Université de Montréal

Étude de la modularité de la synchronisation à la pulsation musicale

Synchronisation sensorimotrice dans l'amusie congénitale

par

Marie-Élaine Lagrois

Département de psychologie

Faculté des arts et des 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

Août, 2018

Résumé

À l'écoute de musique, la plupart des gens ressentent naturellement l'envie de bouger au rythme de celle-ci. Bien que cette activité puisse paraitre anodine, la capacité à émettre un mouvement en synchronie avec une pulsation rythmique (ou le *beat* en anglais) repose sur l'interaction d'un ensemble de mécanismes complexes. L'habileté à percevoir la pulsation dans la musique, et à pouvoir prédire l'occurrence des temps marquant celle-ci, serait entre autres possible grâce à un mécanisme d'entrainement. La spécificité des processus d'entrainement à la pulsation musicale demeure cependant débattue. Afin de mieux cerner la modularité de l'entrainement au beat musical, le travail de recherche présenté dans cette thèse avait pour objectif de caractériser, chez des individus présentant une amusie congénitale, la capacité à se synchroniser à la musique, au chant et à la parole.

L'amusie congénitale est un trouble d'origine neurodéveloppementale qui affecte, selon le cas, l'habileté à percevoir dans la musique les fines variations mélodiques ou la capacité à percevoir et à se synchroniser au beat. L'étude 1 visait d'abord à déterminer l'influence d'une amusie congénitale affectant le traitement des hauteurs sur la perception et la synchronisation au beat. Les résultats ont mis en évidence, dans 50 % des cas, une faible capacité à se synchroniser avec le beat de séquences rythmiques contenant ou non des variations mélodiques. Néanmoins, des cas de dissociation claire ont pu être identifiés. Ainsi, bien que dans la majorité des cas une perception des variations de hauteurs déficiente semble s'accompagner d'une capacité réduite à percevoir la pulsation musicale et s'y synchroniser, ces deux habiletés peuvent également se trouver atteintes de façon isolée. Cette dissociation se constate d'ailleurs chez les individus ayant pris part aux études subséquentes de la thèse, qui ne parviennent pas à synchroniser un mouvement simple avec le beat, tout en demeurant dans les limites de la norme à une épreuve mesurant la perception des hauteurs en contexte mélodique.

Les études 2 et 3 ont été élaborées dans le but d'investiguer la modularité de l'entrainement à la pulsation musicale en regard d'hypothèses suggérant que ce mécanisme soit plutôt partagé avec d'autres domaines tels que le langage. L'étude 2 avait pour objectif de tester l'implication des processus d'entrainement au beat musical dans la synchronisation à la parole.

Pour se faire, un groupe d'individus ayant une amusie congénitale touchant la synchronisation au beat et des participants neurotypiques appariés ont complété une tâche de synchronisation motrice sur des extraits de chant et de parole. Les participants devaient taper du doigt la pulsation rythmique perçue dans des phrases chantées, dites avec un rythme régulier (comparable à du rap), et énoncées de manière naturelle (rythme irrégulier). Dans ces trois conditions, les individus atteints d'amusie sont moins bien parvenus à se synchroniser, comparativement aux participants contrôles. Ce résultat suggère l'existence d'un mécanisme d'entrainement commun à la parole et à la musique.

Enfin, dans l'étude 3, nous avons testé une hypothèse selon laquelle la capacité d'imitation vocale et la synchronisation au beat reposeraient en partie sur un mécanisme commun de couplage sensorimoteur. Dans cette étude, des participants ayant une difficulté à se synchroniser au beat, ainsi que des participants contrôles appariés, ont chanté une mélodie connue avec et sans paroles, d'abord de mémoire, ensuite par imitation d'un modèle, et enfin en synchronie avec le modèle et avec un métronome. Les résultats de cette étude montrent, en premier lieu, que la capacité des participants ayant un trouble de la synchronisation au beat à chanter avec justesse est comparable à la population générale. Certains participants ont tout de même pu être identifiés comme étant de mauvais chanteurs. Cependant, nous n'avons pas pu mettre en évidence d'association claire entre la justesse du chant et l'habileté à se synchroniser à une pulsation rythmique. En revanche, la faible habileté à se synchroniser au beat était généralisée à la difficulté à se synchroniser par le chant.

Dans l'ensemble, les résultats de la thèse suggèrent que, dans les cas d'amusie congénitale étudiés, la capacité à produire un mouvement synchronisé à la musique émergerait d'un mécanisme plus général d'entrainement permettant la synchronisation à d'autres stimuli auditifs, y compris la parole.

Mots-clés : amusie congénitale, pulsation rythmique, synchronisation sensorimotrice, musique, chant, parole

Abstract

Music naturally compels most individuals to engage in rhythmic behaviors. We can think of someone tapping his foot or nodding his head to the beat of music. Synchronizing a movement to the beat may seem simple at first, but it is a complex behavior. At least, someone needs to be able to extract the beat and predict the timing of upcoming beats. This ability to couple movement and music could be achieved through entrainment. The specificity of beat-based entrainment to music is, however, debated. This thesis aimed to assess sensorimotor synchronization to music, singing, and speech in congenital amusia, in order to test the specificity of entrainment mechanism to musical beat.

Congenital amusia refers to a neurodevelopmental disorder that can affect, depending on cases, pitch perception (pitch-based amusia) or beat perception (beat-based amusia). In the first study, individuals with pitch-based amusia were tested on their ability to perceive and to produce musical beat. The results indicated that about fifty percent of pitch-based amusic participants had associated deficits in beat production and beat perception. Still, cases of dissociation, with spared ability to synchronize to and perceive musical beat, were also identified. Therefore, findings highlight a connection between melody and rhythm processing in most cases of pitch-based amusia, although it is possible for these domains to be selectively impaired. Concurrently, cases of beat-based amusia, included in Study 2 and Study 3, had a specific impairment in their ability to synchronize a simple movement with musical beat, while performing within the normal range on a standardized measure of pitch perception.

Study 2 and Study 3 had the objective to test hypotheses regarding the domain specificity of beat-based entrainment. In Study 2, entrainment to speech and music was compared in participants with beat-based amusia. Participants had to align taps to the perceived regularity in the rhythm of naturally spoken, regularly spoken (similar to rap music), and sung sentences. It was found that the amusic participants synchronized less accurately in all conditions. This result suggests that a general entrainment mechanism could maybe drive sensorimotor synchronisation to speech and music.

In Study 3, we assessed the hypothesis according to which vocal imitation and beat-based synchronisation could share a common sensorimotor coupling mechanism. Here, participants had to sing a familiar song from memory, after hearing a model, with the model, and with a metronome. First, results from this study indicate that vocal-pitch abilities are similar to the general population in beat-based amusia, when singing from memory. Cases of poor-pitch singing could still be identified among amusic participants. However, no clear association between synchronization to beat and vocal-pitch abilities was found. Nonetheless, results from synchronous singing and singing with a metronome mirrored results from tapping tasks, indicating a lowered ability to synchronize with the beat across contexts in beat-based amusia.

Overall, by investigating sensorimotor synchronization to music and singing in congenital amusia, this thesis work provides new evidence that sensorimotor synchronization to musical beat may be built on a domain-general entrainment mechanism that could be involved with auditory stimuli that are not beat-based, like speech.

Keywords: congenital amusia, beat, sensorimotor synchronization, music, singing, speech

Table des matières

Liste des tableaux

Chapitre II : Méthodologie et résultats

Liste des figures

Chapitre II : Méthodologie et résultats

Liste des sigles et abréviations

ANOVA	Analyse de variance / *Analysis of variance*
BAT	*Beat Alignment Test*
BPM	Beat par minute / *Beat per minute*
CV	Coefficient de variation / *Coefficient of variation*
d’	d-prime
EEG	Électroencéphalographie
Hz	Hertz
IBI	*Inter-beat interval*
IFG	Gyrus frontal inférieur / *Inferior frontal gyrus*
IOI	*Inter-onset interval*
ITI	*Inter-tap interval*
IVI	*Inter-vocalic interval*
M	*Mean*
MBEA	*Montreal Battery for Evaluation of Amusia*
ms	Millisecondes / *Milliseconds*
SD	*Standard deviation*
SE	*Standard error*
sec	Secondes
SMA	Aire motrice supplémentaire / *Supplemental motor area*
STG	Gyrus temporal supérieur / *Superior temporal gyrus*
VL	*Vector Length*
V-nPVI	*Vocalic Normalized Pairwise Variability Index*
WAIS-III	*Wechsler Adult Intelligence Scale*, 3e édition

À ma famille et mes proches, pour votre soutien continu

Remerciements

Je tiens d'abord à remercier Isabelle Peretz, ma directrice de recherche. Je peux difficilement exprimer toute ma reconnaissance pour le soutien et la confiance que vous avez su me prodiguer au cours des dernières années. Mon parcours ne s'est pas fait sans embuches, mais vous avez toujours su faire preuve de patience et d'écoute, tout en me fournissant l'encadrement nécessaire pour donner le meilleur de moi-même. L'accomplissement de cette thèse n'aurait été possible sans votre soutien et votre engagement. La passion que vous démontrez pour le domaine de la recherche est réellement inspirante et je considère avoir eu une chance incroyable de pouvoir profiter de votre expérience et de votre expertise. Ce fut un plaisir de partager ces dernières années avec vous. Sincèrement, merci pour tout, Isabelle.

Merci à Mihaela Felezeu, pour son aide avec le recrutement et la collecte de données, mais aussi pour toute sa douceur et son écoute. Merci à Nathanäel Lecaudé et Bernard Bouchard pour leur aide précieuse dans la résolution de problèmes techniques. À cet égard, merci également à Mailis Rodrigues pour son aide dans la programmation informatique des tâches incluses dans le projet. Je tiens aussi à remercier les assistants de recherche qui ont contribué à la tâche fastidieuse d'analyse du chant. Un mot également pour les étudiants rencontrés au BRAMS qui ont su contribuer à des discussions inspirantes et me fournir des conseils judicieux, entre autres, Dominique Vuvan, Benjamin Zendel et Simon Rigoulot. Je souhaite remercier plus particulièrement Isabelle Royal et Sébastien Paquette pour leur support, leurs conseils et leurs encouragements tout au long de mon parcours au BRAMS. Un grand merci à Pauline Tranchant pour m'avoir guidée dans l'apprentissage de l'analyse de données de synchronisation et pour m'avoir fourni un regard critique et des conseils précieux à cet égard. Pauline, je me considère chanceuse d'avoir pu partager l'aventure '*beat-deaf*' avec toi. Merci pour ton soutien dans les dernières années. Les nombreux weekends de travail dans les cafés resteront gravés dans ma mémoire. J'aimerais aussi remercier Caroline Palmer pour sa collaboration au deuxième article inclus dans la thèse.

Merci également aux personnes qui m'ont accompagnée dans mon parcours doctoral, entre autres les membres de ma cohorte, Jessica, Catherine, Anne, Émilie & Emilie, Mathieu et Vincent. Cette expérience n'aurait pas été la même sans vous. Une pensée également pour mes superviseures cliniques, Christiane Flessas et Hélène Roy, qui ont su se montrer particulièrement encourageantes et compréhensives dans les dernières semaines entourant le dépôt de la thèse.

Je voudrais également remercier les proches qui m'ont permis de garder un équilibre tout au long de l'accomplissement du doctorat, tout en me fournissant un important soutien. Merci donc à Stéphanie, Catherine et Éva. Plus spécialement, merci à Eugénie et Isabelle d'avoir su agrémenter les dernières années d'expériences inoubliables. Votre amitié, nos fous rires et vos encouragements m'ont été et me sont très précieux.

Je tiens aussi à remercier le plus sincèrement les membres de ma famille, et plus particulièrement mes parents, pour leur compréhension, leurs encouragements sans égal et leur bienveillance. Merci d'avoir pris part à cette aventure avec moi. Je ne peux mesurer l'importance qu'a eu votre soutien dans cet accomplissement. Merci également à Jean-Pierre et Louise, mes beaux-parents. Vos encouragements ont été très appréciés.

Enfin, merci à Olivier d'avoir partagé ces neuf dernières années avec moi. Tu as été mon pilier à travers tous les hauts et les bas. Je ne pourrai assez te remercier pour tout le support et le réconfort que tu as su m'apporter. J'ai hâte de pouvoir entamer le prochain chapitre de ma vie avec toi.

Chapitre I : Introduction

Avant-propos

Dès le plus jeune âge et quelle que soit la culture, la grande majorité des gens ont tendance à initier naturellement des mouvements rythmiques à l'écoute de la musique (Dowling, 1999 ; Mithen, 2006 ; Zentner & Eerola, 2010). On peut, par exemple, s'imaginer quelqu'un taper des mains, taper du pied ou hocher la tête au rythme de la musique. Cette synchronisation au rythme musical dépend de la capacité à pouvoir percevoir une pulsation rythmique, ou *beat* (le terme anglais plus couramment utilisé dans la littérature) dans la musique. Bien que cette activité puisse paraitre simple de prime abord, la capacité à percevoir et à se synchroniser au beat repose sur l'interaction complexe de plusieurs processus cognitifs. Un mécanisme central serait celui « d'entrainement » (*entrainment* de l'anglais). Le concept d'entrainement réfère ici à la synchronisation de rythmes internes, souvent représentés sous forme d'oscillateurs, aux régularités d'un rythme externe, comme le beat de la musique (Large & Jones, 1999).

Le mécanisme d'entrainement permet d'anticiper la survenue d'un évènement récurrent dans une séquence rythmique et ainsi planifier une action motrice synchronisée à cet évènement, qui dans le cas de la musique correspond le plus souvent au beat. Certains auteurs proposent que par la présence d'un beat, la musique puisse bénéficier d'un mécanisme d'entrainement unique favorisant la synchronisation sensorimotrice (Patel, 2008 ; Patel & Iversen, 2014). Par ailleurs, la capacité à pouvoir suivre et décoder la structure temporelle d'un rythme est également importante dans d'autres contextes où le signal est moins périodique, comme lors d'une conversation (Phillips-Silver, Aktipis, & A. Bryant, 2010 ; Wilson & Wilson, 2005). De ce fait, nombreux sont les chercheurs qui suggèrent plutôt l'existence d'un mécanisme global d'entrainement applicable à des signaux auditifs plus réguliers, comme la musique, mais aussi à des signaux moins périodiques, comme la parole (e.g., Cummins, 2009b; Cummins & Port, 1998; Giraud & Poeppel, 2012; Goswami, 2012; O'Dell & Nieminen, 1999; Peelle & Davis, 2012; Port, 2003; Wilson & Wilson, 2005). Patel (2006) avance également la possibilité que le réseau cérébral de couplage sensorimoteur sous-tendant l'entrainement au beat puisse, selon une perspective évolutionniste, être partagé avec la capacité d'imitation vocale, qui demande aussi un haut niveau de précision sur le plan du couplage auditivo-moteur. Afin de mieux cerner la spécificité des processus d'entrainement au beat, les travaux de cette thèse proposent de s'intéresser à la synchronisation sensorimotrice dans l'amusie congénitale.

L'amusie congénitale est un trouble neurodéveloppemental de la cognition musicale, qui se présente malgré une intelligence normale et en l'absence de trouble auditif ou neurologique (Ayotte, Peretz, & Hyde, 2002 ; Peretz & Vuvan, 2017). Dans sa forme la plus étudiée, l'amusie congénitale se caractérise par une incapacité à discriminer les fines variations de hauteurs (Ayotte et al., 2002 ; Vuvan, Nunes-Silva, & Peretz, 2015). Une nouvelle forme d'amusie congénitale, qui se définit par une incapacité à synchroniser un mouvement simple au beat de la musique, en l'absence de difficulté à percevoir les variations de hauteurs, a plus récemment fait l'objet d'études empiriques (Phillips-Silver et al., 2011 ; Sowiński & Dalla Bella, 2013). Une meilleure caractérisation des aptitudes de synchronisation sensorimotrice dans l'amusie congénitale offre l'occasion d'évaluer la spécificité du mécanisme d'entrainement au beat en mesurant l'impact du trouble dans des domaines connexes, comme la synchronisation à la parole et le chant.

L'objectif général de la thèse est de mieux déterminer la spécificité des processus d'entrainement au beat musical, par l'étude de la synchronisation sensorimotrice à la musique, au chant et à la parole dans l'amusie congénitale. Dans un premier temps, la capacité à percevoir le beat et s'y synchroniser a été étayée chez un groupe d'individus présentant une amusie congénitale affectant la perception des hauteurs de notes. L'objectif était de tester la séparabilité entre le traitement de la mélodie et du rythme dans cette population. Dans un deuxième temps, l'hypothèse d'un mécanisme général d'entrainement applicable à la musique et à la parole a été investiguée en comparant l'habileté de sujets amusiques et non-amusiques à synchroniser une réponse motrice à des extraits de chant et de parole. Dans un troisième temps, l'hypothèse selon laquelle l'aptitude à se synchroniser au beat recruterait un réseau de couplage sensorimoteur commun à la capacité d'imitation vocale a été mesurée par l'étude du chant des amusiques ayant un trouble de synchronisation au beat.

La thèse est organisée en trois chapitres. Le premier chapitre consiste en une introduction faisant une revue des éléments pertinents de la littérature scientifique. Cette section s'attarde d'abord à définir différents concepts de bases liés à la synchronisation à la musique, puis offre une description des mécanismes de perception et de synchronisation au beat selon les modèles théoriques actuels et les corrélats neuronaux y étant associés. S'en suit une description de l'amusie congénitale. Pour continuer, les arguments en lien avec la séparabilité du traitement du beat et de la mélodie, la spécificité du mécanisme d'entrainement à la musique comparativement à la parole, ainsi que l'hypothèse d'un lien entre imitation vocale et synchronisation au beat sont développés. Ce chapitre se complète par l'exposition des objectifs et des hypothèses des trois études incluses dans la thèse et qui composent le deuxième chapitre. Enfin, le troisième chapitre offre une discussion générale des résultats des travaux de recherche de cette thèse et ouvre sur des perspectives de recherche futures.

1.1 Concepts de bases de synchronisation à la musique

La structure temporelle de la musique peut se décomposer en différents aspects, les principaux étant le rythme, le beat et la métrique. Nous étayerons ici ces différents concepts, tels qu'ils ont été définis dans le cadre des travaux de recherche de la thèse.

Rythme. Le rythme désigne en fait directement l'organisation temporelle des évènements sonores constituant une séquence auditive (Levitin, Grahn, & London, 2018 ; McAuley, 2010 ; Patel, 2008). Il se définit comme un motif d'intervalles de temps, ou de durées, qui peut être périodique ou non (Large & Jones, 1999 ; Patel, 2008). Le point central étant ici la relation entre les intervalles de temps, le rythme d'une séquence auditive peut demeurer stable que celle-ci soit lente ou rapide. Si l'on s'imagine par exemple une chanson connue, disons *Joyeux anniversaire*, le rythme de la mélodie demeure le même que celle-ci soit chantée lentement ou rapidement. Par contre, sa vitesse en déterminera le beat.

Beat. Le beat, ou la pulsation, se définit comme la récurrence d'un intervalle de temps périodique, ou quasi-périodique, sous-tendant l'organisation de la structure rythmique (Lerdahl & Jackendoff, 1983). Il importe de mentionner que le beat est un percept, c'est-à-dire qu'il s'agit d'une propriété émergente pouvant être perçue dans les rythmes périodiques (Large, 2008). En effet, un beat peut être perçu en l'absence d'une correspondance systématique d'un à un entre le signal acoustique sonore et l'occurrence des beats, soit les points discrets dans le temps marquant le beat (Large, 2008 ; Patel, 2008). Ceci a été démontré par le phénomène de *missing pulse*, qui désigne la capacité à extraire un beat dans une séquence rythmique périodique, en l'absence d'évènements sonores correspondant aux temps des beats (Chapin et al., 2010 ; Drake, Jones, & Baruch, 2000 ; Large, Herrera, & Velasco, 2015 ; Repp, Iversen, & Patel, 2008). Considérant la nature perceptuelle du beat, des facteurs comme l'expérience musicale et le tempo spontané peuvent influencer également la période à laquelle le beat est perçu d'un individu à l'autre (Drake, Penel, & Bigand, 2000 ; Iversen, Patel, & Ohgushi, 2008 ; Martens, 2011). Le tempo spontané correspond ici à la période à laquelle une personne tape naturellement lorsqu'on lui demande de taper du doigt de façon régulière en l'absence d'un stimulus régulier auquel se synchroniser (McAuley, 2010 ; McAuley, Jones, Holub, Johnston, & Miller, 2006). Le tempo, plus spécifiquement, désigne la vitesse à laquelle les évènements d'une séquence se déploient dans le temps (Levitin et al., 2018 ; McAuley, 2010). On décrit souvent le tempo en beats par minute (BPM). Ainsi, un tempo de 90 BPM correspond à un tempo plutôt lent, alors qu'un tempo de 200 BPM serait plutôt rapide.

Métrique. La notion de métrique désigne le regroupement des beats en temps forts et temps faibles (Lerdahl & Jackendoff, 1983 ; Patel, 2008). Par l'alternance des temps forts et des temps faibles, on peut distinguer, par exemple, la structure ternaire d'une valse (UN, deux, trois, UN, deux, trois) de la structure binaire d'une marche (UN, deux, UN, deux). Les temps forts dans la métrique sont aussi appelés des *accents* (Large, 2008 ; Patel, 2008). Les accents perçus aux temps forts sont le plus souvent marqués par des changements dans l'intensité, la hauteur, ou la durée des sons, mais pas nécessairement (Hannon, Snyder, Eerola, & Krumhansl, 2004 ; Lerdahl & Jackendoff, 1983 ; McKinney & Moelants, 2006 ; Palmer & Krumhansl, 1990). En effet, des résultats de recherche montrent que selon la métrique perçue, la position des accents peut changer (Iversen, Repp, & Patel, 2009). Par ailleurs, la présence d'accents crée une hiérarchie dans l'organisation temporelle de la musique, avec le beat au niveau de base et les temps forts, ou accents, au niveau hiérarchique supérieur (Lerdahl & Jackendoff, 1983 ; London, 2002 ; Palmer & Krumhansl, 1990).

Sur la base d'études menées auprès de patients cérébrolésés et d'études en neuroimagerie, il a été démontré que la perception du rythme, d'une part, et la perception du beat et de la métrique, d'autres parts, peut être distinguée, tant sur le plan des processus cognitifs que des bases cérébrales impliquées dans chacun des cas (p. ex., Di Pietro, Laganaro, Leemann, & Schnider, 2004 ; Grahn & Brett, 2007 ; Liégeois-Chauvel, Peretz, Babaï, Laguitton, & Chauvel, 1998 ; Teki, Grube, Kumar, & Griffiths, 2011 ; Thaut, 2003 ; Tierney & Kraus, 2015 ; Wilson, Pressing, & Wales, 2002). Dans la section qui suit, nous nous pencherons plus spécifiquement sur les mécanismes d'entrainement et de synchronisation au beat, le principal sujet d'intérêt de la thèse.

1.2 Mécanismes de synchronisation et d'entrainement au beat

La synchronisation sensorimotrice correspond à la coordination temporelle d'une action avec un évènement externe prévisible (Repp, 2005). Les études montrent que la majorité des gens parviennent à se synchroniser de façon très précise au beat ; dans des tâches utilisant des tapes du doigt (*tapping*), l'action motrice et le beat ne sont souvent séparés que de quelques dizaines de millisecondes (Repp, 2005 ; Repp & Su, 2013 ; Van Der Steen & Keller, 2013), et ce pour des tempi pouvant aller de 94 à 176 BPM (McAuley, 2010 ; McAuley et al., 2006 ; Repp, 2003, 2005). Une synchronisation précise dépendrait de la capacité à : 1) prédire ou anticiper l'occurrence de l'évènement auquel se synchroniser ; 2) produire une action motrice synchronisée, et 3) corriger l'erreur de synchronisation afin de maintenir un couplage précis entre l'action et l'évènement anticipé (Phillips-Silver et al., 2010 ; Van Der Steen & Keller, 2013).

En lien avec la première composante de la synchronisation sensorimotrice, l'anticipation des beats se démontre par la mesure de l'asynchronie moyenne, soit l'écart temporel moyen entre la réponse motrice et le beat. Une asynchronie moyenne inférieure au simple temps de réaction associé au mouvement indique l'anticipation de l'occurrence du beat. Dans le cas de stimuli simples et parfaitement périodiques, comme un métronome par exemple, on observe même que l'action précède typiquement le beat, ce qui est appelé l'asynchronie moyenne négative (Repp, 2005 ; Repp & Su, 2013). L'aptitude à pouvoir prédire le beat découlerait d'un mécanisme d'entrainement, qui se définit comme la tendance à prédire et à synchroniser une réponse avec les régularités temporelles d'un stimulus rythmique externe (Large et al., 2015 ; Large & Snyder, 2009). Actuellement, le modèle prédominant de l'entrainement se fonde sur la théorie de l'attention dynamique (DAT ; Jones, 1976, 1987 ; Jones & Boltz, 1989), selon laquelle des oscillateurs neuraux internes, ayant leur propre période, s'aligneraient avec le rythme d'un signal externe (Large & Jones, 1999).

1.2.1 Entrainement et oscillateurs internes

Selon la théorie de l'attention dynamique, la synchronisation d'oscillateurs internes avec le rythme externe permettrait de centrer l'énergie attentionnelle sur les évènements récurrents dans le signal rythmique, pouvant être ainsi anticipés et faciliter la coordination d'une réponse à ces évènements (Large & Jones, 1999). Cependant, le beat de la musique est rarement parfaitement isochrone (c.-à-d. une période parfaitement identique d'un intervalle à un autre du beat) et, malgré tout, nous parvenons à nous synchroniser aux rythmes présentant de légères déviations de périodicité (Drake, Penel, et al., 2000 ; Large & Palmer, 2002 ; Palmer, 1997). Il a ainsi été proposé que les oscillateurs internes pourraient adapter leur période et leur phase pour permettre la synchronisation à des rythmes quasi-périodiques (Large, 2008 ; Large & Palmer, 2002 ; Large & Snyder, 2009). Nous sommes, par ailleurs, également en mesure de nous synchroniser à des rythmes variant en complexité (Drake, Penel, et al., 2000 ; Large et al., 2015 ; Large & Palmer, 2002) et ne présentant pas systématiquement une correspondance d'un à un entre les beats et le signal acoustique. Pour expliquer cette capacité à se synchroniser au beat, il a été proposé que des populations d'oscillateurs, ayant une période de référence leur étant propre, se synchroniseraient à diverses périodicités dans le signal. L'interaction non linéaire entre ces populations d'oscillateurs, organisées de façon hiérarchique, permettrait de générer une réponse oscillatoire à des périodicités émergentes. Ces périodicités émergentes correspondraient par exemple aux harmoniques des périodicités présentes dans le signal acoustique, expliquant ainsi la possibilité de percevoir un beat dans la musique en l'absence de signal acoustique correspondant aux beats (Large, 2008 ; Large & Palmer, 2002 ; Large & Snyder, 2009).

Ces modèles de l'entrainement basés sur l'activité d'oscillateurs internes sont appuyés par des études en électroencéphalographie (EEG), montrant une plus grande énergie de la réponse électrique du cerveau à la fréquence correspondant à celle du beat d'une séquence rythmique entendue, et parfois aux harmoniques de cette fréquence également (Nozaradan, 2014; Nozaradan, Peretz, & Keller, 2016; Nozaradan, Peretz, Missal, & Mouraux, 2011; Nozaradan, Peretz, & Mouraux, 2012; Stupacher, Wood, & Witte, 2017; Tierney & Kraus, 2014b). Reflétant la perception du beat et non une simple copie du signal acoustique, cette réponse peut également être mesurée en l'absence de signal acoustique au moment des beats (Chapin et al., 2010 ; Large et al., 2015 ; Tal et al., 2017). Ainsi, selon ces modèles de l'entrainement, le processus de couplage entre des oscillateurs neuronaux et le signal acoustique serait suffisant pour expliquer la capacité à percevoir le beat. Par ailleurs, d'autres auteurs suggèrent que l'entrainement d'oscillateur ne permet pas d'expliquer entièrement le processus de perception du beat. Ces auteurs mettent de l'avant des processus descendants (*top-down* de l'anglais) qui proviendraient principalement des aires cérébrales frontales et motrices (Grahn & Rowe, 2009 ; Iversen et al., 2009 ; Patel & Iversen, 2014 ; Van Der Steen & Keller, 2013 ; Zatorre, Chen, & Penhune, 2007).

