

Université de Montréal

**Réactions des bébés au chant et à la parole:
Impacts sur l'attention et l'affect**

par
Mariève Corbeil

Département de Psychologie
Faculté des Arts et Sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de Philosophiae Doctor (Ph.D.)
en Psychologie- recherche et intervention
option neuropsychologie clinique

Février 2016

© Mariève Corbeil, 2016

Résumé

Les parents à travers le monde chantent et parlent à leurs bébés. Ces deux types de vocalisations aux enfants préverbaux partagent plusieurs similarités de même que des différences, mais leurs conséquences sur les bébés demeurent méconnues. L'objectif de cette thèse était de documenter l'efficacité relative du chant et de la parole à capter l'attention des bébés sur de courtes périodes de temps (Étude 1) ainsi qu'à réguler l'affect des bébés en maintenant un état de satisfaction sur une période de temps prolongée (Étude 2).

La première étude a exploré les réactions attentionnelles des bébés exposés à des enregistrements audio non familiers de chant et de parole. Lors de l'expérience 1, des bébés de 4 à 13 mois ont été exposés à de la parole joyeuse s'adressant au bébé (séquences de syllabes) et des berceuses fredonnées par la même femme. Ils ont écouté significativement plus longtemps la parole, qui contenait beaucoup plus de variabilité acoustique et d'expressivité que les berceuses. Dans l'expérience 2, des bébés d'âges comparables n'ont montré aucune écoute différentielle face à une version parlée ou chantée d'une chanson pour enfant turque, les deux versions étant exprimées de façon joyeuse / heureuse. Les bébés de l'expérience 3, ayant entendu la version chantée de la chanson turque ainsi qu'une version parlée de façon affectivement neutre ou s'adressant à l'adulte, ont écouté significativement plus longtemps la version chantée. Dans l'ensemble, la caractéristique vocale joyeuse plutôt que le mode vocal (chanté versus parlé) était le principal déterminant de l'attention du bébé, indépendamment de son âge.

Dans la seconde étude, la régulation affective des bébés a été explorée selon l'exposition à des enregistrements audio non familiers de chant ou de parole. Les bébés ont été exposés à du chant ou de la parole jusqu'à ce qu'ils rencontrent un critère d'insatisfaction exprimée dans le visage. Lors de l'expérience 1, des bébés de 7 à 10 mois ont écouté des enregistrements de paroles s'adressant au bébé, de paroles s'adressant à l'adulte ou du chant dans une langue non familière (turque). Les bébés ont écouté le chant près de deux fois plus longtemps que les paroles avant de manifester de l'insatisfaction. Lors de l'expérience 2, des bébés ont été

exposés à des enregistrements de paroles ou de chants issus d'interactions naturelles entre la mère et son bébé, dans une langue familière. Comme dans l'expérience 1, le chant s'adressant au bébé était considérablement plus efficace que les paroles pour retarder l'apparition du mécontentement. La construction temporelle du chant, avec notamment son rythme régulier, son tempo stable et ses répétitions, pourrait jouer un rôle important dans la régulation affective, afin de soutenir l'attention, rehausser la familiarité ou promouvoir l'écoute prédictive et l'entraînement.

En somme, les études présentées dans cette thèse révèlent, pour la première fois, que le chant est un outil parental puissant, tout aussi efficace que la parole pour capter l'attention et plus efficace que la parole pour maintenir les bébés dans un état paisible. Ces découvertes soulignent l'utilité du chant dans la vie quotidienne et l'utilité potentielle du chant dans des contextes thérapeutiques variés impliquant des bébés.

Mots-clés : bébés, musique, langage, chant, parole, émotion, affect, régulation affective, attention

Abstract

Parents throughout the world sing and speak to infants. These two classes of vocalization to pre-verbal infants have various similarities as well as differences, but their consequences on infant listeners are poorly understood. The goal of this thesis was to document the relative efficacy of singing and speech for capturing infants' attention over short periods (Study 1) and for regulating infant affect or maintaining their composure over extended periods (Study 2).

The first study explored infants' attention to unfamiliar audio samples of speech and singing. In Experiment 1, infants 4 to 13 months of age were exposed to happy-sounding infant-directed speech and to hummed lullabies by the same woman. They listened significantly longer to the speech stimuli, which had considerably greater acoustic variability and expressiveness, than to the lullabies. In Experiment 2, infants of comparable age did not exhibit differential listening to the spoken lyrics of a Turkish children's song or to sung version, both of which featured a joyful/happy manner. Infants in Experiment 3 who heard the happily sung lyrics of the children's song and a spoken version in an adult-directed or affectively neutral manner listened significantly longer to the sung version. Overall, happy voice quality rather than vocal mode (speech or singing) was the principal contributor to infant attention, regardless of age.

In the second study, infant affect regulation was explored by means of exposure to unfamiliar audio samples of speech and singing. Infants were exposed to singing or speech until they met a criterion of distress based on negative facial expression. In Experiment 1, infants 7 to 10 months of age listened to scripted recordings of infant-directed speech, adult-directed speech, or singing in an unfamiliar language (Turkish). They listened to singing for roughly twice as long as speech before meeting the distress criterion. In Experiment 2, infants were exposed to natural recordings of infant-directed speech or singing in a familiar language. As in Experiment 1, infant-directed singing was considerably more effective than speech for delaying the onset of distress. The temporal patterning of singing, with notably its regular beat, stable tempo, and repetition may play an important role in affective regulation, perhaps by sustaining attention, enhancing familiarity or promoting predictive listening and

entrainment.

In sum, the present studies reveal, for the first time, that singing is a powerful parenting tool, as effective as speech in capturing infant attention and more effective than speech in maintaining infants' composure. The findings indicate the utility of singing in everyday life and the potential utility of singing in a variety of therapeutic contexts involving infants.

Keywords: infants, music, language, singing, speech, emotion, affect, affective regulation, attention

Table des matières

Résumé.....	1
Abstract.....	3
Table des matières.....	5
Liste des tableaux.....	7
Liste des figures.....	8
Liste des sigles et abréviations.....	9
Remerciements.....	11
Chapitre I : Contexte théorique.....	13
1. Introduction : le chant et la parole dans l'enfance.....	14
1.1 Similarités.....	14
1.1.1 Style vocal distinctif.....	15
1.2 Différences.....	16
1.2.1 Hauteur et rythme.....	16
1.2.2 Entraînement.....	18
2. Les réactions aux sons dans le développement.....	19
2.1 Le développement de préférences sonores précoces.....	19
2.1.1 Préférence pour le style DB.....	22
2.1.2 Fonctions du style DB.....	23
2.2 Les préférences attentionnelles relatives à la parole et à la musique.....	25
3. Régulation des émotions.....	26
3.1 Les vocalisations DB en fonction de la relation parent-enfant.....	27
3.2 La musique comme outil de régulation affective.....	29
4. Études comparatives entre la parole et le chant.....	31
5. Objectifs et hypothèses.....	32
5.1 Première étude : Capter l'attention.....	32
5.2 Deuxième étude : Réguler l'affect.....	33

Chapitre II : Partie expérimentale	34
Article I.....	35
Article II.....	67
Analyses supplémentaires.....	92
Chapitre III : Discussion générale.....	95
1. Retour sur les objectifs et résumé des résultats	96
1.1 Résumé de l'étude 1 : Les bébés « choisissent » les sons joyeux.....	96
1.2 Résumé de l'étude 2 : Le chant retarde l'apparition d'un affect négatif chez le bébé	97
2. Interprétation et implications des résultats	97
2.1 Identification des facteurs influençant l'attention des bébés	97
2.1.1 Un biais envers la positivité.....	97
2.1.2 L'influence de la parole	98
2.1.3 La distinction des processus attentionnels	99
2.2 Une motivation précoce envers la musique	101
2.2.1 Identification des facteurs contribuant à la régulation affective	101
2.2.2 Le chant versus la musique instrumentale	104
2.2.3 Le chant et la parole en contextes naturels	104
2.2.4 Les biais culturels	106
2.2.5 Les caractéristiques d'une « super-chanson ».....	107
2.3 Applications cliniques potentielles	107
3. Conclusion	109
Bibliographie.....	i
Annexe 1. Questionnaire remis aux parents.	i
Annexe 2. Matériel audio supplémentaire.	iv

Liste des tableaux

Chapitre II : Partie expérimentale

Article 1 : Speech vs. singing: infants choose happier sounds

Tableau I. Caractéristiques acoustiques des stimuli	46
--	----

Liste des figures

Chapitre II : Partie expérimentale

Article 1 : Speech vs. singing: infants choose happier sounds

Figure 1. Les contours de la fréquence fondamentale (F0) contours selon des extraits de 5 secondes provenant de chaque son: (A) une berceuse murmurée (chilienne) et une séquence de syllabes, (B) les paroles de la chanson turque chantées ou parlées de façon joyeuse / DB, (C) les paroles de la chanson turque chantées de façon joyeuse / DB ou parlées de façon neutre / DA..... 76

Figure 2. Temps de regards cummulatifs en secondes (s) pour le chant et la parole. Les barres d'erreurs représentent l'erreur standard (** = $p < .01$; * = $p < .05$) 49

Article 2 : Singing delays the onset of infant distress

Figure 1. Temps jusqu'à l'apparition de l'affect négatif (en minutes) pour le chant et la parole dans l'expérience 1 et 2. Les barres d'erreurs représentent l'erreur standard (** $p < .01$; * $p < .05$)..... 97

Figure 2. La notation musicale et les paroles d'une chanson française pour enfants commune, *Une Souris Verte*, qui faisait partie de l'ensemble de stimuli utilisés dans l'expérience 1. Les répétitions de séquences mélodiques sont indiquées par des boîtes équivalentes 98

Liste des sigles et abréviations

ANOVA	Analyse de la variance / <i>Analysis of variance</i>
dB	Décibels / <i>Decibels</i>
DB / ID	Dirigé vers le bébé / <i>Infant-directed</i>
DA / AD	Dirigé vers l'adulte / <i>Adult-directed</i>
ÉT / SD	Écart-Type / <i>Standard deviation</i>
e.g.	Exemplī grātia
et al.	Et alii
etc.	Et cætera
F0	Fréquence fondamentale / <i>Fundamental frequency</i>
Hz	Hertz
i.e.	Id est
IOI	Intervalles entre le début de chaque stimulus / <i>Inter-onset intervals</i>
IRMf / fMRI	Imagerie par résonance magnétique fonctionnelle / <i>Functional magnetic resonance imaging</i>
ISI	Intervalles entre la fin et le début des stimuli (périodes de silence) / <i>Inter-stimulus intervals</i>
<i>M</i>	Moyenne / <i>Mean</i>
min	Minutes / <i>Minutes</i>
msec	Millisecondes / <i>Milliseconds</i>
MEG	Magnétoencéphalographie
non-DB / non-ID	Non dirigé vers le bébé / <i>non-infant-directed</i>
<i>n</i>	Taille d'échantillon / <i>Sample size</i>
sec	Secondes / <i>Seconds</i>
vs.	Versus

À toutes ces petites boules de lumières qui ont illuminées mon parcours.

À mes supernovas, Lili et Bé.

Remerciements

Cette thèse revient à plusieurs individus sans qui cette épopée n'aurait jamais eu lieu et qui méritent bien plus que des remerciements, quoique ce soit tout ce que je puisse leur offrir.

J'aimerais d'abord remercier mes directrices de recherches, Isabelle Peretz et Sandra Trehub, des femmes plus grandes que nature, nourrissant la passion et le dépassement de soi, qui m'ont fait confiance il y a plusieurs années et dont ma mémoire restera gravée de leurs forces inspirantes. Sandra, you have been as a second mother to me for all those years and I will be forever grateful.

J'aimerais évidemment remercier tous ceux ayant gravité autour de mes recherches, de près ou de loin, notamment à travers le laboratoire BRAMS. J'ai eu la chance de profiter de précieuses collaboratrices qui ont fait avancer mon projet de recherche par leur persévérance et leurs talents : Roxane Campeau, Audrey Morin, Émilie Gilbert, Cynthia Paquin, Camille Rivard et Ali Bergeron-Caron. Je me dois de citer également Axelle Calcus qui a enrichi mon parcours. Merci aux techniciens du BRAMS pour leur soutien: Olivier Piché, Nathanaël Lécaudé, Mihaela Felezeu, tout en soulignant le soutien incommensurable de Bernard Bouchard, compositeur de musique, de sourires et de souvenirs. Un grand merci à Carl-André Corbeil pour son assistance et ses multiples talents en programmation. Je remercie également mes bienfaiteurs de recrutement, m'ayant gracieusement permis de récolter des participants parmi leur clientèle. Un merci particulièrement chaleureux à Anne Corbeil pour son assistance et ses multiples talents en rédaction. Un merci tout spécial à mes voix : Athena Vouloumanos pour avoir prêté ses stimuli de recherches, Jessica Phillips-Silver pour ses berceuses enveloppantes, Beste Kalender pour ses prestations turques, ainsi que toutes ces mamans, qui ont gracieusement prêté leurs paroles et leurs chants à la science et qui ont envouté les oreilles des bébés.

Je remercie de tout mon cœur ces collègues neuropsychologues extraordinaires, amis et membres de ma famille qui ont rehaussé les hauts et facilité les bas, pour leurs épaules. Je serai éternellement reconnaissante de n'avoir pas couru le chemin seule, mais entourée des amours de ma vie : Olivier, les racines; Dalie et Béatrice, ces branches qui touchent ciel.

Enfin, je remercie, les derniers mais non les moindres, les quelques centaines de bébés, jeunes enfants ainsi que leurs parents qui se sont soumis à mes expérimentations, qui ont illuminés mon parcours et qui resteront gravés dans ma mémoire.

Chapitre I : Contexte théorique

1. Introduction : le chant et la parole dans l'enfance

Alors que le développement de la perception du langage durant l'enfance soit largement documenté (e.g. Curtin & Werker, 2007; Gervain & Mehler, 2010; Kuhl, 2004), la littérature concernant le développement perceptif musical est beaucoup moins étendue (Hannon & Trainor, 2007; Trehub, 2001, 2003; Trehub & Hannon, 2006). Quoique que l'intérêt pour le domaine soit en croissance, peu d'études se sont intéressées à la comparaison musique-langage dans l'enfance (Brandt, Gebrian, & Slevc, 2012; Chen-Hafteck, 1997; McMullen & Saffran, 2004; Trehub, Trainor, & Unyk, 1993). Pourtant, de la perspective d'un bébé, la parole dans les quelques mois suivant la naissance est essentiellement une question de sons plutôt que de sens. Le contenu de la parole a peu d'importance, ce qui rend la comparaison chant-parole encore plus intéressante lors de cette période préverbale.

1.1 Similarités

En général, le chant et les paroles sont distincts, mais il existe des versions hybrides partageant des caractéristiques de chacun, par exemple dans le cas de la poésie, les paroles rythmées (ex. le « slam » ou la comptine), ou encore l'utilisation de hauteurs contrastées dans des langues dites « tonales » (ex. le mandarin ou le thaïlandais). Alors que le contenu sémantique de la parole transmet une information explicite, la prosodie de la parole transmet le message implicite ou l'intention affective du locuteur (Frick, 1985). La prosodie, ou les caractéristiques non-verbales des paroles, inclue l'intonation (les contours de hauteurs), le rythme et les accentuations (variations d'intensité ou d'amplitude). Les caractéristiques prosodiques de la parole sont souvent référées comme étant la « musique du langage » (e.g. Wennerstrom, 2001) ou la mélodie de la parole, en contraste avec la musique, désignée comme le « langage des émotions » (e.g. Juslin & Sloboda, 2010). Pour l'enfant préverbal, les chants et les paroles ont des similarités frappantes (Chen-Hafteck, 1997; McMullen & Saffran, 2004; Trehub, Trainor, et al., 1993) poussant certains chercheurs à avancer que la parole est une forme du musique, tous deux définies comme étant des « jeux créatifs sonores » aux oreilles des tout-petits/bébés (Brandt et al., 2012).

1.1.1 Style vocal distinctif

Toutes les cultures humaines utilisent une forme vocale distincte dans les soins aux enfants (Dissanayake, 2000; Trehub, 2000; Trehub & Schellenberg, 1995; Trehub & Trainor, 1998). Le chant et la parole « dirigés vers le bébé –DB » (*infant-directed- ID*) contrastent du chant et de la parole usuels « dirigés vers l'adulte- DA » (*adult-directed –AD*) (Ferguson, 1964; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Jacobson, Boersma, Fields, & Olson, 1983). Des auditeurs adultes peuvent distinguer une chanson interprétée de façon DB de sa contrepartie non-DB, que celle-ci proviennent de leur propre culture ou d'une autre (Trehub, Unyk, & Trainor, 1993; Unyk, Trehub, Trainor, & Schellenberg, 1992). De même, les paroles DB (aussi appelé *motherese* ou *baby talk*) sont distinguables des paroles DA à travers les cultures et langues différentes (Ferguson, 1964; Fernald, 1989; Fernald & Morikawa, 1993; Grieser & Kuhl, 1988; Kuhl et al., 1997; Papousek, Papousek, & Symmes, 1991), incluant les langues des signes (Masataka, 1992a; 1992b; 1996; 1998). Au-delà des similarités transculturelles, des différences s'observent néanmoins entre les cultures, notamment dans le degré d'expressivité (Masataka, 1992b; Shute & Wheldall, 2009).

Les paroles et les chants des adultes s'adressant aux bébés (DB) diffèrent tous deux de leur paroles et chants ne s'adressant pas à des bébés (DA ou non-DB) principalement sur la base de d'une élévation de la hauteur et d'un tempo ralenti (Fernald & Simon, 1984; Trehub, Hill, & Kamenetsky, 1997; Trehub, Unyk, et al., 1997). Même les enfants modifient leur chant pour les nourrissons dès l'âge de 3 ans, avec une hauteur de la voix supérieure et un tempo plus lent (Trehub, Unyk, & Henderson, 1994), de façon donc similaire aux modifications des adultes. En d'autres mots, nous parlons et chantons plus aigu et plus lentement aux bébés.

Puisque le chant est limité par les règles structurelles de la chanson et le registre vocal des chanteurs, il ne peut pas varier autant que la prosodie des paroles. En moyenne, un chant DB implique une élévation de la hauteur d'environ un demi-ton en comparaison avec une performance non-DB de la même chanson (Bergeson & Trehub, 1999, 2002; Trainor, Clark, Huntley, & Adams, 1997; Trehub, Hill, et al., 1997). Par contraste, les paroles DB sont de trois à quatre demi-tons plus élevés que les paroles DA (Fernald & Simon, 1984; Jacobson et al., 1983). Les paroles DB présentent aussi, de façon générale, de plus grandes variations

d'amplitude et de hauteurs, un débit plus lent, plus de répétitions et un rythme plus régulier, notamment atteint par un allongement de la durée des voyelles et une augmentation de la durée et de l'amplitude des mots de fin de phrase (Ferguson, 1964; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Fernald et al., 1989; Grieser & Kuhl, 1988; Stern, Spieker, Barnett, & Mackain, 1983; Stern, Spieker, & Mackain, 1982). Le chant DB présente également un rythme plus régulier et de plus grande variation d'amplitude que le chant conventionnel (Nakata & Trehub, 2011; Trainor et al., 1997; Trehub, Hill, et al., 1997), mais les différences demeurent moins saillantes que pour les paroles. Les aspects les plus distinctifs des paroles DB, par ses contours exagérés et son rythme accentué, impliquent des dimensions centrales de la musique. En résumé, les paroles de l'adulte orientées vers le bébé apparaissent plus musicales (Brandt et al., 2012; Fernald, 1989, 1992; Trainor, Austin, & Desjardins, 2000).

1.2 Différences

1.2.1 Hauteur et rythme

Bien que les interactions vocales avec les bébés présentent des similarités, il existe des différences notables en ce qui a trait à la mélodie et au rythme. D'abord, les hauteurs et leurs variations ont des manifestations bien différentes dans la parole et le chant. Dans toutes les cultures, les hauteurs utilisées tendent à être discrètes, issues d'un groupe de hauteurs relativement petit (i.e. une gamme se répétant à travers des octaves), et les écarts entre deux sons successifs (intervalles) sont en moyenne petits (Savage, Brown, Sakai, & Currie, 2015). Par contre, les hauteurs dans la parole peuvent varier sur de plus grands intervalles et de façon continue (Patel, Peretz, Tramo, & Labreque, 1998). Les hauteurs dépendent ainsi d'une plus grande précision en musique qu'en langage. Par exemple, un changement d'un demi-ton dénature une chanson (comme étant faussée ou *out-of-tune*) alors que la même transformation appliquée à des paroles n'en altère que très peu la perception (Zatorre & Baum, 2012). Bien que les variations de hauteurs ne soient pas essentielles pour la compréhension sémantique, elles contribuent néanmoins à son expressivité (Frick, 1985).

En ce qui a trait à la dimension temporelle de la musique, toutes les cultures possèdent des musiques avec un rythme isochrone (i.e. régulier), avec un minimum de subdivisions et de durées (souvent des ratios 2 :1) ainsi que des répétitions de motifs rythmiques (Savage et al.,

2015), augmentant ainsi leur simplicité et leur prévisibilité par rapport à des rythmes plus complexes (London, 2004). Le chant DB présente encore plus de régularité temporelle que le chant non-DB (Nakata & Trehub, 2011). Fait intéressant, les nouveau-nés sont sensibles au rythme de la musique, l'omission d'un temps fort pouvant être détectée suivant les enregistrements électriques du cerveau (Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Les rythmes réguliers facilitent la discrimination de changements rythmiques tant chez les adultes (Drake & Botte, 1993; Drake, Botte, & Baruch, 1992; Drake, Jones, & Baruch, 2000; Povel & Essens, 1985), les enfants de 5 ans (Drake & Gérard, 1989) que les bébés de 9 mois (Bergeson & Trehub, 2006). Bien que des similarités globales existent dans les constructions rythmiques de la langue et de la musique d'une culture donnée (Patel & Daniele, 2003; Patel, Iversen, & Rosenberg, 2006), les paroles n'ont pas la régularité rythmique de la musique (Grabe & Low, 2002).

Une différence majeure entre le chant et la parole tient du fait que les répétitions sont omniprésentes en musique, contrairement au langage où celles-ci sont rares. En musique, les répétitions ne dépendent pas uniquement de motifs rythmiques, mais également des notes, des contours, des phrases, des thèmes, en faisant un élément central des constructions musicales (Margulis, 2013b, 2014). De telles répétitions dans le discours sont atypiques, à l'exception de répétitions de contenu visant à attirer l'attention de l'auditeur vers un élément important du discours (Johnstone, 1994). En fait, une illusion notoire (appelée la transformation de la parole en chanson ou *speech to song transformation- S2ST*) démontre que la répétition d'une même phrase parlée à plusieurs reprises la transforme en musique aux oreilles de l'auditeur, qui la perçoit soudainement comme étant chantée (Deutsch, Henthorn, & Lapidis, 2011). Les variations de hauteurs ont un impact important sur cette illusion. Plus précisément, les contours se répétant de façon stable dans le temps intensifient la force de l'illusion, alors que des modifications rythmiques, non (Falk, Rathcke, & Dalla Bella, 2014). De plus, la capacité de discrimination fine des hauteurs est meilleure lorsque la phrase est perçue comme chantée que parlée, suggérant des mécanismes de perception spécifiques à la musique (Vanden Bosch der Nederlanden, Hannon, & Snyder, 2015). Dans tous les cas, cette illusion renforce l'importance de la répétition en musique.

Dans les chansons pour enfants, les répétitions sont particulièrement proéminentes (Trehub & Trainor, 1998). Par exemple, dans la chanson *Frère Jacques*, les motifs rythmiques, les mélodies et les paroles se répètent en parallèle (*Frère Jacques, Frère Jacques. Dormez vous? Dormez vous? Sonnez les matines. Sonnez les matines. Ding, dang, dong. Ding, dang, dong.*). De surcroît, les mères amplifient la nature répétitive des chansons en leur chantant presque identiquement, selon un même niveau de hauteur et un même tempo, à différentes occasions (Bergeson & Trehub, 2002). Bien que les mères parlent à leur bébé de façon beaucoup plus variable (Bergeson & Trehub, 2002), elles présentent des répétitions de contours/d'intonation, des « chansons signatures », reconnaissables d'une fois à l'autre (Bergeson & Trehub, 2007), renforçant ainsi la musicalité de leurs paroles.

1.2.2 Entraînement

Percevoir un rythme n'est pas qu'une simple expérience auditive, puisqu'il existerait un lien étroit entre l'expérience auditive et la motricité. L'entraînement (*entrainment*), pouvant être défini comme étant la coordination spatio-temporelle résultant d'une réponse rythmique à un signal rythmique (Phillips-Silver, Aktipis, & Bryant, 2010) se retrouve à travers les cultures humaines en ce qui a trait à la musique (Grahn, 2012; Merker, Madison, & Eckerdal, 2009; Savage et al., 2015). L'entraînement musical implique des manifestations externes telles que taper, jouer des percussions, danser ou bouger sur le rythme (Kirschner & Tomasello, 2009; McAuley, Jones, Holub, Johnston, & Miller, 2006), mais également des manifestations internes de synchronisation à des rythmes ou pulsations auditives (Grahn, 2012; Large & Jones, 1999). La perception du rythme intensifie l'activité cérébrale entre les régions auditives et motrices, présumées responsables de la génération d'un rythme interne (Chen, Zatorre, & Penhune, 2006; Grahn & Rowe, 2009). De surcroît, la perception auditive est affectée par une stimulation vestibulaire (Phillips-Silver & Trainor, 2005; Phillips-Silver & Trainor, 2007; Phillips-Silver & Trainor, 2008; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009). Si ont fait bondir des bébés dans les bras tous les deux temps (marche) ou trois temps (valse), ceux-ci vont par la suite écouter plus longtemps la structure métrique représentant celle induite par le mouvement (Phillips-Silver & Trainor, 2005). Alors que l'alternance de tours de paroles dans une conversation peut être considérée comme une forme de coordination spatio-temporelle

(Phillips-Silver et al., 2010), l'entraînement musical se distingue en étant simultanée, quoique pas nécessairement synchrone, au signal rythmique.

Contrairement aux adultes (voir Repp & Su, 2013, pour une revue), les bébés ne sont pas capables de synchroniser leurs mouvements à un rythme musical (Zentner & Eerola, 2010), possiblement à cause d'un faible contrôle moteur. Ils pourraient cependant démontrer un entraînement interne puisqu'ils possèdent des capacités de discrimination élevées en ce qui a trait au rythme et à la structure métrique de la musique (Hannon & Trehub, 2005a, 2005b; Phillips-Silver & Trainor, 2005; Vanden Bosch der Nederlanden et al., 2015; Winkler et al., 2009). De plus, ils sont également capables de produire certaines réponses motrices corrélées avec un stimulus rythmique (Provasi, Anderson, & Barbu-Roth, 2014). Par exemple, les bébés présentent des mouvements corrélés avec le chant et les mouvements de la mère (Longhi, 2009). Dès 5 mois, ils manifestent une réponse motrice rythmique, i.e. du mouvement quoique non-synchrone, sur de la musique rythmée, et pas sur des paroles (Zentner & Eerola, 2010). La synchronisation motrice n'émergerait que plus tard dans le développement. À l'âge de 2 ans, les enfants démontrent quelques capacités de mouvements périodiques en lien avec un stimulus auditif (Eerola, Luck, & Toivainen, 2006; Kirschner & Tomasello, 2009; Provasi & Bobin-Bègue, 2003), mais les capacités de synchronisation se raffinent ensuite pendant plusieurs années avant d'être précises (Drake, 1997; Drake et al., 2000; Van Noorden & De Bruyn, 2009).

2. Les réactions aux sons dans le développement

2.1 Le développement de préférences sonores précoces

Les bébés deviennent rapidement « experts » en discrimination vocale, de plusieurs aspects de la prosodie de la langue (e.g. Gervain & Mehler, 2010; Kuhl, 2004) et de la musique (e.g. Hannon & Trainor, 2007; Trehub & Hannon, 2006). Des habiletés perceptives rudimentaires s'observent dès le dernier trimestre de gestation (Jardri et al., 2008). Les sons extérieurs sont considérablement altérés dans l'utérus par les fluides et tissus, diminuant notamment les hautes fréquences et plusieurs contrastes phonétiques, mais rehaussant la saillance de la voix de la mère et la prosodie des paroles (Gerhardt & Abrams, 2000; Griffiths,

Brown Jr, Gerhardt, Abrams, & Morris, 1994). L'expérience prénatale peut ainsi moduler les préférences des nouveau-nés, tel que démontré par des apprentissages auditifs fœtaux (Moon & Fifer, 2000). Par exemple, dans les jours suivants la naissance, les nouveau-nés présentent des réponses différentielles (diminution du rythme cardiaque et des mouvements) à une musique entendue durant la grossesse (Hepper, 1991) et préfèrent une histoire lue par leur mère durant la grossesse comparativement à une nouvelle histoire (DeCasper & Spence, 1986).

Bien que les chercheurs fassent généralement référence à des « préférences » chez les bébés, les préférences ne peuvent pas être établies avec de jeunes enfants préverbaux telles qu'il est possible de le faire avec des auditeurs plus âgés. Plutôt, ces préférences sont généralement inférées selon leurs réponses comportementales ou physiologiques à des stimuli externes contrastés. Une méthode, fréquemment utilisée afin de mesurer les préférences auditives des bébés, est de mesurer le temps de regard de l'enfant vers un stimulus visuel apparié à deux stimuli auditifs différents (e.g. Trainor, Tsang, & Cheung, 2002; Vouloumanos & Werker, 2004). Pour de très jeunes bébés, la succion non-nutritive peut être utilisée plutôt que le regard (e.g. Vouloumanos & Werker, 2007). Dans tous les cas, l'enfant doit apprendre la contingence entre sa réponse et la présentation du son. Spécifiquement, la présence du comportement (ex. sucer ou regarder) provoque la présentation du son et l'arrêt de celui-ci résulte en l'arrêt du son. Ainsi, la préférence ou l'intérêt s'opérationnalise comme étant des temps d'écoute plus longs d'un son comparativement à un autre.

L'interprétation de temps d'écoute plus ou moins long comme étant des préférences, soit des évaluations esthétiques affectives (positives ou négatives) des bébés, en l'absence de manifestations comportementales affectives, est questionnable (Trehub, 2012). Néanmoins, la procédure se base sur la capacité de contrôle de l'enfant sur les sons qui lui sont présentés afin d'inférer ce qu'il préfère. Ses temps d'écoute reflèteraient donc l'expression d'un « choix », d'un intérêt ou d'un biais attentionnel momentané et relatif envers un son, plutôt qu'une préférence esthétique telle que chez l'adulte.

