling tones? *Journal of Speech, Language, and Hearing Research*, 41, 576-587.
- Snow, D. (2002). Intonation in the monosyllabic utterances of 1-year-olds. *Infant Behavior and Development*, 24, 393-407.

- Snow, D. (2006). Regression and reorganization of intonation between 6 and 23 months. *Child Development*, 77, 281-296.
- Song, J. Y., Demuth, K., & Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of Acoustical Society of America*, 128(1), 388-400. doi:10.1121/1.3419786
- Spinelli, M., Fasolo, M., & Mesman, J. (2017). Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Developmental Review*, 44, 1-18. doi:10.1016/j.dr.2016.12.001
- Stern, D. N. (1974). Mother and infant at play: The dyadic interaction involving facial, vocal and gaze behaviors. In M. Lewis & L. A. Rosenblum (Eds.), *The effect of the infant on its caregiver* (pp. 187-213). New York: Wiley.
- Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and alternation: two modes of communication within the mother-infant DYAD. *Annals of the New York Academy of Sciences*, 263, 89-100.
- Stern, D. N., Beebe, B., Jaffe, J., & Bennett, S. L. (1977). The infant's stimulus world during social interaction: A study of caregiver behaviors with particular reference to repetition and timing. In H. R. Schaffer (Eds.), *Studies in mother-infant interaction* (pp. 177-202). New York: Academic Press.
- Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation as signals in maternal speech to pre-linguistic infants. *Developmental Psychology*, 18, 727-735.
- Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language*, 10, 1-15.
- Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin & Review*, 9(2), 335-340.
- Trehub, S. E., & Russo, F. A. (2020). Infant-Directed Singing from a Dynamic Multimodal Perspective. In F. Russo, B. Ilari, & A. J. Cohen (Eds.), *The Routledge Companion to Interdisciplinary Studies in Singing* (1st ed., pp. 249-261). Routledge. doi:10.4324/97811315163734-19
- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Adults identify infant-directed music across cultures. *Infant Behavior and Development*, 16, 193-211.
- Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences*, 930, 1-16.
- Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, 6(7), 669-673. doi:10.1038/nn1084
- Trehub, S. E., Schneider, B. A., & Henderson, J. L. (1995). Gap detection in infants, children, and adults. *The Journal of the Acoustical Society of America*, 98(5), 2532-2541. doi:10.1121/1.414396
- Trehub, S. E. (2009). Music lessons from infants. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 229-234). Oxford: Oxford University Press.
- Trehub, S. E. (1990). Human infants' perception of auditory patterns. *International Journal of Comparative Psychology*, 4, 91-110.
- Trevarthen, C. (1977). Descriptive analysis of infant communicative behavior. In H. R. Schaffer (Ed.). *Studies in mother-infant interaction* (pp. 227-270). New York, NY: Academic Press.

- Trevarthen, C. (1999). Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication. *Musicae Scientiae*, 3(1), 155-215. doi:10.1177/10298649000030S109
- Trevarthen, C., & Aitken, K. J. (2003). Intersubjectivité chez le nourrisson : Recherche, théorie et application clinique. *Devenir*, 15(4), 309-428.
- Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American psychologist*, 44(2), 112- 119.
- Tsang, C. D., Falk, S., & Hessel, A. (2017). Infants Prefer Infant-Directed Song Over Speech. *Child Development*, 88(4), 1207-1215. doi:10.1111/cdev.12647
- Vihman, M. M., Nakai, S., & DePaolis, R. A. (2006). Getting the rhythm right: A cross-linguistic study of segmental duration in babbling and first words. In L. Goldstein, D. Whalen, & C. Best (Eds.), *Laboratory Phonology 8* (pp. 341-366). New York: Mouton de Gruyter.
- Vihman, M. M. (2019). First words and prosodic structures: A cross-linguistic perspective. In M. M. Vihman, *Phonological Templates in Development* (pp. 93-121). Oxford University Press. doi:10.1093/oso/9780198793564.003.0004
- Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Oxford: Blackwell.
- Vihman, M. M. & Miller, R (1988). Words and babble at the threshold of lexical acquisition. In M. D. Smith, & J. L. Locke (Eds.), *The emergent lexicon* (pp. 151-183). New York: Academic Press.
- Vosoughi, S., & Roy, D. (2012). A longitudinal study of prosodic exaggeration in child-directed speech. *Proceedings of the 6th International Conference on Speech Prosody*. Shanghai, China.
- Weinberg, M. K., & Tronick, E. Z. (1994). Beyond the face: an empirical study of infant affective configurations of facial, vocal, gestural, and regulatory behaviors. *Child Development*, 65(5), 1503-1515.
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, 43(2), 230-246.
- Whalen, D. H., Levitt, A. G., & Wang, Q. (1991). Intonational differences between the reduplicative babbling of French- and English-learning infants. *Journal of Child Language*, 18(3), 501-516. doi:10.1017/S0305000900011223
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), 2468-2471. doi:10.1073/pnas.0809035106
- Zhao, T. C., & Kuhl, P. K. (2016). Musical intervention enhances infants' neural processing of temporal structure in music and speech. *Proceedings of the National Academy of Sciences*, 113(19), 5212-5217. doi:10.1073/pnas.1603984113