1.2.2 L'implication du système moteur dans la prédiction du beat

Plusieurs données en neuropsychologie et en neuroimagerie montrent que les régions motrices du cerveau seraient non seulement impliquées dans la synchronisation au beat, mais aussi dans des tâches de perception du beat n'impliquant pas de réponse motrice (Chen, Penhune, & Zatorre, 2008 a; Grahn & Brett, 2007 ; Grahn & Rowe, 2009 ; Grahn & Rowe, 2013 ; Kung, Chen, Zatorre, & Penhune, 2013). Le modèle ASAP (*Action Simulation for Auditory Prediction*; Patel & Iversen, 2014) va même jusqu'à proposer un rôle causal du système de planification motrice dans la perception et la prédiction du beat. Selon ce modèle, le système de planification motrice utiliserait la simulation d'une réponse motrice périodique aux évènements acoustiques marquant le rythme d'un signal de façon à entrainer l'activité neuronale des régions motrices au beat. Ce motif d'activité neuronale serait ensuite transmis des aires de planification motrice vers les aires auditives, permettant ainsi d'affiner la prédiction de l'occurrence des beats. L'idée centrale est que la perception du beat dépendrait principalement d'une communication étroite entre les aires auditives et de planification motrice. Des études d'imagerie suggèrent, en ce sens, que le degré de périodicité d'un rythme influence le couplage entre les aires auditives et les régions de planification motrice par une voie dorsale via le cortex pariétal (*dorsal auditory pathway*; Chen, Penhune, & Zatorre, 2008 b ; Chen, Zatorre, & Penhune, 2006 ; Grahn & Rowe, 2009 ; Patel, Iversen, Chen, & Repp, 2005). Grahn & Rowe (2013) proposent également deux systèmes distincts pour la détection et la prédiction du beat. Le système de détection du beat impliquerait un vaste réseau d'aires cérébrales incluant le gyrus temporal supérieur (STG), le lobule pariétal inférieur, le cortex prémoteur, et le cervelet, alors que le système de prédiction impliquerait plus spécifiquement l'aire motrice supplémentaire SMA et les ganglions de la base, dont plus particulièrement le putamen.

1.2.3 Mécanismes de synchronisation au beat

Au-delà de la perception du beat, la possibilité de synchroniser un mouvement au beat de la musique implique un système plus complexe de couplage sensorimoteur. Les principales aires cérébrales associées à la synchronisation au beat sont le gyrus temporal supérieur (STG), le cortex préfrontal, le cortex prémoteur, l'aire motrice supplémentaire (SMA), ainsi que les ganglions de la base et le cervelet (pour des méta-analyses voir Chauvigné, Gitau, & Brown, 2014 ; Witt, Laird, & Meyerand, 2008 ; et pour des revues récentes Haegens & Zion Golumbic, 2018 ; Rajendran, Teki, & Schnupp, 2017).

Le maintien d'une synchronisation précise au beat de la musique dépend en grande partie de l'aptitude à pouvoir corriger sa réponse à la suite d'une perturbation de la synchronisation. Des variations normales dans la précision de la réponse motrice et dans la périodicité du signal auditif externe se produisent lors de la synchronisation à un stimulus. Ainsi, afin d'éviter d'accumuler de l'erreur dans la synchronisation et éventuellement ne plus être synchronisé du tout, le système doit être en mesure de s'ajuster et corriger le décalage entre le mouvement et le stimulus externe. Deux types d'erreurs sont possibles lors de la synchronisation : une erreur de période, donc d'intervalle, et une erreur de phase, impliquant le maintien de la période, mais un décalage entre le beat et la réponse motrice. Chaque type d'erreurs est géré par un mécanisme d'ajustement différent (Repp, 2005 ; Repp & Su, 2013 ; Van Der Steen & Keller, 2013). Le mécanisme de correction de la phase est rapide et automatique (Repp, 2004 ; Repp, London, & Keller, 2008). En revanche, le mécanisme de correction de la période implique de devoir porter attention activement au signal afin de pouvoir détecter le changement de période et ajuster la période de la réponse, demandant ainsi un contrôle actif (Repp & Keller, 2004, 2008 ; Van Der Steen & Keller, 2013). Sur la base d'une boucle de rétroaction intégrant l'asynchronie entre la production motrice et les beats, les personnes peuvent adapter la coordination de la production motrice suivante afin de compenser pour l'erreur et ainsi maintenir la synchronisation (Repp, 2005). Le cervelet jouerait un rôle essentiel dans l'ajustement de la réponse de synchronisation (Rao, Mayer, & Harrington, 2001 ; Schwartze, Keller, & Kotz, 2016 ; Teki et al., 2011 ; Zatorre et al., 2007).

En somme, la synchronisation sensorimotrice avec le beat de la musique dépend principalement de l'habileté à percevoir et à prédire l'occurrence des beats, à synchroniser une action motrice et à pouvoir corriger l'erreur de synchronisation. Le dysfonctionnement de l'un ou l'autre de ces processus pourrait causer, en théorie, un trouble de la synchronisation au beat.

1.3 L'amusie congénitale

L'amusie congénitale est un trouble neurodéveloppemental de la cognition musicale. Le terme *congénital* réfère à l'aspect développemental du trouble et implique que les personnes atteintes présenteraient des difficultés depuis l'enfance ne pouvant être associées à un dommage cérébral connu, différant ainsi des amusies acquises. La prévalence de l'amusie congénitale est estimée à environ 1,5 % de la population (Peretz & Vuvan, 2017). La forme la plus étudiée d'amusie congénitale concerne l'incapacité à percevoir les fines variations de hauteurs dans une mélodie (Ayotte et al., 2002). Une autre forme d'amusie congénitale, plus nouvellement étudiée, est associée plutôt à une inaptitude à extraire le beat musical et à se synchroniser au beat (Phillips-Silver et al., 2011 ; Sowiński & Dalla Bella, 2013). Dans les deux cas, le trouble ne peut être attribuable à un retard mental, un trouble auditif, un trouble neurologique, ou un manque d'exposition à la musique (Ayotte et al., 2002 ; Phillips-Silver et al., 2011). Ainsi, des caractéristiques bien différentes sont propres à ces deux formes d'amusie congénitale (à laquelle nous réfèrerons par *amusie* à partir d'ici afin d'alléger le texte), qui sont dissociées sur la base d'outils standardisés décrits dans la section suivante.

1.3.1 Outils d'identification de l'amusie

MBEA. La Batterie de Montréal d'évaluation de l'amusie (MBEA ; Peretz, Champod, & Hyde, 2003) est l'outil normé le plus communément utilisé pour identifier l'amusie. La batterie se compose de six épreuves s'attardant à des aspects uniques du traitement musical, suivant le modèle de Peretz & Coltheart (2003). Chaque épreuve débute par des exemples pratiques, puis trente essais tests sont présentés. Les trois premières épreuves de la MBEA portent sur le traitement de la mélodie. Pour les quatre premières épreuves, le participant entend deux mélodies présentées l'une à la suite de l'autre et doit détecter si la seconde mélodie présente un changement ou non en comparaison de la première mélodie. La moitié des essais contient un changement de hauteur pour une des notes de la mélodie. Dans la première épreuve, le changement de hauteur implique l'introduction d'une note hors tonalité (alors que le patron mélodique, ou le contour de la mélodie, est préservé). La deuxième épreuve demande de détecter un changement dans le contour de la mélodie (p. ex. : do-ré vers do-si), sans changement de la tonalité. Dans la troisième épreuve, le changement se situe au niveau de l'intervalle de hauteur entre deux notes, sans une modification du contour de la mélodie ni sa tonalité (p. ex. : do-mi vers do-fa).

Les deux épreuves suivantes portent sur l'aspect rythmique de la mélodie. Ainsi, dans la quatrième épreuve, il s'agit de comparer encore des paires de mélodies, alors que, cette fois, la seconde mélodie diffère dans la moitié des cas par un changement de durée entre deux notes consécutives. Cette tâche implique donc d'encoder les intervalles de temps entre les notes de la mélodie et peut être réussie sans avoir recours à la perception du beat (Tranchant & Vuvan, 2015). Dans la cinquième épreuve, une seule mélodie est présentée à chaque essai et le participant doit juger si celle-ci présente la métrique d'une valse ou d'une marche. Cette épreuve se rapproche plus de la capacité à pouvoir extraire le beat, bien qu'il soit possible que des indices acoustiques dans les mélodies permettent de réussir la tâche malgré une pauvre perception du beat (Tranchant & Vuvan, 2015).

Enfin, la sixième épreuve est une tâche de mémoire implicite qui teste la reconnaissance musicale. Trente mélodies sont présentées et le participant doit identifier les mélodies entendues précédemment au cours du test. La moitié des mélodies sont nouvelles alors que l'autre moitié correspond à des mélodies déjà entendues.

En général, le score obtenu aux trois épreuves mélodiques est combiné pour former un score global qui sert de critère pour identifier les cas d'amusie affectant le traitement mélodique (Peretz et al., 2003 ; Vuvan et al., 2017). Certaines études ont également utilisé les épreuves de la MBEA portant sur le traitement du rythme pour identifier des cas d'amusie touchant la perception du beat. Cependant, Tranchant & Vuvan (2015) argumentent que ces épreuves ne correspondent pas directement à la capacité à traiter le beat et devraient donc être évitées comme outil de dépistage pour identifier ces cas.

Le test en ligne d'identification de l'amusie. Ce test est un dérivé de la MBEA, initialement développé pour permettre le dépistage de l'amusie. Il comprend trois épreuves. La première épreuve est identique à la première épreuve de la MBEA. Les seconde et troisième épreuves consistent à écouter une mélodie et à indiquer si celle-ci contient, respectivement, un intervalle de temps déviant dans le rythme ou une fausse note. Ces deux épreuves comprennent 24 essais chacune. En fonction des résultats d'un échantillon normé de 14 686 personnes, Peretz & Vuvan (2017) proposent que des scores inférieurs à des seuils établis aux deux épreuves mélodiques du test en ligne soient une forte indication de la présence d'une amusie touchant la perception des hauteurs.

Le *Beat Alignment Test* (BAT). Le BAT est une batterie de tests ayant été développée pour mesurer plus spécifiquement la perception et la synchronisation au beat (Iversen & Patel, 2008). Dans la tâche de perception du beat, le participant entend une mélodie sur laquelle une séquence de métronome est ajoutée. La tâche du participant consiste à indiquer si le métronome se trouve sur le beat de l'extrait musical ou non. Le métronome peut soit être correctement aligné avec le beat, présenter une erreur de période (tempo du métronome plus lent ou plus rapide que le beat) ou une erreur de phase (bon tempo, mais décalage avec le beat, soit à l'avance ou en retard par rapport au beat). Dans une autre tâche, de synchronisation cette fois, les participants doivent aligner des tapes sur le beat qu'ils perçoivent dans des séquences rythmiques. D'autres versions de ce test ont depuis été développées (p. ex., Fujii & Schlaug, 2013). Dalla Bella et collaborateurs (2017) ont entre autres développé une batterie comprenant plusieurs épreuves reliées à la perception et à la synchronisation au beat. La batterie inclut également une épreuve de comparaison de durée (dire si deux intervalles de temps sont de même longueur ou non) et une épreuve de détection d'anisochronie (c.-à-d. une déviation de la durée d'un intervalle de temps dans une séquence autrement périodique), afin de distinguer entre différents troubles possibles associés au traitement de l'aspect rythmique de la musique.

Pour les travaux de cette thèse, une version adaptée du BAT, développée dans notre laboratoire de recherche, a été utilisée (Tranchant et al., en préparation). La tâche de perception de cette version du BAT comprend 80 essais, dont la moitié sont présentés avec un métronome sur le beat et l'autre moitié avec un métronome ne correspondant pas au beat. Pour les essais où le métronome ne correspond pas au beat, l'erreur en est soit une de période ou de phase. L'erreur de période implique une augmentation ou une diminution de la période du métronome de 5 % par rapport à la période du beat. L'erreur de phase correspond à un décalage entre le son du métronome et le beat de 15 % de la période, le son du métronome étant soit avant ou après le beat. La tâche inclut dix extraits musicaux de genres et tempi variant de 82 à 170 BPM (voir Annexe 1). Chaque extrait est présenté dans huit conditions, dont quatre sur le beat et quatre hors-beat. Dans la tâche de synchronisation au beat, les participants entendent les mêmes dix extraits musicaux, sans métronome. Chaque extrait est présenté dans un ordre aléatoire et répété dans deux blocs séparés, pour un total de 20 essais. La tâche consiste à taper du doigt de façon à se synchroniser avec le beat perçu dans chaque extrait musical. La tâche de synchronisation est toujours présentée avant la tâche de perception et ce afin d'éviter un effet d'indiçage. Un article présentant les données aux épreuves de la M-BAT pour un groupe de référence de quarante adultes sans trouble neurologique ou neurodéveloppemental est actuellement en préparation (Tranchant et al., en préparation). Dans le cadre de la thèse, les participants ont été inclus dans le groupe d'amusiques présentant un trouble de la synchronisation au beat sur la base de leur performance à la tâche de synchronisation de la M-BAT. Une synchronisation réussie à moins de 15 essais était considérée comme l'indication d'un trouble, des données pilotes indiquant qu'en moyenne les gens réussissent à se synchroniser à tous les essais, la performance la plus faible chez le groupe de référence étant de 17 essais réussis sur 20. Le succès de la synchronisation est évalué à partir du test de Rayleigh ($p < .05$), un test de statistique circulaire permettant de déterminer la constance du couplage entre la période des tapes et la période du beat. La performance à la tâche de perception du beat n'a pas été utilisée de façon critériée ici, mais la plupart des participants ayant un trouble de la synchronisation au beat ont également montré une piètre performance à cette tâche (Article 2).

1.3.2 L'amusie touchant le traitement mélodique

L'amusie touchant la perception des hauteurs est la forme la plus étudiée d'amusie. Ce qui caractérise principalement ce trouble est l'incapacité à détecter les fines variations de hauteurs correspondant aux demi-tons utilisés pour construire la tonalité dans le système musical occidental (Ayotte et al., 2002 ; Vuvan et al., 2015). Ces personnes peuvent difficilement identifier qu'une personne chante faux, y compris eux-mêmes, reconnaitre une mélodie connue sans la présence des paroles, discriminer des mélodies sur la base des notes et apprendre des mélodies (Ayotte et al., 2002 ; Peretz, 2016). La sévérité du trouble de discrimination des hauteurs serait par contre variable d'un cas à un autre (Hyde & Peretz, 2004 ; Vuvan et al., 2015). Par ailleurs, environ la moitié des individus présentant ce type d'amusie aurait également une difficulté à traiter le rythme de la musique, un résultat qui sera abordé plus en détail ultérieurement (Ayotte et al., 2002 ; Peretz et al., 2003).

Des études de potentiels évoqués en EEG ont montré que le trouble des personnes amusiques proviendrait de l'incapacité à se représenter les hauteurs, alors que le traitement automatique de bas niveau de discrimination de hauteurs serait fonctionnel (Moreau, Jolicœur, & Peretz, 2009, 2013 ; Peretz, Brattico, Järvenpää, & Tervaniemi, 2009 ; Zendel, Lagrois, Robitaille, & Peretz, 2015). Sur le plan neuroanatomique, le trouble proviendrait d'une connectivité anormale entre le cortex auditif droit, plus précisément le gyrus temporal supérieur, et le gyrus frontal inférieur droit, associé à une diminution de volume du faisceau arqué reliant ces deux aires cérébrales (Hyde et al., 2007 ; Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006 ; Hyde, Zatorre, & Peretz, 2011 ; Loui, Alsop, & Schlaug, 2009).

1.3.3 L'amusie touchant la synchronisation au beat

La première étude empirique systématique d'un cas d'amusie touchant spécifiquement le traitement du beat dans la mélodie a été publiée en 2011 (Phillips-Silver et al.). Cet individu, nommé Mathieu, est incapable de synchroniser par lui-même un mouvement de fléchissement des genoux (*bouncing* en anglais) ou des tapes au beat de la musique. Il produirait également une réponse motrice plus variable que la normale lors de la synchronisation à un métronome (Palmer, Lidji, & Peretz, 2014). Mathieu montre, en revanche, une certaine capacité à ajuster la vitesse de son mouvement à un changement de tempo de 20 % dans la musique. Mathieu présente aussi des scores normaux aux épreuves mélodiques de la MBEA et à une tâche plus simple de discrimination de hauteur entre deux notes. Il obtient toutefois un résultat plus faible à l'épreuve de la métrique, tout en étant dans la norme à l'épreuve de rythme de la MBEA. Par ailleurs, Mathieu chanterait juste.

D'autres études ont depuis décrit de nouveaux cas d'individus n'ayant pas l'aptitude à se synchroniser au beat de la musique. Sowiński & Dalla Bella (2013) ont testé près de cent personnes sur leur capacité à synchroniser des tapes au beat d'extraits musicaux. Ils ont ainsi identifié dix participants ayant une difficulté à se synchroniser, qu'ils ont qualifié de *poor synchronizers*. Ces personnes sont comparables à d'autres participants sans trouble de la synchronisation dans une tâche de détection d'anisochronie et dans une tâche de tempo spontané, éliminant la possibilité d'un trouble plus général de *timing*. Au score mélodique global de la MBEA, en tant que groupe, les *poor synchronisers* ont obtenu de moins bons résultats que les autres participants de l'étude. Par contre, trois des dix cas identifiés seulement avaient des scores sous le seuil indicateur d'un trouble de la perception des hauteurs. En revanche, ces participants ont démontré également une mauvaise synchronisation dans une tâche de *tapping* sur une séquence rythmique mimant la modulation de l'enveloppe sonore de la musique, mais sans variation de hauteurs de sons. Ce dernier résultat suggérait donc, chez ces individus, un trouble primaire de la synchronisation au beat. Les auteurs proposent que, chez les individus présentant un trouble de la synchronisation au beat, le système de couplage auditivo-moteur puisse être fautif. De plus, considérant que l'analyse de la synchronisation, qui indiquait l'absence de correction de la coordination entre les tapes et le beat en fonction de l'asynchronie (lag1 positif), la possibilité de mécanismes de correction déficients était proposée.

Similairement, Palmer et collaborateurs (2014) ont mesuré chez Mathieu, et un nouveau cas, Marjorie, la synchronisation à des perturbations de phase ou de période introduites dans des séquences de sons isochrones. Les deux cas ont montré un ajustement anormal aux deux types de perturbations, suggérant des mécanismes de correction inefficaces. Selon les auteurs, si des mécanismes de correction déficients expliquent le trouble chez les individus montrant une faible aptitude à se synchroniser au beat, ceux-ci devraient montrer une mauvaise synchronisation, peu importe le type de stimulus rythmique présenté.

Tranchant, Vuvan, & Peretz (2016) ont également identifié quatorze individus, dans un échantillon de cent personnes, incapables de se synchroniser au beat de la musique en fléchissant les genoux. Cependant, tous ces participants ne montraient pas systématiquement un trouble de la synchronisation au beat en tapant des mains sur le beat. Ainsi, le mouvement choisi pour se synchroniser au beat pourrait influencer la performance de certaines personnes, qui présenteraient un trouble d'origine plutôt motrice, par exemple.

L'étude de personnes présentant une amusie touchant la perception des variations de hauteurs permet d'examiner l'association entre le traitement de la mélodie et du rythme dans la musique, alors que les personnes présentant une amusie affectant la synchronisation au beat permettent de tester la modularité du mécanisme d'entrainement par rapport à d'autres domaines, comme la parole. Suivant un principe d'ingénierie inversée ou de rétro-ingénierie, l'intérêt d'étudier une population clinique est de tester des modèles en se basant sur les comportements observés pour cheminer vers les mécanismes cognitifs pouvant expliquer la présence des troubles. C'est cette approche qui a été favorisée pour mener les travaux de recherche de la thèse. Dans la section qui suit, les hypothèses ciblées par la thèse concernant la spécificité de l'entrainement au beat musical seront étayées.

1.4 Spécificité des mécanismes d'entrainement au beat

La tendance qu'a la musique à entrainer un mouvement synchronisé est unique en comparaison à d'autres stimuli de l'environnement, et découlerait principalement de la présence d'un beat dans celle-ci (Honing, 2012 ; Iversen, 2016). Cependant, différentes hypothèses suggèrent que le mécanisme d'entrainement et de synchronisation au beat ne serait pas spécifique au domaine musical ou même à l'intérieur du domaine musical. D'abord, l'étude de la synchronisation au beat musical chez les personnes présentant un trouble de la perception des hauteurs présente des résultats mitigés entre l'association possible de ces deux compétences. En effet, selon les études précédentes, seulement la moitié des cas de ce type d'amusie aurait un trouble associé de la perception du beat (Ayotte et al., 2002 ; Peretz et al., 2003), contestant l'indépendance des modules cognitifs traitant le beat et les variations de hauteurs. Ensuite, certains auteurs suggèrent que l'extraction des accents prosodiques et la segmentation des unités temporelles de la parole (soit les phonèmes et les syllabes) se feraient par un mécanisme d'entrainement d'oscillateurs, comme pour l'extraction du beat dans la musique (Ghitza, 2011, 2013 ; Giraud & Poeppel, 2012 ; Goswami, 2012 ; Mai, Minett, & Wang, 2016 ; Meyer, 2017 ; Peelle & Davis, 2012). Wilson & Wilson (2005) proposent également qu'un modèle d'entrainement d'oscillateurs puisse être utilisé pour expliquer l'ajustement des tours de parole entre des interlocuteurs lors d'une conversation. Enfin, considérant l'importance du couplage sensorimoteur dans la synchronisation au beat et dans l'imitation vocale, il est proposé qu'un mécanisme commun pourrait sous-tendre ces deux comportements (Patel, 2008). Dans la section qui suit, ces trois domaines possibles de dissociation ou d'association avec l'entrainement au beat seront exposés, constituant les hypothèses étudiées dans la thèse.

1.4.1 Séparabilité du traitement de la mélodie et du rythme dans la musique

La première étude de la thèse s'attarde à la séparation entre le traitement de la mélodie et du rythme dans la musique. Nous discuterons donc ici de la contribution de la mélodie ou des variations de hauteurs dans la synchronisation au beat.

Plusieurs études montrent que les variations de hauteurs peuvent influencer la perception du beat ou la synchronisation au beat. En effet, la présence d'incohérence entre les accents mélodiques et temporaux contenus dans un rythme et un contexte musical atonal peuvent réduire la capacité à suivre le beat pour s'y synchroniser (Ellis & Jones, 2009 ; Jones & Pfordresher, 1997 ; Pfordresher, 2003 ; Prince, 2011, 2014 ; Prince & Pfordresher, 2012).

Néanmoins, la présence d'accents mélodiques, ou d'une mélodie, n'est pas nécessaire pour engendrer l'entrainement à une séquence rythmique. En effet, la synchronisation à des rythmes contenant ou non de l'information mélodique ne montre pas de différence, quand les stimuli sont équivalents sur le plan de la périodicité (Dalla Bella, Białuńska, & Sowiński, 2013 ; Lidji, Palmer, Peretz, & Morningstar, 2011a ; Snyder & Krumhansl, 2001). Dalla Bella et collaborateurs (2013) ont montré par exemple que de la parole énoncée avec le même degré d'isochronie que des séquences musicales pouvait générer un entrainement similaire au beat.

Ainsi, la présence d'une mélodie n'est pas essentielle à la synchronisation au beat, mais elle peut l'influencer. La présence de variations de hauteurs semble avoir un impact négatif sur la capacité des amusiques à détecter des changements de durée ou une anisochronie dans des séquences rythmiques (Foxton, Nandy, & Griffiths, 2006 ; Hyde & Peretz, 2004 ; Pfeuty & Peretz, 2010). Une étude de Phillips-Silver, Toiviainen, Gosselin, & Peretz (2013) a montré que la perception de la métrique s'améliorait lorsque le jugement était fait sur des séquences percussives plutôt que mélodiques. Des études indiquent aussi que la capacité d'amusiques à se synchroniser avec le beat serait améliorée lorsque des séquences non-mélodiques, comme une séquence isochrone de bruits succincts (Dalla Bella & Peretz, 2003) ou des rythmes percussifs (Phillips-Silver et al., 2013), sont utilisées. Des auteurs proposent ainsi que le trouble observé en lien avec la perception du beat proviendrait, chez ces individus, d'un effet d'interférence causé par un traitement anormal des variations de hauteurs (Dalla Bella & Peretz, 2003 ; Foxton et al., 2006 ; Hyde & Peretz, 2004 ; Pfeuty & Peretz, 2010). Ainsi, un traitement déficient de la hauteur empêcherait soit l'intégration de la mélodie et du rythme en une représentation unifiée ou ajouterait du « bruit » au moment de l'encodage du rythme (Foxton et al., 2006 ; Hyde & Peretz, 2003). Cette vision suppose donc que la perception des hauteurs serait automatique et ne pourrait être dissociée de la perception du rythme en contexte musical.

L'idée d'une interférence possible de la mélodie sur la perception du rythme chez l'amusique présente certaines limites. Notamment, si la difficulté à traiter l'aspect rythmique provient en effet d'une interférence, on pourrait s'attendre à ce que tous les amusiques présentent un certain degré de difficulté avec le traitement du beat et que celle-ci soit associée à la sévérité du déficit de la perception du beat. Or, les études semblent rapporter qu'environ 50 % des cas d'amusie portant sur la perception des hauteurs présentent également une difficulté avec le beat (Ayotte et al., 2002 ; Peretz et al., 2003). Ceci suggère que les troubles concernant le traitement des hauteurs et le rythme, ou le beat, pourraient être distincts chez les amusiques. Des cas de dissociation entre le traitement de la mélodie et du rythme ont d'ailleurs été rapportés par le passé chez des sujets ayant développé une amusie à la suite d'une lésion cérébrale (Liégeois-Chauvel et al., 1998 ; Peretz & Kolinsky, 1993 ; Peretz et al., 1994). Chez l'amusique congénital, la dissociation éventuelle entre une perception des hauteurs déficiente et la synchronisation à la musique a motivé la première étude de la thèse.

1.4.2 Comparaison de l'entrainement à la musique et à la parole

Comme mentionné précédemment, la structure rythmique est une composante importante à la fois du traitement de la musique et de la parole. En effet, la parole, tout comme la musique, présente une organisation hiérarchique sur le plan temporel, avec les phonèmes (unité de son de la parole), les syllabes et les accents prosodiques se produisant sur différentes échelles temporelles, soit un cycle de 0.5 à 4 hertz (Hz) pour les accents prosodiques, 4 à 8 Hz pour les syllabes et plus de 30 Hz pour les phonèmes (Brown, Pfordresher, & Chow, 2017 ; Leong, Stone, Turner, & Goswami, 2014 ; Meyer, 2017 ; Peelle & Davis, 2012). Les locuteurs seraient sensibles à la structure métrique de la parole, basée sur des accents acoustiques le plus souvent associés à l'échelle temporelle des syllabes (Cummins & Port, 1998 ; Kotz & Schwartze, 2010 ; Port, 2003 ; Selkirk, 1986 ; Turk & Shattuck-Hufnagel, 2013). Basés sur ces observations, des chercheurs proposent que le traitement du rythme dans la parole repose également sur un mécanisme d'entrainement (Cummins, 2009 a; Cummins & Port, 1998 ; Giraud & Poeppel, 2012 ; Goswami, 2012 ; O'Dell & Nieminen, 1999 ; Peelle & Davis, 2012 ; Phillips-Silver et al., 2010; Port, 2003 ; Wilson & Wilson, 2005).