Les facteurs qui sous-tendent ces biais attentionnels sont incertains. Certains biais sont acquis par l'expérience, tels que des temps d'écoute plus long pour la voix de leur mère à celle d'une étrangère (DeCasper & Fifer, 1980), pour son prénom comparé à un autre (Mandel, Jusczyk, & Pisoni, 1995), pour la langue familière (maternelle) comparée à une langue étrangère (Moon, Cooper, & Fifer, 1993) ou encore les rythmes familiers comparés à des rythmes non-familiers (Soley & Hannon, 2010). Par exemple, les bébés américains écoutent davantage des rythmes simples, typiques des musiques occidentales, comparativement à des rythmes plus complexes. Les bébés turcs quant à eux, habitués aux deux types de rythmes dans leur culture musicale, ne démontrent aucun intérêt différentiel (Soley & Hannon, 2010). Toutefois, lors de cette étude, les bébés des deux cultures présentaient un intérêt pour un rythme métrique légitime, simple ou complexe, comparativement à un rythme arbitraire beaucoup plus complexe, suggérant une préférence relative envers la simplicité rythmique. Cette préférence envers une la simplicité rythmique se reflète également dans les musiques à travers le monde (Savage et al., 2015).

Certains biais d'écoute pourraient aussi être innés, issus de la saillance naturelle de paramètres acoustiques du son (ex. un son très fort ou contrastant) ou de leur signification biologique (ex. des vocalisations d'une même espèce). Toutefois, au-delà de l'innéité, une attention différentielle à un son pourrait indiquer l'importance du stimulus pour l'enfant. L'exposition, même sans instruction spécifique, conduit l'individu à une expertise perceptuelle des sons musicaux et langagiers de sa propre culture. L'attention envers un son amplifie son encodage sensoriel dans le cortex auditif (Alho, Rinne, Herron, & Woods, 2014; Petkov et al., 2004). Ainsi, un biais attentionnel, qu'il soit d'origine inné ou acquise, pourrait faciliter l'apprentissage (Vouloumanos & Werker, 2004).

Toutefois, les biais attentionnels des bébés peuvent être inconstants, affectés par une série de facteurs. Par exemple, il est possible d'altérer une « préférence » par l'exposition lors d'une phase de familiarisation (e.g. Rose, Gottfried, Melloy-Carminar, & Bridger, 1982; Trehub, Plantinga, & Russo, 2016) ou encore en variant le délai entre la phase de familiarisation et l'expérimentation (e.g. Spence, 1996). Les facteurs régissant l'attention des bébés sont donc fort probablement multiples et complexes (Singh, Morgan, & Best, 2002)..

2.1.1 Préférence pour le style DB

Plusieurs études sur les préférences attentionnelles des bébés se sont intéressées au biais pour le style vocal DB. Les bébés démontrent davantage d'intérêt pour des vocalisations DB comparativement à non-DB, tant parlées (e.g. Cooper, Abraham, Berman, & Staska, 1997; Cooper & Aslin, 1990; Fernald, 1985; Glenn & Cunningham, 1983; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989) que chantées (Masataka, 1999; Trainor, 1996). La préférence pour les paroles DB est présente peu importe le genre du locuteur (Pegg et al., 1992; Werker & McLeod, 1989), la familiarité de la langue (Werker, Pegg, & McLeod, 1994) ou sur des paroles filtrées (atténuation des hautes fréquences ou *low-pass*), obscurant le contenu phonétique tout en gardant la prosodie (Cooper & Aslin, 1990). Les bébés présentent également une préférence pour les gestes et expressions faciales associés aux paroles DB plutôt qu'à des paroles DA (Brand & Shallcross, 2008; Kim & Johnson, 2014) ou à une langue signée DB (Masataka, 1996; 1998). Une récente méta-analyse de 34 études (Dunst, Gorman, & Hamby, 2012) révèle que cette préférence pour les paroles DB serait accentuée lorsque l'interaction est naturelle (i.e. produites par la mère avec son enfant) plutôt que simulée (i.e. sans l'enfant) ou acoustiquement modifiée. Le degré de la préférence augmenterait aussi de la naissance jusqu'à environ 9 mois, bien qu'il existe quelques résultats contradictoires pour les bébés de 9 mois et plus (Hayashi, Tamekawa, & Kiritani, 2001; Newman & Hussain, 2006). Les facteurs sous-tendant ce changement pourraient être en lien avec l'expérience (ex. l'augmentation de la durée d'exposition au style) ou encore la maturation (ex. un meilleur contrôle attentionnel). En ce qui concerne le chant, l'intérêt pour le chant DB versus le chant non-DB observé chez les nouveau-nés de parents sourds, ayant eu une exposition présumée très limitée à la musique, suggère que cet intérêt ne dépend pas uniquement de l'expérience (Masataka, 1999).

Quelques études ont également mesuré des réactions affectives des bébés aux paroles DB, renforçant le concept d'une préférence ou d'une valeur affective (Fernald, 1993; Werker & McLeod, 1989; Werker et al., 1994). Dans deux études (Werker & McLeod, 1989; Werker et al., 1994), des observateurs ont évalué sur des vidéos le comportement de bébés de 4 à 9 mois ayant écouté des paroles DB et DA. Pour chaque vidéo, ils devaient identifier le degré d'envie d'interagir, le degré d'intérêt et le degré de réponse émotionnelle (positifs versus négatif) du

bébé (sur des échelles en 9 points). Les bébés présentaient une plus haute intensité de réponses affectives, consistant en la cumulation des trois échelles, en réaction au registre DB plutôt que DA. Lors d'une autre étude (Fernald, 1993), les bébés sont évalués comme ayant plus d'affect positif sur des paroles DB que DA.

2.1.2 Fonctions du style DB

Bien que certaines caractéristiques acoustiques des paroles DB pourraient soutenir l'acquisition du langage (voir Soderstrom, 2007, pour une revue), par exemple l'hyperarticulation facilitant la discrimination des voyelles chez les bébés (Liu, Kuhl, & Tsao, 2003; Trainor & Desjardins, 2002), plusieurs modifications prosodiques des paroles DB ne faciliteraient pas l'apprentissage de la langue (Gauthier & Shi, 2011) ou nuiraient à la discrimination chez le bébé (Trainor & Desjardins, 2002). De façon générale, les contrastes phonétiques tendent à être en moyenne légèrement moins prononcés dans le style de paroles DB que non-DB (Martin et al., 2015), nuisant ainsi à son intelligibilité. Les mots des chansons sont aussi prononcés moins clairement lorsqu'ils s'adressent à des bébés plutôt qu'à des enfants préscolaires (Bergeson & Trehub, 1999). Ces résultats suggèrent que le style DB vise plusieurs buts au-delà l'acquisition du langage, tels que capter l'attention de l'enfant, lui exprimer une émotion ou encore renforcer des liens interpersonnels (Cristia, 2013).

Ainsi, une raison probable qui sous-tendrait l'intérêt des bébés envers le style DB serait l'expressivité émotionnelle positive amplifiée dans les paroles et les chants DB (Nakata & Trehub, 2011; Trainor et al., 2000; Trehub, Hannon, & Schachner, 2011). Lorsque les paroles DB sont comparées à des paroles DA émotionnelles, les différences acoustiques entre les deux registres diminuent fortement (Trainor et al., 2000). Les paroles DA sont toutefois typiquement émotionnellement neutres.

Les adultes identifient avec une plus grande précision l'intention du locuteur (ex. approbation ou prohibition) dans des paroles filtrées DB (atténuation des hautes fréquences) que dans des paroles comparables DA (Fernald, 1989). Même des auditeurs qui proviennent d'une culture allophone non-industrialisée et non-alphabétisée peuvent identifier les intentions dans les paroles DB (Bryant & Barrett, 2007). De plus, des cultures ou des genres généralement reconnus pour être moins expressifs dans leurs manifestations explicites

d'émotions que les mères nord-américaines (le groupe le plus fréquemment étudié), par exemple les mères japonaises (Fernald & Morikawa, 1993; Masataka, 1992b) ou les hommes nord-américains (Werker & McLeod, 1989), présentent des modifications prosodiques moins prononcées dans leurs paroles DB.

Les nourrissons présentent un intérêt et plus d'affect positif (ex. sourires) en réaction à des manifestations vocales positives de l'adulte (Fernald, 1993; Haviland & Lelwica, 1987; Papousek, Bornstein, Nuzzo, Papousek, & Symmes, 1990). Par exemple, vers 2 mois, les bébés manifestent plus d'expressions faciales joyeuses et regardent plus longtemps leur mère exprimant devant eux de la joie (en mode vocal et facial), comparativement à leur mère exprimant de la tristesse ou de la colère (Haviland & Lelwica, 1987). Vers 4-5 mois, ils vont présenter plus d'affect positif (ex. sourire) et écouter plus longtemps des phrases d'approbation (positives) que des phrases prohibitives (négatives), même dans un contexte de langage non familier (Fernald, 1993) ou en l'absence de contenu sémantique (Papousek et al., 1990). De surcroît, ils écoutent davantage des phrases joyeuses en comparaison avec des phrases tristes ou neutres (Kitamura & Burnham, 1998), et ce, indépendamment du registre DB ou DA (Singh et al., 2002). Notamment, les bébés regardent plus longtemps du langage joyeux DA plutôt que du langage neutre DB, alors qu'ils ne regardent pas plus longtemps ni les paroles DB ni DA lorsque l'affect est constant, suggérant qu'ils sont attirés vers l'affect positif en général (Singh et al., 2002). Il en va de même pour les expressions faciales positives, qui captent davantage leur attention que des expressions neutres (Kim & Johnson, 2013; Kuchuk, Vibbert, & Bornstein, 1986; Serrano, Iglesias, & Loeches, 1995). Faire sourire un bébé pourrait être un puissant motivateur pour le style DB, puisque des aires corticales associées au système dopaminergique de la récompense sont davantage activées dans le cerveau des mères lorsqu'elles voient leur bébé sourire que lorsque celui-ci présente un affect neutre ou triste (Strathearn, Li, Fonagy, & Montague, 2008).

Parmi les caractéristiques acoustiques impliquées dans les paroles DB, les variations de hauteurs pourraient jouer un rôle important dans les réponses attentionnelles des bébés. Ainsi, l'élimination des variations (i.e. une fréquence fondamentale ou F0 moyenne constante) annule la préférence pour les paroles DB chez les bébés de 4 mois (Fernald & Kuhl, 1987). Lorsque la F0 (hauteur moyenne et variations) est appariée approximativement pour deux

stimuli de paroles DB, les bébés préfèrent écouter les paroles DB jugées d'affect positif élevé par des adultes en comparaison avec des paroles DB jugées plus neutres (Kitamura & Burnham, 1998), mais ne démontrent aucune préférence pour les deux sons lorsque l'affect est apparié mais que la F0 (hauteur moyenne et variations) diffère (Kitamura & Burnham, 1998; Singh et al., 2002). Ces résultats suggèrent que l'affect détiendrait un rôle primordial afin de capter l'attention, alors que les variations de hauteurs pourraient détenir un rôle secondaire, possiblement en lien avec l'expressivité affective. En fait, les variations de hauteurs sont généralement importantes pour l'expression et la reconnaissance des émotions en musique ainsi que dans la parole conventionnelle (DA) (Bowling, Sundararajan, Han, & Purves, 2012; Laukka, Juslin, & Bresin, 2005; Scherer, 1995). Toutefois d'autres facteurs auraient leur importance dans l'expression, tels que le tempo/débit, l'amplitude/intensité et la répartition des fréquences spectrales (Juslin & Laukka, 2003).

Au-delà de viser à capter l'attention de l'enfant ou le faire sourire, les paroles DB ont aussi des répercussions sur le comportement social des bébés. Des bébés de 5 mois vont regarder plus longtemps la photo, en silence, d'une personne leur ayant adressé des paroles DB plutôt qu'une nouvelle personne; à l'inverse, ils vont regarder plus longtemps la photo d'une nouvelle personne plutôt que celle d'une personne leur ayant adressé des paroles DA (Schachner & Hannon, 2011). S'adresser aux bébés selon des paroles DB pourrait conséquemment renforcer des liens interpersonnels.

2.2 Les préférences attentionnelles relatives à la parole et à la musique

La majorité des études sur les préférences sonores des bébés se sont intéressées à la comparaison des paroles DB versus DA. Néanmoins, les bébés présenteraient aussi des biais attentionnels pour les paroles naturelles, en comparaison avec des sons artificiels, tels que du bruit blanc (Colombo & Bundy, 1981) ou des fréquences sinusoïdales (Vouloumanos & Werker, 2004, 2007). Bien que les nouveau-nés préfèrent des vocalisations naturelles provenant de singes ou d'humains à des sons artificiels, une préférence pour des vocalisations humaines plutôt que des singes n'émergerait qu'aux alentours de 3 mois (Vouloumanos, Hauser, Werker, & Martin, 2010), indiquant que la familiarité y jouerait un rôle. De plus, les bébés de 3 mois présentent un intérêt pour les paroles en comparaison avec des vocalisations

humaines non-langagières telles que des éternuements ou des toux, suggérant que les bébés favorisent des signaux communicatifs plutôt que non-communicatifs (Shultz & Vouloumanos, 2010). Quelques études ont aussi investigué l'existence de biais attentionnels dans le domaine musical, par exemple pour la consonance vs la dissonance (Trainor & Heinmiller, 1998; Zentner & Kagan, 1998) la simplicité rythmique (Soley and Hannon, 2010) ou des versions vocales (chantées *a capella*) plutôt que des versions vocales-instrumentales d'une chanson (Ilari & Sundara, 2009). De façon générale, il semble que les vocalisations communicatives, humaines ou non, attirent l'attention des bébés.

3. Régulation des émotions

Les bébés ayant des capacités d'autorégulation très limitées (Kopp, 1989; Thompson, 1994), le rôle des donneurs de soins aux enfants (encore à ce jour principalement les mères) devient primordial dans la régulation de leur émotions, non seulement dans l'immédiat, par exemple en calmant l'enfant, mais également à long terme, telles que leur capacité à s'autoréguler ou leurs compétences sociales (Cole, Martin, & Dennis, 2004; DiCorcia & Tronick, 2011; Feldman, Greenbaum, & Yirmiya, 1999; Gunnar & Quevedo, 2007; Kopp, 1989; Trehub, Ghazban, & Corbeil, 2015). Plusieurs comportements peuvent être utilisés afin de réguler le comportement du bébé : le toucher (Jean & Stack, 2009), porter ou bercer l'enfant (Gammie, 2013) ou encore le mode vocal, soit les paroles très mélodieuses dirigées vers l'enfant (Fernald, 1992; Papoušek, 1994) ainsi que le chant (Trehub, Ghazban, et al., 2015; Trehub et al., 2011). Les conséquences régulatrices du toucher sur l'enfant (Feldman, Singer, & Zagoory, 2010; Stack & Muir, 1992; Tronick, 1995) ainsi que de bercer ou de transporter l'enfant ont été documentées (Byrne & Horowitz, 1981; Esposito et al., 2013; Hunziker & Barr, 1986; Vrugt & Pederson, 1973). Par exemple, le fait de transporter le bébé (marcher) en comparaison avec le tenir en étant immobile, calme les bébés humains, mais également les souriceaux, plus rapidement, en réduisant les mouvements moteurs, les vocalisations et le rythme cardiaque (Esposito et al., 2013).

Pourtant, peu d'études se sont attardées aux compétences régulatrices des vocalisations DB sur les bébés. Le système auditif est relativement plus développé que le système visuel à la

naissance, avec comme conséquence que les jeunes bébés distinguent mieux les expressions émotionnelles vocales que faciales (Grossmann, 2010). Dans des situations ambiguës ou menaçantes, les bébés de 12 mois décodent mieux l'intention émotionnelle vocale de la mère plutôt que lorsqu'elle s'exprime seulement par son visage (Mumme, Fernald, & Herrera, 1996; Vaish & Striano, 2004).

3.1 Les vocalisations DB en fonction de la relation parent-enfant

Les vocalisations envers un bébé sont variables, selon le contexte et l'âge de l'enfant, mais également en fonction de mécanismes interactifs entre l'adulte et l'enfant, i.e. que l'adulte influence l'enfant, mais l'enfant influence aussi l'adulte. D'abord, les vocalisations DB peuvent communiquer une variété d'intentions affectives dépendant du contexte. Les paroles DB peuvent évidemment viser une multitude d'intentions ayant chacune leur propre signature acoustique, telles que calmer, capter l'attention, approuver ou encore proscrire un comportement (Fernald, 1989). Typiquement toutefois, les paroles DB sont joyeuses et s'apparentent acoustiquement aux vocalisations DA de joie ou bonheur à haute activation ou d'allégresse plutôt que de basse activation (Banse & Scherer, 1996; Bänziger & Scherer, 2005). Les intentions véhiculées dans les chansons DB sont généralement plus restreintes. Elles se limitent typiquement à deux grandes classes de chansons pour les bébés et les jeunes enfants, soit la chanson joyeuse (*playsong*) visant à amuser ou stimuler, versus la berceuse (*lullaby*) visant à calmer ou endormir l'enfant (Trehub & Trainor, 1998). La berceuse s'apparenterait acoustiquement davantage à une émotion de tendresse ou joie de basse activation alors que la chanson joyeuse s'apparenterait plutôt à une joie de haute activation (Banse & Scherer, 1996; Bänziger & Scherer, 2005).

Les vocalisations DB varient également selon l'âge de l'enfant. Plusieurs caractéristiques des paroles DB telles que la hauteur maximale, l'ampleur de ses variations et la quantité de répétitions augmentent progressivement de la naissance jusqu'à 4 mois, pour diminuer dans les mois subséquents (Stern et al., 1983). Des changements reliés à l'âge pour le chant DB n'ont pas été documentés, autre que de noter que les mères élèvent la hauteur de leur voix et parlent moins clairement lorsqu'elles s'adressent à leur bébé qu'à leur enfant préscolaire (Bergeson & Trehub, 1999).

Ces modifications développementales reflètent possiblement des changements dans les intentions maternelles autant que dans les préférences des bébés. Les mères accentueraient différentes intentions émotionnelles selon l'âge, selon des phrases plus réconfortantes (avec une hauteur moyenne plus grave) pour les bébés de 3 mois, plus affectives (avec une hauteur élevée) à 6 mois et progressivement plus directives (avec une diminution de la hauteur de la voix) à 9 mois (Kitamura & Burnham, 2003). Qui plus est, les bébés démontrent, selon leur âge, des préférences attentionnelles reliées à ces intentions affectives (Lam & Kitamura, 2006), suggérant que, jusqu'à un certain point, les intentions véhiculées par les mères reflèteraient les besoins de l'enfant selon les différents stades de son développement.

Les bébés réagissent à l'intentionnalité dans les paroles de la mère. Dès 2 mois, ils réagissent de façon différentielle à leur mère leur parlant de façon joyeuse, triste ou fâchée (Haviland & Lelwica, 1987). Toutefois, les nouveau-nés ne présentent des réponses différentielles à des expressions émotionnelles que dans leur langue maternelle et non dans une langue étrangère, suggérant une certaine contribution de l'expérience prénatale (Mastropieri & Turkewitz, 1999).

Les bébés réagissent également différemment selon l'intention visée par la chanson. Ils démontrent davantage d'attention vers soi lors de l'écoute d'une chanson interprétée comme une berceuse et davantage d'attention dirigée vers l'extérieur (ex. regarder le parent ou ailleurs) lors de l'écoute de cette même chanson interprétée comme une chanson joyeuse (Rock, Trainor, & Addison, 1999).

De surcroît, les bébés ont aussi un rôle actif dans l'interaction, puisque les mères modifient certains paramètres acoustiques de leurs paroles selon les réactions du bébé. Par exemple, elles ajustent la hauteur de leur voix, en l'augmentant ou la diminuant, selon les réactions positives du bébé par rapport à ces ajustements (Smith & Trainor, 2008). De plus, les mères articulent davantage les voyelles lorsque leur bébé les entend que lorsque celui-ci les entend peu ou pas du tout (Lam & Kitamura, 2012). Ces ajustements seraient indépendants du fait que les mères croient être entendues par l'enfant ou non. Ces résultats suggèrent que certaines modifications prosodiques dans les paroles de la mère ne sont pas volontaires et dépendent de l'enfant.

3.2 La musique comme outil de régulation affective

La musique serait particulièrement appropriée pour réguler l'état affectif de l'enfant. Une des principales fonctions auto-rapportées de l'écoute musicale chez l'adulte dans les cultures occidentales contemporaines consiste à s'autoréguler émotionnellement, notamment en relaxant, prolongeant un état de plaisir ou de bien-être ou encore en évacuant des émotions négatives (Greasley & Lamont, 2006; Saarikallio, 2011; Sloboda, Lamont, & Greasley, 2009). La musique pourrait influencer le bien-être et la santé via plusieurs systèmes neurochimiques incluant le système de la récompense/motivation/plaisir, le système du stress/activation (*arousal*), le système immunitaire et les mécanismes d'affiliations sociales (Chanda & Levitin, 2013). Ces effets confèreraient à la musique d'être hautement bénéfiques, mais les mécanismes sous-tendant ces influences demeurent à être identifiés. L'utilité de nos facultés musicales résideraient possiblement dans la modulation de notre état afin d'atteindre une homéostasie de l'organisme (Habibi & Damasio, 2014), par l'entremise de l'influence sur ces différents systèmes.

Chez l'adulte, la musique peut générer une série d'émotions positives allant de la tranquillité, de la tendresse à l'euphorie ou à des « frissons » de plaisir (Blood & Zatorre, 2001; Menon & Levitin, 2005; Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011). De façon similaire aux adultes, les nouveau-nés présentent une activité cérébrale accrue dans certaines régions du système limbique (e.g. l'amygdale) dédiées au traitement émotionnel lors d'écoute musicale, suggérant que l'activation cérébrale émotionnelle à la musique serait présente dès la naissance (Perani et al., 2010).

Certains auteurs ont avancé l'hypothèse que la musique serait traitée par le cerveau comme une voix super-expressive (Juslin & Västfjäll, 2008). La musique s'apparenterait ainsi au super-stimulus de Tinbergen (1963), avançant qu'il est possible de créer des stimuli artificiels plus forts afin de libérer un instinct ou une réaction que les stimuli naturels à l'origine de cette réaction. Celle-ci pourrait s'immiscer dans les réseaux neuronaux responsables du traitement émotionnel de stimuli biologiquement significatifs tels que la voix (e.g. hypothèse de l'invasion neuronale: Peretz, Aubé, & Armony, 2013). Curieusement, Ilie et Thompson (2006) observent que des extraits musicaux sont en moyenne jugés comme ayant

une plus grande valeur émotionnelle, en termes de valence positive (plaisance) et d'intensité (énergie), que des extraits langagiers. Les auteurs spéculent que ces jugements résulteraient de conséquences affectives différentes chez l'auditeur, concordant avec l'idée que la musique viserait à modifier un état affectif, induire du plaisir ou réguler notre niveau d'énergie, par exemple.

L'écoute musicale avant, pendant ou après un événement stressant réduit les niveaux de cortisol chez l'adulte (Khalifa, Dalla Bella, Roy, Peretz, & Lupien, 2003; Leardi et al., 2007) et des effets modulateurs du chant maternel sur le niveau cortisol s'observent aussi chez les bébés (Shenfield, Trehub, & Nakata, 2003). Le chant maternel, lors d'une interaction face-à-face entre la mère et l'enfant, modulerait le niveau de stress ou d'activation des bébés, tel que mesuré par le niveau de cortisol salivaire, soit en le rehaussant, soit en le diminuant selon le niveau de cortisol préalable de l'enfant.

Dans des contextes sociaux, la synchronisation entre individus (Valdesolo, Ouyang, & DeSteno, 2010; Wiltermuth & Heath, 2009) et aux rythmes de la musique (Kokal, Engel, Kirschner, & Keysers, 2011) favoriserait des comportements pro-sociaux. Chez l'enfant, la synchronisation de celui-ci avec autrui dans des contextes musicaux aurait aussi des répercussions pro-sociales. Lorsque des adultes bougent sur de la musique en synchronie avec des enfants de 14 mois, les enfants sont subséquemment plus susceptibles d'aider cet adulte qu'un adulte ayant bougé en asynchronie (Cirelli, Einarson, & Trainor, 2014). Des enfants de 4 ans, ayant produit de la musique et dansé sur celle-ci en groupe, sont ensuite plus susceptibles d'aider un autre enfant que ceux ayant fait une activité commune en parlant (Kirschner & Tomasello, 2009). Les perceptions et comportements sociaux pourraient être régulés par l'influence de la musique sur les systèmes hormonaux (Chanda & Levitin, 2013; Fukui & Toyoshima, 2014), par exemple par la sécrétion de l'hormone ocytocine qui pourrait réguler le comportement social (Heinrichs, von Dawans & Domes, 2009). Bien que les effets du chant sur les niveaux d'ocytocine n'ont pas à ce jour été investigués, la parole maternelle, très mélodieuse et en mode auditif seulement, serait efficace afin d'augmenter l'ocytocine chez des bébés préalablement exposés à une condition stressante (Seltzer, Ziegler, & Pollak, 2010).

4. Études comparatives entre la parole et le chant

Alors que certains défendent des prédispositions précoces à être attirés envers la parole (Shultz & Vouloumanos, 2010; Vouloumanos et al., 2010) ou des prédispositions précoces à être attirés envers la musique (Trehub & Hannon, 2006), très peu d'études se sont intéressées aux effets différentiels du chant et des paroles sur les bébés. Les études examinant les corrélats neuronaux de la musique et de la parole utilisent typiquement de la musique instrumentale en comparaison avec des paroles, amplifiant ainsi les différences entre les deux modes d'expression et observant des résultats variables (e.g. Dehaene-Lambertz et al., 2010; Kotilahti et al., 2010). Une étude en MEG ayant comparé le chant et les paroles n'observe aucune différence d'activation cérébrale chez les nouveau-nés entre une chanson joyeuse et les paroles de celle-ci récitées (Sambeth, Ruohio, Alku, Fellman, & Huotilainen, 2008). Certains chercheurs ont suggéré que les différences observées dans les cerveaux adultes émergeraient graduellement avec l'âge (Brandt et al., 2012). Toutefois, il n'est pas possible de tirer des conclusions basées sur une seule étude d'imagerie sans différences observables. Même chez l'adulte, un chevauchement des aires corticales activées n'est pas une preuve que les mêmes circuits sont impliqués dans le traitement des deux sons (Peretz, Vuvan, Lagrois, & Armony, 2015).

Une étude a comparé les réactions comportementales des bébés relatives à des extraits audio-visuels de chants et de paroles maternels (dans un design inter-sujets) (Nakata & Trehub, 2004). Les bébés de 6 mois regardent plus longtemps l'écran (basé sur un temps de regard cumulatif sur 3 minutes d'exposition) et plus intensément (réduction des mouvements) lors de vidéos de leur mère qui chantent que les bébés exposés à leur mère qui leur parlent. De plus, leur regard initial sur l'écran est également plus long lors des vidéos de chant que des paroles. Toutefois, puisque que la présentation était audio-visuelle, il est difficile d'identifier les contributions spécifiques des facteurs auditifs et visuels, ainsi que les paramètres musicaux contribuant à cette différence dans l'attention et l'engagement des bébés.

5. Objectifs et hypothèses

En résumé, en dépit de la ressemblance acoustique du chant et la parole pour l'enfant préverbal, peu de comparaisons expérimentales existent visant à mesurer leur impact comportemental chez les bébés. Les deux études incluses dans la présente thèse visent à comparer, pour la première fois, les conséquences comportementales de la parole et du chant dirigés vers le bébé (DB) en mode auditif seulement, spécifiquement en ce qui a trait à capter l'attention et à réguler l'affect, afin de mieux cerner les facteurs d'influence. Dans tous les cas, les voix parlées et chantées étaient non familières pour l'enfant, afin de maintenir une uniformité des stimuli à travers les participants. Les sections subséquentes sont présentées sous forme d'articles publiés.

5.1 Première étude : Capter l'attention

La première étude vise à évaluer l'influence relative du chant et des paroles afin de capter l'attention des bébés de 4 à 13 mois. En fait, il s'agit de la première étude comparant directement l'attention des bébés aux chants et aux paroles. Spécifiquement, elle vise à comparer des paroles selon différents types de chants et à investiguer la possibilité de changements reliés à l'âge du bébé. Pour ce faire, l'attention des bébés a été mesurée grâce à une procédure de regard préférentiel, où leur regard envers un stimulus visuel contrôle la durée du son présenté (e.g. Singh et al., 2002). Une préférence peut être inférée selon la capacité du bébé à réguler son attention relative entre deux stimulations sonores.

L'hypothèse principale est que le chant pourrait être plus efficace que les paroles afin d'attirer l'attention des bébés à travers les âges étudiés, selon une étude précédente démontrant que les bébés accordent plus d'attention (regard initial et regards cumulatifs) au chant maternel qu'à la parole (Nakata & Trehub, 2004). Toutefois, considérant l'impact d'un affect joyeux sur les préférences attentionnelles des bébés pour les paroles (Kitamura & Burnham, 1998; Singh et al., 2002), il est attendu que l'affect joyeux aura aussi un impact sur leurs préférences attentionnelles.

5.2 Deuxième étude : Réguler l'affect

La deuxième étude est la première à explorer les conséquences du chant et des paroles dans la régulation affective de bébés. Pour ce faire, des bébés de 7 à 10 mois ont été exposés soit à du chant joyeux DB, soit à des paroles joyeuses DB ou encore à des paroles neutres DA, jusqu'à ce qu'ils présentent des signes de mécontentement, mesurés par un affect facial négatif. Afin d'inférer une régulation émotionnelle, il importe de démontrer qu'une émotion soit bel et bien activée et qu'un processus de régulation ait lieu indépendamment (Cole et al., 2004). Une émotion peut être conséquemment régulée si un changement est observé dans sa valence, son intensité ou un changement au cours du temps, dans la durée d'activation ou d'apparition par exemple. La régulation est ainsi opérationnalisée en termes de délai sur l'affect. Évidemment, l'augmentation du délai d'apparition d'un affect négatif pourrait également impliquer la prolongation d'un état de contentement.

En se basant principalement sur un intérêt soutenu et plus intense des bébés face à des extraits audio-visuels de chants maternels plutôt que de paroles maternelles (Nakata & Trehub, 2004) ainsi que la capacité du chant maternel à moduler les niveaux de cortisol (l'activation ou l'*arousal*) des bébés (Shenfield et al., 2003), l'hypothèse est que des extraits de chants, en mode auditif seulement, pourraient être plus efficaces que des paroles afin de réguler l'état affectif des bébés.