Des études comportementales suggèrent aussi que l'entrainement au rythme de la parole est possible. Les interlocuteurs dans un échange conversationnel sont chacun entrainés au débit de parole de l'autre, qui correspond à la vitesse de production des syllabes (Schultz et al., 2015 ; Wilson & Wilson, 2005). Le débit de parole d'un locuteur peut également influencer implicitement le débit de parole d'une autre personne (Borrie & Liss, 2014 ; Jungers, Palmer, & Speer, 2002 ; Schwartze et al., 2016). Nous serions aussi en mesure d'adapter notre parole pour nous aligner avec un métronome (Cummins & Port, 1998) et de nous synchroniser avec une autre personne à la lecture d'un texte (Cummins, 2001, 2002a, 2002b, 2009b). Lidji, Palmer, Peretz, & Morningstar (2011 b) ont montré de plus que la variabilité entre les intervalles de tapes produites, lorsque des locuteurs tentent de se synchroniser à des extraits de paroles en français et en anglais, reflète la variabilité dans l'organisation temporelle des stimuli. Des études en EEG rapportent également une réponse oscillatoire à la structure temporelle de la parole (p. ex., Di Liberto, O'Sullivan, & Lalor, 2015 ; Ding & Simon, 2014 ; Ghitza, 2011, 2013 ; Giraud & Poeppel, 2012 ; Gross et al., 2014 ; Peelle, Gross, & Davis, 2013 ; Zhang & Ding, 2017). Cet entrainement aurait pour fonction de concentrer l'attention aux moments associés à l'occurrence d'éléments critiques pour le décodage de la parole, comme la distinction des phonèmes, par exemple (Ding et al., 2017 ; Meyer, 2017 ; Peelle & Davis, 2012). Par ailleurs, on peut se demander si l'entrainement à la parole et au beat de la musique repose sur le même mécanisme ou non.

Une différence importante demeure entre l'organisation temporelle de la musique et de la parole : le beat. En effet, peu d'évidences suggèrent la présence d'un beat dans la parole, les intervalles de temps entre les accents métriques de la parole étant beaucoup moins réguliers que dans la musique (Dauer, 1983 ; Jadoul, Ravignani, Thompson, Filippi, & de Boer, 2016 ; Nolan & Jeon, 2014 ; Patel, 2008 ; Turk & Shattuck-Hufnagel, 2013). Pour cette raison, des auteurs proposent que l'entrainement, ou le système de couplage sensorimoteur, impliqué dans le traitement musical diffère de l'entrainement à la parole ou à d'autres stimuli non-périodiques, ces derniers n'offrant pas la possibilité de générer une réponse prédictive comme celle trouvée en réponse au beat (Patel & Iversen, 2014 ; Zatorre et al., 2007). Cette hypothèse est, de plus, appuyée par des données d'imagerie cérébrale, montrant une distinction entre le réseau de structures cérébrales impliqué dans le traitement de séquences rythmiques périodiques et non-périodiques (Iversen & Balasubramaniam, 2016 ; Leow & Grahn, 2014 ; Schwartze & Kotz, 2013 ; Steen, Schwartze, Kotz, & Keller, 2015 ; Teki et al., 2011).

Malgré ces positions entourant la spécificité des mécanismes d'entrainement à la pulsation musicale, très peu d'études ont comparé directement l'entrainement à la musique et à la parole. Hausen, Torppa, Salmela, Vainio, & Särkämö (2013) ont étudié l'association entre les aptitudes de perception musicale, mesurées par le test en ligne d'amusie, et la perception des accents prosodiques (*stress* en anglais) dans la parole. Ils ont trouvé que de meilleurs scores à la tâche rythmique du test en ligne étaient associés à une meilleure performance à la tâche langagière utilisée. Cette association n'était pas présente avec les épreuves mélodiques du test en ligne. Similairement, une autre étude a rapporté une association entre la performance à une tâche de perception du rythme musical et l'amplitude de la réponse de potentiels évoqués au positionnement incongru d'accents prosodiques dans la parole (Magne, Jordan, & Gordon, 2016). Cependant, ces études ne s'intéressent pas spécifiquement aux mécanismes d'entrainement, aucune réponse sensorimotrice n'étant mesurée.

Lidji et collaborateurs (2011a) ont évalué directement la synchronisation de tapes avec la régularité perçue dans des phrases énoncées selon trois conditions : chantées, dites avec une prosodie régulière (similaire à du rap) et dites avec une prosodie naturelle, donc irrégulière. Les participants de l'étude ont produit des tapes avec un intervalle plus irrégulier pour les phrases énoncées avec une prosodie naturelle que dans les deux autres conditions. De plus, les tapes étaient plus précisément alignées avec le beat des phrases chantées que dites avec un rythme régulier. Les auteurs ont interprété ces résultats comme une indication que le principal facteur permettant l'entrainement d'une réponse synchronisée serait la présence d'un rythme périodique, mais que ce mécanisme d'entrainement serait plus adapté (*attuned*) au signal musical. Cette méthodologie a été reprise dans l'Article 2.

Parallèlement, Dalla Bella et collaborateurs (2013) ont testé dans deux études l'effet d'interférence de musique et de parole sur la synchronisation à un métronome, permettant une mesure implicite de l'entrainement. Dans la première étude, l'effet d'une musique générée par ordinateur (donc ayant un rythme parfaitement périodique) était comparé à l'effet de phrases énoncées avec un rythme régulier. Lorsque la musique et la parole étaient parfaitement alignées avec le métronome, auquel les participants devaient rester synchronisés, aucun effet différencié n'a été trouvé entre les deux types de stimuli. Cependant, lorsque les stimuli étaient décalés par rapport au métronome, la musique avait un plus grand effet d'interférence, générant plus de variabilité dans la synchronisation au métronome. Dans la deuxième étude, les auteurs ont eu recours à des stimuli plus comparables sur le pan de la périodicité, en utilisant des mélodies chantées avec des paroles, chantées sans paroles et les paroles récitées avec un rythme régulier. Dans ce contexte, les trois types de stimuli ont causé un effet d'interférence similaire.

Donc, lorsque des stimuli ont des caractéristiques rythmiques comparables, il semblerait que l'effet de la parole et de la musique sur l'entrainement soit similaire et que la périodicité soit le facteur principal influençant l'entrainement de la réponse sensorimotrice. Les études décrites jusqu'ici ne permettent pas toutefois de déterminer si un mécanisme commun d'entrainement serait impliqué lors de la synchronisation à des stimuli plus ou moins périodiques. C'est pourquoi l'étude de participants présentant un trouble de la synchronisation au beat s'avère intéressante, car elle peut permettre de déterminer l'impact de mécanismes de synchronisation au beat déficients sur la synchronisation dans d'autres contextes, incluant la parole.

1.4.3 Association entre synchronisation au beat et imitation vocale

Patel propose que la capacité de synchronisation au beat aurait dérivé de l'évolution des habiletés d'imitation vocale, les deux comportements nécessitant la production d'un geste moteur précis, qui se fonderait sur un mécanisme de couplage sensorimoteur commun (Patel, 2006 ; Patel, Iversen, Bregman, & Schulz, 2009). Les données provenant d'études en neuroimagerie semblent appuyer un recoupement entre les structures impliquées dans ces deux comportements, incluant principalement le SMA et les ganglions de la base (Patel, 2006 ; Patel et al., 2009). L'hypothèse avait au départ été motivée par l'observation d'animaux capables de se synchroniser au beat de la musique qui présentaient également une bonne aptitude d'imitation vocale, comme le cacatoès par exemple (Benichov, Globerson, & Tchernichovski, 2016 ; Patel et al., 2009 ; Schachner, 2010 ; Schachner, Brady, Pepperberg, & Hauser, 2009). Cependant, de nouvelles données suggèrent que certains animaux seraient en mesure d'apprendre à se synchroniser à un beat, en l'absence d'aptitude d'imitation vocale (Cook, Rouse, Wilson, & Reichmuth, 2013; Large & Gray, 2015; Merchant & Honing, 2014; ten Cate, Spierings, Hubert, & Honing, 2016; Wilson & Cook, 2016).

Seule une étude, jusqu'à maintenant, a mesuré l'association entre les habiletés d'imitation vocale et de synchronisation au beat chez l'humain (Dalla Bella, Berkowska, & Sowiński, 2015). Dans cette étude, des participants ont chanté une mélodie familière après avoir entendu un modèle et ont, dans une seconde tâche, synchronisé des tapes avec un métronome. L'association entre la justesse du chant et de la synchronisation au métronome a par la suite été mesurée. Les résultats de l'étude ont révélé que la justesse du chant était corrélée positivement avec la précision et la constance de la synchronisation au métronome. Plus exactement, la capacité à produire avec précision les intervalles de la mélodie était associée à la précision de la synchronisation entre les tapes et les beats du métronome. Ces résultats semblent donc appuyer l'hypothèse d'une relation entre l'imitation chantée et la synchronisation au beat, qui selon les auteurs de l'étude se fonderait sur un mécanisme commun de couplage sensorimoteur.

Selon les modèles les plus récents, la capacité à chanter avec justesse reposerait principalement sur une boucle sensorimotrice impliquant deux processus : un processus de couplage sensorimoteur et une boucle de rétroaction (Berkowska & Dalla Bella, 2009 a ; Dalla Bella, Berkowska, & Sowinski, 2011 ; Hutchins & Peretz, 2012a ; Hutchins & Peretz, 2012b; Pfordresher, Demorest, et al., 2015 ; Pfordresher, Halpern, & Greenspon, 2015 ; Pfordresher & Mantell, 2014). Le processus de couplage sensorimoteur réfère au processus par lequel une note est associée à une représentation interne du geste vocal nécessaire pour produire cette note. La boucle de rétroaction quant à elle correspond au processus par lequel l'erreur entre la note produite et la note attendue est intégrée de façon à permettre l'ajustement subséquent du geste vocal. Ces processus appliqués au chant trouvent également écho dans certains modèles de synchronisation au beat, dont le modèle ASAP, selon lesquels la justesse de la synchronisation dépendrait de la capacité à générer un modèle interne du beat sous la forme d'une action motrice. Le décalage entre le beat produit et le beat attendu serait ensuite intégré et utilisé pour ajuster la synchronisation (Maes, Leman, Palmer, & Wanderley, 2014 ; Patel & Iversen, 2014 ; Ross, Iversen, & Balasubramaniam, 2016 ; Van Der Steen & Keller, 2013 ; Zatorre et al., 2007). L'étude de personnes présentant une incapacité à se synchroniser au beat offre donc encore ici la possibilité de qualifier plus directement l'association de la synchronisation au beat avec le chant.

Bien que l'hypothèse discutée ici repose principalement sur l'imitation, l'étude du chant choral (à comprendre ici, chanter avec un modèle) chez les individus inaptes à se synchroniser au beat apparait également intéressante, car elle permet de tester la généralisation du trouble à un autre effecteur que les tapes de doigts plus typiquement utilisés. Tremblay-Champoux et collaborateurs (2010) ont trouvé que le fait de chanter une chanson familière à l'unisson avait un effet nul sur la justesse et la variabilité du tempo dans le chant de sujets neurotypiques non-musiciens. Pfordresher & Brown (2007) ont plutôt trouvé que de chanter avec un modèle avait un effet positif sur la production du contour et des intervalles de hauteur de courtes mélodies nouvellement apprises, et avait un effet positif sur la reproduction exacte des notes du modèle chez les bons chanteurs, mais un effet néfaste chez les mauvais chanteurs. À ce jour, aucune étude empirique n'a mesuré le chant chez les personnes présentant un trouble de la synchronisation sensorimotrice et donc la généralisation du trouble au chant choral demeure une question ouverte. Pourtant, la danse et le chant sont deux formes d'expression musicale universelles (Dowling, 1999 ; Mithen, 2006) et qui sont le plus souvent produites dans des contextes sociaux impliquant la synchronisation avec d'autres (Phillips-Silver et al., 2010 ; Phillips-Silver & Keller, 2012 ; Ravignani, Bowling, & Fitch, 2014).

1.5 Objectifs et hypothèses

L'objectif général de cette thèse est de mieux cerner la spécificité du mécanisme d'entrainement au beat dans la musique, par l'évaluation de la synchronisation sensorimotrice à la musique, à la parole et au chant dans l'amusie congénitale. Deux formes d'amusie sont étudiées : La première forme d'amusie est la plus connue et correspond à une incapacité à discriminer les fines variations de hauteurs dans la musique ; la seconde forme d'amusie a commencé plus récemment à être étudiée et implique une inaptitude à synchroniser un mouvement simple avec le beat de la musique.

La première étude de la thèse se penche sur la première forme d'amusie afin d'examiner la séparabilité de la perception des hauteurs et de la synchronisation au beat musical. La seconde étude se concentre sur la deuxième forme d'amusie, mais inclut également des participants présentant les deux formes d'amusie en concomitance, alors que la synchronisation de tapes du doigt au chant et à la parole est comparée. Enfin, la troisième étude porte sur le chant des personnes présentant un trouble de la synchronisation au beat, avec un intérêt particulier pour l'association possible entre ces compétences. Ainsi, l'étude de l'amusie offre la possibilité de mesurer l'impact d'un trouble de la synchronisation au beat ou de la perception des hauteurs, sur le comportement dans des activités connexes, comme la parole et le chant, dont l'interaction avec le mécanisme d'entrainement au beat est débattue.

Objectif et hypothèses de la première étude

L'objectif de la première étude de la thèse était de déterminer l'impact d'un trouble de la perception des hauteurs sur la capacité à percevoir le beat dans la musique et de s'y synchroniser. Pour ce faire, des participants amusiques, ainsi que des participants contrôles sans trouble de la perception musicale, ont pris part à deux expériences. Dans un premier temps, les participants ont accompli une tâche de perception du beat et une tâche de production du beat sur des extraits de pièces de musique variant en genre et en tempo. L'objectif était de déterminer s'il était possible d'établir un lien en la capacité à percevoir le beat et s'y synchroniser et les compétences préalables mesurées de perception des variations des hauteurs. Dans un deuxième temps, les sujets ont participé à une tâche de synchronisation au beat, mais en utilisant cette fois seulement des séquences percussives. L'objectif était ici d'éliminer l'influence possible de variations mélodiques sur la performance et donc un effet possible d'interférence.

Sur la base d'études précédentes qui suggèrent un effet d'interférence du traitement déficient des hauteurs sur la perception et la synchronisation au beat, notre hypothèse était que les participants amusiques auraient de moins bonnes performances que les participants contrôles, à la fois à la tâche de perception et de production du beat, lorsque des extraits musicaux étaient utilisés. De plus, si effectivement les difficultés de synchronisation dans l'amusie découlent d'un effet d'interférence, l'ampleur du trouble de la perception des hauteurs devrait corréler positivement avec l'aptitude à percevoir et à se synchroniser au beat d'extraits musicaux. C'est-à-dire qu'une plus faible capacité à discriminer les variations de hauteurs devrait être associée à une plus faible capacité à détecter le beat et s'y synchroniser. Par contre, nous nous attendions, suivant la même logique d'un effet d'interférence, à ce que la performance des amusiques s'améliore et devienne similaire aux contrôles lorsque des séquences percussives étaient utilisées pour la synchronisation.

Objectif et hypothèses de la seconde étude

La deuxième étude de la thèse visait à déterminer si le trouble de la synchronisation au beat de la musique se généralisait à la synchronisation à la parole, offrant la possibilité de tester la spécificité du mécanisme d'entrainement au beat de la musique. Si le trouble, d'abord mesuré lors de la synchronisation à la musique, s'étend à la synchronisation à la parole, on pourrait alors penser à l'existence d'un mécanisme commun d'entrainement impliqué à la fois dans le traitement de la musique et de la parole. Pour investiguer cette question, les participants amusiques et un groupe de participants neurotypiques appariés ont dû produire des tapes en synchronie avec le beat qu'il percevait dans des extraits de paroles chantées, énoncées avec un rythme régulier (similaire à du rap) et énoncées avec une prosodie naturelle, donc un rythme moins régulier.

Comme des études récentes montrent que dans la population normale la réponse d'entrainement est similaire entre la musique et la parole, lorsque la périodicité du rythme est contrôlée, notre hypothèse est que les participants contrôles devraient produire une synchronisation similaire pour les extraits chantés et énoncés avec un rythme régulier. Parallèlement, les participants amusiques devraient produire une synchronisation plus variable que les contrôles dans ces deux conditions, sans montrer de différence entre leur capacité à se synchroniser au chant et à la parole régulière. Par ailleurs, si la synchronisation à la parole naturelle ne se fonde pas sur les mécanismes d'entrainement au beat, on peut s'attendre à ce que les participants amusiques et les contrôles ait une performance semblable dans la condition de parole naturelle, qui impliquerait une synchronisation plus variable que dans les autres conditions.

Objectifs et hypothèses de la troisième étude

Cette étude avait l'objectif principal de tester si une association existe entre les capacités d'imitation chantée et de synchronisation au beat, en étudiant le chant d'amusiques présentant justement un trouble de la synchronisation au beat. Comme cette étude est également la première étude empirique du chant chez cette population, une investigation plus approfondie des habiletés de chant a été menée. Ainsi, les participants amusiques et un groupe contrôle apparié ont chanté une mélodie familière avec et sans paroles, dans quatre conditions : de mémoire, après un modèle, avec un modèle, et avec un métronome. Les participants ont également accompli une tâche de synchronisation de tapes avec un métronome afin de pouvoir mesurer la corrélation entre la synchronisation au beat et la justesse du chant.

L'étude s'est attardée dans un premier temps au chant produit de mémoire. L'objectif était d'établir un niveau de base de la justesse du chant chez les amusiques présentant un trouble de la synchronisation au beat. La justesse dans la production des intervalles de hauteurs, le respect du rythme de la mélodie, ainsi que la variabilité du tempo étaient les mesures considérées dans cette condition de base. Pour les conditions d'imitation et de chant vocal, nous avons mesuré les erreurs d'intervalles de hauteurs, la variabilité du tempo et la reproduction des durées et du tempo du modèle. En principe, si l'on assume un lien entre l'imitation chantée et les capacités de synchronisation, on devrait s'attendre à ce que les participants amusiques chantent généralement moins juste que les participants contrôles. Par ailleurs, une difficulté à s'adapter au modèle devrait également être retrouvée et ce tant pour les mesures de hauteurs de notes que de rythme et de tempo. Particulièrement, le trouble de synchronisation devrait se manifester également lors du chant choral et de la synchronisation au métronome. Selon l'hypothèse d'une association entre la synchronisation au beat et l'imitation vocale, il devrait y avoir une corrélation positive entre la précision et la constance de l'intervalle de temps entre les tapes lors de la synchronisation au métronome et le nombre d'erreurs d'intervalles de hauteurs produites.

Finalement, l'effet de chanter avec ou sans paroles devrait être similaire entre les groupes, et donc le chant devrait être généralement plus juste pour les productions sans paroles qu'avec paroles, selon les résultats d'études menées auprès d'une population normale (Berkowska & Dalla Bella, 2009 b, 2013 ; Dalla Bella, Tremblay-Champoux, Berkowska, & Peretz, 2012).

Chapitre II : Méthodologie et résultats

Article 1 : The Co-occurrence of Pitch and Rhythm Disorders in Congenital Amusia

Marie-Élaine Lagrois[a, b] & Isabelle Peretz[a, b]

[a] International Laboratory for Brain, Music, and Sound Research (BRAMS)

[b] Département de Psychologie, Université de Montréal

Article sous presse dans *Cortex*

Abstract

The most studied form of congenital amusia is characterized by a difficulty with detecting pitch anomalies in melodies, also referred to as pitch deafness. Here, we tested for the presence of associated deficits in rhythm, beat in particular, processing in pitch deafness. In Experiment 1, participants performed beat perception and production tasks with musical excerpts of various genres. The results show a beat finding disorder in six of the ten assessed pitch-deaf participants. In order to remove a putative interference of pitch variations with beat extraction, the same participants were tested with percussive rhythms in Experiment 2 and showed a similar impairment. Furthermore, musical pitch and beat processing abilities were correlated. These new results highlight the tight connection between melody and rhythm in music processing that can nevertheless dissociate in some individuals.

Keywords: congenital amusia, pitch deafness, musical deficits, melody and rhythm interaction, synchronization

1. Introduction

Musical engagement is ubiquitous and emerges early in life. As soon as they are born, humans respond to abstract properties of musical pitch and time structure, such as changes in tonal key (Perani et al., 2010) and disruptions of musical beat (Winkler, Háden, Ladinig, Sziller, & Honing, 2009). Infants move spontaneously to music (Zentner & Eerola, 2010) and show enhanced pro-social behavior when moved in synchrony with music (Cirelli, Einarson, & Trainor, 2014). In this context, lack of musical skills later in life is puzzling.

Musical deficits are particularly intriguing when they emerge in isolation from speech delay, intellectual disability, acquired brain damage, or music deprivation. These musical deficits are referred to as congenital amusia, pointing to the neurodevelopmental aspect of the disorder. The most common form of congenital amusia concerns the processing of the pitch structure of music and is often referred to as pitch deafness. Individuals with pitch deafness have a normal understanding of speech and prosody in everyday life. They can recognize speakers by their voices and can identify all types of familiar environmental sounds, such as animal cries. What characterizes them behaviorally is their difficulty with detecting out-of-tune singing, including their own, recognizing a familiar tune without the aid of the lyrics, discriminating melodies varying in pitch, and maintaining such melodies in short-term memory (e.g., Ayotte, Peretz, & Hyde, 2002).

Major progress has been made in recent research with regard to the neurobiological etiology of this musical pitch disorder (Peretz, 2016). Pitch deafness is marked by a neural anomaly affecting functional and structural connectivity between the right auditory cortex and inferior frontal cortex. It is also hereditary. Thus, congenital amusia represents a rare chance to examine the neurobiology of music cognition by tracing causal links between genes, brain, and behavior. The logic is essentially one of reverse engineering. An anomaly observed at the behavioral level can be traced back to cognitive processes, then to neurophysiological processes, and ultimately to genes and environment. Accordingly, the identification of associated behavioral deficits is essential. Here, we examine to what extent the pitch deficit characterizing pitch deafness is related to a deficit in abstracting properties of temporal structure from music, namely its beat.

Deficits in beat processing, initially called beat deafness (Phillips-Silver et al., 2011), can occur in isolation (Bégel et al., 2017; Dalla Bella & Sowiński, 2015; Phillips-Silver et al., 2011; Sowiński & Dalla Bella, 2013; Tranchant, Vuvan, & Peretz, 2016), but may also occur in association with pitch deficits. About half of individuals with pitch deafness also show impairments on tasks requiring rhythm discrimination (Ayotte et al., 2002; Peretz, Champod, & Hyde, 2003). Previous studies have shown that, in pitch deafness, the presence of pitch variations interferes with the detection of temporal change (Foxton, Nandy, & Griffiths, 2006; Hyde & Peretz, 2004; Pfeuty & Peretz, 2010). The available research suggests that when pitch variations are removed, discrimination of rhythmic patterns returns to normal (Foxton et al., 2006; Phillips-Silver, Toiviainen, Gosselin, & Peretz, 2013). These findings have led to the conclusion that the rhythmic deficit found in pitch deafness is a cascade effect of inadequate processing of musical pitch (Dalla Bella & Peretz, 2003; Hyde & Peretz, 2004; Pfeuty & Peretz, 2010).

This "pitch interference account" of the associated rhythm deficit in pitch deafness has limitations. If it was the case that rhythm processing is compromised by a faulty pitch processing system, then all individuals with pitch deafness should show a musical rhythm deficit to some extent. As mentioned above, a rhythmic problem does not occur in all cases, but in about half of sampled amusics. Similarly, one would expect to find a correlation between the severity of the pitch impairment and the severity of the associated rhythmic deficit. Foxton et al. (2006) looked at the possible association between perception of pitch intervals and time intervals in pitch-deaf amusics and found no such correlation. This suggests that the pitch and time deficits may be distinct in congenital amusia. Here, we re-examine the co-occurrence of a rhythm deficit in pitch deafness with natural music stimuli, where pitch variations are embedded (Experiment 1) or reduced (Experiment 2).

In Experiment 1, we tested beat perception and synchronization to natural music using an adaptation of the Beat Alignment Test (BAT, Iversen & Patel, 2008). In this test, participants tap to the beat of the musical stimuli (production task) and also judge whether a surimposed metronome track is aligned with the beat of the same stimuli (perception task). About half of the individuals with pitch deafness were expected to perform poorly in these beat perception and production tasks. If rhythm and pitch deficits are distinct in congenital amusia, the beat finding disorder should be unrelated to the severity of the pitch deficit. In order to test these predictions more directly, in Experiment 2, beat finding abilities were assessed in the same participants with drum versions of a subset of the stimuli used in Experiment 1.

2. Experiment 1: Beat alignment tests in natural music

2.1 Method

2.1.1 Participants

Ten participants who met the diagnostic criteria for the pitch-deaf form of congenital amusia (age: 43.6 ± 18.0 years; eight females) and a matched control group of 12 participants (age: 42.4 ± 18.2 years; nine females) took part in the study. Controls were further matched to the pitch-deaf group in education and years of music and dance training. Detailed group characteristics are provided in Table 1. Participants provided written consent and received monetary compensation for their participation. All procedures were approved by the Research Ethics Council for the Faculty of Arts and Sciences at the Université de Montréal.

[Insert Table 1 here]

Prior to being selected for participation in this study, the participants underwent tests of their musical abilities. Pitch-deaf participants were included in this study based on their scores on both the online test (Peretz & Vuvan, 2017) and the Montreal Battery for Evaluation of Amusia (MBEA, Peretz et al., 2003). The online test is composed of three tests: Scale, Off-beat, and Off-key. The Scale test is the same in the online test and in the MBEA; it involves the comparison of 30 pairs of melodies that differ by an out-of-key note in half of the trials. The Off-beat and Off-key tasks require the detection of an out-of-time and out-of-key note in a melody, respectively. All control participants had scores within 2 *SD*s of the mean, indicating normal music perception. The MBEA comprises five additional tests: Contour, Interval, Rhythm, Meter and Memory. The score on the first two tests and the Scale test can be averaged in a melodic composite score, which gives an indication of participants' ability to detect pitch deviations in a melodic context. A score lying 2 *SD*s below the mean for the Melodic Composite score of the MBEA, or both scores on the Scale and Off-key subtests of the online test, indicates the presence of pitch deafness (Peretz & Vuvan, 2017; Vuvan et al., 2017; Table 1).

The Rhythm and Meter tests of the MBEA reflect different aspects of musical rhythm processing. The Rhythm test consists of comparing pairs of melodies where the temporal grouping in the comparison melody differs in half the trials. The Meter test consists of judging if a melody is a march or a waltz. As can be seen in Table 1, two pitch-deaf participants (A1 and A10) had scores below the cut-off on the Rhythm test and two others (A5 and A9) had scores below the cut-off on the Meter test. Thus, four of the 10 pitch-deaf amusics show indications of a rhythm problem in processing music using these tasks as typically observed in previous studies.

In order to get an index of the severity of the pitch deficit experienced by pitch-deaf amusics, they were tested on a pitch-change detection task. In this task, participants hear sequences of five tones and are asked to detect whether the fourth tone changes in pitch. This task is performed as part of the protocol for identification of pitch-deaf individuals in our research group (e.g., Vuvan et al., 2017). Here, we report detection accuracy for pitch changes of a quarter semitone (25 cents), the smallest pitch change included in the task, which is the most discriminant in comparison to neurotypical adults (Hyde & Peretz, 2004; Vuvan, Nunes-Silva, & Peretz, 2015).

All pitch-deaf participants had normal non-verbal reasoning and verbal working memory abilities as assessed by the Matrix Reasoning and the Digit Span tests from the WAIS-III (Wechsler Adult Intelligence Scale; Wechsler, Coalson, & Raiford, 1997).

2.1.2 Materials and Procedure

The Montreal version of the Beat Alignment Test (M-BAT, Tranchant et al., 2018); BAT, Iversen & Patel, 2008) includes a beat tapping task and a beat perception task. In both tasks, our version of the BAT presented the same ten musical excerpts of pop and jazz music at various tempi (range: 82-170 beats per minute). The music excerpts lasted between 23 and 31 seconds and contained at least 24 beats.

In the beat production task, participants were asked to tap along to the beat of the musical stimuli. The 10 excerpts were presented twice, in two distinct blocks, for a total of 20 trials. When the concept of beat was not clear to the participant, it was described as the “tic-toc” of a clock. Participants received four practice trials on musical excerpts that were not part of the test. After each practice trial, the music was presented with a click track surimposed on the beat to make it clear where taps were expected. The presentation order of the stimuli was randomized for each participant. The beat tapping task was always performed first to control for exposure to clicks on the beats of the stimuli in the perception task.

Isochronous clicks were surimposed on the music track for the beat perception task. On half of the trials, the clicks were “on beat” and on the other half “off-beat” by either a phase shift (± 15%) or a period shift (± 5%). The click series started five seconds after each excerpt commenced playing and always included 24 clicks. The presentation order was pseudo-randomized so that no song was presented twice consecutively. The task included 80 trials (eight repetitions of the ten musical excerpts). Participants judged at the end of each stimulus if the clicks were on the beat or not, using four response choices presented on screen: *always on the beat (1), mostly on the beat (2), sometimes on the beat (3)* and *rarely or never on the beat (4).* For the analyses, the first two choices were considered “on beat” responses and the last two “off beat” responses. Before starting the task, participants received six practice trials with feedback on accuracy.

The experiment took place in a large sound-attenuated studio. The stimuli were delivered through headphones (DT 770 PRO, Beyerdynamic) at a comfortable level. The tapping test was programed with MAX-MSP (https://cycling74.com) and the perception test was programed with MATLAB (https://www.mathworks.com). The taps were recorded on a square force sensitive resistor (3.81 cm, Interlink FSR 406) connected to an Arduino Duemilanove microcontroller board (arduino.cc) running an adapted Tap Arduino script (based on fsr_silence_cont.ino; Schultz & van Vugt, 2016; van Vugt & Schultz, 2015) to transmit timing information to a PC (HP ProDesk 600 G1, Windows 7) via the serial USB port.

2.1.3 Data Analysis

A measure of sensitivity (*d'*) of discrimination between "on-beat" and "off-beat" trials was considered for the beat perception test. Correct detection of "off-beat" trials were counted as hits, whereas answering "off-beat" to an "on-beat" trial was considered a false alarm.

For the beat production task, taps were first pre-processed to remove inter-tap intervals (ITIs) that were more than half-smaller or larger than the individual median ITI produced (median ITI ± (median ITI*0.5)). This resulted in one to nine taps per trial being removed. Trials with fewer than eight taps were discarded because the analysis of synchronization is more prone to bias with a small number of data points. However, the number of trials eliminated was low, with at least 18 out of the 20 trials being analyzable for each participant. In order to analyze performance on the same beats across the beat perception and beat production tasks, taps produced during the first ten and last five seconds of each song were discarded. Thus, 24 beats of each song were considered for analysis.

Synchronization with the beat of music was measured with circular statistics using the Circular Statistics Toolbox for MATLAB (Berens, 2009). With this technique, taps are transposed as angles on a circle from 0 to 360 degrees, where a full circle corresponds to the inter-beat interval. The position of the taps on the circle is used to calculate a mean resultant vector. The length of the mean resultant vector indicates how clustered are the points around the circle. Vector length (VL) range from 0 to 1; the larger the value, the more clustered together are the points on the circle, indicating that the period (or time interval) between taps tends to match the inter-beat interval of the stimulus more consistently (see Dalla Bella & Sowiński, 2015, where the same procedure was used). Statistical analyses performed on vector length used a logit transform as vector length distribution is typically skewed in synchronization data (logVL = -1*log(1-VL)). However, for simplicity, untransformed vector length is reported when considering group means and individual data. The Rayleigh z test of periodicity was further used to test if participants' taps had a consistent relationship with the inter-beat period, thus indicating if participants could match the period of the beat with their taps (Wilkie, 1983). A significant Rayleigh test (p-value < .05) indicates successful period matching between taps and the inter-beat interval of the stimulus. Trials with a p-value <.05 on the Rayleigh test were therefore considered as trials with successful period matching, and the percentage of trials with successful period matching was computed for each participant. The inter-beat interval used to generate the mean resultant vector and perform the Rayleigh test was adjusted to fit the metric level (beat period) at which participants tapped on each trial. Three beat periods were considered for each stimulus: the beat period corresponding to the tempo of the song, half the beat period of the tempo, and twice the beat period of the tempo. For example, a song at 120 beats per minute (bpm) would have 500 ms, 250 ms, and 1000 ms as possible beat periods. Based on the mean ITI of a participant for a given song, the closest beat period from that song was chosen to compute circular statistics.

We also computed the coefficient of variation (CV = *SD* ITI/Mean ITI), which is a standard measure of the regularity of the ITI, which does not take into account the period of the stimuli. The smaller the CV, the less variability in the tap intervals.

2.2 Results and Comments

2.2.1 Beat Perception

The average percentage of hits minus false alarms was 72.5% for controls (range: 50.0%–97.5%) and 26.8% for pitch-deaf participants (range: 10.0%-60.0%). The derived *d'* indices were significantly different between the two groups, $t(20) = 4.40$, $p < .001$. Nevertheless, three out of the 10 pitch-deaf (A2, A4, A9) had performances that laid within the controls' range (Figure 1A). All controls were considered to have a performance that lied within the normal range, based on data from Tranchant et al. (2018).

[Insert Figure 1 here]

2.2.2 Beat Production

On average, the control group successfully matched their taps to the inter-beat interval (IBI) of the songs on 96.6% of trials (range: 85%-100%; Figure 1B), whereas most pitch-deaf participants were quite poor at matching their taps to the IBI of the song (M = 57.2% of trials; range: 20% - 95%). Still, four pitch-deaf (A2, A4, A7, A9) could match the period of their taps to the beat of most songs. Three of them (A2, A4, A9) also performed on par with controls for this beat perception task (see Figure 1). No control showed impairment in that task.

The mean vector length (VL) in controls was .90 (*SD* = .04). The average vector length in the pitch-deaf group was .53 (*SD* = .27, range: .22 - .87) and differed significantly from the control group, $t(14.5) = 5.6$, $p < .001$ (comparison with logVL; Table 2). Two of the ten pitch-deaf amusics had vector length similar to controls (A4: VL = .87, and A9: VL = .86).

[Insert Table 2 here]

The mean CV of the control group was .07 (*SD* = .01), while the mean CV for pitch-deaf group was .10 (*SD* = .03, range: .08 - .16). The mean CV differed significantly between groups, $t(20) = -3.86$, $p = .001$. Thus, all but two pitch-deaf individuals (A4 and A9) tapped less regularly and were less consistently aligned with the period of the stimuli than controls (Table 2).

2.2.3 Relation Between Pitch and Beat Deficits

In order to assess a possible relationship between pitch and beat processing abilities, we computed the Pearson correlation coefficient between the scores obtained on the Scale test of the online test (as all participants had completed it) and the *d'* scores obtained in the perception task (Figure 2A). The scores were highly correlated, $r_{(20)} = .68$, $p = .001$. This was also the case for the mean logVL, with $r_{(20)} = .64$, $p < .001$ (Figure 2B). Interestingly, the percentage of accurate pitch-change detection (Table 1) did not predict pitch-deaf performance on the beat perception and production tasks, with $r_{(7)} = -.20$, $p = .61$ and $r_{(7)} = -.17$, $p = .67$ for *d'* and logVL, respectively, using Spearman non-parametric correlation coefficient. This is clearly evident in the cases of A2, A4, A7, and A9, who were quasi-normal at tracking the beat of music for synchronization but impaired in pitch-change detection (Table 1). We say “quasi-normal” because participants A2 and A4 still remained less consistent than controls in the beat production task. A7 was also impaired in the beat perception task.

[Insert Figure 2 here]

3. Experiment 2: Synchronization to Drum Rhythms

The co-occurrence of the pitch deficit with a beat deficit revealed in Experiment 1 in the majority of the pitch-deaf amusics called for a re-examination of beat finding performance in a context where their pitch deficit was unlikely to interfere with beat finding abilities, in case the latter was intact. This was tested in Experiment 2 with percussive music.

3.1 Method

The same participants were tested with percussive renditions of *Suavemente* (by Elvis Crespo), played at a tempo of 112 bpm and 120 bpm, with the audio files lasting 36 s and 33 s, respectively. This procedure has been used previously with a different pool of participants (Phillips-Silver et al., 2013). Each version of the song contained 65 beats that were created with a snare drum, a tenor drum, and a bass drum, so as to reproduce as closely as possible the major instrumental lines of the original song (for a detailed description of these stimuli, see Phillips-Silver et al., 2013). We added a percussive rendition of the song *Brand New Carpet* (by Bodi Bill), similarly created, at 126 bpm and the audio file lasting 16 s. This stimulus had 31 beats. Presentation order of the excerpts was counterbalanced, with *Brand New Carpet* always played in between the two drum versions of *Suavemente*. The original versions of *Suavemente* and *Brand New Carpet* were presented as stimuli in Experiment 1 and could therefore serve here for comparison.

Participants were asked to tap to the beat of the stimuli. A practice trial was performed before starting the task to make sure they understood the instructions. The practice trial used a drum rhythm not included in the test. In addition, the participants were asked to synchronize their taps to a metronome. This control task was included to assess sensorimotor synchronization when there was no need for beat extraction. Participants listened to seven metronome ticks and then had to synchronize their taps to a metronome at the same tempo for 60 taps. The task comprised two trials, one at a tempo of 96 bpm and one at 120 bpm. Each metronome stimulus was composed of 440 Hz sine wave ticks with the duration of 50 ms. The presentation order of the two metronome stimuli was counter-balanced between participants. A practice trial for metronome synchronization at 108 bpm was presented first to make sure participants understood the instructions.

Taps were recorded with the same system described in Experiment 1 section 2.1.2, with the stimuli again presented through headphones.

Circular statistics were used to assess synchronization as described in Experiment 1 section 2.1.3. In order to remove initial variability in synchronization, the first five beats of each drum excerpt were discarded from the analysis. To allow a more direct comparison with the results from the production task of the M-BAT, the next 24 beats were considered for the analysis. For synchronizing to the metronome, the first five beats were discarded to remove initial variability, leaving 55 beats to analyze.

3.2 Results and Comments

3.2.1 Results of the Tapping Task

All but one control participants successfully matched their taps to the period of the three drum trials and so did four of the ten pitch-deaf participants (A2, A4, A7, and A9). The one control participant who failed to synchronize to *Brand New Carpet* (p = .72) was able to synchronize to both *Suavemente* trials. In contrast, one pitch-deaf (A3) participant could only match his taps with the beat period of *Brand New Carpet*.

The four pitch-deaf who could synchronize to all drum trials also succeeded in synchronizing their taps to the beats of both trials of the original songs in Experiment 1 (Table 3). These four "beat-preserved" pitch-deaf individuals could also anticipate the beat with a mean negative asynchrony between taps and beats (M = -39 ms, SD = 30 ms). Controls mean asynchrony was -14 ms (SD = 22 ms). These results indicate that the "beat-preserved" pitch-deaf participants showed a similar phase relationship with the beat to that shown by controls.

[Insert Table 3 here]

Tapping performance obtained here, with the drum rhythms, was compared to the performance obtained with the original versions containing pitch variations (Experiment 1) by looking at the mean vector length (VL) and tapping variability (CV) (Table 4). An ANOVA performed on the mean logVL with Group as the between-subjects variable and Condition (drum vs original) as a within-subject variable revealed a main effect of Group, $F(1,20) = 19.03$, $p < .001$, $\eta^2 = .49$, a main effect of Condition, $F(1,20) = 4.81$, $p = .04$, $\eta^2 = .19$, and no significant Group × Condition interaction, $F(1,20) = 0.62$, $p = .44$. The group of pitch-deaf obtained a smaller VL (.50) than controls (.91) overall. Both groups had smaller VL with the drum versions than the original songs, although the effect was more salient in the pitch-deaf group (Table 4). The mean CVs showed similar trends. These results show that contrary to expectations, pitch-deaf participants synchronized their taps better to the original songs that included pitch variations than to the drum versions. With the latter, the majority of pitch-deaf show evidence of a beat deficit. The correlation between the mean VL obtained for each version was almost significant in the pitch-deaf group, with $r_{(8)} = -.62$, $p = .054$, and clearly significant in controls, $r_{(10)} = .65$, $p = .02$, using Spearman's correlation for nonparametric data.

[Insert Table 4 here]

All control and pitch-deaf participants, except one pitch-deaf (A8), could successfully synchronize their taps to the period of the metronome at both 120 bpm and 96 bpm (Rayleigh test, $p < .05$). A8 successfully synchronized his taps to the metronome at 96 bpm only. Synchronization at 120 bpm was inaccurate because this participant tapped too fast (mean ITI of 427 ms) relative to the 500 ms period of the stimulus. Comparing the length of the resultant vector (logVL) for participants with successful synchronization, we found no significant difference between pitch-deaf and control participants, $F(1,19) = 3.27$, $p = .09$, no effect of Tempo, $F(1,19) = 0.003$, $p = .96$, and no Group × Tempo interaction, $F(1,19) = 0.33$, $p = .57$. Similarly, for the mean asynchrony between the taps and the onsets of the metronome beat, there was no main effect of Group, $F(1,19) = 1.67$, $p = .21$, no effect of Tempo, $F(1,19) = 0.30$, $p = .59$, and no significant interaction, $F(1,19) = 1.28$, $p = .27$. The two groups showed mean negative asynchronies to both tempi: controls' $M = -48$ ms (range: -115 ms to 4 ms), pitch-deaf' $M = -66$ ms (range: -161 to -14 ms). Thus, as shown in previous studies, pitch-deaf amusics could synchronize to the metronome as accurately as controls, suggesting no general sensorimotor synchronization deficits (Dalla Bella & Peretz, 2003; Phillips-Silver et al., 2013).

3.2.2 Relation Between Tapping to Drums and Music Pitch Processing

As in Experiment 1, correlation between the scores obtained on the Scale test and the mean logVLs obtained for drum rhythms in this experiment was significant, $r_{(20)} = .46$, $p = .03$ (Figure 3). This is despite the presence of clear outliers among the pitch-deaf group (A2, A4, A7, A9), who displayed normal synchronization with the drum beat and poor musical pitch perception.

[Insert Figure 3 here]

4. General Discussion

The main finding of the present study is that melody and beat impairments are associated in most cases of pitch deafness. In our sample of ten adults diagnosed as having a deficit in musical pitch processing, at least six also manifest a deficit in finding the musical beat in music and drum rhythms. However, the presence of two to four clear-cut cases of musical pitch disorder with spared beat processing suggests that the pitch and beat deficits are distinct disorders. In what follows, we discuss the possible origins of the frequent co-occurrence of the musical deficits and the implications for the behavioral characterization of congenital amusia.

The attribution of the rhythmic difficulties to the possible interference caused by inadequate processing of pitch variations (i.e., the pitch interference hypothesis) found little support in the present study. The beat finding deficit experienced by the majority of pitch-deaf amusics remains severe whether pitch cues are present or not in the musical stimulus. Moreover, the beat-impaired amusics better align their taps to the original music, which contains pitch variations, than to their percussive renditions, although matched controls do not show such a clear advantage for the original music. Thus, the presence of putative interfering pitch information does not appear to play a significant role in the occurrence of the beat deficit.

Yet, there is a correlation between the severity of the musical pitch disorder and the size of the beat deficit, especially in perception (Figure 2A). This relation holds for amusics and controls alike. The higher the score in discriminating melodies, in which there can be a changed note that violates the pitch structure (on the Scale test of the MBEA), the higher the detection of misalignments of metronome clicks superimposed on music (on the M-BAT test). Obviously, the observed correlation between pitch and beat performance could be due to several factors that are not specific to music structure, such as auditory attention and motivation. Nevertheless, given the presence of correlations across tests of pitch and beat processing and the frequent co-occurrence of deficits in the processing of the two, the possibility of shared processing components should be examined.

Shared mechanisms between pitch and beat processing could occur at several levels, from sensory input through to motor output. Here, we can discard the two end processes since the basic auditory-motor loop appears normal in pitch-deaf amusics. First, there was no correlation between the severity of the sensory impairment observed in pitch-change detection in five-tone isochronous sequences and the tested beat finding abilities, suggesting no direct association between acoustic pitch and beat processing. Secondly, all ten pitch-deaf individuals were able to accurately match their taps to auditory metronome sequences, suggesting intact basic auditory-motor coupling in the context of a tapping task. Thus, shared mechanisms between pitch and beat processing are likely to concern more cognitive components. There is substantial evidence for interaction between pitch and time dimensions in music, although these are separable processing components. For example, a mismatch between pitch and temporal accents, or an atonal melodic context, can lower the capacity to track beats (Ellis & Jones, 2009; Jones & Pfordresher, 1997; Pfordresher, 2003; Prince, 2011, 2014; Prince & Pfordresher, 2012). The question of how information from pitch and time combines in music has been an area of continued interest (see Krumhansl, 2000; Prince, 2011, for reviews), with unfortunately little consensus on the issue of whether the integration of these dimensions is additive (Palmer & Krumhansl, 1987a, 1987b) or interactive (Jones, 1987; Jones & Boltz, 1989) and at what stage in the decision process the two dimensions are integrated. Hence, the identification of a shared processing component will have to await future development of cognitive models.

Identification of the locus for the observed tight association underlying pitch and beat processing might be aided by knowing their neural correlates. Here again, current knowledge gained from neuroimaging studies is not very informative or sufficiently constraining to provide good candidates for shared processing components. In a recent study (Sihvonen et al., 2016), both musical pitch and rhythm processing were examined in 77 brain-damaged patients while using the same screening tests used here, namely the Scale and the Rhythm tests of the MBEA. Deficits in each test were associated with lesions in the auditory cortex, Heschl's gyri, insula, and basal ganglia (putamen, caudate, pallidum) of the right hemisphere. Thus, a common locus for processing both types of structure may lie in that constellation of regions. However, we saw here that our pitch-deaf amusics with a beat finding disorder had normal scores on the rhythm test. Moreover, Grahn and McAuley (2009) found that good beat finders have greater brain activity in the supplementary motor area, left premotor cortex, and left insula, while poor beat-perceivers show relatively greater activation in the left posterior superior and middle temporal gyri and the right premotor cortex. These brain regions do not overlap with the anomalous fronto-temporal network identified in pitch deafness. Pitch deficits in congenital amusia have been linked to anomalies in connectivity between the inferior frontal gyrus (IFG; BA 44/45/47) and the superior temporal gyrus (STG; BA 22). More precisely, deficient connections from the right IFG to the right STG would prevent top-down influence from higher-order cortical regions in pitch processing (for a recent review see Peretz, 2016). Therefore, there is at present no clear indication of how or where in the brain the pitch and beat defects might overlap.

Nevertheless, there is a need to identify the co-occurrence of pitch and time deficits in congenital amusia in order to progress the characterization of the disorder. While there is a large consensus on how to screen for the presence of musical pitch deficits by using, among other tests, the Scale test of the MBEA (Vuvan et al., 2017), there is no equivalent consensus for beat deficits. Here we show that none of the MBEA tests is appropriate, not even the MBEA meter test that is supposed to tap the beat finding abilities. Yet, Phillips-Silver et al. (2013) found a positive correlation between the scores on the Meter test of the MBEA and beat synchronization with the same *Suavemente* song used here for the evaluation of beat finding abilities. We do not corroborate this finding since none of the correlations between the MBEA meter test, and the synchronization measures considered here reached significance. The reasons for this discrepancy between the prior and current studies are unclear. Therefore, in future studies, we propose to use the M-BAT test for its sensitivity to the presence of a beat deficit (see also Tranchant et al., in preparation, for norms on this test) rather than the MBEA meter test. However, the BAT requires the recording of precise motor responses aligned with a stimulus, which can hardly be done outside the lab. One future alternative tool is the BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities, which is currently being developed for the tablet using a touch screen (Dalla Bella et al., 2017; Puyjarinet, Bégel, Lopez, Dellacherie, & Dalla Bella, 2017).

Another area of research that would deserve more attention regarding congenital amusia is whether this population could benefit from musical intervention to improve performance. A few prior studies have been conducted to test if pitch perception could be improved in pitch deafness and results have so far been mostly negative (e.g., Hyde & Peretz, 2004; Liu, Jiang, Francart, Chan, & Wong, 2017; Mignault Goulet, Moreau, Robitaille, & Peretz, 2012). One recent study (Whiteford & Oxenham, 2018) obtained promising results after only five training sessions of pitch-change detection, although the contribution of a practice effect from test-retest could not be excluded since pitch-deaf participants trained on an irrelevant task also improved from pre-test to post-test. So far, training of beat processing abilities has not yet been assessed in amusics. Phillips-Silver et al. (2013) noted that in their group of pitch-deaf participants the accuracy of synchronization to the beat, when bouncing to a musical excerpt, tended to improve between the first and second trial. A follow-up with one of the pitch-deaf participants also showed an improvement in synchronization performance a few months later. In our study, we could not assess practice effect on beat finding abilities since presentation order was randomized for each participant. However, in the synchronization task of the M-BAT, which consists of the repetition of the same songs in successive blocks, we did not find an increase in performance. Future studies should examine more closely the distinct effect of practice and intervention in congenital amusia.

A promising strategy for training rhythmic skills, called Rhythm Workers, has recently been developed (Bégel, Seilles, & Dalla Bella, 2018). The training consists of a beat production (tapping) task or a beat perception task, both implemented on a tablet, using musical excerpts of various beat complexity. The tasks used in the training protocol and to measure pre-post change in performance are very similar to the M-BAT used here. Preliminary testing of the protocol indicates improvement in beat perception assessed before and after training in young neurotypical adults over a two weeks home-based training period (Bégel et al., 2018) as well as in patients with Parkinson’s disease, over a six weeks training period (Dauvergne et al., 2018). Transfer of improvements to different movements than tapping and beat perception in general remains to be addressed.

In summary, we have shown that pitch and time deficits more often co-occur in congenital amusia than they dissociate. This finding highlights the tight connection between melody and rhythm in music processing and invites researchers to systematically test for the joint presence of these deficits to contribute to the understanding of the origins of these neurodevelopmental disorders that are presently considered distinct.

Acknowledgments

We would like to thank Pauline Tranchant for insightful comments on data analysis and Mihaela Felezeu for technical assistance and help with data collection. We would also like to thank Dawn Merrett for help with editing.

Funding

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs program.

References

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238-251. doi: 10.1093/brain/awf028

Bégel, V., Benoit, C.-E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia, 94*, 129-138. doi: https://doi.org/10.1016/j.neuropsychologia.2016.11.022

Bégel, V., Seilles, A., & Dalla Bella, S. (2018). Rhythm Workers: A music-based serious game for training rhythm skills. *Music & Science*, *1*. doi: https://doi.org/10.1177/2059204318794369

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw, 31*(10), 1-21.

Cirelli, L. K., Einarson, K. M., & Trainor, L. J. (2014). Interpersonal synchrony increases prosocial behavior in infants. *Developmental Science, 17*(6), 1003-1011. doi: 10.1111/desc.12193

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods, 49*(3), 1128-1145. doi: 10.3758/s13428-016-0773-6

Dalla Bella, S., & Peretz, I. (2003). Congenital Amusia Interferes with the Ability to Synchronize with Music. *Annals of the New York Academy of Sciences, 999*(1), 166-169. doi: 10.1196/annals.1284.021

Dalla Bella, S., & Sowiński, J. (2015). Uncovering Beat Deafness: Detecting Rhythm Disorders with Synchronized Finger Tapping and Perceptual Timing Tasks. *Journal of Visualized Experiments: JoVE*(97), 51761. doi: 10.3791/51761

Dauvergne, C., Bégel, V., Gény, C., Puyjarinet, F., Laffont, I., & Dalla Bella, S. (2018). Home-based training of rhythmic skills with a serious game in Parkinson's disease: Usability and acceptability. *Annals of physical and rehabilitation medicine*. doi: https://doi.org/10.1016/j.rehab.2018.08.002

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264-280. doi: 10.1037/a0013482

Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in "tone deafness". *Brain and Cognition, 62*(1), 24-29. doi: https://doi.org/10.1016/j.bandc.2006.03.005

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage, 47*(4), 1894-1903. doi: https://doi.org/10.1016/j.neuroimage.2009.04.039

Hyde, K. L., & Peretz, I. (2004). Brains That Are out of Tune but in Time. *Psychological Science, 15*(5), 356-360. doi: 10.1111/j.0956-7976.2004.00683.x

Iversen, J. R., & Patel, A. D. (2008). *The Beat Alignment Test (BAT): surveying beat processing abilities in the general population.* In Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10), Sapporo.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*(6), 621-634. doi: 10.3758/bf03210494

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459-491. doi: 10.1037/0033-295X.96.3.459

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51* (4), 271-291. doi: 10.1037/1196-1961.51.4.271

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin, 126*(1), 159-179. doi: 10.1037/0033-2909.126.1.159

Liu, F., Jiang, C., Francart, T., Chan, A. H. D., & Wong, P. C. M. (2017). Perceptual Learning of Pitch Direction in Congenital Amusia. *Evidence from Chinese Speakers, 34*(3), 335-351. doi: 10.1525/mp.2017.34.3.335

Mignault Goulet, G., Moreau, P., Robitaille, N., & Peretz, I. (2012). Congenital Amusia Persists in the Developing Brain after Daily Music Listening. *PLOS ONE, 7*(5), e36860. doi: 10.1371/journal.pone.0036860

Palmer, C., & Krumhansl, C. L. (1987a). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*(1), 116-126. doi: 10.1037/0096-1523.13.1.116

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception: Effects of harmony, performance timing, and familiarity. *Perception & Psychophysics, 41*(6), 505-518. doi: 10.3758/bf03210485

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., . . . Koelsch, S. (2010). Functional specializations for music processing in the newborn human brain. *Proceedings of the National Academy of Sciences, 107*(10), 4758-4763. doi: 10.1073/pnas.0909074107

Peretz, I. (2016). Neurobiology of Congenital Amusia. *Trends in Cognitive Sciences, 20*(11), 857-867. doi: https://doi.org/10.1016/j.tics.2016.09.002

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of Musical Disorders. *Annals of the New York Academy of Sciences, 999*(1), 58-75. doi: 10.1196/annals.1284.006

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal Of Human Genetics, 25*, 625. doi: 10.1038/ejhg.2017.15

Pfeuty, M., & Peretz, I. (2010). Abnormal pitch-time interference in congenital amusia: Evidence from an implicit test. *Attention, Perception, & Psychophysics, 72*(3), 763-774. doi: 10.3758/app.72.3.763

Pfordresher, P. Q. (2003). The Role of Melodic and Rhythmic Accents in Musical Structure. *Music Perception: An Interdisciplinary Journal, 20*(4), 431-464. doi: 10.1525/mp.2003.20.4.431

Phillips-Silver, J., Toiviainen, P., Gosselin, N., & Peretz, I. (2013). Amusic does not mean unmusical: Beat perception and synchronization ability despite pitch deafness. *Cognitive Neuropsychology, 30*(5), 311-331. doi: 10.1080/02643294.2013.863183

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia, 49*(5), 961-969. doi: https://doi.org/10.1016/j.neuropsychologia.2011.02.002

Prince, J. B. (2011). The Integration of Stimulus Dimensions in the Perception of Music. *Quarterly Journal of Experimental Psychology, 64*(11), 2125-2152. doi: 10.1080/17470218.2011.573080

Prince, J. B. (2014). Pitch structure, but not selective attention, affects accent weightings in metrical grouping. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 2073-2090. doi: 10.1037/a0037730

Prince, J. B., & Pfordresher, P. Q. (2012). The role of pitch and temporal diversity in the perception and production of musical sequences. *Acta Psychologica, 141*(2), 184-198. doi: https://doi.org/10.1016/j.actpsy.2012.07.013

Puyjarinet, F., Bégel, V., Lopez, R., Dellacherie, D., & Dalla Bella, S. (2017). Children and adults with Attention-Deficit/Hyperactivity Disorder cannot move to the beat. *Scientific Reports, 7*(1), 11550. doi: 10.1038/s41598-017-11295-w

Schultz, B. G., & van Vugt, F. T. (2016). Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behavior Research Methods, 48,* 1591–1607.