Chapitre II : Partie expérimentale

Article I

Speech vs. singing: infants choose happier sounds

Mariève Corbeil^{1,2,3}, Sandra E. Trehub⁴, and Isabelle Peretz^{1,2,3}

¹Department of Psychology, Université de Montréal

²The International Laboratory for Brain Music and Sound Research (BRAMS), University of Montréal, Montréal, Québec, Canada

³Centre for Research on Brain, Language and Music (CRBLM), McGill University, Montréal, QC, Canada

⁴Department of Psychology, University of Toronto Mississauga, Mississauga, Ontario, Canada

Article publié : Corbeil, M., Trehub, S. E., & Peretz, I. (2013). Speech vs. singing: infants choose happier sounds. *Frontiers in Psychology*, 4, 372-372. DOI: [10.3389/fpsyg.2013.00372](https://doi.org/10.3389/fpsyg.2013.00372)

Running page heading: INFANTS CHOOSE HAPPIER SOUNDS

Abstract

Infants prefer speech to non-vocal sounds and to non-human vocalizations, and they prefer happy-sounding speech to neutral speech. They also exhibit an interest in singing, but there is little knowledge of their relative interest in speech and singing. The present study explored infants' attention to unfamiliar audio samples of speech and singing. In Experiment 1, infants 4-13 months of age were exposed to happy-sounding infant-directed speech versus hummed lullabies by the same woman. They listened significantly longer to the speech, which had considerably greater acoustic variability and expressiveness, than to the lullabies. In Experiment 2, infants of comparable age who heard the lyrics of a Turkish children's song spoken versus sung in a joyful/happy manner did not exhibit differential listening. Infants in Experiment 3 heard the happily sung lyrics of the Turkish children's song versus a version that was spoken in an adult-directed or affectively neutral manner. They listened significantly longer to the sung version. Overall, happy voice quality rather than vocal mode (speech or singing) was the principal contributor to infant attention, regardless of age.

Keywords: infants, music, language, singing, speech, emotion, attention

Introduction

There is considerable debate about similarities and differences in the processing of language and music (e.g., Pinker, 1997; Patel, 2008; Jackendoff, 2009; Peretz, 2009). Because the greatest differences arise from the presence of propositional meaning in language but not in music, comparisons in the early pre-verbal period are of particular interest (Trehub et al., 1993; Chen-Hafteck, 1997; McMullen & Saffran, 2004; Brandt et al., 2012), notably when both modes of parental communication are used to regulate infant attention and affect (Fernald, 1992; Papoušek, 1994; Kitamura & Burnham, 2003; Trehub et al., 2010). To date, however, the only study comparing young infants' behavioral responsiveness to speech and singing (Nakata & Trehub, 2004) used audiovisual stimuli, obscuring the relative contributions of auditory and visual expressiveness to infants' greater engagement with maternal music. Another study found no difference in newborns' neural responses to happy-sounding speech and singing (Sambeth et al., 2008). The present investigation examined infants' attentiveness to speech and singing on the basis of auditory cues alone.

Whereas verbal aspects of speech convey propositional meaning, non-verbal or prosodic aspects such as intonation and rhythm convey the speaker's affective intent and emotional state (Frick, 1985). Mothers across cultures speak and sing to their pre-verbal infants in the course of providing care (Fernald, 1992; Trehub & Trainor, 1998; Dissanayake, 2000; Trehub, 2000). Their manner of speaking or singing to infants (infant-directed or ID) differs dramatically from their manner in other contexts (adult-directed or AD; self-directed or non-ID) (Ferguson, 1964; Jacobson et al., 1983; Fernald & Simon, 1984; Trainor et al., 1997; Trehub et al., 1997a; Trehub et al., 1997b), with notable variations across cultures (Grieser & Kuhl, 1988; Fernald et al., 1989; Kitamura et al., 2002). In general, ID speech features higher pitch, expanded pitch contours, slower speaking rate, longer vowels, larger dynamic range, and greater rhythmicity and repetition than AD speech (Stern et al., 1982; Stern et al., 1983; Fernald & Simon, 1984; Fernald et al., 1989). These features, especially high pitch, expanded pitch contours, rhythmicity, repetition, and reduced speaking rate, make ID speech sound much more musical than AD speech (Fernald, 1989; 1992). High pitch, expanded pitch contours, and large dynamic range also reflect the heightened affective quality of typical ID speech, which contrasts with the affective restraint of typical AD speech (Trainor et al., 2000). Nevertheless, ID speech is finely tuned to the infant's age and needs, with mothers using relatively more comforting speech for 3-month-olds, more approving speech for 6-

month-olds, and more directive speech for 9-month-olds (Kitamura & Burnham, 2003). Approving speech, with its higher pitch and greater pitch range, receives higher ratings of positive affect by adult listeners (Kitamura & Lam, 2009).

Unlike speech, singing is constrained by the prescribed pitch and rhythmic form of the material (i.e., specific songs). Nevertheless, ID versions of singing are also characterized by higher pitch and slower tempo than non-ID versions of the same songs by the same singers (Trainor et al., 1997; Trehub et al., 1997a; Trehub et al., 1997b). While repetition is an important aspect of ID speech, it is central to music in general (Kivy, 1993; Trainor & Zatorre, 2008) and to songs for young children in particular (Trehub & Trainor, 1998).

The available evidence indicates that infants find ID singing more engaging than non-ID singing (Trainor, 1996; Masataka, 1999) just as they find ID speech more engaging than AD speech (Fernald, 1985; Werker & McLeod, 1989; Pegg et al., 1992). One possible source of infants' enhanced engagement is the heightened positive expressiveness of typical ID speech and singing (Trainor et al., 2000; Trehub et al., 2010; Nakata & Trehub, 2011). In fact, infants exhibit preferential listening to speech that sounds happy rather than sad or inexpressive regardless of the intended audience (Kitamura & Burnham, 1998; Singh et al., 2002). For example, infants listen longer to happy AD speech than to affectively neutral ID speech even when the latter is higher in pitch (Singh et al., 2002). Note, however, that happy ID vocalizations are closer to AD vocalizations described as *high-arousal joy/happiness* or elation rather than to *low-arousal joy/happiness* (Banse & Scherer, 1996; Bänziger & Scherer, 2005). Infants also exhibit more positive affect to ID expressions of approval than to disapproval or prohibition even when the utterances are low-pass filtered (Papoušek et al., 1990) or presented in an unfamiliar language (Fernald, 1993). The general consensus is that positive vocal emotion, especially the high arousal variety, makes a substantial contribution to infants' interest in ID speech. Nevertheless, one cannot rule out alternative explanations such as the attention-getting potential of expanded pitch and dynamic range and the attention-holding potential of repetition. When these acoustic factors are controlled, however, infants exhibit preferences for the happier speech version (Kitamura & Burnham, 1998; Singh et al., 2002), suggesting that these acoustic features make secondary contributions to infant preferences. Infants' interest is also affected by their age and corresponding needs. For example, 3-month-old infants exhibit greater attention to comforting than to approving ID speech (Kitamura & Lam, 2009).

The influence of ID pitch contours is seen in infants' preferential listening for sine-wave replicas of ID speech that preserve the pitch contours (and timing) with uniform amplitude over those that preserve the timing and amplitude with unvarying pitch (Fernald & Kuhl, 1987). Despite the fact that infants display greater positive affect to approving than to disapproving ID utterances, they listen longer to the former only if they exhibit greater F0 modulation (Fernald, 1993). Interestingly, pitch modulation also makes important contributions to the differentiation of emotions in music and in AD speech (Scherer, 1986, 1995; Laukka et al., 2005). Across cultures, happy-sounding speech and music feature high mean pitch, large pitch variability, relatively high mean amplitude, and rapid rate or tempo (Juslin & Laukka, 2003). Smiling elevates pitch and increases amplitude by altering the mouth opening and shape of the vocal tract, contributing to the vocal qualities associated with happiness (Tartter, 1980). Tender speech and music, by contrast, have lower mean pitch, pitch variability, mean amplitude, and slower rate or tempo than happy speech and music (Juslin & Laukka, 2003).

Perhaps the two classes of songs for infants, lullabies and play songs, are caregivers' expressions of tenderness and happiness, respectively, as well as tools for soothing or amusing infants. In line with their soothing function, lullabies feature very slow tempo, low pitch, falling pitch contours, limited amplitude variation, and soothing tone of voice (Unyk et al., 1992; Trehub et al., 1993; Trehub & Trainor, 1998), properties that are shared with soothing ID speech (Papoušek & Papoušek, 1981; Fernald, 1989). Lullabies are also soothing to adult listeners, so it is not surprising that they are used, at times, as laments (Trehub & Prince, 2010) and in palliative care (O'Callaghan, 2008). Although play songs are commonly sung to Western infants, they are not universal, as lullabies are (Trehub & Trainor, 1998).

Maternal speech melodies are considered central to the expression of maternal affect and the regulation of infant attention and arousal (Fernald, 1992; Papoušek, 1994). Is it possible that musical melodies would be equally effective or even more effective in regulating infant attention and arousal? The melodies or pitch contours of expressive speech differ from those in music (Zatorre & Baum, 2012). In music, pitches are discrete and sustained, and steps from one pitch level to another are generally small, most commonly, one or two semitones, with larger pitch jumps much less frequent (Vos & Troost, 1989). By contrast, pitches in speech glide continuously over a larger range (Patel et al., 1998), which is even larger in ID speech (Ferguson, 1964; Stern et al.,

1982; Stern et al., 1983; Fernald & Simon, 1984). Moreover, pitches have precise targets in music but not in speech (Zatorre & Baum, 2012).

If the expanded pitch and dynamic range of ID speech underlies infants' greater attention to ID than to AD speech (e.g., Fernald, 1993), then infants could show more interest in ID speech than ID singing. If rhythmicity and predictability are relevant (e.g., McRoberts et al., 2009), then infants might exhibit more attention to ID singing than to ID speech. If positive emotion is the critical feature (Kitamura & Burnham, 1998; Singh et al., 2002), then infants could show greater interest in the stimulus expressing more positive affect regardless of whether it is speech or music. For adults, music generates a range of positive emotions from tranquillity and tenderness to joy and euphoria (Blood & Zatorre, 2001; Menon & Levitin, 2005; Zentner et al., 2008; Salimpoor et al., 2011). Some scholars contend that the expression of emotion by some form of music (e.g., protomusic) preceded language (Darwin, 1871; Mithen, 2005). Others regard speech, even at present, as a type of music, especially when considered in developmental perspective (Brandt et al., 2012). If the status of speech is privileged, as some contend (Vouloumanos & Werker, 2004, 2007; Shultz & Vouloumanos, 2010; Vouloumanos et al., 2010), then ID speech would be favored over forms of singing that exclude speech. Obviously, the aforementioned factors are not independent. Nevertheless, comparisons of infants' responsiveness to speech and music are a first step toward the long-range goal of identifying the acoustic features that attract and hold infants' attention. Such features may differ for infants of different ages, as reflected in age-related changes in listening biases for ID speech with comforting, approving, or directive tones of voice (Kitamura & Lam, 2009) and for regular or slowed ID speech (Panneton et al., 2006).

It is difficult to assess infants' degree of engagement with music and even more difficult to ascertain their aesthetic preferences. Instead of overt affective responses to music, infants commonly exhibit interest or attention, sometimes accompanied by reduced motor activity (Nakata & Trehub, 2004). The usual assumption is that longer listening to one of two auditory stimuli reflects preference or greater liking for that stimulus (e.g., Fernald & Kuhl, 1987; Trainor, 1996; Vouloumanos & Werker, 2004). In general, such "preferences" are assessed with the head-turn preference procedure, which is used with infants as young as 2 or 3 months of age (e.g., Trainor et al., 2002; Shultz & Vouloumanos, 2010). The procedure involves pairing one auditory stimulus with a visual display and a contrasting auditory stimulus with the same visual display, at the same or different locations, on a series of trials. Infants control the procedure in the sense that looking

away from the visual stimulus terminates the auditory stimulus. In other words, they can *choose* to listen to one stimulus longer than another. The interpretation of longer or shorter listening times as positive or negative aesthetic evaluations is questionable in the absence of positive or negative affective displays (Trehub, 2012). At times, infants listen longer to familiar stimuli and, at other times, to novel stimuli (e.g., Rose et al., 1982; Volkova et al., 2006; Soley & Hannon, 2010). Even when infants show positive affect to one auditory stimulus and negative or neutral affect to another, their listening times to the stimuli may not differ (Fernald, 1993). Unquestionably, looking or listening times indicate infants' listening choice or relative attention to the stimuli, but the factors that contribute to such attention are unclear. Some listening biases may be innate, arising from the salience of biologically significant stimuli (e.g., human vocal sounds) or biologically significant parameters of sound (e.g., loud or unexpected). Other listening biases may arise from acquired salience, as in preferential responding to the sound of one's name (Mandel et al., 1995) or to a stimulus heard previously (Zajonc, 2001). Attention biases, regardless of their origin, are likely to facilitate learning (Vouloumanos & Werker, 2004).

In addition to the well-documented listening bias for ID over AD speech, there are reported biases for vocal over non-vocal sounds (Colombo & Bundy, 1981; Vouloumanos & Werker, 2004; 2007), speech over non-human vocalizations, (Vouloumanos et al., 2010), speech over human nonspeech vocalizations (Shultz & Vouloumanos, 2010), musical consonance over dissonance (Trainor & Heinmiller, 1998; Zentner & Kagan, 1998), and familiar over unfamiliar musical meters (Soley & Hannon, 2010). Infants also exhibit considerable interest in vocal music (Glenn et al., 1981), but their exposure to music is much more limited than their exposure to speech (Eckerdal & Merker, 2009). To date, however, there has been little exploration of infants' relative interest in speech and singing. In the single study that addressed this question directly (Nakata & Trehub, 2004), 6-month-old infants watched audio-visual recordings of their mother singing or speaking from an earlier interaction. Infants showed more intense and more sustained interest in singing than in speech episodes, as reflected in greater visual fixation coupled with reduced body movement. Infants' heightened interest in these maternal singing episodes could stem from mothers' propensity to smile more when singing than when talking to infants (Plantinga et al., 2011). In the present study, we used the head-turn preference procedure to assess infants' interest in speech and singing with unfamiliar materials and voices. As noted above, the procedure provides information about infants' listening choices or relative attention rather than their aesthetic preferences.

In line with age-related changes in infants' attention to the affective tone of ID speech (Kitamura & Lam, 2009), developmental changes might be evident in infants' responsiveness to ID speech and song. Accordingly, infants in the present research, who were 4-13 months of age, were divided into three age groups to explore the possibility of comparable age-related changes. In Experiment 1, infants were exposed to ID or happy-sounding speech syllables and soothing hummed lullabies produced by the same woman. The principal question concerned the relative efficacy of soothing hummed song and happy ID speech for attracting and maintaining infants' attention. In other words, is vocal music compelling for infants, as it is for adults, even in the absence of speech or properties associated with heightened arousal? If infants listened longer to hummed lullabies than to simple ID speech, it would challenge the prevailing view that infants have an innate or early developing preference for speech over any other auditory stimulus (Vouloumanos & Werker, 2004; 2007; Shultz & Vouloumanos, 2010; Vouloumanos et al., 2010). Experiments 2 and 3 narrowed the differences between speech and singing stimuli by comparing the same verbal materials that were spoken or sung with comparable or contrasting affective intentions. Specifically, infants in Experiment 2 heard sung versus spoken renditions of the lyrics of a Turkish children's song, both in an ID/joyful manner. Infants in Experiment 3 heard the ID children's song versus a spoken version of the lyrics in an AD or affectively neutral manner.

All of the stimuli in the present study were portrayed or acted rather than being recorded during actual interactions with infants and adults. Early research on infants' responsiveness to ID and AD speech (e.g., Fernald, 1985) used recordings of women's interactions with their infant and with an adult experimenter. Such stimuli differed dramatically in content as well as expressiveness, making it difficult to identify the factors contributing to infants' responsiveness. Later research used portrayals of ID and AD speech (e.g., Singh et al., 2002; Kitamura & Lam, 2009) so that the content could be carefully controlled across speech registers. When studying infants' responsiveness to ID and non-ID singing (e.g., Trainor, 1996; Masataka, 1999), it is possible to use recordings of mothers' singing the same song in the presence or absence of their infant. Comparisons of natural ID speech and singing (e.g., Nakata & Trehub, 2004), however, necessarily differ in content as well as form. Because the features of ID speech and singing have been described extensively (e.g., Ferguson, 1964; Trainor et al., 1997), it is possible to create relatively natural portrayals of those stimuli. For practical as well as ethical reasons, most of the research on vocal emotion (e.g.,

Scherer, 1986, 1995; Juslin & Laukka, 2003) has used portrayals of various emotions rather than emotional expressions produced in natural contexts.

Experiment 1

The goal of the present experiment was to examine the possibility that infants might be more responsive to vocal music than to happy ID speech even for vocal music lacking the acoustic features (e.g., highly variable pitch and dynamics) and expressive intentions (high-arousal happiness) that have been linked to infant preferences for ID speech (e.g., Fernald, 1985; Singh et al., 2002). By using hummed songs, it was possible to generate vocal music without speech. Humming, usually with closed mouth, can be used to generate melodies with sustained nasal sounds that have low spectral amplitude (Kent et al., 2002). Because humming constrains amplitude modulation, it provides reduced scope for expressing high-arousal emotions. There are speculations, however, that humming played an important role in early hominid evolution, functioning like contact calls in other species (Jordania, 2010). At present, humming may be the most common type of informal, solitary singing.

We considered lullabies the musical genre of choice because of their suitability for humming, their universal use in caregiving (Trehub & Trainor, 1998), and their stark contrast with happy ID speech in acoustic features and affective intentions. As noted, lullabies transmit positive affective qualities such as tranquillity and tenderness both in their musical features and vocal tone. The ID speech stimuli approximated those used in previous research on infants' listening biases for speech (Vouloumanos & Werker, 2004, 2007). They consisted of nonsense syllables with typical exaggerated pitch contours and happy voice quality. For adults, it is likely that the lullabies, although unfamiliar, would have high aesthetic appeal, while the repetitive, high-pitched nonsense syllables would sound boring or worse. Nevertheless, the speech combined the exaggerated pitch contours and joyful expressiveness that have been linked to infant preferences in contemporary urban cultures (Fernald & Kuhl, 1987; Kitamura & Burnham, 1998; Singh et al., 2002). If infants share adults' aesthetic appraisals or favor universal forms, they would listen longer to the hummed versions of traditional lullabies. On the basis of previous research with Western infants, however, one might expect them to listen longer to the arousing and joyfully rendered speech.

Method

Participants

The sample consisted of 50 healthy, full-term infants who were 4.3 to 13.1 months of age ($M = 8.6$ months, $SD = 2.6$) divided into 3 age groups: 4 to 6 months ($M = 5.5$, $SD = .48$; $n = 16$), 7 to 9 months ($M = 8.6$, $SD = .87$; $n = 16$) and 10 to 13 months ($M = 11.5$, $SD = .74$; $n = 18$). No infant had a family history of hearing loss or personal history of ear infections, and all were free of colds or ear infections on the day of testing. An additional five infants failed to complete the test session because of fussiness. This experiment and others in this report were approved by the Arts and Sciences ethics committee of the University of Montreal, and written informed consent was obtained from all participating parents.

Stimuli

The speech stimulus, which was comparable to that used by Vouloumanos and Werker (2004) except for a different speaker, consisted of 12 variations of each of two nonsense syllables (*lif* and *neem*) spoken with ID prosody. Varied repetitions of each syllable had rising, falling, and rising-falling (i.e., bell-shaped) pitch contours. There were two versions of the syllabic sequence, differing only in the order of elements. Each sequence consisted of a semi-random ordering of syllables, with the constraint that any 4 consecutive syllables contained two instances each of *lif* and *neem*. Syllables were separated by silent inter-stimulus intervals (ISIs) of 300 to 500 msec, and the order of ISIs was randomly distributed, with a mean of 450 msec, as in Vouloumanos and Werker (2004). Each sequence was approximately 20 s in duration, and was repeated for an overall duration of 40 s. The music stimulus consisted of a hummed version of a lullaby. There were two traditional lullabies, one Chilean (in duple meter, AA form) and one German (in triple meter, AB form), each approximately 40 sec in duration and each assigned to half of the infants. Hummed and spoken stimuli were produced by a native speaker of English who had considerable music training, singing experience, and experience with children. She was instructed to produce the nonsense syllables in a lively ID manner and to hum the melodies as if lulling an infant to sleep. She listened to many samples of ID speech and singing beforehand (including the Vouloumanos and Werker's syllables) and used pictures of infants to help induce the appropriate mood for her speaking or lulling. Sample stimuli are presented in Supplementary Materials.

Acoustic features of the stimuli, which were measured with Praat software (Boersma & Weenink, 2010), are shown in Table I. Because pitch extraction software is prone to octave errors,

it is common to manually specify a minimum and maximum fundamental frequency (F0 in Hz) or to use a formula for setting the F0 range of each sound such as that suggested by De Looze and Hirsh (2008): floor = $q25*0.75$; ceiling = $q75*1.5$. We used this formula for acoustic analyses in the present study. Mean F0 was higher for singing ($M = 280.2$ Hz) than for speech ($M = 244.2$ Hz, difference of 2.46 semitones), but speech was more variable in F0, amplitude, and timing. The standard deviation (*SD*) of F0, a measure of pitch variability, was 3.81 and 3.40 semitones for speech and singing, respectively. As can be seen in Figure 1, which depicts the F0 contours, changes in pitch were larger and more abrupt for the speech than for the humming stimuli. Amplitude variation (*SD*), measured in the voiced portions of each sound, was 9.31 dB for speech and 4.46 dB for singing. The timing of the syllables was varied deliberately as in Vouloumanos and Werker (2004).

Table I. Acoustic features of stimuli.

	Stimuli				
	Experiment 1		Experiment 2 and 3		
	Humming	Syllables	ID singing	ID speech	AD speech
F0 mean (Hz)	280.20	244.23	351.14	312.28	210.24
F0 SD (semitones)	3.40	3.81	2.34	3.86	2.30
F0 range (semitones)	14.77	18.22	11.41	17.64	11.33
Overall duration	—	—	26.8	24.6	19.02

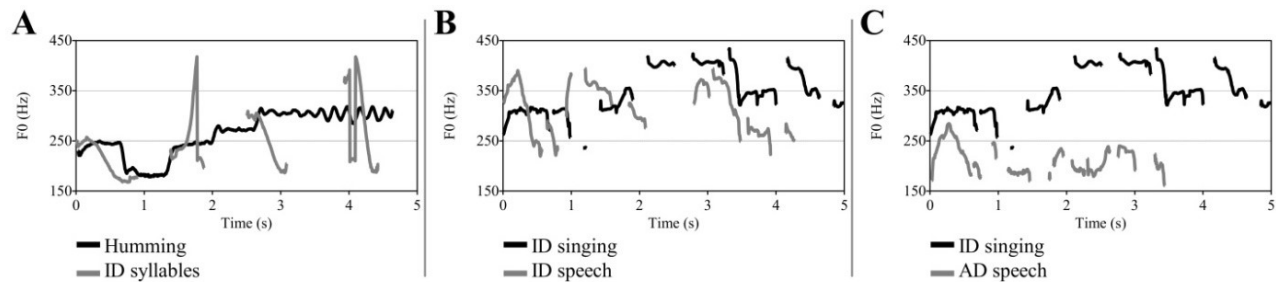


Figure 1. Fundamental frequency (F0) contours of 5-second excerpts from each sound type: (A) hummed lullaby (Chilean) and syllable sequence, (B) ID sung and spoken lyrics of Turkish play song, (C) ID sung and AD spoken lyrics of Turkish play song.

Apparatus

Testing was conducted in a sound-attenuating booth (IAC) 4 by 4 m in size. Infants were seated on their parent’s lap facing a central computer monitor at a distance of 127 cm, with two identical monitors to the right and left side of the central monitor and at a distance of 152 cm from infants. Parents wore earphones (ER-4 MicroPro with reusable ER-4S eartips) with an approximate attenuation of 35 dB and earmuffs (Peltor H10A, Optime 105) with an approximate attenuation of 30 dB. They heard continuous music through the headphones to mask the sounds presented to infants. The walls and table for the monitors were covered with black cloth to reduce visual distraction and optimize attention to the target stimuli. A camera immediately above the central screen provided a continuous record of infant visual behavior on a monitor outside the booth. Two loudspeakers (Genelec 8040A) located behind the lateral monitors transmitted the sounds at a comfortable listening level, approximately 60-65 dB (A). The procedure was controlled by customized software on a computer (Mac Pro 8 cores) located outside the booth.

Procedure

The head-turn preference procedure (Kelmer Nelson et al., 1995) was used. Infants remained seated on their parent’s lap throughout the procedure, and parents were asked to minimize their own movement. Infants were randomly assigned to one of the two speech sequences and one of the two hummed lullabies. The speech and singing stimuli were presented on 10 alternating trials, with order of stimuli (speech or singing first) and side of presentation (left or right) counterbalanced across infants. On each trial, the infant’s attention was attracted to one monitor by a flashing red square. As soon as the infant looked at that monitor, one sound stimulus was presented together

with a visual animation of a carousel. When the infant looked away from the monitor for more than 2 s, the visual and sound stimuli were terminated. The infant's attention was then attracted to the other monitor. Looking at that monitor initiated the same visual stimulus but the contrasting auditory stimulus, which continued until the infant looked away for 2 s. On each trial, the stimulus was always presented from the beginning (i.e., beginning of the lullaby or syllable sequence). The experimenter outside the booth, who had no access to sound (auditory stimuli or infant vocalization) and no information about test conditions, observed the infant's behavior on the external monitor and continuously recorded looking toward or away from each monitor in the booth by means of key codes on a computer keyboard. Looking times during the presentation of each stimulus type were computed automatically. Typically, infants completed the procedure in approximately 5 min.

Results

Infants often look disproportionately long on the initial trial of a novel stimulus, so it is common to exclude the first two trials (i.e., initial exposure of each stimulus) from data analysis (e.g., Vouloumanos & Werker, 2004; Volkova et al., 2006), a procedure followed here. These initial trials can be considered familiarization rather than test trials. Missing values from infants (4 incomplete trials: 1-2 trials from 3 infants) were replaced with the multiple imputation method (Graham, 2012) implemented with NORM software (Schafer, 1999). Substitution or omission of those values yielded similar results. A preliminary analysis of variance revealed that the effect of age (4-6 months, 7-9 months, 10-13 months) on looking time was not significant. Age, considered as a continuous variable in a regression analysis, also made no contribution to looking time. Consequently, age was excluded from the main analysis. A paired sample *t*-test on cumulative looking time across the four trials with each stimulus revealed a significant difference between speech and singing ($t(49) = 3.35, p < .01$, two tailed). Infants looked longer during the syllable sequences ($M = 77.93$ s, $SD = 53.98$ s) than during the hummed lullabies ($M = 50.14$ s, $SD = 29.58$) (see Figure 2). A binomial test revealed that of the 50 infants in the sample, 36 (72.0 %) had longer looking times for speech, $z = 3.11, p < .01$.

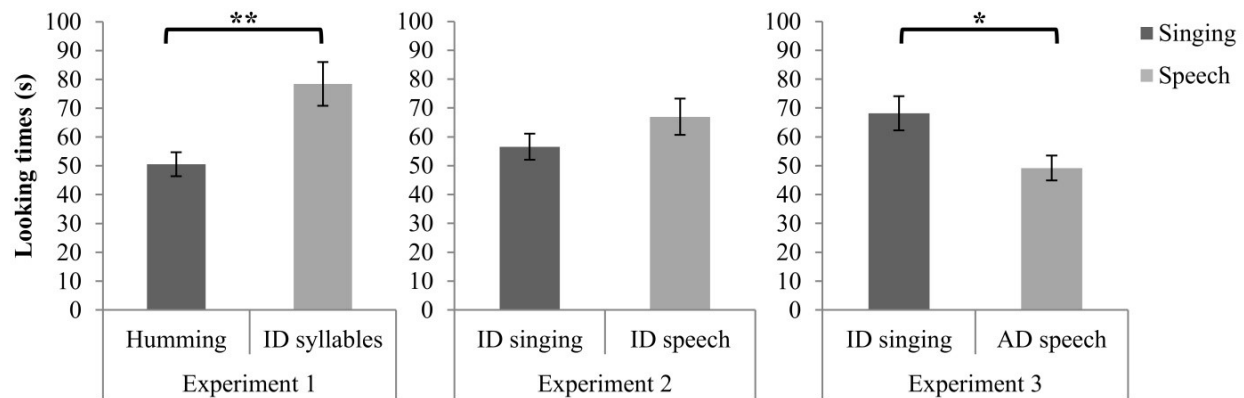


Figure 2. Cumulative looking time in seconds (s) for singing and speech. Error bars are standard errors (** = $p < .01$; * = $p < .05$).

Discussion

Infants exhibited greater attention to the ID speech syllables than to the hummed lullabies despite the greater coherence and continuity of the lullabies. Although our findings are consistent with the speech bias that has been proposed for young infants (Vouloumanos & Werker, 2004, 2007), there are a number of alternative interpretations. The stimuli contrasted in other respects than the presence or absence of speech or syllabic content. For one thing, the speech was considerably more variable than the humming in pitch and amplitude. Typical ID speech has much more continuity than the present sequence of disconnected syllables, each of which had the properties of stressed syllables. Moreover, each of the syllables had the exaggerated pitch contours that are considered critical in capturing infant attention (e.g., Fernald & Kuhl, 1987), and these contours were highly variable. The speech stimuli also had bursts of very high-pitched sound at irregular time intervals (see Figure 1), which could have functioned as salient alerting signals. Hummed speech produces less neural activation than natural speech (Perani et al., 2011), so one would expect hummed music to produce less cortical activation than other types of vocal music.

The affective qualities of the stimuli also differed dramatically, with the speech having the properties of high-arousal happiness or joy and the hummed lullabies being tranquil and soothing. Joyful or happy speech reliably attracts and maintains the attention of Western infants (Singh et al., 2002), and joyful music may do likewise. In contrast to Western mothers, who engage in lively vocal and non-vocal interactions with their infants, mothers in many other cultures interact in ways that are primarily soothing rather than arousing (Toda et al., 1990; Trehub & Schellenberg, 1995;

Trehub & Trainor, 1998). It is possible that infants who are accustomed to soothing vocal interactions would distribute their attention differently from the infant participants in the present experiment. Nevertheless, the youngest infants in the present study, who might experience more soothing interactions than the older infants (Kitamura & Burnham, 2003), responded no differently than the older infants.

Finally, the stimuli in the present study were atypical in a number of respects. The speech stimulus had the usual exaggerated pitch contours and happy affect of Western mothers in the unusual context of two single, non-contiguous syllables that repeated with variable renditions (following Vouloumanos & Werker, 2004; 2007). In other words, it was dramatically different from conventional ID speech. Although lullabies, sung or hummed, are common in non-Western cultures, they are used infrequently in Western cultures (Trehub & Trainor, 1998). If Western infants are exposed to lullabies, such exposure typically occurs when they are sleepy or distressed rather than awake or alert. For those reasons, we used more conventional stimulus materials in subsequent experiments, namely the lyrics of foreign children's songs that were spoken or sung.

Experiment 2

The goal of the present experiment was to ascertain the relative efficacy of speech and singing for maintaining infant attention when verbal or syllabic content and affective intentions are similar across vocal modes. Infants were presented with a sung and spoken version of an unfamiliar Turkish play song, both produced in an ID or joyful manner. The same lyrics ensured comparable phoneme sequences despite their different realization in speech and singing. Although the overall affective intentions were joyful in both cases, the means of achieving those intentions differ in speech and singing, with unknown consequences.

In research with ID and AD speech, the stimuli are often drawn from natural interactions with infants and adults (e.g., Kitamura & Burnham, 1998) so that verbal content and speaking style differ. At other times, actors portray ID and AD speech with the same verbal content (e.g., Singh et al., 2002). No previous study used the texts of play songs, which include words and nonsense syllables that are distinctive and memorable as well as alliteration, assonance, and rhyme. As a result, the spoken ID version was closer to a spoken nursery rhyme than to conventional ID speech,

reducing many of the usual differences between spoken and sung material for infants. Differences between speech and singing still remained, however, with speech being more variable in its pitch patterns and amplitude and also lacking the steady beat of music. If the expanded pitch range and greater pitch variability of speech drive infant attention (e.g., Fernald & Kuhl, 1987; Fernald, 1992), then infants could be expected to attend longer to the spoken lyrics. If happy affect is primarily responsible for infants' listening choices, as is the case for speech style (Singh et al., 2002), then infants might respond no differently to happy ID speech and singing with comparable verbal content.

Method

Participants

The sample included 48 healthy full-term infants who were 4.2 to 12.4 months of age ($M = 8.3$ months, $SD = 2.3$), with the same inclusion criteria as Experiment 1, and the same age groups: 4- 6 months ($M = 5.7$, $SD = .9$; $n = 16$), 7-9 months ($M = 8.5$, $SD = .8$; $n = 16$), and 10-12 months ($M = 10.8$, $SD = .8$; $n = 16$). An additional 6 infants were excluded from the final sample because of experimenter error ($n = 2$) or failure to complete the test session ($n = 4$).

Stimuli

Stimuli consisted of unfamiliar foreign lyrics (Turkish) of a play song (duple meter, AABAA form) that were spoken or sung. The performer was a native Turkish speaker and trained singer who had considerable experience with children. She listened to many samples of ID speech and singing and was instructed to speak and sing as if doing so for an infant. Stimuli are available in Supplementary Materials. Acoustic features of the sounds, as analyzed by Praat software (Boersma & Weenink, 2010) with pitch range settings following Experiment 1, are shown in Table I. Sung versions were slightly longer than spoken versions, 26.8 sec versus 24.6 sec. Mean pitch level was 2.3 semitones higher for sung ($M = 351.14$ Hz) than spoken versions ($M = 312.28$ Hz), but spoken versions had considerably greater pitch range (17.64 vs. 11.41 semitones) and pitch variability (SDs of 3.86 and 2.34 semitones, respectively). The mean pitch of the sung lyrics was substantially higher for the highly trained Turkish singer than for mothers' ID singing of play songs (253.6 Hz) (Trainor et al., 1997), but the pitch level of the spoken lyrics was comparable to that of mothers' ID speech (Fernald et al., 1989). As can be seen in Figure 1, however, there was more overlap of the ID speech and singing contours than was the case for Experiment 1.

Apparatus and Procedure

The apparatus and procedure were identical to Experiment 1.

Results

As in Experiment 1, a preliminary ANOVA revealed no effect of age on looking time, so age was excluded from the main analysis. A paired sample *t*-test on cumulative looking time across four trials with each stimulus (initial two trials omitted, as in Experiment 1) revealed no difference between speech ($M = 66.97$ sec, $SD = 43.24$ sec) and singing ($M = 56.58$ sec, $SD = 31.57$ sec) ($t(47) = 1.30$, $p = .199$, two tailed) (see Figure 2).

Discussion

Infants' attention did not differ for spoken and sung versions of a Turkish play song performed in an ID manner. The absence of differential attention, even in the presence of greater pitch and duration variability of the spoken versions (i.e., lively and rhythmic ID speech), implies that such acoustic variability, in itself, cannot account for the attention differences in Experiment 1 or in previous research (Nakata & Trehub, 2004). The findings raise the possibility that happy vocal affect, which characterized the spoken and sung versions, is primarily responsible for infants' engagement. Affective voice quality may be transmitted, in part, by the acoustic features that were measured but it is also transmitted by vocal timbre (i.e., tone of voice), which is not readily amenable to quantification. Issues of affective intent were addressed in the subsequent experiment.

Experiment 3

In the present experiment, we altered the affective intent of the spoken lyrics of Experiment 2 for comparison with the ID sung lyrics. Infants were exposed to the ID sung version from Experiment 2 and a spoken version in a non-ID style with neutral affect. If infants' attention is driven primarily by the joyful or happy quality of adult vocalizations, then they should exhibit greater attention to the sung versions than to the spoken versions. Just as infants are more engaged by happy speech than by neutral speech regardless of the ID or AD register (Kitamura & Burnham, 1998; Singh et al., 2002), we expected them to be more engaged by happy than by neutral vocal material regardless of whether it was spoken or sung.

Method

Participants

The sample included 48 healthy, full-term infants who were 4.7 to 12.5 months of age ($M = 8.3$ months, $SD = 2.5$). Inclusion criteria were comparable to Experiment 1, as were the age groups: 4-6 months ($M = 5.7$, $SD = .7$; $n = 16$), 7-9 months ($M = 8.0$, $SD = .9$; $n = 16$), and 10-12 months ($M = 11.3$, $SD = .8$, $n = 16$). An additional five infants were excluded from the final sample because of failure to complete the test session ($n = 4$) or parents' interaction with infant during the test session ($n = 1$).

Stimuli

Stimuli consisted of the same sung lyrics of the Turkish play song used in Experiment 2, which was unfamiliar to infants or mothers, and an affectively neutral version of the spoken lyrics. The lyrics were spoken by the same native Turkish speaker from Experiment 2, who was instructed to speak with neutral affective tone as if communicating with an adult. Stimuli are available in Supplementary Materials. Acoustic features of the sounds (analyzed by means of Praat software) are shown in Table I. Pitch range setting followed the procedures described in Experiment 1. The sung version was substantially longer (26.8 sec) than the spoken version (19.02 sec), reflecting the slow pace of singing relative to ordinary speech. Mean pitch level for the sung and spoken versions was 350.14 Hz and 210.24 Hz, respectively, corresponding to a difference of 8.9 semitones. F0 variability (SD) for the spoken and sung lyrics was similar at 2.30 and 2.34 semitones, respectively, as was the pitch range (i.e., difference between minimum and maximum pitch) of 11.33 and 11.41 semitones, respectively (see Figure 1). In short, the singing and speech stimuli differed substantially in pitch level, rate, and vocal tone (happy vs. neutral) but were comparable in pitch variability and pitch range.

Apparatus and procedure

The apparatus and procedure were identical to Experiment 1.

Results

Missing values for one infant on the final trial were handled by the multiple imputation method (Graham, 2012), as in Experiment 1. Data from one outlier ($> 3 SD$ from the mean) were excluded from the data set. Inclusion of the outlier and omission of the missing trial did not alter the results. A preliminary ANOVA revealed no effect of age on looking time, so age was excluded from the

main analysis. A paired sample *t*-test on cumulative looking time across the four trials for each stimulus type revealed a significant difference between speech and singing ($t(46) = 2.34, p < .05$, two tailed). Infants looked longer in the context of singing ($M = 68.17$ sec, $SD = 40.41$ sec) than in the context of neutral speech ($M = 49.20$ sec, $SD = 29.45$) (see Figure 1). A binomial test revealed that, of the 47 infants in the sample, 34 (72.3 %) looked longer during the presentation of singing, $z = 3.016, p < .01$.

Discussion

As predicted, infants exhibited greater attention during the presentation of the happy ID singing than during the neutral AD speech. Despite identical lyrics, similar pitch range (but different pitch register), and similar pitch variability of the sung and spoken versions, singing maintained infants' attention more effectively than did speech. The findings are consistent with a critical role for positive vocal affect, specifically happy or joyful vocalizations. An alternative explanation is that infants responded on the basis of pitch register, with the higher register of ID singing attracting their attention more effectively than the lower register of AD speech (see Figure 1). In speech contexts, however, happy vocal affect makes a greater contribution to infant attention than pitch register does (Kitamura & Burnham, 1998; Singh et al., 2002).

General Discussion

The purpose of the present study was to ascertain infants' relative interest in singing and speech. In Experiment 1, infants showed greater attention to happy ID versions of a series of unconnected nonsense syllables than to soothing hummed lullabies. The soothing humming proved to be no match for the effusively spoken syllables, which combined features of alerting vocalizations and joyful speech as well as high acoustic variability. In general, Western mothers' interactions with infants, whether spoken or sung, are lively and playful, in contrast to the soothing interactions and high levels of body contact that prevail in many non-Western cultures (Morikawa et al., 1988; Fernald, 1992; Trehub & Trainor, 1998). Perhaps infants' listening choices to stimuli such as these would differ in different cultures (e.g., non-Western) and contexts (e.g., when infants are experiencing fatigue or distress).

In Experiment 2, infants heard the lyrics of a Turkish play song that were spoken or sung in a lively, joyful manner. Neither the higher mean pitch of the sung versions nor the greater pitch range

and pitch variability of the spoken version resulted in differential infant attention, as they have in previous studies of ID and AD speech (Fernald & Simon, 1984; Fernald & Kuhl, 1987) or ID and non-ID singing (Trainor, 1996; Trainor & Zacharias, 1998). Obviously, the absence of a difference does not provide definitive evidence of equivalent interest in the stimuli, but it is consistent with the notion that infants' listening preferences are influenced primarily by the joyful or happy expressiveness of speech and singing. It is also consistent with newborns' comparable right hemisphere responses to lyrics that are spoken or sung in a happy manner (Sambeth et al., 2008).

In Experiment 3, infants' greater interest in the joyfully sung lyrics than in the neutrally spoken lyrics is in line with high positive affect driving infant attention. The speech stimuli of Experiment 1, the speech and singing stimuli of Experiment 2, and only the singing stimuli of Experiment 3 had features associated with vocal expressions of high-arousal happiness or joy (Banse & Scherer, 1996; Bänziger & Scherer, 2005). Taken together, the results of the three experiments are consistent with the possibility that features associated with vocal expressions of high-arousal happiness or joy are the principal determinants of infant preferences. Infants' attention to stimuli reflecting high levels of positive affect has been documented in visual (Kuchuk et al., 1986; Serrano et al., 1995) as well as auditory (Papoušek et al., 1990; Fernald, 1993; Kitamura & Burnham, 1998; Singh et al., 2002) contexts.

Although caregivers' expressive intentions are important for regulating infants' attention, other factors such as timing and pitch patterns may play an independent role. Music is much more predictable than speech in its temporal and pitch structure, generating expectations and the fulfillment of those expectations as the music unfolds (Kivy, 1993; Trainor & Zatorre, 2008; Jones, 2010). Such predictability contributes to the appeal of music for mature listeners (Kivy, 1993), and it may do so for infants as well. Maternal sung performances for infants have even greater predictability than other music, with many mothers singing the same songs at the same tempo and pitch level on different occasions (Bergeson & Trehub, 2002). Although maternal speech, with its frequent repetition of phrases and intonation contours, is much more predictable than AD speech, the contours are usually repeated with different verbal content (Bergeson & Trehub, 2002; 2007). The speech in Experiment 1, consisting of variable renditions of two syllables, carried repetition to an extreme from the perspective of adults, but the predictable content in the context of changing pitch contours may have highlighted those contours. The lullabies were also repetitive, as are most

lullabies (Unyk et al., 1992), but repetition occurred on a longer timescale than for the monosyllabic speech sounds.

The slow tempo and minimal amplitude variation of the lullabies de-emphasized the typical rhythmic regularity of music. The Turkish play song was more rhythmic than its spoken counterpart in Experiment 2, but the simple, repetitive lyrics sounded more like a nursery rhyme or poetry than conventional ID speech. Poetry blurs many of the distinctions between speech and singing by its inclusion of rhythm, meter, rhyme, alliteration, and assonance (Tillmann & Dowling, 2007; Obermeier et al., 2013), all of which were featured to varying degrees in the ID spoken and sung versions of the play song. In addition to having several repeated and rhyming syllables, the speech stimuli in Experiment 2 also had wider pitch contours than the sung stimuli. Such pitch contours have been linked to infants' listening bias for ID over AD speech (e.g., Fernald & Kuhl, 1987). Expanded pitch contours may compete with timing regularity for gaining and retaining infants' attention. Differences in pace, timing regularity, and rhythmicity between speech and singing were pronounced in Experiment 3 when singing finally prevailed. Naturally, one would expect infants' attention to be influenced by several factors acting together rather than a single factor (Singh et al., 2002), with some features being more salient than others in different situations.

The acoustic parameters of the speech stimuli in Experiments 2 and 3 conformed to conventional differences between Western ID and AD registers (e.g., Fernald & Simon, 1984), with the ID speech having substantially higher mean pitch, a pitch range that was over 6 semitones greater, and a speaking rate that was substantially slower than the AD or neutral versions (Ferguson, 1964; Stern et al., 1982; Stern et al., 1983). In fact, the ID version of spoken lyrics, with its heightened pitch and slowed rate (see Table I), was much closer to the sung version than it was to the neutral or AD spoken version (see Figure 1).

Obviously, speech and singing are not uniform across persons or contexts, and the differences between them narrow or widen in different situations. ID speech capitalizes on dimensions that are central to music, especially pitch and rhythm, which make it sound more musical than non-ID speech (Fernald, 1992; Trainor et al., 2000). Although maternal speech is more acoustically variable than maternal singing (Bergeson & Trehub, 2002), mothers make their speech more accessible to infants by the use of individually distinctive intonation patterns or tunes (Bergeson & Trehub, 2007).

To the adult ear, speech and singing, even ID speech and singing, are distinct classes. For young infants, however, melodious speech and singing may be variations on a theme. Brandt et al. (2012) suggest that speech is a special form of music, at least from the perspective of pre-verbal infants. Before language achieves referential status, infants may hear human vocal sequences as sound play, which is what music is all about (Brandt et al., 2012). Because speech lacks the constraints of music, it can become music-like without losing the essential properties of speech. Not only does ID speech exaggerate the features of conventional speech; it also incorporates some musical features such as sustained vowels and phrase-final lengthening, exaggerating others such as pitch range expansion (e.g., Fernald et al., 1989). The elevated pitch and slow tempo of ID speech are comparable to the pitch and tempo of ID singing and to music in general. Perhaps ID speech would be misjudged as music in cultures in which vocal music incorporates free rhythm and pitch glides (e.g., Clayton, 2000).

The present study provides support for the view that happy vocalizations or those with high positive affect, whether speech or singing, play an important role in regulating infant attention. The happy talk of Experiment 1 elicited greater infant attention than the soothing humming, and the happy singing of Experiment 3 elicited greater attention than the neutral speech. When speech and singing were both happy, as in Experiment 2, there was no difference in infants' attention. Can one conclude that that there would be no difference in infants' attention to happy speech and singing outside as well as inside the laboratory? Not necessarily. In everyday life, ID vocal interactions typically involve a familiar voice (e.g., parent), familiar content (e.g., frequently sung song, familiar phonemes, repeated syllable sequences), familiar face and facial expressions, as well as physical contact or movement, creating many possibilities for differential responsiveness to multimodal speech and singing. In fact, infants are more attentive to happy maternal singing than to happy maternal speech when the material is presented audiovisually (Nakata & Trehub, 2004).

Finally, the present research examined infants' attention in a series of relatively brief trials, providing insight into the potential of the stimuli for *capturing* their attention rather than *maintaining* it for sustained periods of time. In principle, one stimulus might be better for initial attention capture (e.g., unconnected speech syllables rendered in a happy voice) while another could have greater efficacy for maintaining attention or contentment, preventing distress, or alleviating distress (e.g., coherent passages of speech or singing). Visual fixation, the measure used in the present study, provides a limited perspective on attention and engagement, being imperfectly

correlated with physiological and neural measures of infant attention (Richards et al., 2010) and with infant facial affect (Fernald, 1993). We know, for example, that infants move rhythmically to rhythmic music but not to ID or AD speech (Zentner & Eerola, 2010) and that intense infant attention to vocal music initially leads to reduced body movement (Nakata & Trehub, 2004). Maternal singing also modulates infant cortisol levels (Shenfield et al., 2003). Future research with a wider variety of stimuli and measures may resolve the unanswered questions about infants' responsiveness to expressive speech and singing.

Acknowledgments

We thank Roxane Campeau, Audrey Morin, Émilie Gilbert and Cynthia Paquin for their assistance in recruiting and data collection. We also thank Jessica Phillips-Silver and Beste Kalender for their talking, humming, and singing. Finally, we thank Athena Vouloumanos for providing samples of speech stimuli from her research with infants. This research was supported by grants from the Social Sciences and Humanities Research Council of Canada and Advances in Interdisciplinary Research in Singing (AIRS) to the second author and by a doctoral fellowship from the Natural Sciences and Engineering Research Council of Canada to the first author.

References

- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614-636. doi: 10.1037/0022-3514.70.3.614
- Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Commun.* 46:3, 252-267. doi: 10.1016/j.specom.2005.02.016
- Bergeson, T.R., & Trehub, S.E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychol. Sci.* 13:1, 72-75. doi: 10.1111/1467-9280.00413
- Bergeson, T.R., & Trehub, S.E. (2007). Signature tunes in mothers' speech to infants. *Infant Behav. Dev.* 30:4, 648-654. doi: 10.1016/j.infbeh.2007.03.003
- Blood, A., & Zatorre, R. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U.S.A.* 98:20, 11818. doi: 10.1073/pnas.191355898
- Boersma, P., & Weenink, D. (2010). "Praat: doing phonetics by computer [Computer program], Version 5.1. 44".
- Brandt, A., Gebrian, M., & Slevc, L.R. (2012). Music and early language acquisition. *Front. Psychol.* 3, 327. doi: 10.3389/fpsyg.2012.00327
- Chen-Hafteck, L. (1997). Music and language development in early childhood: Integrating past research in the two domains. *Early Child Dev. Care* 130:1, 85-97. doi: 10.1080/0300443971300109
- Clayton, M. (2000). *Time in Indian Music*. New York, NY: Oxford University Press.
- Colombo, J., & Bundy, R.S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behav. Dev.* 4, 219-223. doi: 10.1016/S0163-6383(81)80025-2
- Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. London, UK: John Murray.
- De Looze, C., & Hirst, D.J. (2008). Detecting changes in key and range for the automatic modelling and coding of intonation, *Proceedings of International Conference on Speech Prosody* 4, Campinas, Brazil, 135-138.
- Dissanayake, E. (2000). "Antecedents of the temporal arts in early mother-infant interaction", in *The Origins of Music*, eds. N.L. Wallin, B. Merker, & S. Brown (Cambridge, MA: MIT Press), 389-410.

- Eckerdal, P., & Merker, B. (2009). "Music and the "action song" in infant development: An interpretation", in *Communicative Musicality: Exploring the Basis of Human Companionship*, eds. S. Malloch and C. Trevarthen (New York, NY: Oxford University Press), 241-262.
- Ferguson, C. (1964). Baby talk in six languages. *Am. Anthropol.* 66:6, 103-114. doi: 10.1525/aa.1964.66.suppl_3.02a00060
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8:2, 181-195. doi: 10.1016/S0163-6383(85)80005-9
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Dev.* 60, 1497-1510. doi: 10.2307/1130938
- Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.* 16:3, 477-501. doi: 10.1017/S0305000900010679
- Fernald, A. (1992). "Meaningful melodies in mothers' speech to infants", in *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, eds. H. Papoušek, & U. Jürgens (New York, NY: Cambridge University Press), 262-282.
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Dev.* 64:3, 657-674. doi: 10.2307/1131209
- Fernald, A., & Kuhl, P.K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behav. Dev.* 10:3, 279-293. doi: 10.1016/0163-6383(87)90017-8
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Dev. Psychol.* 20:1, 104-113. doi: 10.1037//0012-1649.20.1.104
- Frick, R. (1985). Communicating emotion: The role of prosodic features. *Psychol. Bull.* 97:3, 412-429. doi: 10.1037//0033-2909.97.3.412
- Glenn, S., Cunningham, C., & Joyce, P. (1981). A study of auditory preferences in nonhandicapped infants and infants with Down's syndrome. *Child Dev.* 52, 1303-1307. doi: 10.2307/1129520
- Graham, J. W. (2012). *Missing Data*. New York, NY: Springer, 73-94.
- Grieser, D.A.L., & Kuhl, P.K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Dev. Psychol.* 24, 14-20. doi: 10.1037//0012-1649.24.1.14

- Jackendoff, R. (2009). Parallels and nonparallels between language and music. *Music Percept.* 26, 195-204. doi: 10.1525/mp.2009.26.3.195
- Jacobson, J., Boersma, D., Fields, R., & Olson, K. (1983). Paralinguistic features of adult speech to infants and small children. *Child Dev.* 54, 436-442. doi: 10.2307/1129704
- Jones, M.R. (2010). "Attending to sound patterns and the role of entrainment", in *Attention and Time*, eds. K. Nobre, & J. T. Coull (New York, NY: Oxford University Press), 317-330.
- Jordania, J. (2010). Music and emotions: Humming in human prehistory. *Proceedings of the International Symposium on Traditional Polyphony*, Tblisi, Georgia, 41-49.
- Juslin, P., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychol. Bull.* 129:5, 770-814. doi: 10.1037/0033-2909.129.5.770
- Kelmer Nelson, D.G.K., Jusczyk, P.W., Mandel, D.R., Myers, J., Turk, A., & Gerken, L. (1995). The headturn preference procedure for testing auditory perception. *Infant. Behav. Dev.* 18:1, 111-116.
- Kent, R.D., Read, C., & Kent, R.D. (2002). *The Acoustic Analysis of Speech*. San Diego, CA: Singular Publishing Group.
- Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. *Adv. Inf. Res.* 12, 221-236.
- Kitamura, C., & Burnham, D. (2003). Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy* 4:1, 85-110. doi: 10.1207/S15327078IN0401_5
- Kitamura, C., & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy* 14:1, 77-100. doi: 10.1080/15250000802569777
- Kitamura, C., Thanavishuth, C., Burnham, D., & Luksaneeyanawin, S. (2002). 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 Behav. Dev.* 24:4, 372-392. doi: 10.1016/S0163-6383(02)00086-3
- Kivy, P. (1993). *The fine art of repetition: Essays in the philosophy of music*. Cambridge, UK: University press.

- Kuchuk, A., Vibbert, M., & Bornstein, M.H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child Dev.* 57, 1054-1061. doi: 10.2307/1130379
- Laukka P., Juslin P.N., & Bresin R. (2005). A dimensional approach to vocal expression of emotion. *Cogn. Emot.* 19:5, 633–653. doi: 10.1080/02699930441000445
- Mandel, D.R., Jusczyk, P.W., & Pisoni, D.B. (1995). Infants' recognition of the sound patterns of their own names. *Psychol. Sci.* 6, 314-317. doi: 10.1111/j.1467-9280.1995.tb00517.x
- Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Dev. Psychol.* 35:4, 1001-1005. doi: 10.1037//0012-1649.35.4.1001
- McMullen, E., & Saffran, J.R. (2004). Music and language: A developmental comparison. *Music Percept.* 21:3, 289-311. doi: 10.1525/mp.2004.21.3.289
- McRoberts, G. W., McDonough, C., & Lakusta, L. (2009). The role of verbal repetition in the development of infant speech preferences from 4 to 14 months of age. *Infancy* 14:2, 162-194. doi: 10.1080/15250000802707062
- Menon, V., & Levitin, D.J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage* 28:1, 175-184. doi: 10.1016/j.neuroimage.2005.05.053
- Mithen, S. J. (2005). *The Singing Neanderthals: The Origins of Music, Language, Mind, & Body*. Cambridge, MA: Harvard University Press.
- Morikawa, H., Shand, N., & Kosawa, Y. (1988). Maternal speech to prelingual infants in Japan and the United States: Relationships among functions, forms and referents. *J. Child Lang.* 15:2, 237-256. doi: 10.1017/S0305000900012356
- Nakata, T., & Trehub, S.E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behav. Dev.* 27:4, 455-464. doi: 10.1016/j.infbeh.2004.03.002
- Nakata, T., & Trehub, S.E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain* 21, 45. doi: 10.1037/h0094003
- Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Front. Psychol.* 4, 10. doi: 10.3389/fpsyg.2013.00010

- O'Callaghan, C. (2008). Lullament: Lullaby and lament therapeutic qualities actualized through music therapy. *American J. of Hospice and Palliat. Medicine* 25:2, 93-99. doi: 10.1177/1049909107310139
- Panneton, R., Kitamura, C., Mattock, K., & Burnham, D. (2006). Slow speech enhances younger but not older infants' perception of vocal emotion. *Research in Human Development* 3:1, 7-19. doi: 10.1207/s15427617rhd0301_2
- Papoušek, M., Bornstein, M.H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behav. Dev.* 13:4, 539-545. doi: 10.1016/0163-6383(90)90022-Z
- Papoušek, M., & Papoušek, H. (1981). Musical elements in the infant's vocalization: Their significance for communication, cognition, & creativity, in *Advances in Infancy Research, vol. 1*, eds. L. P. Lipsitt and C. K. Rovee-Collier (Norwood, NJ: Ablex), 163-224.
- Papoušek, M. (1994). Melodies in caregivers' speech: A species-specific guidance towards language. *Early Dev. Parenting* 3:1, 5-17. doi: 10.1002/edp.2430030103
- Patel, A.D. (2008). *Music, language, & the brain*. Oxford ; Toronto: Oxford University Press.
- Patel, A.D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain Lang.* 61:1, 123-144. doi: 10.1006/brln.1997.1862
- Pegg, J.E., Werker, J.F., & Mcleod, P.J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behav. Dev.* 15:3, 325-345. doi: 10.1016/0163-6383(92)80003-D
- Perani, D., Saccuman, M.C., Scifo, P., Awander, A., Spada, D., Baldoli, C., Poloniato, A., Lohmann, G., & Friederici, A.D. (2011). Neural language networks at birth., *Proc. Natl. Acad. Sci. U.S.A.*, 108:38, 16056-16061. doi: 10.1073/pnas.1102991108
- Peretz, I. (2009). Music, language and modularity framed in action. *Psychol. Belg.* 49:2-3, 157-175.
- Pinker, S. (1997). *How the Mind Works*. New York, NY: Norton.
- Plantinga, J., Trehub, S. E., & Russo, F. (2011, June). Multimodal aspects of maternal speech and singing. Presented at *Neurosciences and Music IV*, Edinburgh, Scotland.
- Richards, J.E., Reynolds, G.D., & Courage, M.L. (2010). The neural bases of infant attention. *Curr. Dir. Psychol. Sci.* 19:1, 41-46. doi: 10.1177/0963721409360003

- Rose, S.A., Gottfried, A.W., Melloy-Carminar, P., & Bridger, W.H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Dev. Psychol.* 18:5, 704-713. doi: 10.1037//0012-1649.18.5.704
- Salimpoor, V.N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R.J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nat. Neurosci.* 14:2, 257-262. doi: 10.1038/nn.2726
- Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clin. Neurophysiol.* 119:2, 332-341. doi: 10.1016/j.clinph.2007.09.144
- Schafer, J. L. (1999). NORM: Multiple imputation of incomplete multivariate data under a normal model (Version 2.02 for Windows 95/98/NT) [Software]. Retrieved from <http://www.stat.psu.edu/~jls/misoftwa.html>.
- Scherer, K.R. (1986). Vocal affect expression: A review and a model for future research. *Psychol. Bull.* 99:2, 143-165. doi: 10.1037//0033-2909.99.2.143
- Scherer, K.R. (1995). Expression of emotion in voice and music. *Journal of Voice* 9, 235-248. doi: 10.1016/S0892-1997(05)80231-0
- Serrano, J.M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behav. Dev.* 18, 477-482. doi: 10.1016/0163-6383(95)90036-5
- Shenfield, T., Trehub, S.E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music* 31:4, 365-375. doi: 10.1177/03057356030314002
- Shultz, S., & Vouloumanos, A. (2010). Three-month-olds prefer speech to other naturally occurring signals. *Lang. Learn. Dev.* 6:4, 241-257. doi: 10.1080/15475440903507830
- Singh, L., Morgan, J.L., & Best, C.T. (2002). Infants Listening Preferences: Baby talk or happy talk? *Infancy* 3:3, 365-394. doi: 10.1207/S15327078IN0303_5
- Soley, G., & Hannon, E.E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Dev. Psychol.* 46:1, 286. doi: 10.1037/a0017555
- Stern, D.N., Spieker, S., Barnett, R.K., & Mackain, K. (1983). The prosody of maternal Speech: Infant age and context related changes. *J. Child Lang.* 10:1, 1-15. doi: 10.1017/S0305000900005092
- Stern, D.N., Spieker, S., & Mackain, K. (1982). Intonation contours as signals in maternal speech to pre-linguistic infants. *Dev. Psychol.* 18:5, 727-735.

- Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Percept. Psychophys.* 27:1, 24-27. doi: 10.3758/BF03199901
- Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Mem. Cognition*, 35: 4, 628-639. doi: 10.3758/BF03193301
- Toda, S., Fogel, A., & Kawai, M. (1990). Maternal speech to three-month-old infants in the United States and Japan. *J. Child Lang.* 17, 279-294. doi: 10.1017/S0305000900013775
- Trainor, L.J. (1996). Infant preferences for infant-directed versus non infant-directed playsongs and lullabies. *Infant Behav. Dev.* 19:1, 83-92. doi: 10.1016/S0163-6383(96)90046-6
- Trainor, L.J., Austin, C.M., & Desjardins, R.N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11:3, 188-195. doi: 10.1111/1467-9280.00240
- Trainor, L.J., Clark, E.D., Huntley, A., & Adams, B.A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behav. Dev.* 20:3, 383-396. doi: 10.1016/S0163-6383(97)90009-6
- Trainor, L.J., & Heinmiller, B.M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behav. Dev.* 21:2, 77-88. doi: 10.1016/S0163-6383(98)90055-8
- Trainor, L.J., Tsang, C.D., & Cheung, V.H.W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Percept.* 20:2, 187-194. doi: 10.1525/mp.2002.20.2.187
- Trainor, L. J., & Zacharias, C. A. (1998). Infants prefer higher-pitched singing. *Infant Behav. Dev.* 21:4, 799-805. doi: 10.1016/S0163-6383(98)90047-9
- Trainor, L.J., & Zatorre, R.J. (2008). "The neurobiological basis of musical expectations: from probabilities to emotional meaning", in *Oxford Handbook of Music Psychology*, eds. S. Hallam, I. Cross, & M. Thaut (New York, NY: Oxford University Press), 171–183.
- Trehub, S.E. (2000). "Human processing predispositions and musical universals", in *The Origins of Music*, eds. N.L. Wallin, B. Merker, & S. Brown (Cambridge, MA: MIT Press), 427–448.
- Trehub, S.E. (2012). Behavioral methods in infancy: pitfalls of single measures. *Ann. N.Y. Acad. Sci.* 1252:1, 37-42. doi: 10.1111/j.1749-6632.2012.06448.x
- Trehub, S. E., Hannon, E. E., & Schachner, A. (2010). "Perspectives on music and affect in the early years", in *Handbook of music and emotion: Theory, research, applications*, eds. P. N. Juslin and J. A. Sloboda (Oxford: Oxford University Press), 645-668.