Sihvonen, A. J., Ripollés, P., Leo, V., Rodríguez-Fornells, A., Soinila, S., & Särkämö, T. (2016). Neural Basis of Acquired Amusia and Its Recovery after Stroke. *The Journal of Neuroscience, 36*(34), 8872-8881. doi: 10.1523/jneurosci.0709-16.2016

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51*(10), 1952-1963. doi: https://doi.org/10.1016/j.neuropsychologia.2013.06.027

Tranchant, P., Lagrois, M., Bellemare Pépin, A., G. Schultz, B., & Peretz, I. (2018). *Beat alignment test of the motor origin of musical entrainment deficits*. Manuscript in preparation

Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the Beat: A Large Sample Study of Bouncing and Clapping to Music. *PLOS ONE, 11*(7), e0160178. doi: 10.1371/journal.pone.0160178

van Vugt, F. & Schultz, B. G. (2015). Taparduino v1.01. Zenodo, 16178. https://doi.org/10.5281/zenodo.16178

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex, 69*, 186-200. doi: https://doi.org/10.1016/j.cortex.2015.05.002

Vuvan, D. T., Paquette, S., Mignault Goulet, G., Royal, I., Felezeu, M., & Peretz, I. (2017). The Montreal Protocol for Identification of Amusia. *Behav Res Methods*. doi: 10.3758/s13428-017-0892-8

Wechsler, D., Coalson, D. L., & Raiford, S. E. (1997). *WAIS-III: Wechsler adult intelligence scale*: Psychological Corporation San Antonio, TX.

Wilkie, D. (1983). Rayleigh test for randomness of circular data. *Applied statistics, 32*(3), 311-312.

Whiteford, K. L., & Oxenham, A. J. (2018). Learning for pitch and melody discrimination in congenital amusia. *Cortex, 103*, 164-178. doi: https://doi.org/10.1016/j.cortex.2018.03.012

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn Infants Detect the Beat in Music. *Proceedings of the National Academy of Sciences of the United States of America, 106*(7), 2468-2471.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences, 107*(13), 5768-5773. doi: 10.1073/pnas.1000121107

Tables

Table 1. Characteristics of Amusic and Matched Control Participants

Characteristic	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Control Group (Range)
Gender	F	F	F	F	M	M	F	F	F	F	9F/3M
Age (years)	60	25	32	55	31	30	59	58	18	68	42.4 (23-72)
Education (years)	19	16	19	19	19	21	20	15	14	18	17.1 (12-25)
Music Training (years)	2	1	0	5	0	0	0	0	0	1	0.5 (0-3)
Dance Training (years)	0	0	0	0	0	0	0	0	0	1	0.7 (0-6)
Online Test[a]											
Scale (22/30)	**20**	**15**	**19**	23	**21**	22	**14**	22	22	**20**	27 (22-29)
Off-beat (17/24)	19	**15**	18	20	17	19	18	17	19	18	20 (17-21)
Off-key (16/24)	19	**12**	**13**	**15**	16	**13**	16	**9**	**14**	**15**	20 (17-22)
MBEA[a]											
Melodic Composite (21.4/30)	**16.3**	**17.8**	**18**	**20**	**20.3**	**20.3**	**21**	22*	22*	22.7*	-
Rhythm (22/30)	22	**18**	25	25	**18**	22	22	25	24	22	-
Meter (17/30)	20	25	22	25	20	**15**	**16**	26	27	22	-
25 cents pitch-change detection (% accuracy)	30.0	21.1	63.3	53.3	N/A	81.7	10.4	40.6	77.8	57.8	92.1 (75.6-100)[b]

Note: M = male; F = female; MBEA = Montreal Battery of Evaluation of Amusia. [a] Scores in parentheses indicate the cut-off score for each test from Peretz & Vuvan (2017, online test) and Vuvan et al. (2017, MBEA). Score in bold indicates a deficit. * Below cut-off according to earlier norms (Peretz et al., 2003). [b] From an additional control group (n = 30, mean age: 52.5 years-old)

Table 2. Mean Vector Length (VL) and Coefficient of Variation (CV) in the M-BAT Production Task

	A1	A2	A3	**A4**	A5	A6	A7	A8	**A9**	A10	Controls Mean (Range)
VL	0.50	0.78	0.29	0.87	0.22	0.40	0.80	0.25	0.86	0.35	.90 (.83-.95)
CV	0.10	0.10	0.11	0.09	0.14	0.09	0.16	0.09	0.08	0.10	.07 (.06-.09)

Note: Amusic participants in bold were comparable to controls for both the VL and the CV.

Table 3. Number of Trials with Successful Period Matching in the Drum and Original Versions of the Songs in the Amusic Group

Version	A1	**A2**	A3	**A4**	A5	A6	**A7**	A8	**A9**	A10
Drums	0/3	3/3	1/3	3/3	0/3	0/3	3/3	0/3	3/3	0/3
Original	2/4	4/4	1/4	4/4	2/4	2/4	4/4	1/4	4/4	2/4

Note: The four "beat-preserved" participants with pitch deafness are indicated in bold.

Table 4. Individual Pitch-deaf Participants' Mean Vector Length (VL) and Mean Coefficient of Variation (CV) of Tapping Performance to the Drum and Original Versions of the Songs.

Group values for controls are included for comparison.

	VL		CV	
	Drums	Original	Drums	Original
A1	.09	.59	.15	**.10**
A2	.66	.79	.12	.14
A3	.20	.22	.09	.13
A4	**.87**	**.96**	**.07**	**.05**
A5	.10	.27	.14	.12
A6	.17	.58	.09	**.08**
A7	**.95**	**.93**	.10	.15
A8	.23	.21	.12	**.10**
A9	**.81**	**.90**	**.08**	**.10**
A10	.08	.31	.09	**.09**
Controls Mean (Range)	.89 (.67 - .98)	.92 (.77 - .97)	.06 (.04 -.10)	.07 (.05 -.13)

Note: Normal performance in pitch-deaf participants is in bold.

Figure Legends

Figure 1. Participants performance on the M-BAT. A. *d*' scores on the beat perception task of the M-BAT. Error bars represent two standard deviations from the mean. **B**. Percentage of trials with successful period matching on the beat production task of the M-BAT. Each dot represents a participant.

Figure 2. Illustration of Correlation Between the Scale Test Score from the Online Test of Amusia and Performance in the M-BAT. A. Correlation between the Scale test score and *d'* on the beat perception task. **B**. Correlation between the Scale score and the mean logVL on the beat production task. Controls are marked by black dots and pitch-deaf amusics by white dots. Pitch-deaf participants A2, A4, and A9 performed like controls on the M-BAT.

Figure 3. Illustration of the Correlation Between the Scale Test Score and Mean LogVL When Tapping to Drum Rhythms. Controls are marked with black dots and pitch-deaf amusics white dots. Marked pitch-deaf participants A2, A4, A7, and A9 exhibited normal synchronization with the drums' beat.

Figures

Figure 1. Participants' Performance on the M-BAT

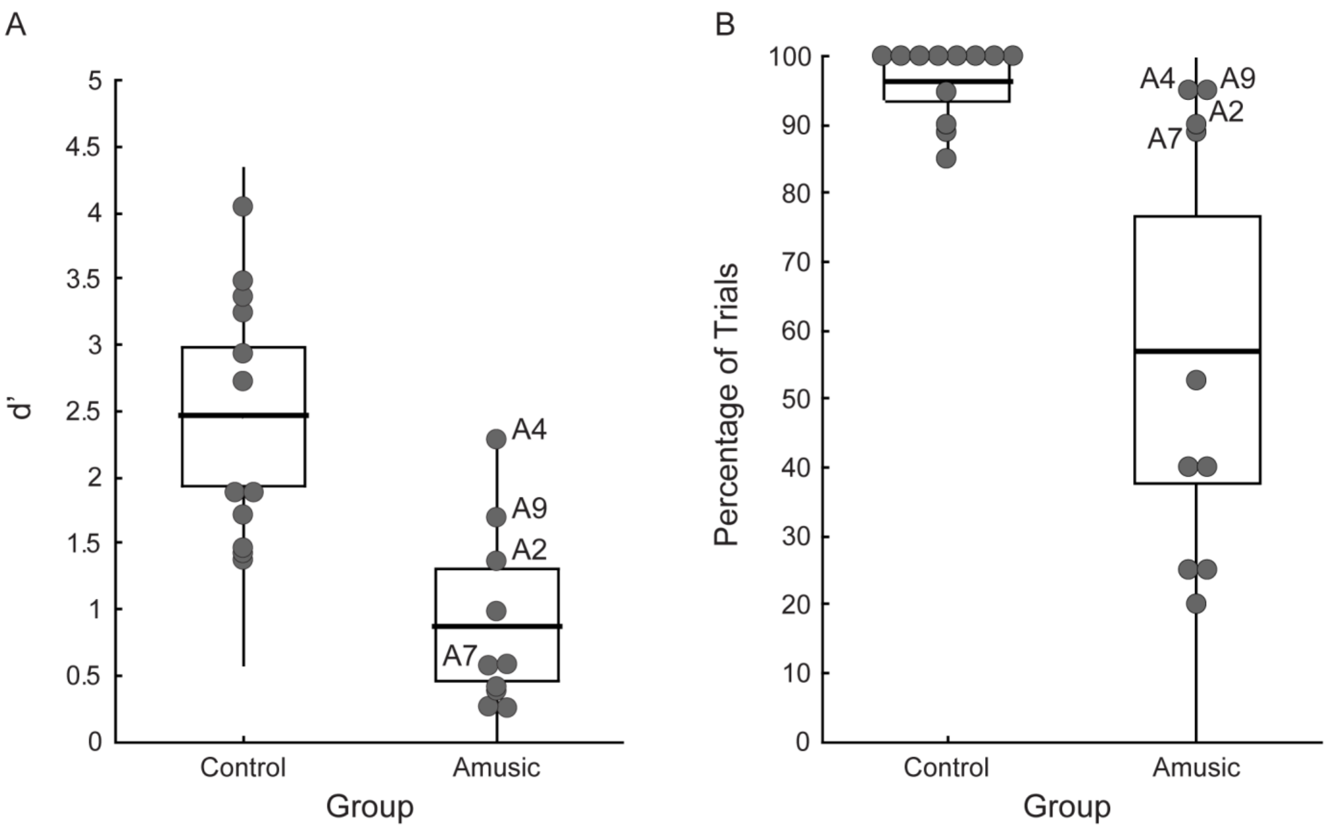


Figure 2. Illustration of the Correlations Between the Scale Test Score from the Online Test of Amusia and Performance in the M-BAT

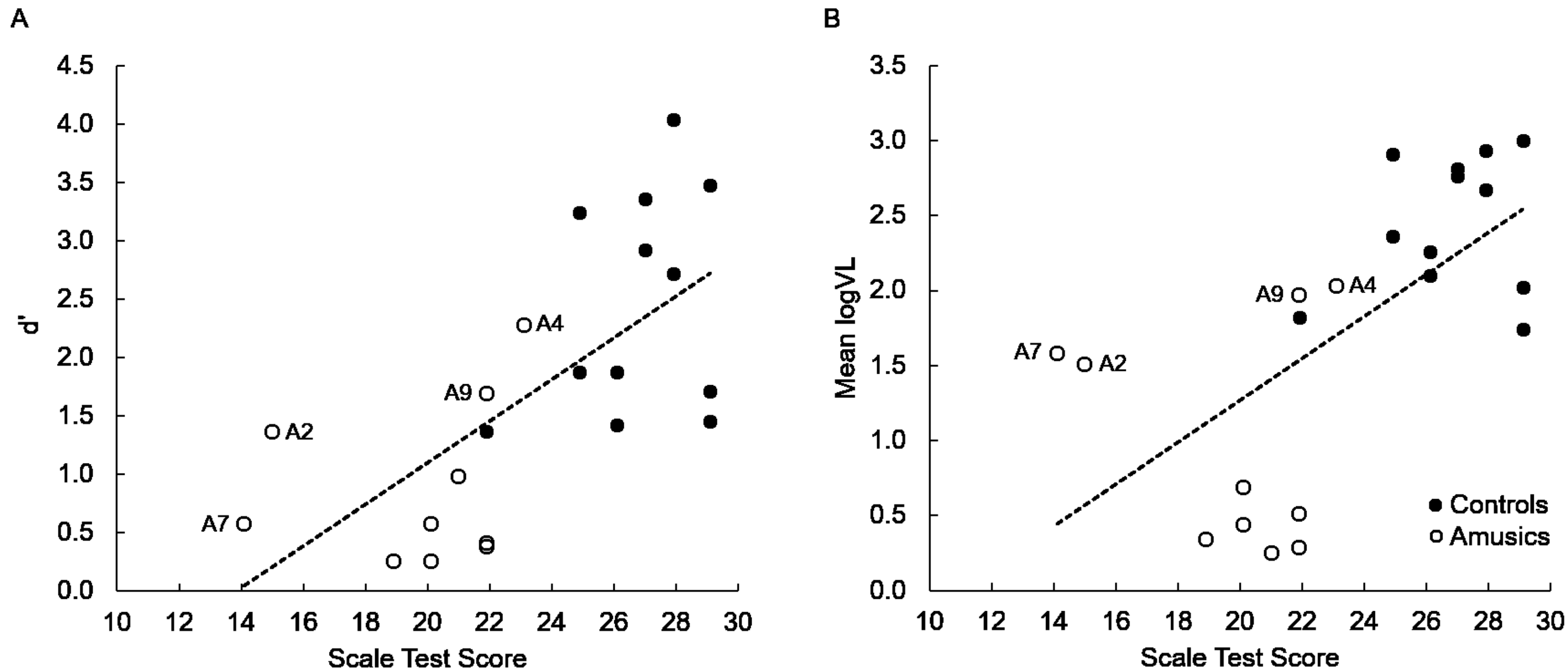
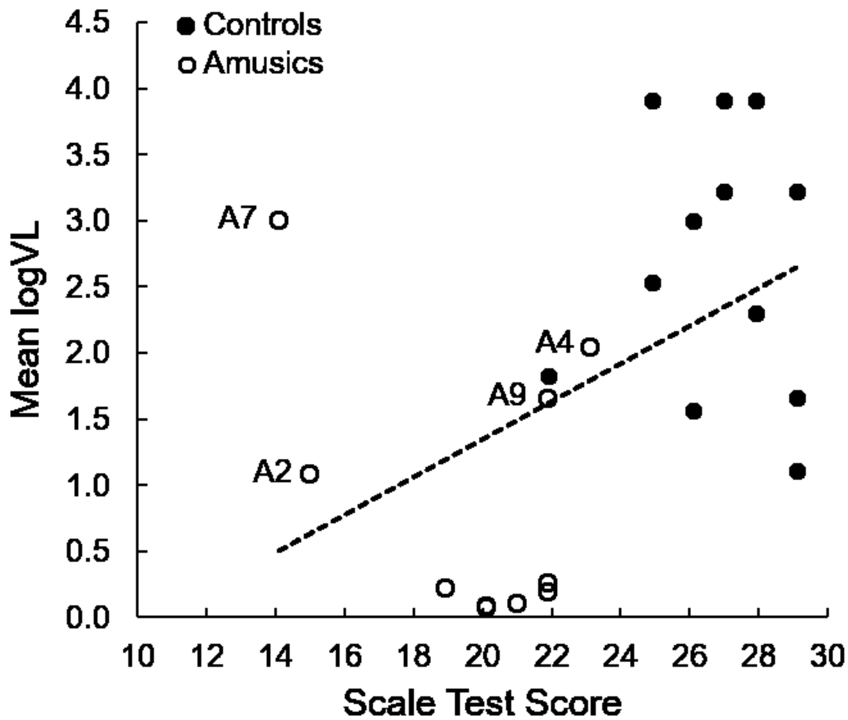


Figure 3. Illustration of the Correlation Between the Scale Test Score and Mean LogVL When Tapping to Drum Rhythms



Article 2 : Musical Beat Finding Deficiencies Generalize to Speech

Marie-Élaine Lagrois[a,b], Caroline Palmer[a,c], and Isabelle Peretz[a,b]

[a] International Laboratory for Brain, Music, and Sound Research

[b] Département de Psychologie, Université de Montréal

[c] Department of Psychology, McGill University

Article en préparation

Abstract

The rhythmic nature of speech may recruit entrainment mechanisms in a manner similar to music. Here we tested the possibility that individuals who display a severe deficit in synchronizing their taps to a musical beat also experience difficulties in entraining to speech. These beat-deaf participants and their matched controls were required to align taps with the perceived regularity in the rhythm of naturally spoken, regularly spoken and sung sentences. The results showed that beat-deaf individuals synchronized their taps less accurately than the control group across conditions. In addition, participants from both groups exhibited more inter-tap variability to natural speech than to regularly spoken and sung sentences. The findings support the idea that acoustic periodicity is a major factor of entrainment in both music and speech. Therefore, a beat-finding deficiency may affect auditory rhythms in general, not just music.

Keywords: music, speech, entrainment, sensorimotor synchronization, beat-finding impairment

1. Introduction

Music is quite unique in the way it compels us to engage in rhythmic behaviors. Indeed, most people will spontaneously nod their head, tap their feet or clap hands when listening to music. Already in infancy, children show spontaneous movements to music (Zentner & Eerola, 2010). This coupling between movements and music is achieved through entrainment. Entrainment can be broadly defined as the tendency of behavioral or brain responses to synchronize with external rhythmic signals (Large & Jones, 1999; Phillips-Silver & Keller, 2012). Currently, the predominant models of entrainment are based on the dynamic attending theory (DAT; Jones, 1976, 1987; Jones & Boltz, 1989; Large & Jones, 1999). According to this theory, the alignment between internal neural oscillators and external rhythms enables listeners to anticipate recurring acoustic events in the signal, allowing for maximum attentional energy to occur at the onset of these events, and thus facilitating the response to these events (Large & Jones, 1999). Building on the DAT, authors have proposed the inclusion of multiple internal oscillators that are hierarchically organized in terms of their natural frequency or period. The possibility for these oscillators to interact permits the extraction of regularities in complex rhythms that can be periodic or quasi-periodic in nature, such as music (Large, 2008; Large & Palmer, 2002; Large & Snyder, 2009). It has been argued that entrainment mechanisms modeled by oscillators would apply not only to music but also to speech (e.g., Cummins, 2009; Cummins & Port, 1998; Giraud & Poeppel, 2012; Goswami, 2012; Lidji, Palmer, Peretz, & Morningstar, 2011a; O'Dell & Nieminen, 1999; Peelle & Davis, 2012; Port, 2003; Wilson & Wilson, 2005).

Tracking rhythmic structure is an important component of both music and speech, where rhythm is defined as the temporal organization of events' duration into complex patterns (Patel, 2008; Large, 2008; Large & Palmer, 2002). However, one major difference between music and speech is the periodic nature of the rhythmic patterns underlying their temporal structure. Accordingly, musical organization may rely on unique entrainment mechanisms compared to speech processes (Haegens & Zion Golumbic, 2018; London, 2012; Patel, 2008; Patel & Iversen, 2014).

The periodicities contained in musical rhythms typically induce the perception of a beat, that is, the sensation of a regular pulsation, on which timed behaviors are built (Lerdahl &

Jackendoff, 1983). Simple movements, like taps, are usually produced within a few tens of milliseconds from the beat onset, indicating the precision of the temporal predictions made about the timing of upcoming beats (Repp, 2005; Repp & Su, 2013; Van Der Steen & Keller, 2013). Listeners can extract the beat from various complexity of rhythms, without the need for a one-to-one correspondence between acoustic events and beat occurrences (Chapin et al., 2010; Drake, Jones, & Baruch, 2000; Large, Herrera, & Velasco, 2015; Repp, Iversen, & Patel, 2008), and across a range of tempi (~94–174 beats per minute; London, 2002; McAuley, 2010; McAuley, Jones, Holub, Johnston, & Miller, 2006; Repp, 2003; Repp, 2005). Beat extraction is also robust to moderate tempo fluctuations (Drake, Penel, & Bigand, 2000; Large & Palmer, 2002; Palmer, 1997). Beat induction from music has in fact been proposed as one of the fundamental and universal traits of music (Honing, 2012; Iversen, 2016).

Musical meter, which corresponds to the hierarchical organization of beat, where some beats are perceived as stronger than others, leads to higher-order periodicities of strong and weak beats (for example, a march versus a waltz; Lerdahl & Jackendoff, 1983). Speech, like music, has a hierarchically organized temporal structure, with phonemes, syllables and prosodic cues occurring each at different time scales (Brown, Pfordresher, & Chow, 2017; Leong, Stone, Turner, & Goswami, 2014; Meyer, 2017; Peelle & Davis, 2012). Similarly to music, a metric hierarchy in speech may rely on the occurrence of stressed or accented acoustic events, typically associated with syllables (Cummins & Port, 1998; Kotz & Schwartze, 2010; Port, 2003; Selkirk, 1986; Turk & Shattuck-Hufnagel, 2013). Stress pattern in speech additionally varies and depends on different acoustic cues according to language. The metric of "stress" languages, like English for example, is usually clearer than the metric of "syllabic" languages like French (Liberman & Prince, 1977; Lidji et al., 2011a). Nevertheless, temporal intervals between stressed syllables are not as regular in speech as in music (Dauer, 1983; Jadoul, Ravignani, Thompson, Filippi, & de Boer, 2016; Nolan & Jeon, 2014; Patel, 2008; Turk & Shattuck-Hufnagel, 2013).

There is behavioral support to entrainment to speech. Speakers entrain to one another's syllable rate in conversational turn taking (Schultz et al., 2015; Wilson & Wilson, 2005); they adapt their speech rate to match another speaker (Borrie & Liss, 2014; Jungers, Palmer, & Speer, 2002); they can adjust the timing of their speech to fit a metronome (Cummins & Port, 1998); In a prior study (Lidji et al., 2011a) using a similar experimental design as the present one, French and English monolingual speakers and French-English bilingual speakers were invited to tap their finger along with the beat they perceived in French and English sentences spoken with natural prosody. It was found that the utterances' variability in intervocalic intervals (that is, the variability of interval duration between vowel of syllables) predicted interval variability between the taps produced by the participants.

While there is evidence of entrainment to speech, a puzzling difference exists between the absence of synchronous ("chorus") speech and the widespread coordination of movements to music. To address this issue, Cummins (2011, 2013) proposed that synchronous speech should be possible because speakers of the same language have mastered the association between motor actions and speech sounds of their language and share their knowledge of speech timing. He supports his claim by evidence showing that speakers can synchronize while reading an unfamiliar text without prior practice, which the author considerer an indication of aperiodic synchronization (Cummins, 2002a, 2002b, 2009; Cummins, Li, & Wang, 2013). According to this perspective, sensorimotor coordination, whether in the context of synchronous speech or head bopping to music, would reflect a general ability of humans to time, or adapt, their action with the rhythm of an external event.

Entrainment to speech and music has rarely been compared. Only two prior studies have done so. In one of these studies (Dalla Bella, Białuńska, & Sowiński, 2013), the influence of music and speech on entrainment has been assessed through interference. The main task was to synchronize finger taps to a metronome while hearing highly isochronous computer-generated music or regularly spoken poems. When the metronome tones and musical beats or stressed syllables were perfectly aligned, higher variability in the asynchronies between taps and metronome was found with the speech distractor compared to the musical one. When misaligned, both music and speech led to synchronization interference by increasing the asynchrony between taps and metronome onsets compared to when the metronome and the target were aligned. Still, music induced larger asynchronies and more variability in synchronization accuracy (variability in asynchronies). In a second experiment, the stimuli were better matched: songs, either sang with lyrics, sang with a single syllable, or spoken with a regular pace were presented. In this condition, whether the stimuli were spoken or sung had identical detrimental effect on tapping to the metronome. Therefore, when isochrony is equated between music and speech, entrainment occurs. However, natural speech is typically not isochronous. In the second study (Lidji, Palmer, Peretz, & Morningstar, 2011b), using the same paradigm as used here, native French and English speakers tapped along with French and English sentences in three conditions: naturally spoken, regularly spoken and sung with a simple melody. The inter-taps intervals were more variable in the naturally spoken sentences compared to the other conditions. The taps were also more closely aligned to the beat (the nearest implied metronome click to which the singer had synchronized to produce the stimuli) of sung than regularly spoken sentences. Tapping was also less regular to English than French stimuli. These results show an overall effect of regularity on entrainment, with music being more attuned to elicit entrainment than regular speech.

Here, we tested the same material as the one used by Lidji and collaborators (2011b) with individuals who have a documented deficit in finding the beat in music. This disorder is characterized by an inability to synchronize a whole-body movement, clapping or tapping, to the beat of music (Bégel et al., 2017; Dalla Bella & Sowiński, 2015; Phillips-Silver et al., 2011; Sowiński & Dalla Bella, 2013; Tranchant, Vuvan, & Peretz, 2016), musical rhythms (Palmer et al., 2014) or to amplitude-modulated noise derived from music (Sowiński & Dalla Bella, 2013), in the absence of intellectual disability or acquired brain damage. The study of this population provides an opportunity to test the domain specificity of entrainment. If the beat-finding disorder, initially diagnosed with music also disrupts entrainment to speech, that association will provide support for the domain-general nature of the entrainment mechanism to auditory rhythms.

More precisely, thirteen beat-deaf individuals and 13 neurotypical controls were invited to tap to spoken and sung sentences. If entrainment abilities are domain general, then neurotypical participants' tapping period should align less well with the intervocalic period between syllables in natural speech than in rhythmically regular speech and sung sentences, whereas beat-deaf participants should show deficits in tapping to both rhythmically regular speech and songs. If, on the contrary, entrainment is domain-specific, it should affect the perception of regularity in both speech and music but with a more pronounced effect with sung than with regularly spoken sentences, leaving tapping to natural speech intact compared to the neurotypical participants.

2. Method

2.1 Participants

Thirteen beat-deaf French-speaking adults (10 females) and thirteen French-speaking matched control participants (11 females) took part in the study. The groups were matched for age, education, and years of music and dance training (detailed in Table 1). One beat-deaf participant was completing an undergraduate degree in contemporary dance at the time of testing. Accordingly, a trained contemporary dancer was also included in the control group. All participants had normal verbal auditory working memory and non-verbal reasoning abilities, as assessed by the Digit Span and the Matrix Reasoning tests from the WAIS-III (Wechsler Adult Intelligence Scale; Wechsler, Coalson, & Raiford, 1997; see Table1). All participants were non-musicians and had no history of neurological, cognitive, hearing or motor disorders. Participants provided written consent to take part in the study and received monetary compensation for their participation. All procedures were approved by the Research Ethics Council for the Faculty of Arts and Sciences at the Université de Montréal.