- Trehub, S.E., Hill, D.S., & Kamenetsky, S.B. (1997a). Parents' sung performances for infants. *Can. J. Exp. Psychol.* 51:4, 385-396. doi: 10.1037/1196-1961.51.4.385
- Trehub, S.E., & Prince, R.L. (2010). Turkish Lullabies. *UNESCO Observatory E-journal* 2:1, 1-19.
- Trehub, S. E. and Schellenberg, E. G. (1995). Music: Its Relevance to Infants. *Annals of Child Development* 11, 1–24.
- Trehub, S.E., & Trainor, L.J. (1998). Singing to infants: Lullabies and play songs. *Adv. Inf. Res.* 12, 43-78.
- Trehub, S.E., Trainor, L.J., & Unyk, A.M. (1993). Music and speech processing in the first year of life. *Adv. Child Dev. Behav.* 24, 1-35. doi: 10.1016/S0065-2407(08)60298-0
- Trehub, S.E., Unyk, A.M., Kamenetsky, S.B., Hill, D.S., Trainor, L.J., Henderson, J.L., & Saraza, M. (1997b). Mothers' and fathers' singing to infants. *Dev. Psychol.* 33, 500-507. doi: 10.1037//0012-1649.33.3.500
- Unyk, A.M., Trehub, S.E., Trainor, L.J., & Schellenberg, E.G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music* 20:1, 15-28. doi: 10.1177/0305735692201002
- Volkova, A., Trehub, S.E., & Schellenberg, E.G. (2006). Infants' memory for musical performances. *Dev. Sci.* 9:6, 583-589. doi: 10.1111/j.1467-7687.2006.00536.x
- Vos, P., & Troost, J. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Percept.* 6, 383-396. doi: 10.2307/40285439
- Vouloumanos, A., Hauser, M.D., Werker, J.F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Dev.* 81:2, 517-527. doi: 10.1111/j.1467-8624.2009.01412.x
- Vouloumanos, A., & Werker, J.F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Dev. Sci.* 7:3, 270-276. doi: 10.1111/j.1467-7687.2004.00345.x
- Vouloumanos, A., & Werker, J.F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev. Sci.* 10:2, 159-164. doi: 10.1111/j.1467-7687.2007.00549.x
- Werker, J., & Mcleod, P. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Can. J. Psychol.* 43:2, 230-246. doi: 10.1037/h0084224

- Zajonc, R.B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychol. Sci.* 10:6, 224-228. doi: 10.1111/1467-8721.00154
- Zatorre, R.J., & Baum, S.R. (2012). Musical Melody and Speech Intonation: Singing a Different Tune. *PLoS Biology* 10:7, e1001372. doi: 10.1371/journal.pbio.1001372
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. U.S.A.*, 107:13, 5768-5773.
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, & measurement. *Emotion* 8:4, 494-421. doi: 10.1073/pnas.1000121107
- Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* 21:3, 483-492. doi: 10.1016/S0163-6383(98)90021-2

Article II

Singing delays the onset of infant distress

Mariève Corbeil^{1,2,3}, Sandra E. Trehub⁴, and Isabelle Peretz^{1,2,3}

¹Department of Psychology, Université de Montréal

²The International Laboratory for Brain Music and Sound Research (BRAMS), University of Montréal, Montréal, Québec, Canada

³Centre for Research on Brain, Language and Music (CRBLM), McGill University, Montréal, QC, Canada

⁴Department of Psychology, University of Toronto Mississauga, Mississauga, Ontario, Canada

Article publié : Corbeil, M., Trehub, S. E., & Peretz, I. (2016). Singing Delays the Onset of Infant Distress. *Infancy*, 21(3), 373-391. DOI: [10.1111/infa.12114](https://doi.org/10.1111/infa.12114)

Running page heading: SINGING DELAYS INFANT DISTRESS

Abstract

Much is known about the efficacy of infant-directed (ID) speech and singing for capturing attention, but little is known about their role in regulating affect. In Experiment 1, infants 7–10 months of age listened to scripted recordings of ID speech, adult-directed speech, or singing in an unfamiliar language (Turkish) until they met a criterion of distress based on negative facial expression. They listened to singing for roughly twice as long as speech before meeting the distress criterion. In Experiment 2, they were exposed to natural recordings of ID speech or singing in a familiar language. As in Experiment 1, ID singing was considerably more effective than speech for delaying the onset of distress. We suggest that the temporal patterning of ID singing, with its regular beat, metrical organization, & tempo, plays an important role in inhibiting distress, perhaps by promoting entrainment and predictive listening.

Keywords: infants, music, singing, speech, attention, affect

Introduction

The most ubiquitous and most important function of music listening in contemporary Western cultures is emotional self-regulation aimed at relaxation, prolonging positive moods, or discharging negative emotions (Saarikallio, 2011; Sloboda, Lamont, & Greasley, 2009). Because infants' self-regulation skills are limited (Kopp, 1989; Thompson, 1994), caregivers' role in regulating their arousal and emotion is critical (Gunnar & Quevedo, 2007). Maternal regulatory behaviors include carrying (Gammie, 2013), touch (Jean & Stack, 2009), melodious speech (Fernald, 1992; Papoušek, 1994), & singing (Trehub, Hannon, & Schachner, 2011; Trehub & Trainor, 1998). The emotion-regulatory consequences of touch (Feldman, Singer, & Zagoory, 2010; Stack & Muir, 1992; Tronick, 1995), rocking, & carrying (Byrne & Horowitz, 1981; Esposito et al., 2013) have been documented, but there is no comparable evidence for infant-directed (ID) speech or singing.

Instead, studies of ID speech and singing have focused largely on their acoustic features (e.g. Delavenne, Gratier, & Devouche, 2013; Longhi, 2009; Nakata & Trehub, 2011; Trainor, Clark, Huntley, & Adams, 1997) or efficacy in capturing infant attention (e.g. Masataka, 1999; Trainor, 1996), with little consideration of their efficacy in emotion regulation (e.g., promoting displays of positive affect, inhibiting displays of negative affect).

ID speech typically has higher pitch, greater pitch and dynamic (amplitude) range, slower speaking rate, & greater rhythmicity than adult-directed (AD) speech (Ferguson, 1964; Fernald & Simon, 1984; Jacobson, Boersma, Fields, & Olson, 1983). The distinctive pitch contours and slow tempo of ID speech make it sound rather more musical than conventional speech (Brandt, Gebrian, & Slevc, 2012; Fernald, 1989, 1992). In fact, the acoustic features of ID speech (e.g., pitch level, tempo) are closer to those of ID singing than to AD speech (Corbeil, Trehub, & Peretz, 2013).

Like ID speech, ID singing has higher pitch, slower tempo, more regular timing, & a greater dynamic (i.e., amplitude) range than non-ID singing (Nakata & Trehub, 2011; Trainor et al., 1997; Trehub, Hill, & Kamenetsky, 1997; Trehub, Unyk, et al., 1997). Unlike speech, however, each song (e.g., *Itsy Bitsy Spider*, *Frère Jacques*) has a specified pitch range and rhythm. As a result, ID speech is more acoustically variable than ID singing (Corbeil et al.,

2013). Music, whether vocal or instrumental, differs from speech in other important respects. Most music has a regular pulse or beat and elements (e.g., notes, pitch contours, phrases, themes) that repeat at small and large timescales (Margulis, 2014). Repetitive motifs are especially prominent in children's songs. In *Frère Jacques*, for example, motifs and lyrical fragments repeat in parallel (*Frère Jacques, Frère Jacques. Dormez vous? Dormez vous? Sonnez les matines. Sonnez les matines. Ding, dang, dong. Ding, dang, dong.*). Mothers amplify the repetitive nature of children's songs by singing them almost identically on different occasions (Bergeson & Trehub, 2002).

Newborns are sensitive to the musical beat (Winkler, Háden, Ladinig, Sziller, & Honing, 2009), & by 5 months, infants are sensitive to the metrical structure, or pattern of strong and weak beats (Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005; Soley & Hannon, 2010). At 5 months, they also move rhythmically to rhythmic music, but not to speech (Zentner & Eerola, 2010), highlighting the importance of temporal regularity. Temporally regular music also captures infant attention more effectively than temporally irregular music (Soley & Hannon, 2010).

Infants' attention is captured more readily by ID speech than by AD speech (Cooper & Aslin, 1990; Fernald, 1985; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989), by ID hand and facial gestures than by AD gestures (Brand & Shallcross, 2008; Kim & Johnson, 2014), & by ID singing than by non-ID singing (Masataka, 1999; Trainor, 1996). Positive emotional expressiveness, such as happy-sounding speaking or singing voices (Corbeil et al., 2013; Kitamura & Burnham, 1998; Singh, Morgan, & Best, 2002) or happy-looking faces (Kim & Johnson, 2013), seems to underlie the attention-getting consequences of these stimuli. Indeed, happy-sounding speech and singing induce comparable attention capture (Corbeil et al., 2013; Costa-Giomi & Ilari, 2014; Trehub et al., 2016), as do happy-looking ID and AD faces (Kim & Johnson, 2013). Attention capture is greater, however, to audiovisual renditions of maternal singing than to maternal speech (Nakata & Trehub, 2004), perhaps because mothers smile more when they sing than when they speak to infants (Trehub et al., 2015).

In the single study that examined the effects of ID singing on infant arousal (Shenfield, Trehub, & Nakata, 2003), multimodal (live) maternal singing modulated salivary cortisol

levels, which index arousal or stress. Such modulation occurred within modest bounds, however, because infants were attentive or absorbed without exhibiting overt signs of pleasure or displeasure.

In the present study, we asked whether ID speech and singing (audio-only), which are comparable in capturing infant attention (Corbeil et al., 2013; Costa-Giomi & Ilari, 2014; Trehub et al., 2015), differ in their efficacy of inhibiting distress. The onset of distress is readily amenable to measurement because it is typically marked by negative vocalization and facial expression. Accordingly, we compared the efficacy of ID speech and singing for delaying infant distress. Experiment 1 used scripted portrayals of Turkish speech and singing, which ensured comparable content across vocal modes and unfamiliarity with the sung and spoken materials. Experiment 2 used natural ID speech and singing in the ambient language of infant listeners.

Infants in both experiments were tested in a minimally stimulating, non-social environment designed to highlight the target auditory stimulus and limit distraction. They listened to a continuous recording of speech or singing until they met a criterion of distress based on facial displays alone, specifically, the cry face. The cry face, which is characterized by lowered brows, lip corners pulled to the side, & variable degrees of mouth opening and cheek raising, is the most common expression of negative affect in infants (Messinger, 2002). Infant vocalizations also provide cues to distress, but it was necessary to block access to such cues so that the observer remained blind to the nature of auditory stimuli (speech or singing) presented to infants. On the basis of high temporal regularity in singing, especially ID singing (Nakata & Trehub, 2011), & much lesser regularity in speech, we predicted greater latency to infant distress for singing than for speech.

Entrainment to music, a human universal (Grahn, 2012; Merker, Madison, & Eckerdal, 2009), involves synchronized *internal* rhythmic processes to regular or pulsed auditory sequences (Grahn, 2012; Large & Jones, 1999). Nevertheless, the typical focus of study has been on its *external* manifestation in tapping, drumming, or moving to the beat (Kirschner & Tomasello, 2009; McAuley, Jones, Holub, Johnston, & Miller, 2006). Infants do not entrain their movements to the beat of music (Zentner & Eerola, 2010), perhaps because of immature

motor control, but they may be capable of internal entrainment in view of their sensitivity to the beat and to the metrical structure of music (Hannon & Trehub, 2005; Phillips-Silver & Trainor, 2005; Winkler et al., 2009). Entrainment of internal rhythmic processes would confer a distinct advantage for music over speech with respect to forestalling distress.

Experiment 1

We aimed to ascertain the relative efficacy of recorded speech and singing—unfamiliar song and language—in delaying visible signs of distress in pre-verbal infants. Infants were presented with one of three audio-recordings: (1) a rhythmic Turkish play song (unfamiliar melody in Western tonality) sung in a lively ID manner, (2) the words of the Turkish song spoken in an ID manner, with the spoken version having greater variability in pitch, dynamics, & timing than the sung version, & (3) the same words spoken in a neutral or AD manner, with smaller pitch and dynamic range than the version spoken in an ID style. The designated recording played continuously until infants met a criterion of mild distress based on facial expression alone. On the basis of the greater rhythmicity of ID singing, with its implications for the entrainment of internal responses (Grahn, 2012; Large & Jones, 1999), we expected infants to exhibit longer latency to distress for singing than for speech. Moreover, on the basis of greater attention capture for ID than AD speech (Fernald, 1985; Pegg et al., 1992; Werker & McLeod, 1989) and for happy-sounding than neutral or sad-sounding speech (Kitamura & Burnham, 1998; Singh et al., 2002), we expected longer latency to distress for ID speech than for AD speech.

Method

Participants

The sample consisted of 30 healthy, full-term infants (16 girls, 14 boys) who were 6.9 to 9.7 months of age ($M = 8.5$ months, $SD = 0.8$). Infants had no family history of hearing loss or personal history of ear infections, & they were free of recent colds, according to parents' report. An additional five infants were excluded from the sample because of unwillingness to sit in the high chair ($n = 2$), computer failure ($n = 2$) and experimenter error ($n = 1$). Parents reported that the ambient language in infants' environment was exclusively French ($n = 20$), 50% or more French combined with one or more other languages (English, Arabic, Spanish,

Polish, German, or Hebrew; $n = 7$), Italian-only ($n = 2$) or English-only ($n = 1$). No infant had any exposure to Turkish.

Apparatus

Testing was conducted in a dimly lit sound-attenuating booth (IAC) 4 by 4 m in size. Infants sat in a high chair facing a central computer monitor at a distance of 100 cm, with parent seated behind and out of view. The walls and table (for monitors) were covered with black cloth to reduce visual distraction and optimize attention to the target stimuli. A camera immediately above the central screen provided a continuous record of infant visual behavior on a monitor outside the booth. Two loudspeakers (Genelec 804A) equidistant from the central monitor transmitted the sounds at the same comfortable listening level for speech and singing, approximately 60-65 dB (A). The procedure was controlled by customized software on a computer (Mac Pro 8 cores, Apple Inc., Cupertino, CA, USA) located outside the booth.

Stimuli

The auditory stimuli (from Corbeil et al., 2013) consisted of unfamiliar foreign (Turkish) lyrics of a play song spoken in conversational style or sung in duple meter (i.e., accented on every second beat). The performer, a native Turkish speaker and trained singer, had considerable experience with children. She listened to many samples of ID speech and singing and attempted to speak and sing as if interacting with an infant (aided by an infant photograph). Subsequently, she attempted to speak with neutral affective tone as if conversing with an adult. Overall durations of the ID singing, ID speech and AD speech versions were 26.8 sec, 24.6 sec, & 19.02 sec, respectively, indicating a similar rate of vocalization for ID speech and singing and faster AD speech. Each stimulus was repeated, without interruption, until infants met the distress criterion. Excerpts of the three stimuli (ID speech, ID singing, & AD speech) are available in Supplementary Materials.

Procedure

Parents, who were out of infants' view throughout the test phase, were asked to refrain from interacting with infants (e.g., vocalization, touch) during testing and to minimize their movement. They remained in infants' line of sight only until the central screen flashed red, which was their cue to sit behind the infant because testing was about to begin. Custom

software randomly assigned infants to one of the three conditions—ID speech, AD speech or ID singing—without the knowledge of the observer. During the presentation of auditory stimuli (all conditions), a monitor in front of the infant displayed slow-motion movement of shaded animal forms. The visual display functioned to illuminate the infant’s face and maintain it in full view for transmission to the observer. Auditory and visual stimuli continued until infants met the criterion of mild distress, which consisted of continuous negative facial expressions or the cry-face (Messinger, 2002) for 4 s or two briefer episodes of negative facial expression within any 10-s period. The observer, who had no access to sound (auditory stimuli or infant vocalization) and no knowledge of infants’ condition, watched the monitor outside the test room and maintained a continuous record of negative facial expression by means of key codes on a computer keyboard¹.

Results and discussion

The distress criterion was achieved considerably later for singing ($M = 8.96$ min, $SD = 3.18$, range: 2.95-15.00 min) than for speech (ID speech: $M = 4.20$ min, $SD = 3.42$, range: 0.72-9.27 min; AD speech: $M = 3.91$ min, $SD = 2.14$, range: 1.09-6.85 min). An analysis of variance (ANOVA) revealed a significant effect of condition (ID singing, ID speech, AD speech) on latency to the distress criterion, $F(2, 27) = 7.70$, $p < .01$, $\omega^2 = .30$ (large effect size) (see Figure 1). Omega-squared was used instead of eta-squared because the latter overestimates effect sizes, especially for small samples (Olejnik & Algina, 2003). Post-hoc Tukey tests revealed significant differences between ID singing and each speech condition ($p < .05$) but no difference between ID and AD speech conditions.

¹ Afin de limiter une influence de l’âge entre les conditions, les bébés ont été distribués de façon quasi-aléatoire selon trois groupes d’âge: 7 mois (3 bébés par condition), 8 mois (3 bébés par condition) et 9 mois (4 bébés par condition).

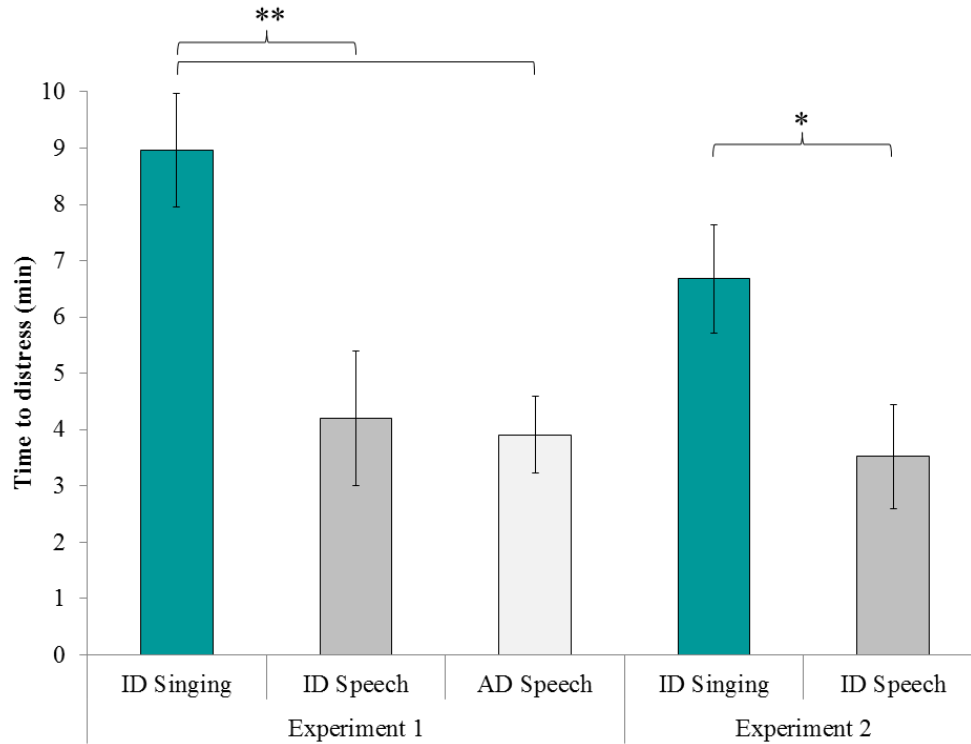


Figure 1. Time to distress (in minutes) for singing and speech in Experiments 1 and 2. Error bars are standard errors (** = $p < .01$; * = $p < .05$).

When listening to unfamiliar music sung in an unfamiliar language by an unfamiliar singer, infants maintained their composure for almost 9 min before meeting the distress criterion, roughly twice as long as for unfamiliar ID or AD speech by an unfamiliar speaker. Contrary to expectations, ID speech was no more effective than AD speech in delaying the onset of distress, even though ID speech in familiar and unfamiliar languages is more effective than AD speech in capturing infant attention (Fernald, 1985; Pegg et al., 1992).

Unfamiliar sound sequences may not detract from the engaging qualities of singing, but they may detract from the impact of speech, as least for infants in the age range tested (6.9-9.7 months). It is also possible that the engaging qualities of singing arose, in part, from the use of a professional singer. Moreover, the spoken lyrics of a song lacked some of the qualities of conventional maternal speech. To address these issues, we tested infants with natural speech and singing samples from their ambient language.

Experiment 2

The present goal was to compare the efficacy of natural samples of ID speech and singing in a familiar language for delaying the onset of distress. Infants were presented with one of two audio-recordings: an ID version of a rhythmic French play song or ID speech. Instead of scripted portrayals, as in Experiment 1, which are used frequently in research on infant listening preferences (e.g., Cooper & Aslin, 1990; Kitamura & Lam, 2009; Singh et al., 2002), the present stimuli were drawn from natural maternal interactions with 6-month-old infants. For infant listeners, the language was familiar, but the singing or speaking voice was unfamiliar. As in Experiment 1, the speech or singing stimuli were presented continuously until infants met the same distress criterion based exclusively on facial expression. On the basis of the rhythmicity and regularity of singing (e.g., Margulis, 2014; Nakata & Trehub, 2011) and its presumed consequences for internal entrainment (Grahn, 2012; Large & Jones, 1999), we expected the advantage of ID singing over ID speech to be evident with materials in a familiar language, as with those in an unfamiliar language in Experiment 1.

Method

Participants

The sample consisted of 28 healthy, full-term infants (11 girls, 17 boys) who were 7.3 to 10.0 months of age ($M = 8.5$ months, $SD = 0.8$). Infants' ambient language was French only ($n = 16$), 70% or more French combined with another language (English: $n = 7$, Creole: $n = 1$, Cambodian: $n = 1$, Spanish: $n = 1$), equal French and English ($n = 1$), & 60% German combined with 40% French ($n = 1$), based on parental estimates. Two additional infants were excluded from the final sample because parents interacted with infants during the test session ($n = 1$) or because of technical problems ($n = 1$).

Apparatus and Stimuli

The apparatus was the same as in Experiment 1. French-speaking mothers were recorded speaking and singing to their infants in a sound-attenuating booth. For the 2-min period of singing, mothers were asked to sing play songs that they usually sang to their infants. For the speech interactions, mothers were asked to speak to infants as they usually did but to refrain from using the child's name. Mothers placed infants in the high chair, held infants, or

remained together on the floor, as they chose. All mothers wore lapel microphones, & the sessions were recorded in stereo 44.1-kHz, 16-bit uncompressed WAV audio files (Adobe Audition Software, Adobe Systems Inc., San Jose, CA, USA), then converted to monaural files. Some maternal recordings were excluded because of poor sound quality (e.g., breathing noises that interfered with the audibility of vocalizations, singing substantially out of tune, audible infant hiccups). Recordings of singing had 20–45 sec continuous segments of good sound quality with a maximum of 5 infant vocalizations. Speech samples of equal duration were selected, usually from the beginning of the recording (to avoid selection biases), as long as they excluded the infant’s name or cues to child gender. Non-vocal sounds (e.g. audible microphone movement) and infant sounds (e.g., sighs, laughs, cries) were removed if they occurred during maternal pauses. Otherwise, they were attenuated (using iZotope RX2 Advance software, iZotope Inc., Cambridge, MA, USA) or equalized by adding a similar vocalization to a similar location in the comparison stimulus (i.e., singing from the same mother). The final singing and speech stimuli were of equal overall duration, with an equal number of infant vocalizations. There were 14 songs and corresponding speech samples ($M = 33.10$ sec, $SD = 8.52$ sec) generated from 11 mothers². In three instances, the same mother generated two different singing and speech samples. Repeated motifs are readily apparent in *Une Souris Verte*, a common French children’s song that was included in the current stimulus set (see Figure 2 and Supplementary Materials).

² Les bébés des mères ayant enregistré les stimuli avaient un âge moyen de 8.6 mois ($\acute{E}T = .9$) selon une étendue de 2.8 mois (min = 7.3, max = 10.0).

U - ne sou - ris ver - te qui cou - rait dans l'her - be.
 5 Je l'at - tra - pe par la queue, je la montre à ces mes-sieurs.
 9 Ces mes-sieurs me di - sent: trem - pez - la dans l'hui - le,
 13 trem-pez - la dans l'eau ça fe - ra un es - car - got tout chaud.
 18 Je la mets dans mon ti - roir, elle me dit: il fait trop noir.
 22 Je la mets dans mon cha-peau, elle me dit: il fait trop chaud.

Figure 2. Musical notation and lyrics for a common French children’s song, *Une Souris Verte*, which was part of the stimulus set in Experiment 2 (audio version in Supplementary Material). Repeating melodic sequences are depicted in matching boxes.

Procedure

To preclude the possibility of age differences across conditions, each infant was matched to another infant of similar age (age difference: $M = 6.43$ days, $SD = 4.33$ days, maximum = 16 days) and randomly assigned to a specific mother’s ID speech or singing³. A new observer (i.e., not the one from Experiment 1) who was unaware of the conditions and had no access to the stimulus or infant sounds judged negative facial expressions from the monitor outside the booth.

³ Chaque bébé était exposé à l’enregistrement d’une mère qu’il ne connaissait pas.

Results and Discussion

The distress criterion was achieved significantly later for ID singing ($M = 6.68$ min, $SD = 3.60$, range: 1.62-14.45 min) than for ID speech ($M = 3.52$ min, $SD = 3.46$, range: 0.12-12.76 min), $t(26) = 2.37$, $p < .05$, $\omega^2 = .18$ (large effect size) (see Figure 1). When the data from the ID speech and singing conditions from Experiment 1 were compared with those from the present experiment, there was a main effect of condition, $F(1, 44) = 13.06$, $p < .001$, partial $\omega^2 = .11$ (moderate effect size), reflecting the greater latency to distress for singing ($M = 7.63$ min, $SD = 3.55$, range: 1.62-15.00 min) than for speech ($M = 3.80$ min, $SD = 3.38$, range: 0.12-12.76 min), but there was no significant effect of experiment (familiar/unfamiliar language and portrayal/natural), $F(1, 44) = 2.86$, $p = .098$, & no interaction between experiment and condition, $F(1,44)=.29$, $p = .592$ ⁴.

As in Experiment 1, infants exhibited longer latency to distress for ID singing than for ID speech. Infants' ability to remain visibly content for longer periods while hearing French play songs rather than French ID speech confirms the generality of ID singing as an affect regulator. Language familiarity had no discernible benefit for infants who listened to ID singing or speech.

General Discussion

The present study is the first to assess the relative efficacy of unfamiliar audio excerpts of speech and singing for inhibiting infant distress rather than merely capturing their attention (Corbeil et al., 2013; Costa-Giomi & Ilari, 2014). In Experiment 1, a Turkish play song was more effective than spoken versions of the lyrics, in ID or AD style, in delaying signs of distress in 6.9- to 9.7-month-old infants ($M = 8.5$ months) from largely French-speaking homes. Despite previous research revealing the greater efficacy of ID over AD speech for capturing infant attention (e.g., Cooper & Aslin, 1990; Fernald, 1985), ID speech was no better than AD speech at delaying distress. Moreover, despite evidence that happy-sounding

⁴ Une analyse de variance préliminaire confirmait qu'il n'y avait aucun effet significatif de l'âge (7 mois, 8 mois, 9 mois) ou d'interaction entre la condition et l'âge sur la durée avant l'affect, ni de corrélation entre l'âge et la durée avant l'affect. Conséquemment, l'âge des bébés n'a pas été inclus dans les analyses.

ID speech and singing are equally effective at capturing infant attention (Corbeil et al., 2013; Costa-Giomi & Ilari, 2014), ID singing was more effective at delaying distress.

The use of identical verbal phrases across the three conditions of Experiment 1 (ID singing, ID speech, AD speech) ensured that the stimuli differed primarily in their pitch and temporal patterning, highlighting the key differences between speech and singing. In principle, the greater efficacy of ID singing could have stemmed from the use of an unfamiliar language—a frequent occurrence in songs (e.g., *Frère Jacques*)—as well as highly repetitive text with alliteration and rhyme combined with beat-based timing, which is typical of song (e.g., *Skidamarink a dink a dink, Skidamarink a do, I love you*) but not speech, even ID speech. In addition, the proficiency of a professional singer could have contributed to the efficacy of singing. In Experiment 2, however, French ID songs were more effective than French ID speech—both derived from natural maternal interactions with infants—in delaying the onset of distress in 7.3- to 10-month-old infants ($M = 8.5$ months) from French-speaking homes.

The findings leave little doubt about the efficacy of ID singing for maintaining infants' composure for extended periods. Even in a relatively sterile environment—black walls, dim illumination, no toys, & no human visual or tactile stimulation—the sound of a woman singing prolonged infants' positive or neutral states and inhibited distress. In fact, infants listened to the Turkish play song for roughly 9 min before meeting the cry-face criterion. For the French play songs, latency to the cry-face criterion exceeded 6 min. The temporal pattern of ID singing, with its regular beat, metrical organization, & tempo, facilitates predictive listening, involving expectations for the timing of subsequent events (Jones & Boltz, 1989), & the entrainment of internal, if not external, systems (Grahn, 2012; Large & Jones, 1999). In addition to the repetitive nature of music in general (Margulis, 2014) and children's songs in particular, the continuously repeating songs increased the salience of the motifs and rhythms, enhancing their familiarity and appeal, as for adults (e.g., Thompson, Balkwill, & Vernescu, 2000). Familiarity also enhances adults' liking of specific musical selections (Szpunar, Schellenberg, & Pliner, 2004; Topolinski & Strack, 2009), & it is likely to do so for infants as well (e.g., Trehub et al., 2015).

Perceiving the beat of music is not merely an auditory experience. Rhythm processing intensifies neural activity between auditory and motor regions presumed to underlie internal beat generation (Chen, Zatorre, & Penhune, 2006; Grahn & Rowe, 2009). Moreover, the perception of rhythm is affected by vestibular stimulation (Phillips-Silver & Trainor, 2007, 2008; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009), even in infancy (Phillips-Silver & Trainor, 2005). Perhaps it is no accident, then, that singing to infants is often accompanied by vestibular stimulation arising from rocking or carrying, both of which are potent affect regulators (Byrne & Horowitz, 1981; Esposito et al., 2013). The efficacy of movement for affect regulation may depend on its rate, as indicated by greater calming at the highest bouncing rate tested (1.5 Hz, corresponding to 90 bpm) (Vrugt & Pederson, 1973). In principle, the consequences of rhythmic singing—internal entrainment—could be similar to those of rocking or carrying, a question that awaits further investigation.

The success of ID singing in inhibiting or delaying infant distress implies that it was also successful in sustaining infant attention even though sustained attention was not measured directly. Nevertheless, infant distress onset indicated diminished attention or disengagement in relation to the stimulus, so that latency to distress provided an indirect index of sustained attention to that stimulus.

To date, comparisons of infant attention to ID and AD speech (e.g., Cooper & Aslin, 1990; Fernald, 1985; Singh et al., 2002), to ID and non-ID singing (e.g., Masataka, 1999; Trainor, 1996), & to ID speech and singing (Corbeil et al., 2013; Costa-Giomi & Ilari, 2014) have indexed such attention by infants' initial fixation of a visual stimulus on a series of trials in which the presentation of auditory stimuli was contingent on fixation of the visual stimulus (see Trehub, 2012). Such measures of initial visual fixation reveal the relative efficacy of two auditory stimuli for *capturing* infant attention, but they reveal nothing about their efficacy in *sustaining* or holding infant attention. Nakata and Trehub (2004) examined infants' responsiveness to audiovisual recordings of their own mother's speech and singing over the course of 3 minutes (Nakata & Trehub, 2004). Maternal singing was considerably more effective than maternal speech in capturing infants' attention, as indicated by longer initial fixations, & in sustaining their attention during, as indicated by cumulative fixation over the 3-minute period. Interestingly, attention capture and attention holding have different autonomic,

neural, & behavioral signatures (Colombo, Richman, Shaddy, Greenhoot, & Maikranz, 2001; Courage, Reynolds, & Richards, 2006; Richards, Reynolds, & Courage, 2010). Although previous research revealed that ID speech and singing (audio only) are equally effective in capturing infant attention (Corbeil et al., 2013; Costa-Giomi & Llari, 2014; Trehub et al., 2015), the present research suggests that their efficacy in sustaining attention differs substantially.

Despite speech being less effective than singing in preventing infant distress in a relatively unstimulating environment, it still managed to keep infants reasonably content for over 3 min. Mothers—Western mothers, in particular—speak to pre-verbal infants much more than they sing (Eckerdal & Merker, 2009). It is notable, however, that ID speech is musical in many respects (Brandt et al., 2012; Corbeil et al., 2013; Fernald, 1992) even though it lacks the temporal regularity and predictability that are likely to enhance the efficacy of music. Interestingly, mothers in some cultures speak directly to infants only after infants exhibit rudimentary verbal skills (Ochs & Schieffelin, 1995), but there is no evidence of a culture in which caregivers do not sing to infants.

In everyday life, infants typically experience singing multimodally, with mothers' rhythmic melodies accompanied by rhythmic movement and smiling (Trehub et al., 2015; Trehub & Gudmundsdottir, 2015). Caregivers' multimodal singing is likely to be more effective than their multimodal speech in sustaining attention and forestalling distress in pre-verbal infants. Live maternal singing in face-to-face contexts, even without touch, modulates infant arousal (Shenfield et al., 2003). Moreover, maternal touch, on its own, moderates the distress induced by maternal unresponsiveness in the still-face procedure (Feldman et al., 2010; Stack & Muir, 1992). Although recorded singing in the present study had greater efficacy than speech (audio-only, unfamiliar voices) for delaying distress, audio-recordings of singing are unlikely to be effective in reducing infant distress. We predict, however, that multimodal maternal singing would have such efficacy.

Maternal speech and singing share the advantages of a familiar voice, familiar face, familiar movement patterns, & behavior attuned to infant feedback (Braarud & Stormark, 2008; Smith & Trainor, 2008; Trehub et al., 2015). Maternal singing has the additional

advantages of familiar songs sung in a familiar style with familiar pitch level and tempo (Bergeson & Trehub, 2002) aside from temporal regularity (Nakata & Trehub, 2011). Future studies of the affect-regulatory consequences of maternal speech and singing could benefit from physiological measures (e.g., skin conductance, heart rate) and additional behavioral measures such as facial expression. Such measures may provide an early warning system, revealing changes in infant arousal levels that precede overt manifestations of distress.

Future research could also include measures of infant temperament. Just as irritable temperament influences self-soothing and recovery from distress induced by an unresponsive mother in the still-face procedure (Gunning, Halligan, & Murray, 2013), so temperament may influence infant composure when mother is out of view. In principle, infants with positive temperament and a repertoire of self-soothing strategies could remain content for longer periods than more irritable and less resourceful infants.

The present findings have implications for those with parenting challenges by virtue of adverse socio-economic or emotional circumstances. Although infant distress signals typically prompt parental comforting interventions, they induce frustration and anger in some at-risk parents, leading to insensitive responding and, in the worst cases, to infant neglect or abuse (Barr, Trent, & Cross, 2007; Frodi & Lamb, 1980; Leerkes & Siepak, 2006). At-risk parents within the purview of social service agencies could be encouraged to play vocal music to infants and, better still, to sing to them.

References

- Barr, R.G., Trent, R. B., & Cross, J. (2007). Age-related incidence curve of hospitalized Shaken Baby Syndrome cases: Convergent evidence for crying as a trigger to shaking. *Child Abuse and Neglect, 31*, 601–602. doi: 10.1016/j.chiabu.2005.06.009
- Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science, 13*, 72–75. doi: 10.1111/1467-9280.00413
- Braarud, H. C., & Stormark, K. M. (2008). Prosodic modification and vocal adjustments in mothers' speech during face-to-face interaction with their two- to four-month-old infants: A double video study. *Social Development, 17*, 1074–1084. doi: 10.1111/j.1467-9507.2007.00455.x
- Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology, 3*, 327. doi: 10.3389/fpsyg.2012.00327
- Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science, 11*, 853–861. doi: 10.1111/j.1467-7687.2008.00734.x
- Byrne, J. M., & Horowitz, F. D. (1981). Rocking as a soothing intervention: The influence of direction and type of movement. *Infant Behavior and Development, 4*, 207–218. doi: 10.1016/S0163-6383(81)80024-0
- Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *NeuroImage, 32*, 1771–1781. doi: 10.1016/j.neuroimage.2006.04.207
- Colombo, J., Richman, W., Shaddy, D., Greenhoot, A., & Maikranz, J. (2001). Heart rate-defined phases of attention, look duration, & infant performance in the paired-comparison paradigm. *Child Development, 72*, 1605–1616. doi: 10.1111/1467-8624.00368
- Cooper, R., & Aslin, R. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595. doi: 10.1111/j.1467-8624.1990.tb02885.x
- Corbeil, M., Trehub, S. E., & Peretz, I. (2013). Speech vs. singing: Infants choose happier sounds. *Frontiers in Psychology, 4*, article 372. doi: 10.3389/fpsyg.2013.00372
- Costa-Giomi, E., & Ilari, B. (2014). Infants' preferential attention to sung and spoken stimuli. *Journal of Research in Music Education, 62*, 188–194. doi: 10.1177/0022429414530564

- Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development, 77*, 608–695. doi: 10.1111/j.1467-8624.2006.00897.x
- 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. doi: 10.1016/j.infbeh.2012.10.004
- Eckerdal, P., & Merker, B. (2009). Music and the 'action song' in infant development: An interpretation. In N. L. Wallin, B. Merker & S. Brown (Eds.), *Communicative musicality: Exploring the basis of human companionship* (pp. 241–262). New York, NY: Oxford University Press.
- Esposito, G., Yoshida, S., Ohnishi, R., Tsuneoka, Y., Rostagno, M. d. C., Yokota, S., Okabe, S., Kamiya, K., Hoshino, M., & Shimizu, M. (2013). Infant calming responses during maternal carrying in humans and mice. *Current Biology, 23*, 739–745. doi: 10.1016/j.cub.2013.03.041
- Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science, 13*, 271–278. doi: 10.1111/j.1467-7687.2009.00890.x
- Ferguson, C. (1964). Baby talk in six languages. *American Anthropologist, 66*, 103–114. doi: 10.1525/aa.1964.66.suppl_3.02a00060
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*, 181–195. doi: 10.1016/S0163-6383(85)80005-9
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497–1510. doi: 10.2307/1130938
- Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papoušek & U. Jürgens (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp. 262–282). New York, NY: Cambridge University Press.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*, 104–113. doi: 10.1037//0012-1649.20.1.104
- Frodi, A.M., & Lamb, M. E. (1980). Child abusers' responses to infant smiles and cries. *Child Development, 51*, 238–241. doi: 10.2307/1129612

- Gammie, S. C. (2013). Mother-infant communication: Carrying understanding to a new level. *Current Biology*, *23*, R341–R343. doi: 10.1016/j.cub.2013.03.051
- Grahn, J. A. (2012). Neural mechanisms of rhythm perception: Current findings and future perspectives. *Topics in Cognitive Science*, *4*, 585–606. doi: 10.1111/j.1756-8765.2012.01213.x
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*, 7540–7548. doi: 10.1523/JNEUROSCI.2018-08.2009
- Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. *Annual Review of Psychology*, *58*, 145–173. doi: 10.1146/annurev.psych.58.110405.085605
- Gunning, S. N., Halligan, S. H., & Murray, L. (2013). Contribution of maternal and infant factors to infant responding in the Still Face paradigm: A longitudinal study. *Infant Behavior and Development*, *36*, 319–328.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *16*, 48–55. doi: 10.1111/j.0956-7976.2005.00779.x
- Jacobson, J., Boersma, D., Fields, R., & Olson, K. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, *54*, 436–442. doi: 10.2307/1129704
- Jean, A. D., & Stack, D. M. (2009). Functions of maternal touch and infants' affect during face-to-face interactions: New directions for the still-face. *Infant Behavior and Development*, *32*, 123–128. doi: 10.1016/j.infbeh.2008.09.008
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491. doi: 10.1037/0033-295X.96.3.459
- Kim, H. I., & Johnson, S. P. (2013). Do young infants prefer an infant-directed face or a happy face? *International Journal of Behavioral Development*, *37*, 125–130. doi: 10.1177/0165025413475972
- Kim, H. I., & Johnson, S. P. (2014). Detecting 'infant-directedness' in face and voice. *Developmental Science*, *17*, 621–627. doi: 10.1111/desc.12146
- Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*, 299–314. doi: 10.1016/j.jecp.2008.07.005

- Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. In C. Rovee-Collier, L. Lipsitt & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 221–236). Stamford, CT: Ablex Publishing Corp.
- Kitamura, C., & Lam, C. (2009). Age-specific preferences for infant-directed affective intent. *Infancy, 14*, 77-100.
- Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology, 25*, 343–354. doi: 10.1037/0012-1649.25.3.343
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159. doi: 10.1037/0033-295X.106.1.119
- Leerkes, E.M. & Siepak, K. J. (2006). Attachment linked predictors of women's emotional and cognitive responses to infant distress. *Attachment and Human Development, 8*, 11–32. doi: 10.1080/14616730600594450
- Longhi, E. (2009). 'Songese': Maternal structuring of musical interaction with infants. *Psychology of Music, 37*, 195–213. doi: 10.1177/0305735608097042
- Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford: OxfordUniversityPress,
- Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental Psychology, 35*, 1001–1005. doi: 10.1037//0012-1649.35.4.1001
- McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology: General, 135*, 348–367. doi: 10.1037/0096-3445.135.3.348
- Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex, 45*, 4–17. doi: 10.1016/j.cortex.2008.06.011
- Messinger, D. S. (2002). Positive and Negative: Infant Facial Expressions and Emotions. *Current Directions in Psychological Science, 11*, 1–6. doi: 10.1111/1467-8721.00156
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development, 27*, 455–464. doi: 10.1016/j.infbeh.2004.03.002
- Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain, 21*, 45–53. doi: 10.1037/h0094003

- Ochs, E., & Schieffelin, B. B. (1995). The impact of language socialization on grammatical development. In P. Fletcher & B. MacWhitner (Eds.), *Handbook of child language* (pp. 73–94). New York, NY: Basil Blackwell.
- Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*, 434–447. doi: 10.1037/1082-989X.8.4.434
- Papoušek, M. (1994). Melodies in caregivers' speech: A species-specific guidance towards language. *Early Development and Parenting, 3*, 5–17. doi: 10.1002/edp.2430030103
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development, 15*, 325–345. doi: 10.1016/0163-6383(92)80003-D
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*, 1430–1430. doi: 10.1126/science.1110922
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105*, 533–546. doi: 10.1016/j.cognition.2006.11.006
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition, 67*, 94–102. doi: 10.1016/j.bandc.2007.11.007
- Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science, 19*, 41–46. doi: 10.1177/0963721409360003
- Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music, 39*, 307–327. doi: 10.1177/0305735610374894
- Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music, 31*, 365–375. doi: 10.1177/03057356030314002
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants' listening preferences: Baby talk or happy talk? *Infancy, 3*, 365–394. doi: 10.1207/S15327078IN0303_5
- Sloboda, J., Lamont, A., & Greasley, A. (2009). Choosing to hear music: Motivation, process and effect. In S. Hallam, Cross, I., & Thaut, M. (Ed.), *Oxford handbook of music psychology* (pp. 431–440). New York: Oxford University Press.

- Smith, N. A., & Trainor, L. J. (2008). Infant-directed speech is modulated by infant feedback. *Infancy, 13*, 410–420. doi: 10.1080/15250000802188719
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology, 46*, 286–292. doi: 10.1037/a0017555
- Stack, D. M., & Muir, D. W. (1992). Adult tactile stimulation during face-to-face interactions modulates five-month-olds' affect and attention. *Child Development, 63*, 1509–1525. doi: 10.1111/j.1467-8624.1992.tb01711.x
- Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 30*, 370–381. doi: 10.1037/0278-7393.30.2.370
- Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development, 59*, 25–52. doi: 10.1111/j.1540-5834.1994.tb01276.x
- Thompson, W. F., Balkwill, L. L., & Vernescu, R. (2000). Expectancies generated by recent exposure to melodic sequences. *Memory & Cognition, 28*, 547–555. doi: 10.3758/BF03201245
- Topolinski, S., & Strack, F. (2009). Motormouth: Mere exposure depends on stimulus-specific motor simulations. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 35*, 423–433. doi:10.1037/a0014504
- Trainor, L. J. (1996). Infant preferences for infant-directed versus non infant-directed playsongs and lullabies. *Infant Behavior and Development, 19*, 83–92. doi: 10.1016/S0163-6383(96)90046-6
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development, 20*, 383–396. doi: 10.1016/S0163-6383(97)90009-6
- Trainor, L. J., Gao, X., Lei, J.-j., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex, 45*, 35–43. doi: 10.1016/j.cortex.2007.10.014

- Trehub, S. E. (2012). Behavioral methods in infancy: Pitfalls of single measures. *Annals of the New York Academy of Sciences*, 1252, 37–42. doi: 10.1111/j.1749-6632.2012.06448.x
- Trehub, S. E., & Gudmundsdottir, H. R. (2015). Mothers as singing mentors for infants. In G. F. Welch, D. M. Howard, & J. Nix (Eds.), *Oxford handbook of singing*. Oxford: Oxford University Press. Advance online publication. doi: 10.1093/oxfordhb/9780199660773.013.25
- Trehub, S. E., Hannon, E. E., & Schachner, A. (2011). Perspectives on music and affect in the early years. In P. N. Juslin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 645–668.). Oxford: Oxford University Press.
- Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Canadian Journal of Experimental Psychology*, 51, 385–396. doi: 10.1037/1196-1961.51.4.385
- Trehub, S. E., Plantinga, J., & Russo, F. R. (2016). Maternal vocal interactions with infants: Reciprocal visual influences. Manuscript under review.
- Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. In C. Rovee-Collier, L. Lipsitt & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 43–78). Stamford, CT: Ablex Publishing Corp.
- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Developmental Psychology*, 33, 500–507. doi: 10.1037//0012-1649.33.3.500
- Tronick, E. Z. (1995). Touch in mother-infant interaction. In M. Field (Ed.), *Touch in early development* (pp. 53–65). New York, NY: Psychology Press.
- Vrugt, D. T., & Pederson, D. R. (1973). The effects of vertical rocking frequencies on the arousal level in two-month-old infants. *Child Development*, 44, 205–209. doi: 10.2307/1127705
- Werker, J., & McLeod, P. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, 43, 230–246. doi: 10.1037/h0084224
- 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*, 106, 2468–2471. doi: 10.1073/pnas.0809035106

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, *107*, 5768–5773. doi: 10.1073/pnas.1000121107

Analyses supplémentaires

Analyses supplémentaires - Article I

Une ANOVA à mesures répétées a été effectuée pour l'ensemble des trois expériences ($n = 145$) sur l'impact du mode vocal (musique et parole), en tant que variable intra-sujet, et l'expérience, en variable inter-sujet, sur les temps de regards cumulatifs. Tel qu'attendu, l'analyse révèle une interaction significative entre le mode vocal et l'expérience ($F(2, 142) = 8.47, p < .001, \eta^2 = .12$), les effets du mode vocal ou de l'expérience pris isolément n'étant pas significatifs. Des tests post-hocs de Tukey indiquent des différences significatives (différence critique = 13.61 sec à $p = .05$) pour la comparaison de l'expérience 1 (berceuses vs. syllabes : différence = 27.79 sec) et de l'expérience 3 (chant DB vs. parole DA : différence = 18.97 sec), mais aucune différence significative pour l'expérience 2 (chant DB vs. parole DB : différence = 10.39 sec). Ainsi, les résultats concordent avec les analyses effectuées séparément (tests t) pour chacune des expériences.

Analyses supplémentaires - Article II

Puisqu'une étude précédente démontrait des temps de regards plus long envers des vidéos de chant maternel que des vidéos de paroles maternelles (Nakata & Trehub, 2004), des analyses supplémentaires post-expérimentation visant à mesurer le regard des bébés ont donc été réalisées pour les deux expériences combinées.

Méthode

Chaque bébé fut filmé lors de l'expérimentation à l'aide d'une caméra fixée au-dessus central de l'écran. Les enregistrements vidéos, sans sons, ont été codés par un expérimentateur, naïf aux hypothèses de recherche, à l'aide d'un logiciel de cotation en continu (E-prime 2.0, Psychology Software Tools). Le regard a été divisé en trois possibilités : (1) *regard dirigé vers l'écran*, (2) *regard dirigé vers soi* (se regarde soi-même ou un objet sur ou près de soi), (3) *regard dirigé ailleurs* (la salle ou vers l'arrière). Afin de mesurer la fiabilité de la mesure, un deuxième observateur a codé les regards pour 15% de l'échantillon. Le coefficient de corrélation intra-classe moyen s'avère de .87, suggérant un très haut niveau d'accord inter-juges.

Résultats

Puisque la durée des enregistrements varie pour chacun des participants, trois mesures ont été compilées pour les analyses : la proportion de temps de regard sur l'écran (% de temps de regard sur l'écran vs % de temps de regards total), le nombre de regards vers l'écran durant une fenêtre de temps initiale définie (1 min), et le temps du premier regard sur l'écran (mesuré en continue jusqu'à ce que l'enfant regarde ailleurs pendant plus de 2 sec)⁵ (voir Tableau 1). L'analyse préliminaire des données révélant des temps très courts de regards dirigés vers soi (temps moyen < 5 sec) et l'absence de différence significative entre les conditions (chant, paroles), la catégorie (2) *regard vers soi* n'a pas été maintenue dans les analyses et fut intégrée à la catégorie (3) *regard ailleurs*.

Une analyse de variance multivariée a été réalisée afin d'évaluer les effets de la condition (chant, parole DB) et de l'expérience (1 et 2) sur la proportion de temps de regards sur l'écran, le nombre de regards vers l'écran en 1 minute ainsi que le temps du premier regard. Les résultats indiquent une différence significative entre les conditions ($F(3, 42) = 3.39, p < .05, \eta^2$ partiel = .19) sur les variables de regards. Des analyses de variance univariées sur chacune des variables (avec correction de Bonferroni), révèlent toutefois que seul le temps du premier regard diffère significativement selon la condition ($F(1,44) = 10.05, p < .01, \eta^2$ partiel = .19).⁶ Une tendance envers des proportions de temps de regard plus élevées pour la condition chant que la condition parole s'observe, mais cette différence n'est pas significative avec correction de Bonferroni. De surcroît, une régression multiple linéaire fut calculée afin de prédire la durée jusqu'à l'affect, selon la proportion de temps de regards sur l'écran ainsi que le temps du premier regard. Le R du modèle était significativement différent de zéro, $F(2, 45) = 16.79,$

⁵ Afin de respecter le postulat de normalité, deux variables ont été transformées selon les formules suivantes : $2 * \text{ARSIN}(\text{SQRT}(\text{proportion de temps de regard}))$; $\text{LN}(\text{temps du premier regard})$.

⁶ Il est à noter que 4 participants (2 dans chaque condition) ont maintenu leur premier regard tout au long de l'expérimentation, jusqu'à l'apparition de l'affect négatif. De plus, 4 participants présentaient des temps de regards inférieurs à 1 minute. Toutefois, même exclus des analyses, les résultats demeurent significatifs pour le premier regard.

$p < .000$, avec un R^2 de .43 et un R^2 ajusté de .40. La durée prédit équivaut à $4.076 - .533$ (proportions en %) + 0.019 (premier regard en sec). Seul le temps du premier regard serait un prédicteur significatif de la durée avant l'affect ($r = .65$, $p < .000$). La durée avant l'affect augmenterait ainsi de .02 minutes (1.14 sec) par secondes du premier regard. Bien que la corrélation entre la proportion de temps de regards sur l'écran et la durée avant l'affect soit significativement différente de zéro ($r = .28$, $p < .05$), la proportion ne contribuerait pas de façon significative au modèle. En somme, les bébés exposés au chant regardaient plus longtemps l'écran lors du premier regard ($M = 2.78$ min, $ÉT = 2.51$ min), au début de l'expérience, que les bébés exposés à la parole ($M = 1.07$ min, $ÉT = 1.61$ min).

Discussion

La corrélation entre le temps du premier regard sur l'écran et le retardement de l'affect suggère une relation entre l'attention visuelle de l'enfant au stimulus, inférée par le regard, et la régulation émotionnelle. Toutefois, le regard des bébés dans cette étude ne constitue pas nécessairement un bon indicateur de l'attention auditive de l'enfant. D'abord, la présentation du son n'était pas contingente au regard parce que les bébés pouvaient porter attention aux sons sans regarder l'écran. De plus, l'animation visuelle utilisée (des ombres d'animaux en mouvements) n'était pas en lien avec les sons présentés. Il est néanmoins possible que l'apparition de l'animation visuelle simultanément au son suggère une contingence à l'enfant. Ainsi, le regard au début de l'expérience reflèterait un intérêt attentionnel envers les stimuli auditifs et le chant aurait intéressé davantage les bébés que la parole. Toutefois, il est également possible que les enfants présentant une meilleure régulation regardaient davantage l'écran que les enfants présentant une moins bonne régulation émotionnelle.

Chapitre III : Discussion générale

1. Retour sur les objectifs et résumé des résultats

L'objectif de cette thèse était de documenter l'efficacité relative du chant et de la parole à capter l'attention des bébés sur de courtes périodes de temps (Étude 1) ainsi qu'à réguler l'affect des bébés en maintenant un état de satisfaction ou de neutralité sur une période de temps prolongée (Étude 2). Dans les pages qui suivent, un résumé des principaux résultats des études sera présenté. L'apport scientifique de ces résultats, les limites inhérentes de ceux-ci ainsi que des pistes de recherches futures seront discutés à la lumière d'études précédemment et subséquentement publiées. Finalement, quelques applications cliniques potentielles seront proposées.

1.1 Résumé de l'étude 1 : Les bébés « choisissent » les sons joyeux

La première étude visait à mesurer l'efficacité du chant et de la parole à capter l'attention des bébés, avec des voix féminines non familières pour l'enfant. Chacune des trois expériences comportait deux stimuli variant selon plusieurs aspects. Lors de l'expérience 1, des syllabes dépourvues de sens et prononcées selon des intonations joyeuses étaient contrastées à des berceuses murmurées visant à calmer, résultant en des intentions affectives différentes pour chaque mode vocal (haute activation / positif / joyeux versus basse activation / positif / calmant). Lors de l'expérience 2, des paroles d'une chanson non familière dans une langue non familière (turque) étaient chantées ou récitées de façon joyeuse / DB, de façon à ce que l'intention affective soit similaire. Finalement, lors de l'expérience 3, la chanson turque chantée de façon joyeuse / DB (haute activation / positif) était contrastée aux paroles récitées de façon DA (basse activation / neutre). Les résultats démontrent une préférence attentionnelle pour les sons joyeux, indépendamment du mode chanté ou parlé. Bien que, lors d'études précédentes, les bébés de 6 mois démontraient des préférences attentionnelles pour des stimuli plus joyeux, les bébés de 3 mois présentaient un biais envers des stimuli vocaux plus réconfortants (Lam & Kitamura, 2006). Aucune différence développementale n'a été observée lors de l'étude ci-présente, mais le plus jeune groupe d'âge étudié avait à 4 mois. Il est possible que des bébés de 3 mois écouterait davantage des berceuses murmurées que des syllabes joyeuses dans l'expérience 1, en lien avec leur préférence pour des paroles rassurantes plutôt que stimulantes (Lam & Kitamura, 2006). Finalement, lorsque l'affect fut apparié

(joyeux) dans l'expérience 2, il n'y avait aucune attention différentielle envers le chant ou les paroles.

1.2 Résumé de l'étude 2 : Le chant retarde l'apparition d'un affect négatif chez le bébé

La deuxième étude visait à mesurer l'efficacité du chant comparativement aux paroles afin de réguler l'affect des bébés, avec des voix féminines non familières pour l'enfant. Chaque bébé était exposé à du chant ou des paroles jusqu'à ce qu'il présente un affect facial négatif. Lors de l'expérience 1, une chanson pour enfants dans une langue non familière (turque) était chantée ou parlée de façon joyeuse / DB (prestations simulées en laboratoire) ou parlée de façon neutre / DA. Lors de l'expérience 2, les enregistrements présentés provenaient de mères parlant ou chantant à leur bébé, donc dans une langue familière pour l'enfant. Les résultats démontrent que les bébés écoutent environ deux fois plus longtemps le chant (environ 7 minutes, moyenné selon les deux expériences) que la parole (environ 3 minutes) avant de présenter un affect négatif, indépendamment que la langue soit familière ou non.

De plus, des analyses subséquentes sur les deux expériences indiquent une corrélation entre le regard des bébés et le délai sur l'affect. En effet, les bébés exposés au chant regardent plus longtemps l'écran lors de leur fixation visuelle initiale que les bébés exposés à la parole.

2. Interprétation et implications des résultats

2.1 Identification des facteurs influençant l'attention des bébés

2.1.1 Un biais envers la positivité

Les sons vocaux joyeux, chantés ou parlés, étaient plus efficaces afin de capter l'attention des bébés que des sons vocaux neutres ou calmants (Étude 1), concordant avec les biais attentionnels précédemment rapportés envers l'affect positif dans la parole (Kitamura & Burnham, 1998; Singh et al., 2002) et dans l'expression faciale (Kim & Johnson, 2013; Kuchuk et al., 1986; Serrano et al., 1995). Les résultats corroborent ainsi un biais attentionnel envers l'expressivité positive qui serait domaine-général (*domain-general*), partagé par le domaine vocal (chanté / parlé) et visuel (facial). Toutefois, les résultats soutiennent également un biais domaine-général pour les stimuli à haute intensité ou activation (*arousal*), que celle-ci

soit positive ou négative. Selon des études en IRMf, le cerveau accentuerait l'importance d'information à haute intensité émotionnelle, peu importe sa valence (pour une revue, voir Vuilleumier, 2006). Il est ainsi possible que des stimuli à haute intensité négative, tels que la colère ou la peur, élicitent un biais attentionnel comparable aux expressions joyeuses (haute intensité positive). Par exemple, les bébés regardent parfois plus longtemps des visages exprimant la peur que la joie (Nelson & Dolgin, 1985) et ils désengagent plus difficilement leur regard de visages exprimant la peur que des visages joyeux ou neutres (Peltola, Leppänen, Palokangas, & Hietanen, 2008). Toutefois, certains chercheurs ont suggéré qu'un biais envers la négativité serait présent dès l'enfance et qu'il différerait du biais envers la positivité tant dans son traitement que dans ses conséquences pour l'enfant (pour une revue, voir Vaish, Grossmann, & Woodward, 2008). Bien que des stimuli à valence négative puissent capter l'attention des bébés car ils véhiculent une information pertinente ou dû à leur rareté, ils tendent aussi à induire un affect négatif chez le bébé (e.g. Haviland & Lelwica, 1987; Mumme et al., 1996). Il demeure donc possible que les bébés « préfèrent » écouter des vocalisations positives, malgré la saillance du négatif.

2.1.2 L'influence de la parole

L'absence de résultats différentiels dans l'attention des bébés, lorsque le chant et la parole DB sont comparés (Étude 1, Expérience 2), pourrait être problématique si ce n'était de la réplication de ce résultat dans de multiples études publiées subséquemment à l'Étude 1 (2013). L'absence de différence dans les temps d'écoute des bébés entre la parole et le chant (en mode auditif seulement) a été observée pour des prestations simulées en laboratoire (Costa-Giomi, 2014; Costa-Giomi & Ilari, 2014) ou des enregistrements issus d'interactions entre la mère et l'enfant (Trehub et al., 2016).

Plusieurs chercheurs considèrent la parole comme étant un stimulus spécial, élicitant des traitements distinctifs et spécifiques à l'humain (e.g. Liberman, 1996; Pinker, 1997), alors il est particulièrement intéressant et important de constater que le chant DB était aussi efficace pour capter l'attention et supérieur dans sa capacité à maintenir le bébé dans un état calme. Certains aspects de la parole DB, tels que son élévation générale de la hauteur moyenne ainsi que sa grande variabilité d'amplitude et de hauteurs, ont été considérés comme étant

possiblement responsables de sa capacité à capter l'attention (e.g. Fernald & Kuhl, 1987). Ils sont également typiques d'une haute intensité émotionnelle dans la voix (Banse & Scherer, 1996; Scherer, 1995). Bien que les stimuli parlés et chantés dans l'étude ci-présente aient été appariés en intensité globale, les paroles DB présentaient de plus grandes variations d'amplitude et de hauteurs (étendue de la F0 et ÉT) que le chant (Étude 1, Expérience 2). Comme Singh et al. (2002) suggèrent toutefois, une forte variabilité de hauteurs ne serait ni nécessaire ni suffisante pour générer les préférences attentionnelles relatives à la parole chez le bébé.

Il serait possible de considérer le chant avec des paroles comme étant une forme de parole améliorée ou embellie, c'est-à-dire une parole musicale. De cette perspective, la parole pourrait être responsable des aspects du chant suscitant l'attention. Toutefois, la parole ne peut en elle seule être responsable de la capacité du chant, dans une langue familière ou non, à maintenir les bébés calmes pour une beaucoup plus longue période de temps (Étude 2). Malgré tout, il n'est pas possible d'exclure une contribution potentielle de la parole. Pour ce faire, il faudrait comparer la parole à un chant sans paroles, tel que des chansons joyeuses murmurées plutôt que les berceuses utilisées dans l'Étude 1. Malheureusement, il est plus difficile de faire varier l'expressivité dans une chanson murmurée qu'il ne l'est pour le chant régulier et la parole. Puisque le murmure s'exécute la bouche fermée, il réduit les modulations d'amplitude spectrale (Kent, Read, & Kent, 2002). Puisqu'il est également possible d'argumenter que le murmure équivaut au son d'une consonne (m), une autre option serait de chanter la bouche ouverte selon une voyelle (ex. i) afin de permettre plus d'expressivité, notamment pour représenter une joie de haute activation.