[Insert Table 1 here]

2.1.1 Procedure Prior to Inclusion of Participants in the Study

Participants in the beat-deaf group had taken part in previous studies in our lab (Lagrois & Peretz, 2018, Tranchant et al., 2016 or Tranchant et al., 2018) and were identified as being unable to synchronize simple movements to the beat of music or self-declared as unable to synchronize to the beat of music (participants B6 and B8 only, in Table 2). Control participants had either taken part in previous studies in the lab or were recruited via online advertisements in Montreal's general population and on campus advertisement at the University of Montreal.

Inclusion in the current study was confirmed by performance on the Montreal Beat Alignment Test (M-BAT; Tranchant et al., 2018). This test comprises two tasks, a beat production task and a beat perception task. In the beat production task, participants are asked to align taps to the beat of 10 song excerpts from various musical genres and tempi. Tempo varies from 82 beats per minute (bpm) to 170 bpm. Each song is presented twice, for a total of 20 trials. In the beat perception task, participants hear the same song excerpts superimposed with an isochronous click track and have to decide if the clicks are aligned with the beats of each song or not. Each song is presented in eight conditions, for a total of eighty trials. On half of the conditions the clicks are "on beat" and on the other half "off-beat" by either a phase shift or a period shift.

In the beat production task, control participants successfully matched the period of their taps to the songs' beat in at least 85% of the trials (*M* = 96.9%, *SD* = 5.2%). Successful period matching is determined with a p-value smaller than .05 on the Rayleigh z test of periodicity. In the beat-deaf group the average percentage of trials with successful tempo matching was 39.2% (*range of mean values*: 10–65%, *SD* = 18.3%). As shown in Figure 1A, there was no overlap between the groups' performance on this task, confirming the deficit of participants in the beat-deaf group to synchronize their taps to the beat of music. For the beat perception task, a measure of sensitivity of discrimination (*d'*) was computed, with hits corresponding to correctly detected "off-beat" trials and false alarms as off-beat responses to "on-beat" trials. Higher *d'* values indicate better discrimination. Participants from the control group had on average a *d'* of 2.5 (*SD* = 0.8, *range of individual values*: 1.4-3.5). Individuals in the beat-deaf group had on average a *d'* of 0.9 (*SD* = 0.6, *range of individual values*: 0.2-2.0). Scores on the perceptual task are shown in Figure 1B. Only three out of the thirteen beat-deaf participants had a *d'* score lying in the low but normal range of the control participants. Thus, ten beat-production-impaired participants were also impaired in musical beat perception. Thus, a deficit in beat perception may be at the origin of the impairment in sensorimotor synchronization displayed by most beat-deaf participants.

[Insert Figure 1 here]

Prior to their participation in the current study, participants completed the online test of amusia to screen for the presence of a musical pitch perception impairment (Peretz & Vuvan, 2017). The online test is composed of three tests: Scale, Off-beat, and Off-key. The Scale test requires the comparison of 30 pairs of melodies that differ by an out-of-key note in half of the trials. The Off-beat and Off-key tests consist in the detection of either an out-of-time or an out-of-key note, respectively. A score lying 2 *SD* below the mean of a large population on both the Scale and Off-key tests indicates the likely presence of pitch deafness (Peretz & Vuvan, 2017; Vuvan et al., 2017). Based on the data from Peretz and Vuvan (2017), a cut-off score of 22 out of 30 was used for the Scale test, and of 16 out of 24 for the Off-key test. Table 2 indicates the individual scores of beat-deaf participants on the online test. Seven participants in the beat-deaf group were below the cut-off on both the Scale and the Off-key tests. As these cases of beat-deaf participants could also be considered pitch-deaf, the influence of musical pitch perception will be taken into account in the analysis and interpretation of the results. All control participants had scores above the cut-offs, as this was an inclusion criterion.

[Insert Table 2 here]

2.2 Stimulus Materials

The 12 French sentences used in this experiment were taken from Lidji et al. (2011b). Each sentence contained 13 monosyllabic words and was recorded in three conditions as depicted in Figure 2. The recordings were made by a native Quebec French female speaker in her twenties who had singing training. Recordings were made with a Neumann TLM 103 microphone in a sound-attenuated studio. In the *naturally spoken* condition, the sentences were recorded with a natural prosody (generating non-periodic pattern of stressed syllables). In the *regularly spoken* condition, sentences were recorded by the speaker to align every other syllable with the beat of a metronome set to 120 bpm, heard over headphones. In the *sung* condition, the sentences were sung by the speaker, again with every other syllable aligned to a metronome at 120 bpm, heard over headphones.

In the sung condition, the speaker sang each sentence to a simple melody, aligning each syllable with one note of the melody. Twelve unique melodies composed in the Western tonal style in binary meter, in major or minor modes, were taken from Lidji et al. (2011b) and were novel to all participants. Each sentence was paired with two different melodies. Participants only heard one melody version of each sung sentence, which was counterbalanced between subjects.

Additional trials for all three conditions (natural speech, regular speech, sung) were then created from the same utterances at a slower rate (80% of stimuli's original rate, i.e. around a tempo of 96 bpm) by using software Reaper (v4.611, 2014; time stretch mode élastique 2.28 SOLOIST: speech). This was done to ensure that the beat-impaired participants could adapt their taps to the rate of each stimulus and could comply with the task. Since preliminary analyses indicated that all participants from both groups adapted the general rate of their taps between the original stimulus rate and slowed stimuli, those data for the slower stimuli are not reported here for the sake of clarity. All the stimuli were edited to have a 400 ms silent period before the beginning of the sentence and a 1000 ms period at the end of the sentence. Stimuli were also equalized in root mean square intensity (RMS).

[Insert Figure 2 here]

Table 3 describes the acoustic features related to the rhythmic structure of the stimuli in each condition. Phonemes boundaries were marked by hand using Pratt (Boersma & Weenink, 2017) and classified as vowels or consonants based on criteria defined by Ramus, Nespor, & Mehler (1999). Note that the analyses reported below include the stimuli at the original rate only. Once the segmentation completed, a MATLAB script was used to export the onset, offset, and duration of vocalic (a vowel or a cluster of vowels) and consonantal (a consonant or a cluster of consonants) intervals. The Normalized Pairwise Variability Index for Vocalic Intervals (V-nPVI), an indication of duration variability between successive vowels (formula from Grabe & Low, 2002), was used to measure the rhythmic characteristics of the stimuli. Higher vocalic nPVI indicates higher contrast in duration between consecutive vocalic intervals. Comparisons of sentences in the naturally spoken, regularly spoken, and sung conditions showed a significant difference between conditions, $F(2, 22) = 21.6$, $p <.001$, $\eta^2 = .66$. The nPVI was higher in the naturally and regularly spoken conditions compared to the sung condition (see Table 3). The coefficient of variation (CV; *SD*/mean) of intervocalic intervals (IVI; vowel onset to onset) is another indication of rhythmic variability (Lidji et al., 2011a). Small CV(IVI) indicates similar time intervals between vowel onsets across the sentence. Here we indicate the CV between every other syllable. There was also a significant difference between conditions in this measure, $F(2, 22) = 64.6$, $p < .001$, $\eta^2 = .85$. Naturally spoken sentences had the largest timing variations between vowel onsets ($M = .21$), followed by regularly spoken sentences ($M = .08$), while sung sentences showed the smallest variability ($M = .05$). To ensure the female performer was as accurate in timing the sentences with the metronome in the regularly spoken and sung conditions, the relative asynchrony between each vowel onset and the closest metronome pulsation it was supposed to match was measured. In this context, a negative mean asynchrony indicates that the vowel onset preceded the metronome tone onset, while a positive asynchrony means that the vowel onset followed the metronome tone (see Table 3). There was no significant difference between conditions, indicating similar timing with the metronome in the regularly spoken and sung conditions, $t(11) = 1.1$, $p = .28$.

[Insert Table 3 here]

2.3 Design and Procedure

Participants performed three tasks. First, they performed a spontaneous tapping task to assess their natural tapping rate (mean and variance) in the absence of a pacing stimulus. They were asked to tap as regularly as possible for 30 seconds, as if they were a metronome or the "tic-toc" of a clock (as used by McAuley et al., 2006). Participants were asked to tap with the index finger of their dominant hand. Next, participants performed the tapping task with the spoken/sung sentences, as described below. Then the participants repeated the spontaneous tapping task to determine whether their spontaneous rate had changed, and finally, they tapped at a fixed rate with a metronome set to 120 bpm (inter-beat interval of 500 ms) and 96 bpm (inter-beat interval of 625 ms), chosen to match the tempi of the spoken/sung stimuli used in the experiment. The experimentation had a total duration of approximately 60 minutes.

In the spoken/sung tapping blocks, each participant was presented with 12 naturally spoken sentences at the original rate (120 bpm) and 6 naturally spoken sentences at the slowest rate (96 bpm), 12 regularly spoken sentences at the original rate and 6 at the slower rate, 12 sung sentences at the regular rate and 6 sung sentences at the slower rate. These stimuli were mixed and divided in three blocks of 18 trials each. Two pseudo-random orders were created such that not more than two consecutive sentences from the same condition could occur and that the same sentence was never repeated. On each trial, participants first listened to the stimulus; then, for two additional presentations of the same stimulus, they were asked to tap along to the beat that they perceived in the stimulus (as in Lidji et al., 2011b). The action to perform (listen or tap) was prompted by instructions displayed on a computer screen. Participants pressed a key to start the next trial. Before the task, a demonstration video was presented to participants, which showed an individual finger tapping on the sensor with one example stimulus from each condition. In the demonstration, a different sentence was used for each condition, and was presented at a different rate (84 bpm or 108 bpm) than the ones used in the experiment. The example sung sentence was also presented with a different melody than the one heard by the participant in the task. After the demonstration, participants completed a practice trial for each type of sentence.

For the metronome task, there were two trials at each metronome tempo (120 bpm and 96 bpm) and the presentation order of the two metronome tempi was counterbalanced across participants. Each metronome stimulus contained sixty 50 ms 440 Hz sine wave tones. Before tapping to the metronome, participants listened to seven tones, to have priming comparable to the spoken/sung tapping task. A practice trial was also first performed with a metronome set to 108 bpm. Since all participants could adapt their tapping rate to the stimuli at 120 bpm as well as 96 bpm, only the results to the metronome at 120 bpm (rate of the original speech stimuli) are reported here.

The experiment took place in a large sound-attenuated studio. The tasks were programmed with MAX/MSP (https://cycling74.com). Taps were recorded on a square force-sensitive resistor (3.81 cm, Interlink FSR 406) connected to an Arduino UNO (R3; arduino.cc) running the Tap Arduino script (fsr_silence_cont.ino; Schultz & van Vugt, 2016; van Vugt & Schultz, 2015) transmitting timing information to a PC (HP ProDesk 600 G1, Windows 7) via the serial USB port. The stimuli were delivered at a comfortable volume level through closed headphones (DT 770 PRO, Beyerdynamic) controlled by an audio interface (RME Fireface 800).

3. Data Analyses

3.1 Tapping Data Preprocessing

The first five taps produced in the spontaneous tapping task were discarded and the 30 following ITIs were used, following McAuley et al.'s (2006) procedure. If participants produced fewer than 30 taps in the spontaneous tapping task, the measures included all taps produced (the smallest number of taps produced was 16 taps in this task). Due to recording problems, taps were missing from one beat-impaired participant's first spontaneous tapping trial.

Recorded taps were first pre-processed to remove inter-tap intervals (ITIs) smaller than 100 ms in the spontaneous tapping task and ITIs smaller than 150 ms in the spoken/sung tapping task and the metronome task. In the three tasks, taps were also considered outliers and were removed if they were more than 50% smaller or larger than the median ITI produced by each participant (median ITI ± [median ITI*0.5]), similar to the difference between musical beat duration categories or different levels of beat period. As a result, 1.6% of the taps were removed across groups (*range*: 0.0-6.4%) in the spontaneous tapping task. In the spoken/sung tapping task, 0.85% of taps per trial were removed (range: 0 - 36.4% taps/trial). In the metronome task, 5.27% of taps were removed on average (*range*: 3.4-8.1%), leaving between 54 and 76 taps per trial and participant for analysis. In the metronome task, the first 50 taps produced by each participant were used for the analysis.

3.2 Analysis of Tapping Data

The mean ITI was calculated for all tapping tasks. In the spoken/sung tapping task, since each participant tapped twice on each utterance in succession, the mean ITIs per stimulus were averaged across the two presentations. However, in 0.16% of the trials participants did not tap at the same hierarchical level in the two presentations of the stimulus (for example they tapped on every syllable in the first presentation, and every other syllable in the second presentation). It was decided not to include these trials in the calculation related to mean ITI and CV, to avoid averaging together groups of taps with differing mean ITI. Nevertheless, at least 11 of the 12 trials at 120 bpm per participant per condition could be included in the analyses. For the comparison of mean ITI between groups and conditions, the mean ITIs were scaled to the ITI that would correspond to tapping once every two words. In the metronome task, data were also averaged across the two presentations.

In the spoken/sung tapping task, inter-tap variability (CV; *SD* of ITI/mean ITI) was computed for each condition. Table 3 indicates that the CVs of taps to naturally spoken sentences should be larger than the CVs to regular stimuli. To assess this, we examined how produced inter-tap intervals matched the stimulus intervocalic intervals (as done previously by Chen, Penhune, & Zatorre, 2008b; Giovannelli et al., 2014; Leow, Parrott, & Grahn, 2014). ITI deviation was calculated by averaging the absolute difference between each ITI and the corresponding IVI of the stimulus. To control for differences in IVI for each stimulus, the ITI deviation was normalized to the mean IVI of that stimulus and converted to a percentage of deviation (% ITI deviation). The formula used was:

$$\% \text{ ITI deviation} = ([[\Sigma\, |ITI_x - IVI_x|]/\text{nb ITI}]/\text{mean IVI}) \times 100$$

This measure of period deviation gives an indication of how participants' taps matched the rhythmic structure of the stimuli whether regular or not.

Period-matching between spoken/sung sentences and taps was further assessed for the stimuli that contained regular beat periods (i.e., regularly spoken and sung stimuli, and the metronome stimuli), with circular statistics using the Circular Statistics Toolbox for MATLAB (Berens, 2009). With this technique, taps are transposed as angles on a circle from 0 to 360 degrees, where a full circle corresponds to the period of the intervocalic interval of the stimulus. The position of each tap on the circle is used to compute a mean resultant vector. The length of the mean resultant vector (Vector Length, VL) indicates how clustered the data points are around the circle. Values of VL range from 0 to 1; the larger the value, the more clustered together are the points on the circle, indicating that the time interval between taps tends to match the IVI of the stimulus with more consistency. For statistical analyses, we used a logit transform of vector length, as this measure is typically skewed in synchronization data ($\text{logVL} = -1*\log(1-\text{VL})$). However, for simplicity, untransformed vector length is reported when considering group means and individual data. The Rayleigh z test of periodicity was employed to test if a participant's taps period-matched the IVI of each stimulus consistently (Wilkie, 1983). A significant

Rayleigh test (p-value <.05) demonstrate successful period matching. An advantage of the Rayleigh test is that it considers the number of taps available in determining if there is a significant direction in the data or not (Berens, 2009). Using linear statistics, the accuracy of synchronization was measured by the mean relative asynchrony between taps and beats' onset time in milliseconds. Note that this measure only included trials for which participants could successfully match the inter-beat interval of the stimuli, as assessed by the Rayleigh test, since the asynchrony would otherwise be meaningless.

The period used to perform the Rayleigh test was adjusted to fit the hierarchical level at which participants tapped on each trial. Since the stimuli had a tempo of 120 bpm, this meant that if a participant tapped to every word, the period used was of 250 ms, if a participant tapped every two words 500 ms, and every four words 1000 ms. This was done to avoid having artificial bimodal distributions and variance. Given this adaptation, in the spoken/sung tapping task, we first looked at the closest hierarchical level at which participant tapped. This was approximated based on the tapping level that fitted best the majority of ITI within a trial (i.e. the modal tapping level based on ITIs).

3.3 Correlation between pitch perception, musical beat finding and tapping to spoken/sung sentences

In order to assess the contribution of musical pitch perception and musical beat processing to synchronization with the sentences, the scores from the online test of amusia, which include the Scale, Off-key and Off-beat test, and from the M-BAT perception task, were correlated with measures of tapping variability (CV) and period-matching (% of ITI deviation) from the spoken/sung tapping task.

3.4 Statistical Analyses

Statistical analyses were performed in SPSS (IBM SPSS Statistics, version 24, 2016). A mixed repeated-measures ANOVA with Group as the between-subjects factor was used whenever the two groups were compared on a dependent variable with more than one condition.

Because of the small group sample size, a statistical approach based on sensitivity analysis was applied. This was done to ensure that significant effects were reliable when assumptions on residuals' normality distribution and homogeneity of variance were violated (Thabane et al., 2003). When these assumptions were violated, the approach employed was as follows: 1) inspect residuals to identify outliers (identified using Q-Q plot and box plot), 2) re-run the mixed-design ANOVA without the outliers and assess the consistency of the previous significant results, 3) confirm the results with a non-parametric test of the significant comparisons (Thabane et al., 2003). If the effect was robust to this procedure, the original ANOVA run was reported. Bonferroni correction was used for post-hoc comparisons. Otherwise, group comparisons were performed with the Welch's test, which corrects for unequal variance. Paired t-tests were utilized for within-group comparisons on a repeated measure with only two conditions. Effect sizes are reported for all comparisons with p-value smaller than .50 (Kover & Atwood, 2013). To indicate the estimated effect sizes, partial eta-squared values are reported for repeated-measures ANOVA and Hedge's *g* statistic was computed for the other comparisons.

4. Results

4.1 Spontaneous Tapping

The mean ITI of the spontaneous tempo ranged between 365 and 1109 ms in control participants and from 348 to 1443 ms in the beat-impaired group. There was no significant group difference in the mean ITIs, $F(1,23) = 1.2$, $p = .27$, $\eta^2 = .05$, including spontaneous tapping performed before and after the spoken/sung tapping task. On the CV of spontaneous tapping, a main effect of Group emerged, $F(1,23) = 18.2$, $p < .001$, $\eta^2 = .44$, with no effect of Time, $F(1,23) = 0.30$, $p = .59$, nor interaction, $F(1,23) = 0.19$, $p = .67$. The CV of spontaneous tapping was higher for beat-impaired participants than for control subjects (see Table 4). Thus, beat-impaired individuals showed more inter-tap variability than control participants when trying to tap regularly without a pacing stimulus.

[Insert Table 4 here]

4.2 Tapping to Speech and Song

As expected, participants' inter-tap variability (CV) for the naturally spoken sentences was higher than the CV in the other two conditions. Figure 3A depicts the mean CV for the stimulus materials and Figure 3B depicts the mean CV for both groups in each condition. The CV of participants' taps was larger for the naturally spoken sentences (M = .13) than for the regularly spoken (M = .10) and sung (M = .10) sentences. The groups did not differ, $F(1,24) = 2.8$, $p = .10$, $\eta^2 = .11$, and exhibited a main effect of Condition, $F(1.5,35.4) = 15.2$, $p < .001$, $\eta^2 = .39$, with no interaction between these factors, $F(1.5, 35.4) = 0.6$, $p = .52$. However, there were no significant correlations between the CV of taps and the CV of the stimuli across conditions ($r_{(154)}$ = -.05-.12, $ps > .15$). One control participant had a larger CV than the rest of the group in natural speech. Three beat-impaired participants also had larger CVs across conditions. Removing the outliers did not change the results of the analysis. Thus, the inter-tap variability only discriminated natural speech from the regularly paced stimuli, for both listener groups.

[Insert Figure 3 here]

In general, control participants better matched the stimulus period than did beat-impaired participants, whether the stimuli were regular or not, as indicated by a smaller percentage of deviation between the inter-tap period produced and the corresponding IVI in the stimuli (% ITI deviation; see Figure 4). This was supported by a main effect of Group, $F(1,24) = 8.2$, $p = .008$, $\eta^2 = .26$, a main effect of Condition, $F(1.4,32.5) = 95.9$, $p < .001$, $\eta^2 = .80$, and no interaction, $F(1.4,32.5) = 0.2$, $p = .74$. Post-hoc comparisons showed a significant difference between all conditions: the % ITI deviation was the largest for naturally spoken sentences (20.9% and 27.2% for the control and beat-deaf group, respectively), followed by regular speech (11% and 18.2%) and sung sentences (8.5%, and 15.7%; see Figure 4). Removing outliers did not change the results.

[Insert Figure 4 here]

In order to measure synchronization more precisely, we first examined the hierarchical level at which participants tapped. A chi-squared analysis on the number of participants who tapped at each hierarchical level (1, 2, or 4 words) by Condition and Group indicated a main effect of Group, $\chi^2(2, 78) = 7.4$, $p = .02$. There was no effect of Condition and no interaction between Group and Condition, *ps*> .1. In both groups, participants tapped preferentially every two words (see Figure 5) and were consistent in the hierarchical level chosen for tapping across conditions. The hierarchical level at which a participant tapped determined the period used in the following analysis of synchronization to the regular stimuli.

[Insert Figure 5 here]

The average percentage of trials with successful period matching (by Rayleigh's test) for the control group was 91.7% (range: 58–100%) for regularly spoken sentences and 90.4% (range: 50–100%) for sung ones. In the beat-impaired group, the mean percentage of successful period-matched trials was much lower, with 30.4% (range: 0–75%) and 23.8% (range: 0–66.7%) for regularly spoken and sung sentences, respectively. The percentage of trials with successful period matching did not differ between the regular and sung conditions, $t(25) = 1.3$, $p = .21$, $g = .10$.

We next examined if synchronization was more consistent and accurate for sung than regularly spoken sentences. These analyses were conducted on trials for which participants were able to synchronize successfully with the beat period (i.e. Rayleigh p-value <.05). Because most beat-impaired participants failed to synchronize with the stimuli, the analyses are limited to the control group, which revealed no significant difference between conditions, $t(12) = -0.8$, $p = .46$, $g = .09$. Thus, tapping was as regular or constant with regular spoken sentences as sung ones. Accuracy was assessed with the mean relative asynchrony between taps and beat in milliseconds. Control participants anticipated the periods' onset of sung sentences significantly earlier ($M = -14$ ms, range: -51 ms to 19 ms) than the regularly spoken sentences ($M = 1$ ms, range: -30 ms to 34 ms), $t(12) = 3.8$, $p = .003$, $g = .74$. This result suggests that beat onsets were better anticipated in sung sentences than in regularly spoken ones, corroborating results found by Lidji et al. (2011b). The two beat-impaired participants who could successfully period match the stimuli on more than 50% percent of the trials showed similar consistency and accuracy of synchronization as control participants.

4.3 Tapping to Metronome

All participants could successfully match their taps to the period of the metronome, as assessed by the Rayleigh test, except for one beat-impaired participant who tapped too fast compared to the 120 bpm tempo (mean taps IOI = 409 ms for a metronome IOI of 500 ms). Thus, this participant and his matched control were removed from subsequent analyses in this task. As in previous analyses, control participants had smaller inter-tap variability than beat-impaired participants. This was confirmed by a group comparison with Welch's test on the CV, $t(14.0) = 11.7$, $p = .004$, $g = 1.35$ (Control group's CV: $M = .06$, $SE = .003$; Beat-impaired group's CV: $M = .09$, $SE = .01$). Period-matching consistency, using the log transform of the mean vector length, also showed a significant group difference, $t(22.0) = 9.3$, $p = .006$, $g = 1.20$. The difference between groups was not significant, however, for the mean relative asynchrony between taps and metronome tones, $t(20.4) = 0.1$, $p = .80$ (Control group's mean asynchrony: $M = -56$ ms, range: -120 ms to 0 ms; Beat-impaired group's mean asynchrony: $M = -53$ ms, range: -104 ms to -11 ms).

4.4 Relationship Between Music Perception and Entrainment to Utterances

To assess the impact of music perception on tapping performance, we correlated the online test scores and M-BAT *d'* perception test with tapping variability (CV) and period matching (% of ITI deviation) since these measures were computed for all conditions (Table 5). No indication of correlation between CV and musical tests was found. In contrast, the % of ITI deviation was significantly related to beat perception scores, but not to the Off-beat test, which is an anisochrony detection task (Tranchant & Vuvan, 2015). Therefore, beat perception, and not anisochrony detection, could be predictors of entrainment, regardless of the stimulus type. Musical pitch perception was also associated with the % of ITI deviation in the regular speech and sung conditions, although less strongly than beat perception.

[Insert Table 5]

Looking at correlations within each group, the correlations with the M-BAT beat perception test were stronger in the control group ($r = -.36$ to $-.77$), than in the beat-deaf group

(r = -.27 to .26). On the other hand, the significant association between the Scale test and the %ITI deviation was mostly driven by the beat-deaf group (r = -.61) compared to the control group (r = -.05). It should be mentioned, however, that the score on the M-BAT perception test and the score on the Scale test were highly correlated, r = .81, p <.001 (control group: r = .37; beat-deaf group: r = .56). These results question whether beat-deaf individuals with an additional deficit in pitch perception may have a more severe beat perception deficit. If we compare beat-deaf participants with beat-deaf individuals having a concomitant musical pitch perception deficit, the difference between groups is indeed significant on the M-BAT beat perception test, t(6.8) = 7.5, p = .03, g = 1.5. However, comparisons on measures of CV and %ITI deviation in the spoken/sung tapping task indicated no significant difference between these groups ($ps \geq .17$). Thus, musical pitch perception seems to have little impact on synchronization to musical and non-musical stimuli. Maybe the beat perception task adds more weight onto pitch perception abilities than synchronization tasks.

5. General Discussion

This study investigated to what extent a beat-based entrainment deficit uncovered with music generalizes to speech. The main finding from this study is that beat-deaf participants were less consistent than neurotypical adults to match the period between their taps with the intervocalic period of the sentences across conditions, including for naturally spoken utterances. This result indicates that the beat-finding disorder, characterized by poor synchronization to musical beat, is not specific to music but also generalizes to the coordination with temporal regularities in other auditory contexts. However, the beat-deaf group also displayed more tapping variability than the control group in spontaneous tapping and tapping to the beat of a metronome. This result rather points to a deficit in internal timing mechanisms, which could account for the tendency toward higher tapping variability within the beat-deaf group across conditions, including tapping to a metronome (Palmer et al., 2014; Tranchant & Peretz, 2018).