2.1.3 La distinction des processus attentionnels

L'absence de résultats différentiels entre le chant et la parole sur l'attention des bébés (Étude 1, Expérience 2) contraste avec l'étude 2, résultant en des temps d'écoute ainsi que des fixations visuelles initiales plus longs pour le chant que pour la parole. De plus, une étude précédente de Nakata et Trehub (2004) avait également démontré que les bébés regardaient plus longtemps les vidéos de leur mère qui chante que les bébés exposés à leur mère qui parle, tant au niveau de la fixation visuelle initiale que les regards cumulatifs. Toutefois, des

processus attentionnels distincts pourraient être impliqués afin de capter l'attention des bébés pendant quelques secondes et maintenir les bébés attentifs pendant plusieurs minutes.

Il est suggéré que l'attention des bébés impliquerait différents processus. Le premier peut se définir par les propriétés du stimulus à capter l'attention de l'enfant (*attention-getting*), selon si celui-ci s'oriente vers le stimulus et à quelle vitesse, alors que le deuxième processus est déterminé par les propriétés du stimulus à soutenir l'attention (*attention-holding*), fréquemment mesuré par la durée du regard (Cohen, 1973; voir aussi Richards, Reynolds, & Courage, 2010, pour une différente terminologie). Ce processus implique un traitement de l'information présentée : il est lié aux apprentissages et est influencé par la nouveauté, la complexité et la compréhension du stimulus présenté (Courage & Setliff, 2010; Richards et al., 2010). Il comprend ainsi des changements comportementaux (ex. fixation du regard, réduction des mouvements), physiologiques (ex. diminution du rythme cardiaque) et cérébraux (ex. mesurés par EEG) distinctifs d'une durée de quelques secondes chez le bébé. Par la suite, l'enfant peut regarder le stimulus alors que les changements précédemment associés ont diminués ou disparus, suggérant qu'il n'est plus attentif (Colombo, Richman, Shaddy, Follmer Greenhoot, & Maikranz, 2001; Courage, Reynolds, & Richards, 2006). En d'autres mots, l'attention est corrélée, quoique qu'imparfaitement, avec le regard.

Dans la procédure de regard préférentielle, dans laquelle la présentation des stimuli dépend du regard de l'enfant, l'attention est estimée selon de très brèves fixations du regard lors de chacun des essais. Bien que ces regards corrélerent sans doute à de courtes périodes d'attention, la différence de regards cumulatifs pour deux stimuli concurrentiels pourrait refléter davantage des propriétés du stimulus à capter l'attention. Lors de périodes prolongées d'exposition, les deux processus attentionnels seraient aussi probablement impliqués, tel que la capacité des émissions de télévision à capter et soutenir l'attention des bébés à de multiples reprises, selon plusieurs regards successifs (Courage & Setliff, 2010). Toutefois, la présence d'un seul stimulus pendant une longue période de temps pourrait refléter davantage des propriétés du stimulus à soutenir l'attention. Par exemple, dans l'étude de Nakata et Trehub (2004), la présentation de vidéos de chants ou de paroles maternelles n'était pas contingente au comportement de l'enfant, les fixations du regard cumulatives étaient mesurées sur une période de 3 minutes et un seul stimulus était présenté par enfant, donc sans effet

concurrentiel. La mesure est donc possiblement un meilleur reflet d'attention soutenue chez le bébé. De façon similaire, dans l'Étude 2, la présentation du stimulus était aussi non-concurrentielle. Bien que la présentation du stimulus auditif n'était pas contingente au regard mais à l'affect de l'enfant, les deux variables apparaissent corrélées. Plus la durée de la fixation visuelle initiale était grande, plus la durée de l'expérience, avant l'affect négatif, l'était aussi. Ainsi, l'apparition d'un affect négatif pourrait également refléter une diminution de l'attention ou un désengagement au stimulus, conséquemment être une mesure indirecte d'attention soutenue.

2.2 Une motivation précoce envers la musique

2.2.1 Identification des facteurs contribuant à la régulation affective

Le chant apparaît plus efficace que la parole dans la régulation affective des bébés, ceux-ci demeurant calmes considérablement plus longtemps en écoutant du chant que des paroles (Étude 2). Ces résultats supportent l'expression précoce d'une motivation distinctivement humaine envers la musique (Trehub & Hannon, 2006). L'utilisation de stimuli très similaires, soit les mêmes paroles chantées ou parlées (Étude 2, Expérience 1), permet de circonscrire plus aisément les différences acoustiques possiblement responsables de cet effet différentiel. Les différences les plus saillantes entre ces modes vocaux sont la régularité rythmique, la stabilité des hauteurs, et les répétitions (de contour et de rythme). Ci-dessous, quelques facteurs pouvant contribuer au potentiel de régulation différent entre le chant et la parole sur les bébés sont discutés.

La grande régularité rythmique (Nakata & Trehub, 2011) ainsi les nombreuses répétitions (Trehub & Trainor, 1998) dans les chansons pour enfants pourraient faciliter l'engagement de l'enfant. Les rythmes réguliers et les répétitions augmentent la prévisibilité de la musique (London, 2004) et facilitent le traitement de l'information (Bergeson & Trehub, 2006; Drake & Botte, 1993; Drake et al., 2000; Vanden Bosch der Nederlanden et al., 2015), comparativement aux rythmes irréguliers des paroles (Grabe & Low, 2002) et leurs répétitions de contenu relativement peu fréquentes (Johnstone, 1994). Bien que les paroles DB soient plus répétitives que les paroles DA (Bergeson & Trehub, 2007), elles sont néanmoins beaucoup moins répétitives que le chant DB (Bergeson & Trehub, 2002). De surcroît, la stabilité des

hauteurs, prolongées et discrètes, confère au chant un signal acoustique quasi continu, en comparaison avec l'intermittence du signal et une grande variabilité des hauteurs dans la parole. La continuité et stabilité relative du chant pourraient amplifier sa capacité à maintenir l'intérêt et un état calme chez les bébés.

La familiarité pourrait aussi être un facteur d'influence dans l'attrait aux chansons lors de l'Étude 2, intimement liée aux répétitions omniprésentes en musique (Margulis, 2014). En plus des répétitions intra-chanson (notes, phrases musicales, motifs rythmiques), les répétitions d'une même chanson en entier (répétée environ à chaque 30 sec dans l'étude 2) peuvent augmenter rapidement la familiarité de la chanson pour l'enfant, pour qui la chanson n'était pas nécessairement familière au départ. Bien que les extraits de paroles répètent en entier selon une même fréquence, ils contenaient beaucoup moins de répétitions internes, même si les mots étaient identiques. Même pour les adultes, des répétitions cumulatives d'extraits musicaux accroissent leur jugement de plaisance (Schellenberg, Peretz, & Vieillard, 2008; Szpunar, Schellenberg, & Pliner, 2004) et l'ajout artificiel de répétitions à l'intérieur de pièces musicales amplifie leur valeur esthétique (Margulis, 2013a). La familiarité musicale amplifie l'activation cérébrale dans des zones liées au traitement émotionnel et au système de la récompense (Pereira et al., 2011). Ainsi, plus une musique serait répétée, plus elle serait familière, plus elle serait appréciée.

La répétition pourrait aussi réduire le stress de l'enfant (Dissanayake, 2009), particulièrement dans l'environnement asocial et non familial de l'Étude 2. Il est spéculé que les comportements répétitifs auraient une fonction d'autorégulation, visant à réduire les niveaux d'activation reliés au stress chez l'animal (Cabib, 2006) et chez les enfants ayant l'autisme (Gabriels et al., 2013). Or, justement, la musique réduit la réponse hormonale de stress (niveaux de cortisol) chez l'adulte (Khalifa et al., 2003; Leardi et al., 2007) et le chant maternel module les niveaux de cortisol du bébé (Shenfield et al., 2003).

L'entraînement à la musique pourrait aussi influencer le traitement et la perception attirante du chant. La régularité rythmique de la musique facilite la synchronisation à la musique (e.g. Patel, Iversen, Chen, & Repp, 2005), créant des conditions favorables à un entraînement simultané en comparaison avec la coordination à tour de rôle dans la

conversation (Phillips-Silver et al., 2010). Les bébés sont incapables de synchroniser leurs mouvements à des rythmes musicaux, mais ils bougent plus vite sur de la musique plus rapide (Zentner & Eerola, 2010) et ils synchronisent possiblement intérieurement à des rythmes, tels que les adultes (Grahn, 2012; Large & Jones, 1999). Étant donné les liens établis entre la perception auditive et une stimulation du système vestibulaire présents dès l'enfance (Phillips-Silver & Trainor, 2005), il est possible que les mécanismes en jeu dans l'efficacité du chant afin de maintenir les bébés calmes (Étude 2) soient similaires à ceux en lien avec l'efficacité de bercer ou de transporter l'enfant pour le calmer (e.g. Esposito et al., 2013; Vrugt & Pederson, 1973). Esposito et collègues (2013), ayant étudié l'effet du transport chez les bébés humains et les souriceaux, ont spéculé que la capacité des bébés mammifères à se calmer rapidement augmenterait leurs chances de survie lorsque la mère doit fuir.

Le tempo ou le débit pourrait jouer un rôle important afin de calmer un bébé. Bien qu'Esposito et collègues n'ont pas abordé cette question, une étude a précédemment démontré que bercer le bébé était plus efficace pour le calmer au plus haut débit testé, 667 msec d'intervalles inter-événements (*inter-onset intervals*, IOI), comparé à des débits plus lents de 1000 msec ou 2000 msec (Vrugt & Pederson, 1973). Fait intéressant, le tempo de la chanson turque (Étude 2, Expérience 1) était comparable, 679 msec IOI (sur les pulsations accentuées). Il serait possible de spéculer que des tempi plus lents seraient moins efficaces. Les adultes et les enfants ont davantage de difficultés à percevoir la régularité rythmique des musiques lentes (1000 msec ou plus) que des musiques plus rapides (Drake et al., 2000). En fait, le tempo de traitement optimal chez les adultes (qu'ils discriminent et produisent le mieux) oscillerait autour de 600 msec IOI, alors qu'il serait plus rapide, 400 msec chez les enfants (Drake et al., 2000) et possiblement aussi les bébés (Baruch, 1994; Baruch & Drake, 1997). Si les bébés sont meilleurs à traiter des tempi rapides, il est étonnant de constater que les débits/tempi des paroles et chants DB sont plus lents que les registres DA ou non-DB (e.g. Fernald & Simon, 1984; Trehub, Unyk, et al., 1997) et probablement plus lents que le tempo optimal estimé chez le bébé (Drake et al., 2000). En fait, le chant DB est même de façon générale plus lent que les paroles DB (Falk, 2011). Drake et al. (2000) ont suggéré que le débit ralenti des paroles DB pourrait servir à attirer l'attention des bébés à d'autres niveaux structurels accentués par la prosodie des mères. D'autre part, le tempo des paroles et du chant des adultes pourrait

simplement provenir de leur tempo préféré (Baruch, Panissal-Vieu, & Drake, 2004), plutôt que d'être une accommodation pour les bébés. Baruch et al. (2004) ont découvert que les enfants de 6 et 10 ans préféraient le tempo le plus rapide présenté, mais ils n'ont détecté aucune préférence chez les bébés. En continuité avec la proposition de Drake et al. (2000) que le tempo préféré ralenti avec l'âge, un tempo plus rapide que celui utilisé dans l'étude ci-présente serait peut-être plus efficace pour les bébés.

2.2.2 Le chant versus la musique instrumentale

Puisque la voix était commune au chant et aux paroles utilisées dans l'étude 2, la recherche actuelle ne peut pas départager si les facteurs contributifs proposés – la régularité, la répétition, la familiarité, l'entraînement- induiraient des conséquences comparables dans le contexte de musique instrumentale. Le chant pourrait être un stimulus particulièrement efficace parce qu'il combine la voix avec ces autres facteurs. Pour les adultes, les sons vocaux ont un statut spécial, générant des activations corticales distinctives comparés à des sons non-vocaux (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Gunji et al., 2003; Levy, Granot, & Bentin, 2003). En principe, ce statut particulier pourrait être présent chez les bébés qui écoutent plus longtemps des sons vocaux naturels que des sons artificiels (Colombo & Bundy, 1981; Vouloumanos et al., 2010; Vouloumanos & Werker, 2004, 2007), ou du chant sans accompagnement instrumental (a capella) plutôt que du chant avec accompagnements (Ilari & Sundara, 2009). De plus, les enfants et les adultes mémorisent mieux les mélodies présentées vocalement qu'instrumentalement (Weiss, Schellenberg, Trehub, & Dawber, 2015; Weiss, Trehub, & Schellenberg, 2012; Weiss, Vanzella, Schellenberg, & Trehub, 2015). Un traitement différentiel des sons vocaux pourrait être imputable à une attribution attentionnelle plus grande envers des sons biologiquement significatifs tels que la voix humaine (Levy et al., 2003). Par ailleurs, la musique vocale est omniprésente à travers les cultures humaines (Nettl, 2000). Elle constitue donc possiblement un stimulus musical optimal pour le bébé.

2.2.3 Le chant et la parole en contextes naturels

Dans la vie de tous les jours, les bébés vivent l'expérience de la parole ou du chant de façon multimodale, impliquant des stimulations visuelles (expressions faciales et gestuelles), kinesthésiques (ex. porter ou bercer) et/ou tactiles, provenant d'une personne hautement

familiale et à l'écoute de leur réactions (Braarud & Stormark, 2008; Smith & Trainor, 2008; Trehub et al., 2016). Alors que les capacités régulatrices du toucher (e.g. Feldman et al., 2010) ainsi que de bercer ou de porter l'enfant (e.g. Esposito et al., 2013) ont été documentées, l'étude actuelle ajoute le chant à ces régulateurs. D'ailleurs, une combinaison de ces régulateurs amplifierait probablement l'impact sur l'enfant. Par exemple, l'ajout de la modalité visuelle augmente l'efficacité du chant par rapport à la parole sur l'attention des bébés. Avec du matériel audio-visuel, les enfants écoutent plus longtemps une chanson folk chantée plutôt que récitée (Costa-Giomi, 2014), ce qui concorde avec une plus grande attention des bébés envers des versions audio-visuelles de chants maternels que de paroles maternelles (Nakata & Trehub, 2004). Cette dernière étude utilisait des enregistrements provenant des mères et donc hautement familiers pour les bébés, tant par le visage que la voix, la chanson et la langue. Avec des vidéos en silence de personnes qui chantent ou parlent, les bébés démontrent une attention comparable pour des prestations simulées de chant ou de paroles expressives (Costa-Giomi, 2014), mais présentent plus d'attention pour le chant que les paroles lorsque ceux-ci dérivent d'interactions mamans-bébés en face-à-face (Trehub et al., 2016). En fait, des auditeurs adultes jugent les enregistrements de vocalisations maternelles provenant d'interactions face-à-face avec l'enfant comme étant plus expressives que les vocalisations maternelles lorsque l'enfant est présent, mais hors de vue de la mère (Trehub et al., 2016). De plus, les mères sourient de façon presque continue en chantant, contrastant à des sourires intermittents lorsqu'elles parlent (Trehub et al., 2016). En général, les bébés regardent plus longtemps les expressions faciales de joie plutôt que de tristesse, mais seulement si le sourire implique une bouche ouverte (avec les dents visibles) (Oster, 1981). Or, sourire pendant des vocalisations, chanter ou parler, requièrent généralement d'ouvrir la bouche. Le fait de sourire influence aussi les propriétés acoustiques des vocalisations. Il modifie l'ouverture de la bouche et la forme du tractus vocal au niveau de la gorge, forçant ainsi une élévation de la hauteur moyenne, des fréquences harmoniques et une augmentation de l'amplitude (Tartter, 1980). Toutefois, l'absence de préférences attentionnelles des bébés pour des prestations de chant versus des paroles en mode auditif seulement (Étude 1, Expérience 2; Costa-Giomi, 2014; Trehub et al., 2016) indique que la différence de sourires entre le chant et la parole aurait une contribution limitée afin de capter l'attention des bébés, mais elle pourrait néanmoins jouer un rôle dans le maintien de l'attention et d'un affect calme. De surcroît, dans

des contextes naturels, le chant et la parole impliquent également des mouvements du corps, de la tête et des bras, qui pourraient aussi être des facteurs d'influence sur le comportement des bébés. Chez les adultes, les mouvements de la tête, lorsque vus isolément (sans visage), véhiculent l'émotion de façon similaire dans le chant et la parole (Livingstone & Palmer, 2015). Les mouvements liés au chant maternel tendent à être coordonnés aux pulsations accentuées (Longhi, 2009), amplifiant potentiellement la saillance du rythme des chansons pour enfants et l'entraînement interne. Conséquemment, un chant *live* (face-à-face et multimodal) d'une chanson familière interprétée par une personne familière serait probablement des plus efficaces afin de réguler l'attention et l'affect des bébés.

2.2.4 Les biais culturels

Les bébés participants aux études ci-présentes provenaient en grande majorité de cultures occidentales, où les mères tendent à parler à leur bébé beaucoup plus qu'elles ne chantent (Eckerdal & Merker, 2009). Toutefois, dans certaines cultures, les mères maintiennent des interactions verbales minimales avec leurs enfants préverbaux, ne commençant à leur parler que lorsque ceux-ci démontrent une production verbale suffisante vers l'âge de 1 ou 2 ans, que l'on présume acquise selon une exposition indirecte à des conversations (Ochs & Schieffelin, 1995). À l'inverse, il n'existe aucune culture rapportée à ce jour dans laquelle les personnes s'occupant des enfants préverbaux ne chantent pas à ceux-ci (Trehub & Trainor, 1998), bien qu'elles diffèrent sans doute selon ce qu'elles chantent et comment elles le chantent. Par exemple, les mères dans les sociétés occidentales chantent davantage de chansons joyeuses que de berceuses, en lien avec leur préférence pour les interactions visant à stimuler le bébé, à l'inverse de plusieurs autres cultures où les mères tendent à favoriser des interactions calmantes avec leur bébé et chanter des berceuses (Trehub, Unyk, et al., 1997; Trehub, Unyk, et al., 1993). Il n'est pas certain que l'écoute des rythmes réguliers et rapides des chansons joyeuses (Étude 2) aurait les mêmes conséquences pour des bébés provenant de cultures non-occidentales. Toutefois, il est possible de présumer qu'une pulsation régulière dans un chant aurait les mêmes conséquences à travers les cultures, au même titre que de bercer un bébé.

2.2.5 Les caractéristiques d'une « super-chanson »

En somme, bien que les facteurs spécifiques responsables de conséquences affectives différentes pour le chant et la parole ne puissent pas être identifiés spécifiquement par la recherche actuelle, plusieurs aspects des stimuli pourraient être manipulés systématiquement et leurs effets mesurés sur les bébés dans des recherches subséquentes. Par exemple, l'illusion de la transformation de la parole en chanson décrit par Deutsch (2011), où les adultes perçoivent une phrase parlée comme étant chantée après plusieurs répétitions du même stimulus, indique la puissance de la répétition dans la perception musicale et pourrait être présentée aux bébés. Il serait possible de s'attendre à ce qu'une « super-chanson » pour les bébés, visant une régulation affective, soit hautement répétitive, rythmiquement simple et régulière (pulsations régulières), avec un tempo intermédiaire (autour de 600 msec IOI) ou possiblement plus rapide (autour de 400 msec IOI). Enfin, il est aussi fort probable qu'une chanson *live*, donc multimodale mais principalement vocale, et sensible aux rétroactions du bébé, serait des plus efficaces.

2.3 Applications cliniques potentielles

Bien qu'il existe une variété de régulateurs possibles du comportement de l'enfant, tels que toucher, tenir, transporter, le chant est un outil de régulation particulièrement utile puisqu'il est efficace tant à distance que de près. Conséquemment, les implications du chant pourraient être explorées dans une multitude de contextes thérapeutiques, notamment ceux comprenant de l'inconfort physique ou émotionnel pour le bébé. Pour les bébés nés prématurément aux unités de soins intensifs néonataux, la présence régulière de musique corrèlerait à de nombreux bénéfices incluant une augmentation de la tétée ainsi qu'une réduction des effets négatifs et de la durée d'hospitalisation, tel que révélé par une méta-analyse de 30 études (Standley, 2012). Toutefois, peu d'études ont utilisé le chant comme outil thérapeutique avec les bébés, et encore moins ont utilisé du chant parental, qu'il soit enregistré ou *live* (multimodal). Lors d'une étude récente (Arnon et al., 2014), l'association du chant maternel au contact peau-à-peau (méthode Kangourou) augmentait la stabilité physiologique des bébés prématurés et réduisait l'anxiété maternelle. Dans une autre étude (Filippa, Devouche, Arioni, Imbert, & Gratier, 2013), le chant et la parole maternelles avaient des

effets bénéfiques sur les bébés, en augmentant les niveaux d'oxygénation, le rythme cardiaque et en diminuant l'occurrence des interventions médicales.

L'efficacité des mères à réguler l'affect de leur enfant a des conséquences sur le lien d'attachement et dans la capacité des enfants à s'autoréguler (DiCorcia & Tronick, 2011; Gunnar & Quevedo, 2007; Kopp, 1989; Trehub, Ghazban, et al., 2015). De plus, le succès des parents à réguler leur bébé influence leurs perceptions d'efficacité et leur bien-être psychologique (Jones & Prinz, 2005). Les mères dépressives (Feldman & Eidelman, 2007; Field, 1998), les mères anxieuses (Feldman, Granat, & Gilboa-Schechtman, 2005) ou les femmes avec un historique personnel de rejet parental (Leerkes & Siepak, 2006), pourraient bénéficier d'interventions visant à soutenir la relation parent-enfant à travers le chant. Des bénéfices de ces interventions pourraient résulter tant de l'autorégulation musicale (e.g. Saarikallio, 2011) pour le parent, que de l'efficacité à réguler l'enfant.

Chanter aux bébés ou aux jeunes enfants pourrait aussi faciliter la perception de la parole et, par extension, l'acquisition du langage. Par exemple, les bébés distinguent plus aisément un changement dans l'ordre de syllabes lorsque la séquence est chantée plutôt que parlée (Lebedeva & Kuhl, 2010). Possiblement, la parole DB facilite le traitement langagier par l'entremise de ses propriétés « musicales », telles que les contours de hauteurs exagérés (Liu et al., 2003; Trainor & Desjardins, 2002) et les répétitions de « chansons signatures » (Bergeson & Trehub, 2007). Une méta-analyse récente révèle que l'entraînement musical améliore la performance à des tâches de conscience phonologique chez l'enfant (Gordon, Fehd, & McCandliss, 2015). De plus, une thérapie par le chant (thérapie d'intonation mélodique ou *melodic intonation therapy*) facilite la réhabilitation du langage chez l'adulte aphasique, avec les hauteurs et le rythme contribuant aux conséquences bénéfiques (Zumbansen, Peretz, & Hébert, 2014). Il est probable que la thérapie par le chant soit aussi bénéfique aux jeunes enfants présentant des retards langagiers.

3. Conclusion

La musique est universelle, intimement liée aux événements significatifs de la vie tels que les célébrations ou les deuils, et joue également un rôle important dans la régulation émotionnelle personnelle, particulièrement lors de l'adolescence et la vie adulte. La musique est aussi significativement imbriquée dans la vie des bébés, qui sont exposés aux chants des personnes qui en prennent soins, aux jouets musicaux et aux diverses musiques enregistrées leur étant destinées ou non. Bien que les parents à travers le monde chantent aux bébés, notamment pour les calmer ou les amuser, peu de documentation existe sur les conséquences affectives et attentionnelles des types de musiques auxquelles ils sont généralement exposés. Il est important d'étudier les effets de la musique dans la petite enfance car cela peut nous donner des indices sur les bases biologiques de nos réactions face à la musique et ouvrir la voie, peut-être, sur ce qui rend la musique si spéciale à nos oreilles.

Peu de comparaisons existent sur le chant et la parole dans la petite enfance, malgré leur ressemblance probablement beaucoup plus saillante pour des bébés préverbaux que pour des auditeurs plus âgés. Les résultats de la présente thèse mettent en lumière le statut du chant dans l'enfance. Ils démontrent que le chant est un outil parental puissant, tout aussi efficace que la parole afin de capter l'attention et plus efficace encore afin de réguler l'affect des bébés. Ils renforcent l'utilité du chant dans la vie de tous les jours et l'utilité potentielle du chant dans des contextes thérapeutiques impliquant des bébés. Bien des questionnements demeurent néanmoins sans réponse, incluant les aspects du chant contribuant à son efficacité dans la régulation de l'affect et l'ampleur par laquelle le chant peut réguler le comportement chez des bébés au développement typique ou atypique.

L'efficacité du chant pourrait s'étendre bien au-delà de la régulation affective du bébé. Le chant procure un moyen par lequel les parents peuvent partager leurs états affectifs ou leurs émotions à l'enfant, réguler leurs propres émotions, intensifier le lien d'attachement parent-enfant, faciliter les apprentissages du bébé et amplifier leur bien-être personnel et celui de leur bébé. Ces conséquences ne sont sans doute pas spécifiques au chant, mais aisément atteintes par celui-ci. Chanter nécessite un minimum d'efforts et de capacités de la part du chanteur, tout en ayant des effets puissants pour celui qui l'entend et celui qui le produit.

Bibliographie

- Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: a meta-analysis of fMRI studies. *Hearing research, 307*, 29-41.
- Arnon, S., Diamant, C., Bauer, S., Regev, R., Sirota, G., & Litmanovitz, I. (2014). Maternal singing during kangaroo care led to autonomic stability in preterm infants and reduced maternal anxiety. *Acta Paediatrica, 103*(10), 1039-1044.
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology, 70*(3), 614.
- Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech communication, 46*(3), 252-267.
- Baruch, C. (1994). Discrimination de tempo chez des bébés de 4 mois. *Le Journal de Physique IV, 4*(C5), C5-367-C365-370.
- Baruch, C., & Drake, C. (1997). Tempo discrimination in infants. *Infant Behavior and Development, 20*(4), 573-577.
- Baruch, C., Panissal-Vieu, N., & Drake, C. (2004). Preferred perceptual tempo for sound sequences: comparison of adults, children, & infants *Perceptual and motor skills, 98*(1), 325-339.
- Belin, P., Zatorre, R., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature, 403*(6767), 309-312.
- Bergeson, T., & Trehub, S. (2006). Infants perception of rhythmic patterns. *Music Perception, 23*(4), 345-360.
- Bergeson, T. R., & Trehub, S. E. (1999). Mothers' singing to infants and preschool children. *Infant Behavior & Development, 22*(1), 51-64.

- Bergeson, T. R., & Trehub, S. E. (2002). Absolute pitch and tempo in mothers' songs to infants. *Psychological Science, 13*(1), 72-75.
- Bergeson, T. R., & Trehub, S. E. (2007). Signature tunes in mothers' speech to infants. *Infant Behav Dev, 30*(4), 648-654.
- Blood, A., & Zatorre, R. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America, 98*(20), 11818.
- Bowling, D. L., Sundararajan, J., Han, S. e., & Purves, D. (2012). Expression of emotion in Eastern and Western music mirrors vocalization. *PLoS One, 7*(3), e31942.
- Braarud, H. C., & Stormark, K. M. (2008). Prosodic Modification and Vocal Adjustments in Mothers' Speech During Face-to-face Interaction with Their Two-to Four-month-old Infants: A Double Video Study. *Social Development, 17*(4), 1074-1084.
- Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science, 11*(6), 853-861.
- Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology, 3*, 327.
- Bryant, G., & Barrett, H. (2007). Recognizing intentions in infant-directed speech. *Psychological Science, 18*(8), 746.
- Byrne, J. M., & Horowitz, F. D. (1981). Rocking as a soothing intervention: The influence of direction and type of movement. *Infant Behavior and Development, 4*, 207-218.
- Cabib, S. (2006). The neurobiology of stereotypy II: The role of stress. Dans G. Mason & J. Rushen (dir.), *STEREOTYPIC ANIMAL BEHAVIOUR, Fundamentals and applications to welfare, second edition* (p. 227-255). Trowbridge, UK: Cromwell Press.
- Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in cognitive sciences, 17*(4), 179-193.

- Chen-Hafteck, L. (1997). Music and language development in early childhood: Integrating past research in the two domains. *Early Child Development and Care, 130*(1), 85-97.
- Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage, 32*(4), 1771-1781.
- Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental science, 17*(6), 1003-1011.
- Cohen, L. B. (1973). A two process model of infant visual attention. *Merrill-Palmer Quarterly of Behavior and Development, 19*(3), 157-180.
- Cole, P. M., Martin, S. E., & Dennis, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development, 75*(2), 317-333.
- Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development, 4*, 219-223.
- Colombo, J., Richman, W. A., Shaddy, D. J., Follmer Greenhoot, A., & Maikranz, J. M. (2001). Heart rate-defined phases of attention, look duration, & infant performance in the paired-comparison paradigm. *Child development, 72*(6), 1605-1616.
- Cooper, R., Abraham, J., Berman, S., & Staska, M. (1997). The development of infants' preference for motherese. *Infant Behavior and Development, 20*(4), 477-488.
- Cooper, R., & Aslin, R. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*(5), 1584-1595.
- Costa-Giomi, E. (2014). Mode of Presentation Affects Infants' Preferential Attention to Singing and Speech. *Music Perception: An Interdisciplinary Journal, 32*(2), 160-169.
- Costa-Giomi, E., & Ilari, B. (2014). Infants' preferential attention to sung and spoken stimuli. *Journal of Research in Music Education, 62*(2), 188-194.

- Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child development, 77*(3), 680-695.
- Courage, M. L., & Setliff, A. E. (2010). When babies watch television: Attention-getting, attention-holding, & the implications for learning from video material. *Developmental review, 30*(2), 220-238.
- Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass, 7*(3), 157-170.
- Curtin, S., & Werker, J. F. (2007). The perceptual foundations of phonological development. Dans G. Gaskell (dir.), *The Oxford Handbook of Psycholinguistics* (p. 579-599): Oxford University Press.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*, 1174-1176.
- DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*(2), 133-150.
- Dehaene-Lambertz, G., Montavont, A., Jobert, A., Alliol, L., Dubois, J., & Hertz-Pannier, L. (2010). Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain and Language, 114*(2), 53-65.
- Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America, 129*(4), 2245-2252.
- DiCorcia, J. A., & Tronick, E. (2011). Quotidian resilience: exploring mechanisms that drive resilience from a perspective of everyday stress and coping. *Neuroscience & Biobehavioral Reviews, 35*(7), 1593-1602.
- Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction. Dans N. L. Wallin & B. Merker (dir.), *The origins of music* (p. 389-410): MIT Press.