Unexpectedly, participants with a beat-finding impairment showed higher variability than control participants when tapping regularly without a pacing stimulus, although the mean tempo of spontaneous tapping did not differ between groups. This was not reported in previous studies of participants with a beat-finding impairment (Bégel et al., 2017; Dalla Bella & Sowiński, 2015; Phillips-Silver et al., 2011; Sowiński & Dalla Bella, 2013). In Palmer et al. (2014), two cases of beat-deaf individuals seemed to have higher inter-tap variability in unpaced tapping, however, the difference was not statistically significant in comparison to a control group. Only recently Tranchant & Peretz (2018, in preparation) found, based on a more systematic investigation of spontaneous tapping in a group of eight participants, that beat-deaf individuals produced higher tapping variability when tapping regularly without a pacing stimulus. They also found that these participants were less consistent when tapping to a metronome. Interestingly, in our study, the only participant who failed to synchronize to the metronome had the highest interval variability in spontaneous tapping. These results suggest that internal timing mechanisms may best explain the beat finding impairment in some cases of beat deafness. Palmer et al. (2014) found that beat-deaf participants had more difficulty to adapt their tapping to phase and period perturbations in a metronome sequence. Sowiński & Dalla Bella (2013) also reported that poor beat synchronizers had more difficulty to correct their synchronization error when tapping to a metronome beat, as reflected in lag 1 analyses. Therefore, deficient error correction mechanisms in auditory-motor coupling may explain the generalized deficit of some beat-deaf individuals to synchronize with complex auditory rhythms (Palmer et al., 2014; Sowiński & Dalla Bella, 2013; Van Der Steen & Keller, 2013)

While this study is the first to investigate entrainment to speech in individuals identified based on their beat-finding deficiency, the inverse association of speech or reading skills and beat synchronization has received more attention. School age children with dyslexia have more difficulties than age-matched children to precisely synchronize taps to a metronome beat (Flaugnacco et al., 2014; Thomson & Goswami, 2008), and this group difference would persist in adulthood (Thomson, Fryer, Maltby, & Goswami, 2006). Similar results have also been reported in children with a speech language impairment (SLI; Corriveau & Goswami, 2009; Goswami, 2012; Cumming, Wilson, Leong, Colling, & Goswami, 2015). The timing deficits in dyslexia have primarily been associated with impairments in phonological awareness and more specifically rise time perception, which corresponds to the perception of change in the amplitude of the sound envelope usually marking phoneme onset (Corriveau, Pasquini, & Goswami, 2007; Goswami, 2011, 2012; Leong & Goswami, 2014). Studies have also shown that phonological awareness and reading skills in preschool children, first-grade children, and adolescents are associated with variability in synchronization with a metronome beat (Bonacina, Krizman, White-Schwoch, & Kraus, 2018; Tierney & Kraus, 2013; Woodruff Carr, Fitzroy, Tierney, White-Schwoch, & Kraus, 2014). Children and adults who stutter would also be less consistent than neurotypical control participants when synchronizing taps to musical beat (Falk, Müller, & Dalla Bella, 2015; van de Vorst & Gracco, 2017). In terms of mechanisms, deficient neural oscillatory entrainment to speech rhythm has been hypothesized as a possible cause of the common deficit for speech and music rhythm (Corriveau, Pasquini, & Goswami, 2007; Goswami, 2011, 2012; Leong & Goswami, 2014). In a similar perspective, the PATH model (*Precise Auditory Timing Hypothesis*) from Tierney & Kraus (2014) propose that phonological skills and auditory-motor entrainment are related by their shared reliance on precise neural imprinting of auditory rhythms' timing in auditory areas and the integration of this timing information in the motor areas. It should be noted, still, that previous studies did not find an association between reading skills and unpaced tapping (Corriveau & Goswami, 2009; Tierney & Kraus, 2013; Thomson & Goswami, 2008). Also, most of these studies have been conducted with children. Thus, parallels between results obtain from beat-deaf participants and the studies mentioned above must be done cautiously.

An aspect of the results that needs further consideration is that while most beat-deaf participants failed to match the period of the stimuli, they showed the same general pattern of performance than control participants; their tapping was more regular with regular speech and sung stimuli. It seems that the beat-deaf individuals could match, up to some extent, the more global rhythmic characteristics of the stimuli. This might reflect that while timing mechanisms associated with entrainment, or beat-based timing, are deficient in this group, other processes involved when processing the rhythmic structure of both speech and music, like grouping (the tendency to cluster into patterns events in a rhythm based on temporal proximity) or the perception of accents (events made prominent in the rhythm by changed in pitch or intensity for example; Fitch, 2013; Patel, 2008), might be preserved. This would accord well with results obtained from the online test of amusia (Peretz & Vuvan, 2017) and from the Montreal Battery for Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003) in this group. Indeed, most beat-deaf participants have scores within the normal range in the Off-beat test from the online test and the Rhythm test of the MBEA, which are anisochrony detection tasks (Tranchant & Vuvan, 2015). The majority of beat-deaf participants also have scores within the normal range on the Meter test from the MBEA, where participants must judge if melodic excerpts have the meter of a waltz or a march. Study of children with dyslexia points to poor meter perception, in either music or speech, in this population (Huss, Verney, Fosker, Mead, & Goswami, 2011; Goswami, Huss, Mead, Fosker, & Verney, 2013; Flaugnacco et al., 2014). Thus, the origin of the timing deficits might differ between some individuals with dyslexia and our beat-deaf participants. Further investigation of metric perception in speech and music with beat-deaf participants may help to uncover the cause of their beat synchronization deficit.

In contrast, neurotypical adults showed reduced tapping variability for the regular stimuli, with little difference between regularly spoken and sung sentences, the difference lying in the higher anticipation of beat onsets in the sung condition. This advantage of sung sentences could, however, solely reflect the lower intervocalic variability and Normalized Pairwise Variability Index for Vocalic Intervals (nPVI) for sung sentences compared to regular speech, indications of higher consistency in the time interval between vowels, which would facilitate the prediction of beat occurrences, and therefore synchronization. These results corroborate previous studies proposing that temporal regularity, or beat, is the main factor supporting entrainment to music (Dalla Bella et al., 2013; Lidji et al., 2011b). On the other hand, better anticipation of beats in the sung condition may highlight the contribution of other characteristics, like musical pitch, as additional factors contributing to a more precise prediction of beat occurrence. Some evidence suggests that pitch could influence meter perception and entrainment in music (Ammirante, Thompson, & Russo, 2011; Boasson & Granot, 2012; Cummins, Li, & Wang, 2013; Ellis & Jones, 2009; Hannon, Snyder, Eerola, & Krumhansl, 2004; Jones & Pfordresher, 1997; McKinney & Moelants, 2006; Pfordresher, 2003; Prince, 2011, 2014; Prince & Pfordresher, 2012). The possible contribution of musical pitch in beat extraction is suggested by the correlations we found between perception of musical pitch and period-matching performance, although beat-deaf participants with and without a concomitant musical-pitch perception deficit did not differ from other beat-deaf participants in their synchronization to regularly spoken and sung sentences.

The Action Simulation for Auditory Prediction (ASAP) hypothesis suggests that the periodic nature of musical beats promotes the coupling of the auditory system with the motor planning system which would be central to the making of temporal predictions (Patel & Iversen, 2014). The enhanced predictability of beat occurrences in music would convey an advantage for music to entrain precise synchronization. In accordance with this proposition, it has been shown that the coupling of activation between the motor and auditory cortices depends on the salience of the beat and metrical complexity (Chen et al., 2008b; Chen, Penhune, & Zatorre, 2009; Chen, Zatorre, & Penhune, 2006; Grahn & Rowe, 2009). Many studies now also show the implication of motor areas in beat extraction (e.g., Chen, Penhune, & Zatorre, 2008a; Grahn & Brett, 2007; Grahn & Rowe, 2009; Grahn & Rowe, 2013; Kung, Chen, Zatorre, & Penhune, 2013). Ross, Iversen, and Balasubramaniam (2016), and Maes, Leman, Palmer, and Wanderley (2014) provide reviews of embodied theories of auditory perception, or the perception-action link in music cognition. Embodied theories of speech perception have also been proposed (see Skipper, Devlin, & Lametti, 2017, for a recent review), although their correspondence with embodied theories of music perception has been argued (see Ross et al., 2016). We would propose, based on the current results, that the same auditory-motor coupling network may be recruited when trying to synchronize a discrete movement to an auditory rhythm, whether it is periodic or not. More periodic stimuli, like music would, however, convey an advantage in entraining synchronization by facilitating temporal predictions. This could account for the generalized deficit found in beat-impaired individuals, as for the more accurate synchronization found to sung stimuli in the control group. A next step would be to test the generalization of the current results to other types of auditory-motor coupling behaviors, such as choral speaking and singing, for example.

The current study has still some limitations that should be considered. Although it includes one of the largest sample of beat-deaf participants yet studied, the sample size remains modest. The beat-impaired group also showed more variability than the control group. This could have lowered the power of some statistical analyses, which could have prevented some group differences to emerge. Still, the group differences found were quite robust, with medium to large effect sizes. The current study also has the limitation of using a tapping task, which has little ecological value for entrainment with speech. Indeed, most studies of entrainment to natural speech use shadowing tasks where the natural tendency of speakers to entrain to another speaker's speech rate is measured. The reduced ability of beat-deaf participants to period-match natural speech compared to neurotypical adults should be validated in this kind of context.

Another factor that should be taken into consideration in the generalization of the current results is participants' native language. In this study, participants were French speakers listening to French sentences. French is usually considered a less "rhythmic" language than others like English (Grabe & Low, 2002; Ramus et al., 1999); the metric of "stress" languages, like English, is usually clearer than the metric of "syllabic" languages like French (Liberman & Prince, 1977; Lidji et al., 2011a). One's native language has also been shown to influence perception of speech rhythm (e.g., Cutler, 2000; Iversen, Patel, & Ohgushi, 2008; Lidji et al., 2011b). For example, Lidji et al. (2011a) found that tapping was more variable to French sentences than English sentences, and that English speakers tapped more regularly to sentences of both languages. However, using the same protocol as we did here while including both French and English speakers, Lidji et al. (2011b) found that tapping was more variable to English than French stimuli, irrespective of participants' native language. Despite the lack of consensual results, these two studies raise awareness on the possible influence of participants' native language and stimuli's language when tapping to speech. For example, one could expect that beat-deaf participants maybe show a clearer deficit to entrain to natural speech, since English stimuli could be more prone to elicit entrainment in native English neurotypical adult speakers. Still, here we found that French-speaking beat-deaf participants were also poorer than control participants to entrain to natural speech in French. This suggests that less rhythmic languages would also recruit a domain-general entrainment mechanism, providing additional support to the proposition of a domain-general entrainment mechanism for auditory signals with various degrees of periodicity.

In summary, our results indicate that beat-deaf individuals' deficit in beat synchronization is not specific to music but extends to complex rhythmic stimuli of speech and song. Furthermore, as proposed in previous studies (Dalla Bella et al., 2013; Lidji et al., 2011b), regularity, or isochrony, of the period seems to be the core feature through which entrainment is possible. The current findings suggest that a domain-general entrainment mechanism may support sensorimotor synchronization to both speech and music. However, the origin of the beat synchronization deficit in the beat-deaf participants deserves more investigation considering the unexpected results of higher tapping variability in spontaneous tapping. A next step would be to look at the generalization of the current results to other behaviors, by looking for example at how beat-deaf individuals entrain in social context, like conversational turn taking or joint action.

Acknowledgments

We would like to thank Pauline Tranchant for help with recruitment and insightful comments on data analysis, and Mailis Rodrigues for help with programming the task.

Funding

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada and the Canada Research Chairs program.

References

Ammirante, P., Thompson, W. F., & Russo, F. A. (2011). Ideomotor effects of pitch on continuation tapping. *The Quarterly Journal of Experimental Psychology, 64*(2), 381-393. doi:10.1080/17470218.2010.495408

Bégel, V., Benoit, C.-E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia, 94*, 129-138. doi:https://doi.org/10.1016/j.neuropsychologia.2016.11.022

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw, 31*(10), 1-21.

Boasson, A. D., & Granot, R. (2012). *Melodic direction's effect on tapping.* Paper presented at the 12th International Conference on Music Perception and Cognition, Thessaloniki, Greece.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.2). Retrieved from http://www.praat.org

Bonacina, S., Krizman, J., White-Schwoch, T., & Kraus, N. (2018). Clapping in time parallels literacy and calls upon overlapping neural mechanisms in early readers. *Annals of the New York Academy of Sciences, 1423*(1), 338-348. doi: doi:10.1111/nyas.13704

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a Coordinating Device: Entrainment With Disordered Speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815-824. doi:10.1044/2014_JSLHR-S-13-0149

Brown, S., Pfordresher, P. Q., & Chow, I. (2017). A musical model of speech rhythm. *Psychomusicology: Music, Mind, and Brain, 27*(2), 95-112. doi:10.1037/pmu0000175

Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: the role of attending. *Frontiers in Psychology, 1*, 224.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844-2854.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008b). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience, 20*(2), 226-239.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences, 1169*(1), 15-34.

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. doi:https://doi.org/10.1016/j.neuroimage.2006.04.207

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic Auditory Processing Skills and Specific Language Impairment: A New Look at an Old Hypothesis. *Journal of Speech, Language, and Hearing Research, 50*(3), 647-666. doi: 10.1044/1092-4388(2007/046)

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex, 45*(1), 119-130. doi: https://doi.org/10.1016/j.cortex.2007.09.008

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of Rhythm Patterns in Speech and Music in Children with Specific Language Impairments. *Frontiers in Human Neuroscience, 9*, 672. doi: 10.3389/fnhum.2015.00672

Cummins, F. (2002a). Entraining speech with speech and metronomes. *Cad. Est. Ling., 43*, 55-70.

Cummins, F. (2002b). On synchronous speech. *Acoustics Research Letters Online, 3*(1), 7-11.

Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics, 37*(1), 16-28.

Cummins, F. (2011). Periodic and Aperiodic Synchronization in Skilled Action. *Frontiers in Human Neuroscience, 5*(170). doi:10.3389/fnhum.2011.00170

Cummins, F. (2013). Joint speech: The missing link between speech and music? *PERCEPTA-Revista de Cognição Musical, 1*(1), 17.

Cummins, F., Li, C., & Wang, B. (2013). Coupling among speakers during synchronous speaking in English and Mandarin. *Journal of Phonetics, 41*(6), 432-441. doi:https://doi.org/10.1016/j.wocn.2013.07.001

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*(2), 145-171.

Cutler, A. (2000). Listening to a second language through the ears of a first. *Interpreting, 5*(1), 1-23.

Dalla Bella, S., Białuńska, A., & Sowiński, J. (2013). Why movement is captured by music, but less by speech: Role of temporal regularity. *PLoS ONE, 8*(8), e71945.

Dalla Bella, S., & Peretz, I. (2003). Congenital Amusia Interferes with the Ability to Synchronize with Music. *Annals of the New York Academy of Sciences, 999*(1), 166-169. doi:10.1196/annals.1284.021

Dalla Bella, S., & Sowiński, J. (2015). Uncovering Beat Deafness: Detecting Rhythm Disorders with Synchronized Finger Tapping and Perceptual Timing Tasks. *Journal of Visualized Experiments: JoVE*(97), 51761. doi:10.3791/51761

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*(1), 51-62.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition, 77*(3), 251-288. doi:https://doi.org/10.1016/S0010-0277(00)00106-2

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in Time with Mechanically and Expressively Performed Music. *Music Perception: An Interdisciplinary Journal, 18*(1), 1-23. doi:10.2307/40285899

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264-280. doi:10.1037/a0013482

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*(847). doi: 10.3389/fpsyg.2015.00847

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience, 7*(68). doi:10.3389/fnsys.2013.00068

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., … Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience, 8*(392). doi: 10.3389/fnhum.2014.00392

Giovannelli, F., Innocenti, I., Rossi, S., Borgheresi, A., Ragazzoni, A., Zaccara, G., . . . Cincotta, M. (2014). Role of the Dorsal Premotor Cortex in Rhythmic Auditory–Motor Entrainment: A Perturbational Approach by rTMS. *Cerebral Cortex, 24*(4), 1009-1016. doi:10.1093/cercor/bhs386

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15*(1), 3-10. doi: https://doi.org/10.1016/j.tics.2010.10.001

Goswami, U. (2012). Entraining the brain: applications to language research and links to musical entrainment.

Goswami, U., Huss, M., Mead, N., Fosker, T., & Verney, J. P. (2013). Perception of patterns of musical beat distribution in phonological developmental dyslexia: Significant longitudinal relations with word reading and reading comprehension. *Cortex, 49*(5), 1363-1376. doi: 10.1016/j.cortex.2012.05.005

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in laboratory phonology, 7*(515-546).

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience, 19*(5), 893-906. doi:10.1162/jocn.2007.19.5.893

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*(23), 7540-7548.

Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex, 23*(4), 913-921. doi:10.1093/cercor/bhs083

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews, 86*, 150-165. doi:https://doi.org/10.1016/j.neubiorev.2017.12.002

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 956.

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*(1), 85-91. doi:doi:10.1111/j.1749-6632.2011.06402.x

Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: Perception of musical meter predicts reading and phonology. *Cortex, 47*(6), 674-689. doi: https://doi.org/10.1016/j.cortex.2010.07.010

Iversen, J. R. (2016). In the beginning was the beat. In R. Hartenberger (Ed.), *The Cambridge Companion to Percussion* (pp. 281-295). Cambridge: Cambridge University Press.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *J Acoust Soc Am, 124*(4), 2263-2271. doi:10.1121/1.2973189

Jadoul, Y., Ravignani, A., Thompson, B., Filippi, P., & de Boer, B. (2016). Seeking Temporal Predictability in Speech: Comparing Statistical Approaches on 18 World Languages. *Frontiers in Human Neuroscience, 10*(586). doi:10.3389/fnhum.2016.00586

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*(6), 621-634. doi:10.3758/bf03210494

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459-491. doi:10.1037/0033-295X.96.3.459

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(4), 271-291. doi:10.1037/1196-1961.51.4.271

Jungers, M. K., Palmer, C., & Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 1*(2), 21-35.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392-399. doi:https://doi.org/10.1016/j.tics.2010.06.005

Kover, S. T., & Atwood, A. K. (2013). Establishing equivalence: Methodological progress in group-matching design and analysis. American Journal on Intellectual and Developmental Disabilities, 118(1), 3-15.

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience, 25*(3), 401-420. doi:10.1162/jocn_a_00325

Lagrois, M-E, & Peretz, I. (*in press*). The Co-occurrence of Pitch and Rhythm Disorders in Congenital Amusia. *Cortex*.

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *The psychology of time*, 189-232.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience, 9*. doi:10.3389/fnsys.2015.00159

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26*(1), 1-37. doi:http://dx.doi.org/10.1016/S0364-0213(01)00057-X

Large, E. W., & Snyder, J. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences, 1169*(1), 46-57. doi:doi:10.1111/j.1749-6632.2009.04550.x

Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *J Acoust Soc Am, 136*(1), 366-381. doi:10.1121/1.4883366

Leow, L.-A., Parrott, T., & Grahn, J. A. (2014). Individual Differences in Beat Perception Affect Gait Responses to Low- and High-Groove Music. *Frontiers in Human Neuroscience, 8*(811). doi:10.3389/fnhum.2014.00811

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception: An Interdisciplinary Journal, 1*(2), 229-252.

Liberman, M., & Prince, A. (1977). On Stress and Linguistic Rhythm. *Linguistic Inquiry, 8*(2), 249-336.

Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011b). Entrainment to speech and song. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science* (pp. 123-128). Utrecht, The Netherlands: EAC.

Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011a). Listeners feel the beat: Entrainment to English and French speech rhythms. *Psychonomic Bulletin & Review, 18*(6), 1035-1041. doi:10.3758/s13423-011-0163-0

London, J. (2002). Cognitive Constraints on Metric Systems: Some Observations and Hypotheses. *Music Perception: An Interdisciplinary Journal, 19*(4), 529-550. doi:10.1525/mp.2002.19.4.529

London, J. (2012). Three Things Linguists Need to Know About Rhythm and Time in Music. *Empirical Musicology Review, 7*(1/2), 5-11.

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. (2014). Action-based effects on music perception. *Frontiers in Psychology, 4*(1008). doi:10.3389/fpsyg.2013.01008

McAuley, J. D. (2010). Tempo and Rhythm. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music Perception* (pp. 165-199). New York, NY: Springer New York.

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.

McKinney, M. F., & Moelants, D. (2006). Ambiguity in Tempo Perception: What Draws Listeners to Different Metrical Levels? *Music Perception: An Interdisciplinary Journal, 24*(2), 155-166. doi:10.1525/mp.2006.24.2.155

Meyer, L. (2017). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *Eur J Neurosci, 0*(0). doi:10.1111/ejn.13748

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: a metaphor? *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi:10.1098/rstb.2013.0396

O'Dell, M., & Nieminen, T. (1999). *Coupled oscillator model of speech rhythm.* Paper presented at the Proceedings of the XIVth international congress of phonetic sciences.

Palmer, C. (1997). Music Performance. *Annual Review of Psychology, 48*(1), 115-138. doi:10.1146/annurev.psych.48.1.115

Palmer, C., Lidji, P., & Peretz, I. (2014). Losing the beat: deficits in temporal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi:10.1098/rstb.2013.0405

Patel, A. D. (2008). *Music, Language, and the Brain*. New York, United States of America: Oxford University Press.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57.

Peelle, J. E., & Davis, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology, 3*, 320. doi:10.3389/fpsyg.2012.00320

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of Musical Disorders. *Annals of the New York Academy of Sciences, 999*(1), 58-75. doi:10.1196/annals.1284.006

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal Of Human Genetics, 25*, 625. doi:10.1038/ejhg.2017.15

Pfordresher, P. Q. (2003). The Role of Melodic and Rhythmic Accents in Musical Structure. *Music Perception: An Interdisciplinary Journal, 20*(4), 431-464. doi:10.1525/mp.2003.20.4.431

Phillips-Silver, J., & Keller, P. (2012). Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience, 6*(26). doi:10.3389/fnhum.2012.00026

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia, 49*(5), 961-969. doi:https://doi.org/10.1016/j.neuropsychologia.2011.02.002

Port, R. F. (2003). Meter and speech. *Journal of Phonetics, 31*(3), 599-611. doi:https://doi.org/10.1016/j.wocn.2003.08.001

Prince, J. B. (2011). The Integration of Stimulus Dimensions in the Perception of Music. *Quarterly Journal of Experimental Psychology, 64*(11), 2125-2152. doi:10.1080/17470218.2011.573080

Prince, J. B. (2014). Pitch structure, but not selective attention, affects accent weightings in metrical grouping. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 2073-2090. doi:10.1037/a0037730

Prince, J. B., & Pfordresher, P. Q. (2012). The role of pitch and temporal diversity in the perception and production of musical sequences. *Acta Psychologica, 141*(2), 184-198. doi:https://doi.org/10.1016/j.actpsy.2012.07.013

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292. doi:https://doi.org/10.1016/S0010-0277(99)00058-X

Repp, B. H. (2003). Rate Limits in Sensorimotor Synchronization With Auditory and Visual Sequences: The Synchronization Threshold and the Benefits and Costs of Interval Subdivision. *Journal of motor behavior, 35*(4), 355-370. doi:10.1080/00222890309603156

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969-992.

Repp, B. H., Iversen, J. R., & Patel, A. D. (2008). Tracking an Imposed Beat within a Metrical Grid. *Music Perception: An Interdisciplinary Journal, 26*(1), 1-18. doi:10.1525/mp.2008.26.1.1

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*(3), 403-452.

Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase, 22*(6), 558-565. doi:10.1080/13554794.2016.1242756

Schultz, B. G., O'Brien, I., Phillips, N., McFarland, D. H., Titone, D., & Palmer, C. (2015). Speech rates converge in scripted turn-taking conversations. *Applied Psycholinguistics, 37*(5), 1201-1220. doi:10.1017/S0142716415000545

Schultz, B. G., & van Vugt, F. T. (2016). Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behavior Research Methods, 48,* 1591–1607.

Selkirk, E. O. (1986). *Phonology and syntax: the relationship between sound and structure*: MIT press.

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain and Language, 164*, 77-105. doi:https://doi.org/10.1016/j.bandl.2016.10.004

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51*(10), 1952-1963. doi:https://doi.org/10.1016/j.neuropsychologia.2013.06.027

Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., ... & Debono, V. B. (2013). A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC medical research methodology, 13(1), 92.

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading, 29*(3), 334-348. doi: doi:10.1111/j.1467-9817.2006.00312.x

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris, 102*(1), 120-129. doi: https://doi.org/10.1016/j.jphysparis.2008.03.007

Tierney, A. T., & Kraus, N. (2013). The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. *Brain and Language, 124*(3), 225-231. doi: 10.1016/j.bandl.2012.12.014

Tierney, A., & Kraus, N. (2014). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience, 8*, 949. doi: 10.3389/fnhum.2014.00949

Tranchant, P., Lagrois, M., Bellemare Pépin, A., G. Schultz, B., & Peretz, I. (2018). *Beat alignment test of the motor origin of musical entrainment deficits*. Manuscript in preparation

Tranchant, P., & Vuvan, D. T. (2015). Current conceptual challenges in the study of rhythm processing deficits. *Frontiers in Neuroscience, 9*(197). doi:10.3389/fnins.2015.00197

Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the Beat: A Large Sample Study of Bouncing and Clapping to Music. *PLoS ONE, 11*(7), e0160178. doi:10.1371/journal.pone.0160178

Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong *Laboratory Phonology* (Vol. 4, pp. 93).

van de Vorst, R., & Gracco, V. L. (2017). Atypical non-verbal sensorimotor synchronization in adults who stutter may be modulated by auditory feedback. *Journal of Fluency Disorders, 53*, 14-25. doi: https://doi.org/10.1016/j.jfludis.2017.05.004

Van Der Steen, M. C., & Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience, 7*, 253.

Van Vugt, F. & Schultz, B. G. (2015). Taparduino v1.01. Zenodo, 16178. https://doi.org/10.5281/zenodo.16178

Vuvan, D. T., Paquette, S., Mignault Goulet, G., Royal, I., Felezeu, M., & Peretz, I. (2017). The Montreal Protocol for Identification of Amusia. *Behav Res Methods*. doi:10.3758/s13428-017-0892-8

Wechsler, D., Coalson, D. L., & Raiford, S. E. (1997). *WAIS-III: Wechsler adult intelligence scale*: Psychological Corporation San Antonio, TX.

Wilkie, D. (1983). Rayleigh test for randomness of circular data. *Applied statistics, 32*(3), 311-312.

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review, 12*(6), 957-968.

Woodruff Carr, K., Fitzroy, A. B., Tierney, A., White-Schwoch, T., & Kraus, N. (2017). Incorporation of feedback during beat synchronization is an index of neural maturation and reading skills. *Brain and Language, 164*, 43-52. doi: 10.1016/j.bandl.2016.09.005

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences, 107*(13), 5768-5773. doi:10.1073/pnas.1000121107

Tables

Table 1. Groups' Characteristics

Variables	Beat-deaf (*SD*) *n* = 13	Control (*SD*) *n* = 13
Age (years)	37.4 (*17.6*)	38.7 (*17.8*)
Education (years)	18.2 (*2.2*)	17.6 (*3.2*)
Musical Training (years)	1.0 (*2.3*)	1.1 (*2.1*)
Dance Training (years)	1.3 (*3.0*)	1.8 (*3.1*)
WAIS-III Digit Span (ss)	10 (*3*)	11 (*3*)
WAIS-III Matrix Reasoning (ss)[a]	13 (*3*)	14 (*1*)

Note: ss: standard score. [a] Scores from 12 beat-deaf and 10 control participants. Some participants did not complete the Matrix Reasoning test because they were students in a clinical neuropsychology Ph.D. program and were too familiar with the test.

Table 2. Individual scores of the beat-deaf participants and group average of their matched controls in the online test of amusia

	Group														
	Beat-deaf													Control (n = 13)	
Participant	BI1	BI2	BI3	BI4	BI5	BI6	**BI7**	**BI8**	**BI9**	**BI10**	**BI11**	**BI12**	**BI13**	*M*	*M*	*SD*
Online test																
Scale (22/30)	23	24	23	23	23	24	21	21	20	22	19	18	22	21.8	27.7	2.2
Off-key (16/24)	20	14	19	14	16	14	13	16	15	9	13	14	13	14.6	19.8	2.2
Off-beat (17/24)	23	21	19	17	20	16	15	17	18	17	18	18	19	18.3	19.8	1.4

Note: Scores in parentheses beside each condition of the online test represent the cut-off scores based on Peretz and Vuvan (2017). Participants with concomitant pitch deafness are marked in bold.

Table 3. Stimuli Characteristics Related to Rhythm

	Naturally Spoken Sentences (*SE*)	Regularly Spoken Sentences (*SE*)	Sung Sentences (*SE*)
Mean IVI [ms]	458 (*10*)	503 (*3*)	501 (*1*)
V-nPVI	49.4 (*2.5*)	42.3 (*2.0*)	31.1 (*1.8*)
CV(IVI)	0.21 (*0.02*)	0.08 (*0.01*)	0.05 (*0.004*)
Beat Asynchrony from Vowel Onset [ms]	-	14 (*11*)	-2 (*5*)

Note: Values indicate means; standard errors appear in parentheses. IVI: intervocalic interval (in ms); V-nPVI: normalized Pairwise Variability Index for Vocalic Intervals; CV: coefficient of variation (SD IVI/Mean IVI); Beat asynchrony corresponds to the average of signed values from subtracting metronome tone onset from the closest spoken/sung vowel onset, in milliseconds).