- Dissanayake, E. (2009). Root, leaf, blossom, or bole: Concerning the origin and adaptive function of music. *Communicative musicality: Exploring the basis of human companionship*, 17-30.
- Drake, C. (1997). Motor and perceptually preferred synchronisation by children and adults: binary and ternary ratios. *Polish Quarterly of Developmental Psychology*, 3(1), 41-59.
- Drake, C., & Botte, M. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model *Perception & Psychophysics*.
- Drake, C., Botte, M., & Baruch, C. (1992). Tempo sensitivity in auditory sequences. *The Journal of the Acoustical Society of America*, 92, 2318.
- Drake, C., & Gérard, C. (1989). A psychological pulse train: how young children use this cognitive framework to structure simple rhythms. *Psychological Research*, 51(1), 16-22.
- Drake, C., Jones, M., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3), 251-288.
- Dunst, C., Gorman, E., & Hamby, D. (2012). Preference for infant-directed speech in preverbal young children. *Center for Early Literacy Learning*, 5(1).
- Eckerdal, P., & Merker, B. (2009). Music and the 'action song' in infant development: An interpretation. Dans N. L. Wallin, B. Merker & S. Brown (dir.), *Communicative musicality: Exploring the basis of human companionship* (p. 241-262). New York, NY: Oxford University Press.
- Eerola, T., Luck, G., & Toiviainen, P. (2006). *An investigation of pre-schoolers' corporeal synchronization with music*. Communication présentée Proceedings of the 9th international conference on music perception and cognition.
- Esposito, G., Yoshida, S., Ohnishi, R., Tsuneoka, Y., Rostagno, M. d. C., Yokota, S., . . . Shimizu, M. (2013). Infant calming responses during maternal carrying in humans and mice. *Current Biology*, 23(9), 739-745.

- Falk, S. (2011). Melodic versus intonational coding of communicative functions: A comparison of tonal contours in infant-directed song and speech. *Psychomusicology: Music, Mind and Brain*, 21(1-2), 54.
- Falk, S., Rathcke, T., & Dalla Bella, S. (2014). When speech sounds like music. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1491.
- Feldman, R., & Eidelman, A. I. (2007). Maternal postpartum behavior and the emergence of infant–mother and infant–father synchrony in preterm and full-term infants: The role of neonatal vagal tone. *Developmental psychobiology*, 49(3), 290-302.
- Feldman, R., Granat, A., & Gilboa-Schechtman, E. (2005). *Maternal anxiety and depression, mother–infant synchrony, & infant regulation of negative and positive emotions*. Communication présentée biennial meeting of the Society for Research in Child Development, Atlanta, GA.
- Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental psychology*, 35(1), 223.
- Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, 13(2), 271-278.
- Ferguson, C. (1964). Baby talk in six languages. *American Anthropologist*, 66(6), 103-114. doi: 10.1525/aa.1964.66.suppl_3.02a00060
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8(2), 181-195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60(6), 1497-1510.
- Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. Dans H. Papoušek & U. Jürgens (dir.), *Nonverbal vocal communication: Comparative and developmental approaches* (p. 262-282). New York, NY: Cambridge University Press.

- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development, 64*(3), 657-674.
- Fernald, A., & Kuhl, P. K. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development, 10*(3), 279-293.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27*(2), 209-221.
- Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*(3), 637-656.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20*(1), 104-113.
- Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of child language, 16*(03), 477-501.
- Field, T. (1998). Maternal depression effects on infants and early interventions. *Preventive medicine, 27*(2), 200-203.
- Filippa, M., Devouche, E., Arioni, C., Imberty, M., & Gratier, M. (2013). Live maternal speech and singing have beneficial effects on hospitalized preterm infants. *Acta Paediatrica, 102*(10), 1017-1020.
- Frick, R. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*(3), 412-429.
- Fukui, H., & Toyoshima, K. (2014). Music increase altruism through regulating the secretion of steroid hormones and peptides. *Medical hypotheses, 83*(6), 706-708.
- Gabriels, R. L., Agnew, J. A., Pan, Z., Holt, K. D., Reynolds, A., & Laudenslager, M. L. (2013). Elevated repetitive behaviors are associated with lower diurnal salivary cortisol levels in autism spectrum disorder. *Biological psychology, 93*(2), 262-268.

- Gammie, Stephen C. (2013). Mother–Infant Communication: Carrying Understanding to a New Level. *Current Biology*, 23(9), R341-R343.
- Gauthier, B., & Shi, R. (2011). A connectionist study on the role of pitch in infant-directed speech. *The Journal of the Acoustical Society of America*, 130(6), EL380-EL386.
- Gerhardt, K., & Abrams, R. (2000). Fetal exposures to sound and vibroacoustic stimulation. *Journal of perinatology: official journal of the California Perinatal Association*, 20(8 Pt 2), S21.
- Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual review of psychology*, 61, 191-218.
- Glenn, S., & Cunningham, C. (1983). What do babies listen to most? A developmental study of auditory preferences in nonhandicapped infants and infants with Down's syndrome. *Developmental Psychology*, 19(3), 332-337.
- Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in psychology*, 6.
- Grabe, E., & Low, E. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology*, 7(515-546).
- Grahn, J. A. (2012). Neural mechanisms of rhythm perception: current findings and future perspectives. *Topics in cognitive science*, 4(4), 585-606.
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience*, 29(23), 7540-7548.
- Greasley, A. E., & Lamont, A. M. (2006). *Music preference in adulthood: Why do we like the music we do*. Communication présentée Proceedings of the 9th international conference on music perception and cognition.

- Grieser, D. A. L., & Kuhl, P. K. (1988). Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Developmental Psychology*, 24(1), 14-20.
- Griffiths, S. K., Brown Jr, W., Gerhardt, K. J., Abrams, R. M., & Morris, R. J. (1994). The perception of speech sounds recorded within the uterus of a pregnant sheep. *The Journal of the Acoustical Society of America*, 96(4), 2055-2063.
- Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restorative neurology and neuroscience*, 28(2), 219-236.
- Gunji, A., Koyama, S., Ishii, R., Levy, D., Okamoto, H., Kakigi, R., & Pantev, C. (2003). Magnetoencephalographic study of the cortical activity elicited by human voice. *Neuroscience letters*, 348(1), 13-16.
- Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. *Annu. Rev. Psychol.*, 58, 145-173.
- Habibi, A., & Damasio, A. (2014). Music, feelings, & the human brain. *Psychomusicology: Music, Mind, & Brain*, 24(1), 92.
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends Cogn Sci*, 11(11), 466-472.
- Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychol Sci*, 16(1), 48-55.
- Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: infants learn more readily than adults. *Proc Natl Acad Sci U S A*, 102(35), 12639-12643.
- Haviland, J. M., & Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Developmental Psychology*, 23(1), 97.
- Hayashi, A., Tamekawa, Y., & Kiritani, S. (2001). Developmental change in auditory preferences for speech stimuli in Japanese infants. *Journal of Speech, Language, & Hearing Research*, 44(6), 1189-1200.

- Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, & human social behavior. *Frontiers in Neuroendocrinology*, *30*, 548-557.
- Hepper, P. (1991). An examination of fetal learning before and after birth. *The Irish Journal of Psychology*, *12*(2), 95-107.
- Hunziker, U. A., & Barr, R. G. (1986). Increased carrying reduces infant crying: a randomized controlled trial. *Pediatrics*, *77*(5), 641-648.
- Ilari, B., & Sundara, M. (2009). Music Listening Preferences in Early Life: Infants' Responses to Accompanied Versus Unaccompanied Singing. *Journal of Research in Music Education*, *56*(4), 357.
- Ilie, G., & Thompson, W. F. (2006). A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception: An Interdisciplinary Journal*, *23*(4), 319-330.
- Jacobson, J., Boersma, D., Fields, R., & Olson, K. (1983). Paralinguistic features of adult speech to infants and small children. *Child Development*, *54*(2), 436-442. doi: 10.2307/1129704
- Jardri, R., Pins, D., Houfflin-Debarge, V., Chaffiotte, C., Rocourt, N., Pruvo, J., . . . Thomas, P. (2008). Fetal cortical activation to sound at 33 weeks of gestation: a functional MRI study. *Neuroimage*, *42*(1), 10-18.
- Jean, A. D., & Stack, D. M. (2009). Functions of maternal touch and infants' affect during face-to-face interactions: New directions for the still-face. *Infant Behavior and Development*, *32*(1), 123-128.
- Johnstone, B. (1994). *Repetition in discourse: Interdisciplinary perspectives*. Ablex Publishing Corporation.
- Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical psychology review*, *25*(3), 341-363.

- Juslin, P., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770-814.
- Juslin, P., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(05), 559-575.
- Juslin, P. N., & Sloboda, J. (2010). *Handbook of music and emotion: Theory research, applications*.
- Kent, R. D., Read, C., & Kent, R. D. (2002). *The acoustic analysis of speech*. Singular San Diego, CA.
- Khalifa, S., Dalla Bella, S., Roy, M., Peretz, I., & Lupien, S. J. (2003). Effects of relaxing music on salivary cortisol level after psychological stress. *Neurosciences and Music*, *999*, 374-376.
- Kim, H. I., & Johnson, S. P. (2013). Do young infants prefer an infant-directed face or a happy face? *International Journal of Behavioral Development*, *37*(2), 125-130. doi: 10.1177/0165025413475972
- Kim, H. I., & Johnson, S. P. (2014). Detecting 'infant-directedness' in face and voice. *Developmental science*, *17*(4), 621-627.
- Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of experimental child psychology*, *102*(3), 299-314.
- Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. Dans C. Rovee-Collier, L. Lipsitt & H. Hayne (dir.), *Advances in infancy research* (Vol. 12, p. 221-236). Stamford, CT: Ablex Publishing Corp.
- Kitamura, C., & Burnham, D. (2003). Pitch and Communicative Intent in Mother's Speech: Adjustments for Age and Sex in the First Year. *Infancy*, *4*(1), 85-110.

- Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment-if the rhythm comes easily. *PLoS One*, 6(11), e27272.
- Kopp, C. B. (1989). Regulation of distress and negative emotions: A developmental view. *Developmental Psychology*, 25(3), 343.
- Kotilahti, K., Nissilä, I., Näsi, T., Lipiäinen, L., Noponen, T., Meriläinen, P., . . . Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human brain mapping*, 31(4), 595-603.
- Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child Development*, 1054-1061.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nat Rev Neurosci*, 5(11), 831-843.
- Kuhl, P. K., &ruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., . . . Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science*, 277(5326), 684.
- Lam, C., & Kitamura, C. (2006, December 6-8). *Developmental trends in infant preferences for affective intent in mothers' speech*. Communication présentée Proceedings of the 11th Australian International Conference on Speech Science & Technology, University of Auckland, New Zealand.
- Lam, C., & Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science*, 15(2), 212-221.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review*, 106(1), 119.
- Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, 19(5), 633-653.

- Leardi, S., Pietroletti, R., Angeloni, G., Necozone, S., Ranalletta, G., & Del Gusto, B. (2007). Randomized clinical trial examining the effect of music therapy in stress response to day surgery. *British journal of surgery*, *94*(8), 943-947.
- Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: Infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant behavior and development*, *33*(4), 419-430.
- Leerkes, E. M., & Siepak, K. J. (2006). Attachment linked predictors of women's emotional and cognitive responses to infant distress. *Attachment & Human Development*, *8*(01), 11-32.
- Levy, D. A., Granot, R., & Bentin, S. (2003). Neural sensitivity to human voices: ERP evidence of task and attentional influences. *Psychophysiology*, *40*(2), 291-305.
- Liberman, A. M. (1996). *Speech: A special code*. Cambridge, MA: MIT press.
- Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science*, *6*(3), F1-F10.
- Livingstone, S. R., & Palmer, C. (2015). Head Movements Encode Emotions During Speech and Song.
- London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.
- Longhi, E. (2009). Songese': Maternal structuring of musical interaction with infants. *Psychology of Music*, *37*(2), 195-213.
- Mandel, D. R., Jusczyk, P. W., & Pisoni, D. B. (1995). Infants' recognition of the sound patterns of their own names. *Psychological Science*, 314-317.
- Margulis, E. H. (2013a). Aesthetic responses to repetition in unfamiliar music. *Empirical Studies of the Arts*, *31*(1), 45-57.

- Margulis, E. H. (2013b). Repetition and emotive communication in music versus speech. *Frontiers in psychology, 4*.
- Margulis, E. H. (2014). *On repeat: How music plays the mind*. Oxford University Press.
- Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers speak less clearly to infants than to adults a comprehensive test of the hyperarticulation hypothesis. *Psychological science, 26*(3), 341-347.
- Masataka, N. (1992a). Motherese in a signed language. *Infant Behavior and Development, 15*(4), 453-460.
- Masataka, N. (1992b). Pitch characteristics of Japanese maternal speech to infants. *Journal of child language, 19*(2), 213-223.
- Masataka, N. (1996). Perception of motherese in a signed language by 6-month-old deaf infants. *Developmental Psychology, 32*(5), 874-879.
- Masataka, N. (1998). Perception of motherese in Japanese sign language by 6-month-old hearing infants. *Developmental psychology, 34*(2), 241-246.
- Masataka, N. (1999). Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Developmental psychology, 35*(4), 1001-1005.
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental psychobiology, 35*(3), 204-214.
- McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: life span development of timing and event tracking. *Journal of Experimental Psychology: General, 135*(3), 348.
- McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception: An Interdisciplinary Journal, 21*(3), 289-311.
- Menon, V., & Levitin, D. J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage, 28*(1), 175-184.

- Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45(1), 4-17.
- Moon, C., Cooper, R., & Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495-500.
- Moon, C., & Fifer, W. (2000). Evidence of transnatal auditory learning. *Journal of Perinatology*, 20(8; PART 2), 37-44.
- Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development*, 67(6), 3219-3237.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior and Development*, 27(4), 455-464.
- Nakata, T., & Trehub, S. E. (2011). Expressive timing and dynamics in infant-directed and non-infant-directed singing. *Psychomusicology: Music, Mind and Brain*, 21(1-2), 45.
- Nelson, C. A., & Dolgin, K. G. (1985). The generalized discrimination of facial expressions by seven-month-old infants. *Child development*, 58-61.
- Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. *The origins of music*, 463-472.
- Newman, R. S., & Hussain, I. (2006). Changes in Preference for Infant-Directed Speech in Low and Moderate Noise by 4.5-to 13-Month-Olds. *Infancy*, 10(1), 61-76.
- Ochs, E., & Schieffelin, B. B. (1995). The impact of language socialization on grammatical development. Dans P. Fletcher & B. MacWhitner (dir.), *Handbook of child language* (p. 73-94). New York, NY: Basil Blackwell.
- Oster, H. (1981). Recognition of emotional expression in infancy. *Infant social cognition: Empirical and theoretical considerations*, 85-125.

- Papoušek, M. (1994). Melodies in caregivers' speech: A species-specific guidance towards language. *Early Development and Parenting*, 3(1), 5-17.
- Papousek, M., Bornstein, M. H., Nuzzo, C., Papousek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, 13(4), 539-545.
- Papousek, M., Papousek, H., & Symmes, D. (1991). The meanings of melodies in motherese in tone and stress languages. *Infant Behavior and Development*, 14(4), 415-440.
- Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87(1), B35-B45.
- Patel, A. D., Iversen, J. R., Chen, Y. Q., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, 163(2), 226-238.
- Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of the Acoustical Society of America*, 119(5), 3034-3047.
- Patel, A. D., Peretz, I., Tramo, M., & Labreque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language*, 61(1), 123-144.
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, 15(3), 325-345.
- Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental science*, 11(1), 60-68.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., & Reolli, G., Rovelli, R., . . . Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758-4763.

- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PloS one*, 6(11), e27241.
- Peretz, I., Aubé, W., & Armony, J. L. (2013). Towards a neurobiology of musical emotions. *The evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man*, 277.
- Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1664), 20140090.
- Petkov, C. I., Kang, X., Alho, K., Bertrand, O., Yund, E. W., & Woods, D. L. (2004). Attentional modulation of human auditory cortex. *Nature neuroscience*, 7(6), 658-663.
- Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception: An Interdisciplinary Journal*, 28(1), 3-14.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308(5727), 1430-1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, 105(3), 533-546.
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67(1), 94-102.
- Pinker, S. (1997). *How the mind works*. New York: Norton.
- Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, 2(4), 411-440.
- Provasi, J., &erson, D. I., & Barbu-Roth, M. (2014). Rhythm perception, production, & synchronization during the perinatal period. *Front Psychol*, 5(1048), 1-16.

- Provasi, J., & Bobin-Bègue, A. (2003). Spontaneous motor tempo and rhythmical synchronisation in 2½-and 4-year-old children. *International Journal of Behavioral Development, 27*(3), 220-231.
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403-452.
- Richards, J. E., Reynolds, G. D., & Courage, M. L. (2010). The neural bases of infant attention. *Current Directions in Psychological Science, 19*(1), 41-46.
- Rock, A. M. L., Trainor, L. J., & Addison, T. L. (1999). Distinctive messages in infant-directed lullabies and play songs. *Developmental Psychology, 35*(2), 527-534.
- Rose, S. A., Gottfried, A. W., Melloy-Carminar, P., & Bridger, W. H. (1982). Familiarity and novelty preferences in infant recognition memory: Implications for information processing. *Developmental Psychology, 18*(5), 704-713.
- Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music, 39*(3), 307-327.
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature neuroscience, 14*(2), 257-262.
- Sambeth, A., Ruohio, K., Alku, P., Fellman, V., & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology, 119*(2), 332-341.
- Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences, 112*(29), 8987-8992.
- Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental psychology, 47*(1), 19.
- Schellenberg, E. G., Peretz, I., & Vieillard, S. (2008). Liking for happy- and sad-sounding music: Effects of exposure. *Cognition & Emotion, 22*(2), 218-237.

- Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice*, 9(3), 235-248.
- Seltzer, L. J., Ziegler, T. E., & Pollak, S. D. (2010). Social vocalizations can release oxytocin in humans. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1694), 2661-2666.
- Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behavior and Development*, 18(4), 477-482.
- Shenfield, T., Trehub, S. E., & Nakata, T. (2003). Maternal singing modulates infant arousal. *Psychology of Music*, 31(4), 365-375.
- Shultz, S., & Vouloumanos, A. (2010). Three-month-olds prefer speech to other naturally occurring signals. *Language Learning and Development*, 6(4), 241-257.
- Shute, B., & Wheldall, K. (2009). Pitch alterations in British motherese: Some preliminary acoustic data. *Journal of child language*, 16(03), 503-512.
- Singh, L., Morgan, J. L., & Best, C. T. (2002). Infants Listening Preferences: Baby Talk or Happy Talk? *Infancy*, 3(3), 365-394.
- Sloboda, J., Lamont, A., & Greasley, A. (2009). Choosing to hear music: Motivation, process and effect. Dans S. Hallam, Cross, I., & Thaut, M. (dir.), *Oxford handbook of music psychology* (p. 431–440). New York: Oxford University Press.
- Smith, N. A., & Trainor, L. J. (2008). Infant-Directed Speech Is Modulated by Infant Feedback. *Infancy*, 13(4), 410-420.
- Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, 27(4), 501-532.
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: a cross-cultural comparison. *Developmental Psychology*, 46(1), 286.

- Spence, M. (1996). Young infants' long-term auditory memory: Evidence for changes in preference as a function of delay. *Developmental Psychobiology*, 29(8), 685-695.
- Stack, D. M., & Muir, D. W. (1992). Adult Tactile Stimulation during Face-to-Face Interactions Modulates Five-Month-Olds' Affect and Attention. *Child Development*, 63(6), 1509-1525.
- Standley, J. (2012). Music therapy research in the NICU: an updated meta-analysis. *Neonatal Network*, 31(5), 311-316.
- 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), 1-15.
- Stern, D. N., Spieker, S., & Mackain, K. (1982). Intonation Contours as Signals in Maternal Speech to Pre-Linguistic Infants. *Developmental Psychology*, 18(5), 727-735.
- Strathearn, L., Li, J., Fonagy, P., & Montague, P. R. (2008). What's in a smile? Maternal brain responses to infant facial cues. *Pediatrics*, 122(1), 40-51.
- Szpunar, K. K., Schellenberg, E. G., & Pliner, P. (2004). Liking and memory for musical stimuli as a function of exposure. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 30(2), 370-381.
- Tartter, V. C. (1980). Happy talk: Perceptual and acoustic effects of smiling on speech. *Attention, Perception, & Psychophysics*, 27(1), 24-27.
- Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the society for research in child development*, 59(2-3), 25-52.
- Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20(4), 410-433.
- Trainor, L. J. (1996). Infant preferences for infant-directed versus non infant-directed playsongs and lullabies. *Infant Behavior and Development*, 19(1), 83-92.

- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science, 11*(3), 188-195.
- Trainor, L. J., Clark, E. D., Huntley, A., & Adams, B. A. (1997). The acoustic basis of preferences for infant-directed singing. *Infant Behavior and Development, 20*(3), 383-396.
- 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.
- Trainor, L. J., Gao, X., Lei, J.-j., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *cortex, 45*(1), 35-43.
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development, 21*(1), 77-88.
- Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2-and 4-month-old infants. *Music Perception, 20*(2), 187-194.
- Trehub, S. E. (2000). Human processing predispositions and musical universals. Dans L. Wallin, B. Merker & S. Brown (dir.), *The origins of music* (p. 427-448). Cambridge, MA: MIT Press.
- Trehub, S. E. (2001). Musical predispositions in infancy. *Ann N Y Acad Sci, 930*, 1-16.
- Trehub, S. E. (2003). The developmental origins of musicality. *Nat Neurosci, 6*(7), 669-673.
- Trehub, S. E. (2012). Behavioral methods in infancy: pitfalls of single measures. *Annals of the New York Academy of Sciences, 1252*(1), 37-42. doi: 10.1111/j.1749-6632.2012.06448.x
- Trehub, S. E., Ghazban, N., & Corbeil, M. (2015). Musical affect regulation in infancy. *Annals of the New York Academy of Sciences, 1337*(1), 186-192.

- Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: domain-general or domain-specific mechanisms? *Cognition*, *100*(1), 73-99.
- Trehub, S. E., Hannon, E. E., & Schachner, A. (2011). Perspectives on music and affect in the early years. Dans P. N. Juslin & J. A. Sloboda (dir.), *Handbook of music and emotion: Theory, research, applications* (p. 645-668.). Oxford: Oxford University Press.
- Trehub, S. E., Hill, D. S., & Kamenetsky, S. B. (1997). Parents' sung performances for infants. *Can J Exp Psychol*, *51*(4), 385-396.
- Trehub, S. E., Plantinga, J., & Russo, F. R. (2016). Maternal vocal interactions with infants: Reciprocal visual influences. *Social Development*.
- Trehub, S. E., & Schellenberg, E. G. (1995). Music: Its relevance to infants. Dans R. Vasta (dir.), *Annals of child development* (Vol. 11, p. 1-24). New York: Jessica Kingsley Publishers.
- Trehub, S. E., & Trainor, L. J. (1998). Singing to infants: Lullabies and play songs. Dans C. Rovee-Collier, L. Lipsitt & H. Hayne (dir.), *Advances in infancy research* (Vol. 12, p. 43-78). Stamford, CT: Ablex Publishing Corp.
- Trehub, S. E., Trainor, L. J., & Unyk, A. M. (1993). Music and speech processing in the first year of life. *Advances in child development and behavior*, *24*, 1.
- Trehub, S. E., Unyk, A. M., & Henderson, J. L. (1994). Children's songs to infant siblings: parallels with speech. *J Child Lang*, *21*(3), 735-744.
- Trehub, S. E., Unyk, A. M., Kamenetsky, S. B., Hill, D. S., Trainor, L. J., Henderson, J. L., & Saraza, M. (1997). Mothers' and fathers' singing to infants. *Dev Psychol*, *33*(3), 500-507.
- Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, *16*(3), 285-295.
- Tronick, E. Z. (1995). Touch in mother-infant interaction. Dans M. Field (dir.), *Touch in early development* (p. 53-65). New York, NY: Psychology Press.

- Unyk, A. M., Trehub, S. E., Trainor, L. J., & Schellenberg, E. G. (1992). Lullabies and simplicity: A cross-cultural perspective. *Psychology of Music, 20*(1), 15-28.
- Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: the negativity bias in social-emotional development. *Psychological bulletin, 134*(3), 383.
- Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science, 7*(3), 261-269.
- Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46*(4), 693-695.
- Van Noorden, L., & De Bruyn, L. (2009). *The development of synchronisation skills of children 3 to 11 years old*. Communication présentée Proceedings of ESCOM—7th Triennial Conference of the European Society for the Cognitive Sciences of Music. Jyväskylä, Finland: University of Jyväskylä.
- Vanden Bosch der Nederlanden, C. M., Hannon, E. E., & Snyder, J. S. (2015). Finding the music of speech: Musical knowledge influences pitch processing in speech. *Cognition, 143*, 135-140.
- Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The Tuning of Human Neonates' Preference for Speech. *Child Development, 81*(2), 517-527.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Dev Sci, 7*(3), 270-276.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: evidence for a bias for speech in neonates. *Dev Sci, 10*(2), 159-164.
- Vrugt, D. T., & Pederson, D. R. (1973). The effects of vertical rocking frequencies on the arousal level in two-month-old infants. *Child development, 205-209*.

- Vuilleumier, P. (2006). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, 9(12), 585-594.
- Weiss, M. W., Schellenberg, E. G., Trehub, S. E., & Dawber, E. J. (2015). Enhanced processing of vocal melodies in childhood. *Developmental psychology*, 51(3), 370.
- Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings enhanced memory for vocal melodies. *Psychological Science*, 23(10), 1074-1078.
- Weiss, M. W., Vanzella, P., Schellenberg, E. G., & Trehub, S. E. (2015). Pianists exhibit enhanced memory for vocal melodies but not piano melodies. *The Quarterly Journal of Experimental Psychology*, 68(5), 866-877.
- Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.
- Werker, J., & McLeod, P. (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.
- Werker, J., Pegg, J., & McLeod, P. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, 17(3), 323-333.
- Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological science*, 20(1), 1-5.
- 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*, 106(7), 2468-2471.
- Zatorre, R. J., & Baum, S. R. (2012). Musical Melody and Speech Intonation: Singing a Different Tune. *PLoS Biology*, 10(7).
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, 107(13), 5768-5773.

Zentner, M., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development*, 21(3), 483-492.

Zumbansen, A., Peretz, I., & Hébert, S. (2014). The combination of rhythm and pitch can account for the beneficial effect of melodic intonation therapy on connected speech improvements in Broca's aphasia. *Front. Hum. Neurosci*, 8(592), 10.3389.

Annexe 1. Questionnaire remis aux parents.



Laboratoire international
de recherche sur le cerveau,
la musique et le son



Questionnaire sur l'histoire familiale (l'information divulguée demeurera confidentielle)

Nom du parent/tuteur: _____ Date: _____
 Courriel et/ou Téléphone: _____
 Adresse : _____

Prénom de l'enfant	Sexe (M/F)	Date naissance (jj/mm/aa)	#sem de grossesse	Poids à la naissance

1. Audition

- (i) Y a-t-il des doutes ou une histoire familiale d'un déficit auditif **précoce** dans la famille? Oui ___ Non ___
- (ii) L'audition de l'enfant a-t-elle déjà été testée en audiologie ? Oui ___ Non ___
 Si oui, des déficits auditifs ont-ils été décelés ? Oui ___ Non ___
- (iii) Votre enfant a-t-il déjà eu:
- a) plus de 3 infections aux oreilles en une année? Oui ___ Non ___
- b) des tubes dans les oreilles? Oui ___ Non ___
- c) Votre enfant a-t-il déjà eu ou a un problème significatif qui pourrait affecter la participation à l'étude?
 Oui ___ Non ___ Si oui, spécifiez : _____

2. Langues :

Parent/Tuteur

Autre adulte dans la maison

(ex : autre parent, conjoint(e), grand-parent)

Pays de naissance (des adultes) _____

Langue maternelle (des adultes) _____

Langue(s) de l'enfant:

Le français est-il la seule langue parlée à l'enfant(s)? Oui ___ Non ___

Si non, spécifiez ci-dessous, les autres langues utilisées:

Langue	Par qui?	% de tps parlée à l'enfant

	Parent/Tuteur	Autre adulte (si pertinent)
<u>3. Éducation des parents:</u>		
Un peu de secondaire	_____	_____
Diplômé du secondaire	_____	_____
Un peu de Cégep	_____	_____
Diplômé du Cégep	_____	_____
Un peu d'Université	_____	_____
Diplômé de l'Université	_____	_____
Études cycles supérieures	_____	_____

	Parent/Tuteur	Autre adulte (si pertinent)
<u>4. Musique:</u>		
Votre expérience musicale (excluant les cours scolaires)		
Aucune	_____	_____
Jusqu'à 4 ans	_____	_____
De 4 à 8 ans	_____	_____
Plus de 8 ans	_____	_____

Si expérience musicale, vous vous décrivez comme un musicien :		
Débutant	_____	_____
Intermédiaire	_____	_____
Avancé	_____	_____

Si expérience musicale, cochez si l'affirmation s'applique:		
Apprentissage par soi-même	_____	_____
Capable de jouer à l'oreille	_____	_____
Programme musical (cégep et/ou université)	_____	_____
Vous en jouez présentement (dernière année)	_____	_____

Vos instrument(s) (incluant la voix)	_____	_____
	_____	_____

Pour les énoncés suivants, reportez-vous à l'échelle ci-dessous:

Jamais	Rarement	Parfois	Moyennement	Souvent	Très souvent	Presque toujours
1	2	3	4	5	6	7

Jouez-vous (ou faites-vous jouer) de la musique à la maison? 1 2 3 4 5 6 7
 Jouez-vous (ou faites-vous jouer) de la musique dans l'auto? 1 2 3 4 5 6 7
 Jouez-vous (ou faites-vous jouer) de la musique directement pour l'enfant?
 1 2 3 4 5 6 7

Nombre d'heures (ou minutes) de musique par jour : _____
 Chantez-vous à votre enfant? Oui _____ Non _____ Si oui : 1 2 3 4 5 6 7
 Nombre d'heures (ou minutes) de chant par jour : _____
 Type de musique jouée :

Pop		1	2	3	4	5	6	7	
Rock		1	2	3	4	5	6	7	
Classique	1	2	3	4	5	6	7		
Jazz		1	2	3	4	5	6	7	
rap/hip-hop		1	2	3	4	5	6	7	
pour enfant		1	2	3	4	5	6	7	
autre		1	2	3	4	5	6	7	Spécifiez: _____
autre		1	2	3	4	5	6	7	Spécifiez: _____

Décrivez toute formation musicale que votre enfant a ou a eu (incluant les classes musicales pour bébés):

Merci!

Annexe 2. Matériel audio supplémentaire.

Exemples de stimuli utilisés dans l'article 1

Expérience 1

audio1.wav : Berceuse chilienne

audio2.wav : Séquence de syllabes

Expérience 2

audio3.wav : Chant DB (chanson turque)

audio4.wav : Paroles DB (chanson turque)

Expérience 3

audio3.wav : Chant DB (chanson turque)

audio5.wav : Paroles DA (chanson turque)

Exemples de stimuli utilisés dans l'article 2

Expérience 1

audio3.wav : Chant DB (chanson turque)

audio4.wav : Paroles DB (chanson turque)

audio5.wav : Paroles DA (chanson turque)