Table 4. Mean Inter-tap interval (ITI) and Coefficient of Variation (CV) of Spontaneous Tapping

	Spontaneous Tapping- Pre		Spontaneous tapping- Post	
Group	Mean ITI[a]	CV	Mean ITI[a]	CV
Control (*SE*)	603 (56)	.06 (.003)	565 (57)	.06 (.004)
Beat-impaired (*SE*)	680[b] (47)	.08[b] (.01)	659 (80)	.08 (.01)

Note: Groups' mean and standard error of the mean in parentheses.
[a] Values are in milliseconds. [b] n = 12, otherwise, n = 13.

Table 5. Spearman Correlations Between Tapping Measurements and Music Perception

Variable	Scale test	Off-key test	Off-beat test	M-BAT perception test *d'*
CV - natural utterances[a]	-.20	-.22	-.12	-.30
CV - regular utterances[a]	-.24	-.29	-.03	-.38
CV - sung utterances[a]	-.26	-.35	-.24	-.22
%ITI deviation - natural utterances[b]	-.30	-.34	.11	$-.47^{*}$
%ITI deviation - regular utterances[b]	-.33	$-.55^{**}$	-.19	$-.56^{**}$
%ITI deviation - sung utterances[b]	$-.54^{**}$	$-.43^{*}$	-.21	$-.66^{***}$

Note: Outliers from the beat-impaired group were removed [a] $n = 24$, [b] $n = 23$.
$^{*}p \leq .05$, $^{**}p \leq .01$, $^{***}p \leq .001$.

Figure Legends

Figure 1. Performance of Control and Beat-deaf Participants on the M-BAT. A the beat production task, and **B** the beat perception task. Each dot represents a participant. Boxes corresponds to a 95% confidence interval from the mean based on the standard error mean (SEM). The black line within each box indicates the group mean. The vertical lines show two standard deviations from the mean.

Figure 2. Example of a Sentence in the Three Conditions. IVI refers to the intervocalic interval between stressed syllables.

Figure 3. Coefficient of Variation of the Stimuli and Participants' Tap in Each Condition. A Coefficient of variation (CV) of the intervocalic interval (IVI). Each dot represents a sentence. **B** Mean CV of the inter-tap interval (ITI) produced by each participant in each condition. Each dot represents a participant. Boxes corresponds to a 95% confidence interval from the mean based on standard error mean (SEM). The black line within each box indicates the group's mean. The vertical lines show two standard deviations from the mean.

Figure 4. Mean Percentage of Deviation Between the Inter-tap Intervals Produced by Each Participant and the IVI of the Sentences. Each dot represents a participant. Boxes corresponds to a 95% confidence interval from the mean based on standard error mean (SEM). The black line within each box indicates the group's mean. The vertical lines show two standard deviations from the mean.

Figure 5. **Frequency Chart of Participants' Preferred Tapping Level in Each Group for Each Condition**. Number of participants, in each group, who tapped at every word, every two words or every four words.

Figures

Figure 1. Performance of Control and Beat-deaf Participants on the M-BAT

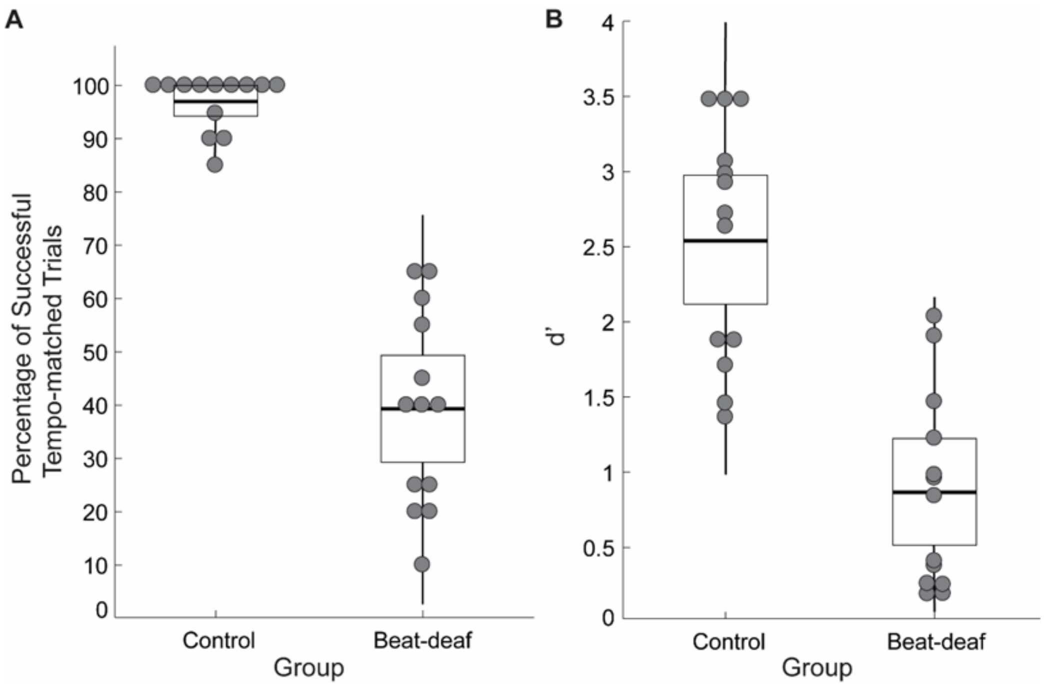


Figure 2. Example of a Sentence in the Three Conditions

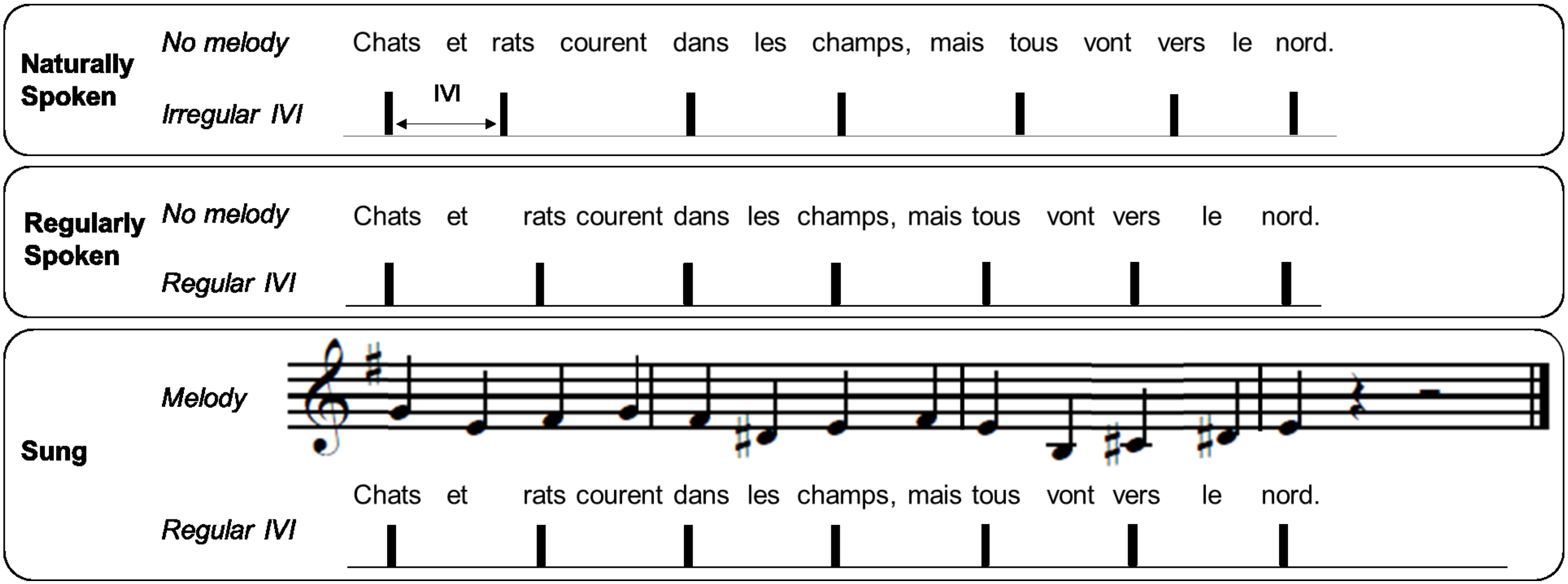


Figure 3. Coefficient of Variation of the Stimuli and Participants' Tap in Each Condition

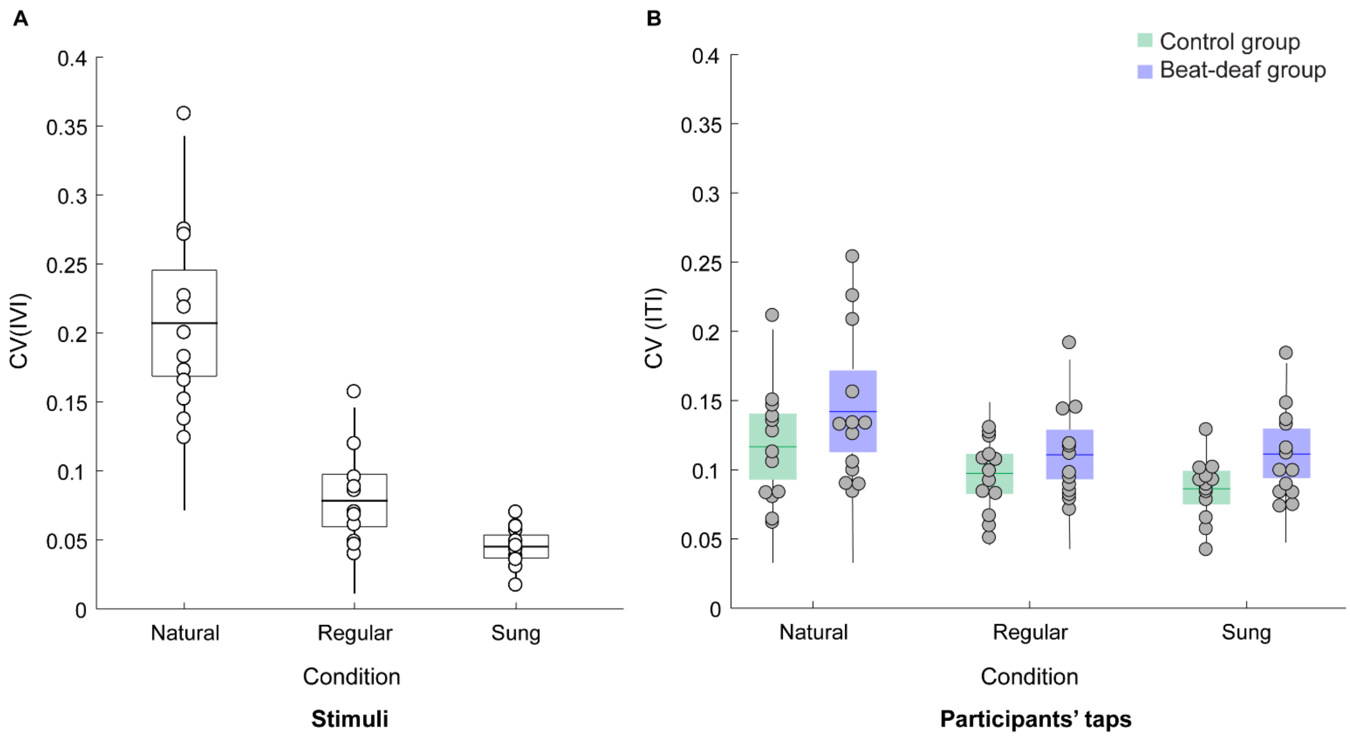


Figure 4. Mean Percentage of Deviation Between the Inter-tap Intervals Produced by Each Participant and the IVI of the Sentences

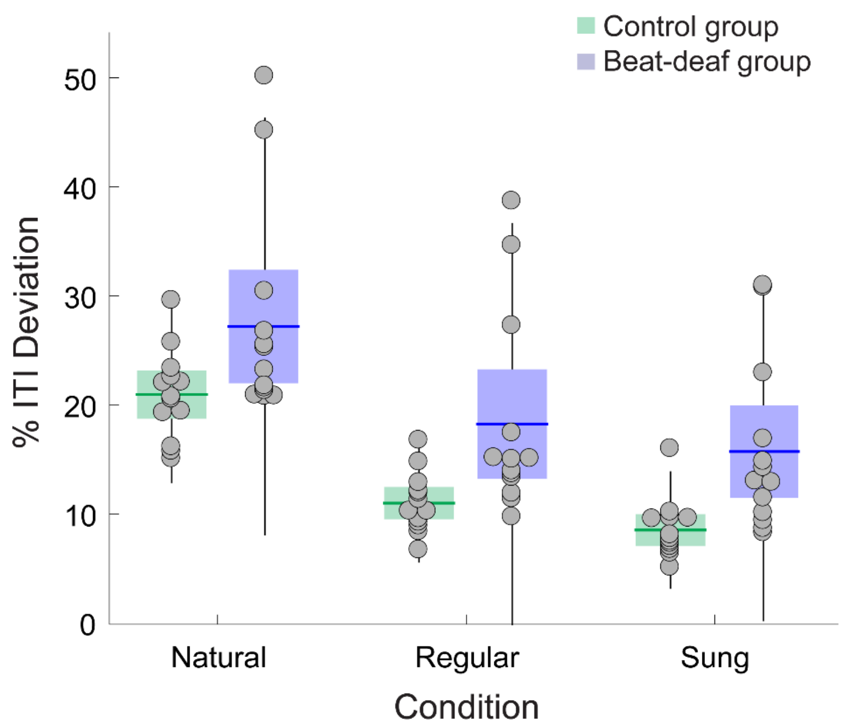
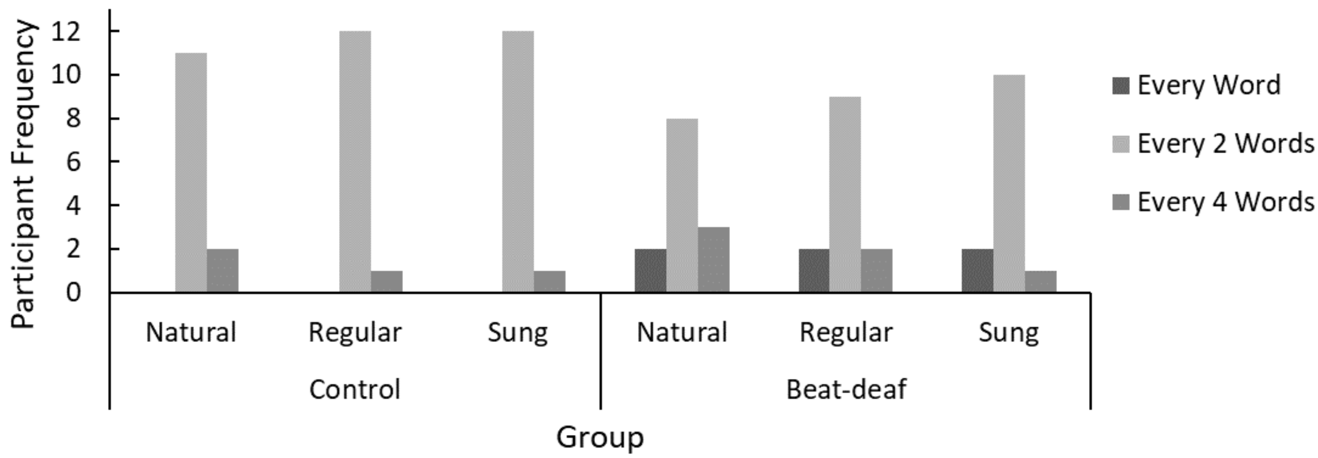


Figure 5. Frequency Chart of Participants' Preferred Tapping Level in Each Group for Each Condition



Article 3 : Singing Alone and Along in Beat Deafness

Marie-Élaine Lagrois[a, b] & Isabelle Peretz[a, b]

[a] International Laboratory for Brain, Music, and Sound Research (BRAMS)

[b] Département de Psychologie, Université de Montréal

Article en préparation

Abstract

Based on an evolutionary perspective, since both beat synchronization and vocal imitation require tight auditory-motor coupling, it has been hypothesized that these two fundamental traits of musical expression could be rooted in a common sensorimotor coupling mechanism. To test the assumption of an association between vocal and synchronization abilities, we measured singing accuracy of beat-deaf participants when singing with a model. Beat-deaf individuals are unable to synchronize a simple movement to the beat of music and therefore should be less accurate to match a model in a singing context. Beat-deaf participants and a group of matched neurotypical adults had to sing a familiar song from memory, after hearing a model (singing alone), in synchrony with the model (synchronous singing), and to a metronome. Singing productions were made with lyrics and with the melody only (singing with the syllable /la/). We found that when singing from memory, groups were mostly similar. Some group differences emerged, however, when singing to a model. Beat-deaf individuals were less consistent than control participants when matching the tempo of the model in synchronous singing. They also made more absolute-pitch errors, but not more pitch-interval errors, when singing alone and in synchrony. However, there was no clear association between performance on a metronome tapping task and singing pitch proficiency. Results from this study provide mitigated support to the existence of a common mechanism of auditory-motor coupling between vocal pitch and beat synchronization. Still, beat-deaf individuals showed a general impairment in synchronization both in the context of singing and tapping to a metronome. This suggests a general sensorimotor mechanism that could support synchronization across tasks and effectors, within the auditory domain.

Keywords: beat deafness, beat, singing, vocal imitation, synchronization, sensorimotor coupling

1. Introduction

Spontaneous singing and dance-like behaviors when listening to music are ubiquitous forms of musical production. These musical responses are found across cultures and most individuals produce them effortlessly, without formal musical training (Dowling, 1999; Mithen, 2006). Indeed, the ability to sing is widespread. The vast majority of the population is able to sing familiar songs and imitate simple melodies accurately in both pitch and time (Dalla Bella & Berkowska, 2009a; Dalla Bella, Giguère, & Peretz, 2007; Pfordresher & Brown, 2007). In the case of congenital amusia, characterized by impaired musical pitch perception, singing is strikingly out-of-tune while leaving singing timing intact (Tremblay-Champoux et al., 2010; Dalla Bella, Giguère, & Peretz, 2009). Here, we examine to what extent amusic individuals, so-called beat-deaf adults, who have marked difficulties to keep in time with music are similarly affected in singing.

The possibility of dissociating rhythm from pitch deficits in singing has theoretical significance. According to the *vocal learning and rhythmic synchronization hypothesis* proposed by Patel (2006), beat synchronization relies on a sensorimotor network that has evolved to support vocal learning (i.e. the learned reproduction of an external auditory signal based on auditory experience and sensory feedback). By this view, vocal performance and synchronization skills can hardly be dissociable. Both require tight auditory-motor coupling in order to precisely match the motor action (vocal pitch and timing) with an external sound event (Patel, 2006; Patel, Iversen, Bregman, & Schulz, 2009). In support of this hypothesis, accurate and precise singers are found to synchronize taps to a metronome beat more precisely than less accurate and precise singers (Dalla Bella, Berkowska, & Sowiński, 2015).

In the pitch-based form of congenital amusia, the deficit is not associated with any difficulty in maintaining the rhythm of a song in singing alone nor in singing with a pre-recorded voice (singing along) (Tremblay-Champoux et al. 2010). Further, in pitch-based amusia, synchronized singing helps in reducing pitch errors (Tremblay-Champoux et al. 2010). Such a dissociation suggests that the origin of rhythmic synchronization is distinct from vocal pitch learning. It may be that synchronization evolved for other reasons than vocal learning, such as social cooperation (Merker, Madison, & Eckerdal, 2009; Valdesolo & DeSteno, 2011). Note, however, that in the prior study of pitch-deaf individuals synchronization per se was not properly measured since temporal variability was not assessed relative to the external sung model. Singing with a model may reduce uncertainty in both pitch and time but not synchronization, which is extremely precise and demanding in humans. Few studies have looked at singing in synchrony with someone else, which provide the opportunity to test sensorimotor coupling in singing in a context that resembles more synchronization to beat.

Here, we tested the effect of a deficit in beat synchronization on singing performance. Beat-deafness, a form of congenital amusia which is characterized by the inability to synchronize a simple movement to the beat of music (Palmer, Lidji, & Peretz, 2014; Phillips-Silver et al., 2011), offers a good opportunity to test the idea that beat synchronization relies on shared mechanisms with pitch proficiency and vocal learning in general. Beat-deaf individuals, also called poor synchronizers, are unable to synchronize a whole-body movement, clapping or tapping to the beat of music or amplitude-modulated noise derived from music (Bégel et al., 2017; Dalla Bella & Sowiński, 2015; Phillips-Silver et al., 2011; Sowiński & Dalla Bella, 2013; Tranchant, Vuvan, & Peretz, 2016). Some authors suggest that one cause of this beat synchronization deficit might be poor auditory-motor mapping (Palmer et al., 2014; Sowiński & Dalla Bella, 2013). Accordingly, vocal imitation and synchronous singing should be less accurate in individuals with a beat finding impairment. This prediction has, to our knowledge, never been assessed. Only a description of the first case of beat deafness (Phillips-Silver et al., 2011) reported correct singing in this individual, without further details.

We compared beat-deaf and matched neurotypical adults on their singing abilities with a special focus on their ability to sing along with a model. Singing from memory was assessed via acoustic and perceptual measures of singing performance. Participants also imitated a familiar song either alone or together with the recorded model, following the same procedure as used with pitch-deaf amusic participants (Tremblay-Champoux et al., 2010). If vocal performance and beat synchronization are linked by a common auditory-motor mechanism, beat-deaf individuals are expected to exhibit poor singing abilities on all dimensions. Alternatively, if synchronization is distinct from vocal imitation, we may expect to find a deficit in beat-deaf individuals for singing along only and not when singing alone.

2. Method

2.1 Participants

Eight participants with a documented beat synchronization deficit (6 females) who took part in prior studies (Tranchant et al., 2016; Tranchant et al., 2018) and a matched group of ten control participants (8 females) were enrolled in the present study. The groups were matched for age, education, and musical training (see Table 1 for details). The beat-deaf participants were identified as being unable to synchronize simple movements to the beat of music. The control participants were recruited via online advertisement in Montreal's general population and on campus advertisement at the University of Montreal. All participants were non-musicians and had no history of neurological, cognitive, hearing or motor disorders. They all had normal verbal auditory working memory and non-verbal reasoning abilities, as assessed by the Digit Span and the Matrix Reasoning tests from the WAIS-III (Wechsler Adult Intelligence Scale; Wechsler, Coalson, & Raiford, 1997).

Beat-deaf participants were identified as such with the Montreal Beat Alignment Test (M-BAT; Tranchant et al., under review). The M-BAT includes two tasks, a beat production task and a beat perception task. In the beat production task, participants are asked to align finger taps to the beat of musical excerpts. The percentage of trials to which the taps are tempo-matched with the beat of the musical excerpts is computed as an indication of synchronization performance. Successful period matching is determined with a p-value smaller than .05 on the Rayleigh z test of periodicity. There was no overlap between groups on the percentage of trials with accurate tempo matching, $t(7.7) = 44.2$, $p < .001$ (Table 1).

For spontaneous tapping (tapping regularly without an external stimulus) the mean tempo and variability of inter-tap intervals (CV) were compared between groups. In the control group, the spontaneous tapping tempo was highly variable, with a mean of 111 bpm ($SE = 10$, range: 61–166). In the beat-deaf group, the mean tempo was of 90 bpm ($SE = 7$, range: 61–118). Welch's test between groups was not significant, $t(15.1) = 3.3$, $p = .90$. The difference between groups was also not significant for tempo variability, $t(13.4) = 2.0$, $p = .18$. Thus, groups were similar on a simple paced motor task.

Participants also completed the online test of amusia to screen for musical pitch perception impairments (Peretz & Vuvan, 2017). The online test is composed of three tests: Scale, Off-beat, and Off-key. In the scale test, participants have to compare 30 pairs of melodies and identifies the pairs that differ by an out-of-key note. In the off-beat and off-key tasks participants have to detect melodies that contain either an out-of-time or an out-of-key note, respectively. Scores from each group on the online test are indicated in Table 1. None of the participants met the criteria for pitch deafness (based on Peretz & Vuvan, 2017). However, as a group, beat-deaf participants had lower scores than control participants on both the scale and off-key tests. Groups performed similarly on the off-beat test, suggesting that auditory attention is normal.

[Insert Table 1 here]

Participants provided written consent to their inclusion in the study and received monetary compensation for their participation. All procedures were approved by the Research Ethics Council for the Faculty of Arts and Sciences at the Université de Montréal.

2.2 Procedure and Material

2.2.1 Singing Task

The singing task involved singing a familiar song in eight conditions. The song was either the chorus of *Gens du pays* (Vigneault & Rochon, 1978), typically sung at birthdays in Québec or *Happy Birthday*. Song notations are presented in Figure 1. *Gens du Pays* includes 32 notes, with one syllable on each note. *Happy Birthday* includes 25 notes, with one syllable on each note. *Happy Birthday* may be considered more difficult to sing because of the octave jump on the 14th note interval which is not found in *Gens du pays*. However, it is a well-known song that most adults feel comfortable to sing (Pfordresher & Brown, 2017). In the first singing condition, participants had to sing the song from memory, with the lyrics (*Memory* condition). They sang the song at the tempo and starting pitch of their choice. Then, they heard a recorded model singing the song and were asked to start singing once the model had terminated (*Solo* condition). There was no specific instruction about having to "imitate" the model. Right after, participants had to sing along with the same model (*Synchronous* condition). These three conditions (*Memory*, *Solo*, and *Synchronous*) were again performed but this time without the lyrics, using the syllable /la/. At the end, participants had to synchronize their singing with a metronome, singing first with lyrics and then on /la/ (*Metronome* condition). The eight conditions were presented in that specific order for all participants. Therefore, the performance of participants was expected to be generally better when singing on /la/ since it was always performed after singing with the lyrics.

Before the task, participants did a warm-up consisting of gliding a sustained pitch from the lowest note they could sing to the highest note they could reach and go back to their lowest note, three times. The experimenter did the warm-up with each participant and stayed in the room for practical reasons. The model was sung either by a male or a female singer to match the pitch range of the participant. It was either sung in English or French, based on the language they felt most comfortable with. Six participants (including 3 beat-deaf participants) sang *Gens du pays* (French only) and twelve participants sang *Happy Birthday* (ten sang it in French, including four beat-deaf participants).

Models were pre-recorded and produced by untrained singers to avoid excessive use of vibrato in the performance (as in Tremblay-Champoux et al., 2010). Songs were performed by different singers, except for the French and English female versions of *Happy Birthday* that were performed by the same singer. The starting pitch for *Gens du pays* with lyrics was F3 for the male version and A#3 for the female version, and F3 and A3 for the versions on /la/. For the French version of *Happy Birthday* the starting pitch was G#2 and G#3 for the male and female version, respectively, with and without lyrics. For the English version, it was A3 with lyrics and A#3 on /la/. Models singers heard a metronome while recording the songs. *Gens du pays* was recorded at 120 beat per minute (bpm) and *Happy Birthday* was recorded at 118 bpm. The metronome tempi used for recording the models were also used in the *Metronome* condition of the singing task.

The participants were tested in a sound-attenuated booth. They sat in front of a Neumann TLM 103 or an AKG cgn521e microphone about 10 centimeters away. They heard the model and their own voice through headphones, both presented in stereo. Audio output and input were controlled by an audio interface (RME Fireface 800) with a sampling rate of 44.1 kHz and were recorded as an Adobe Audition session (v.3.0) containing all the stimuli and recordings of the participants that could be exported in .wav files for offline analyses.

[Insert Figure 1 here]

2.2.2 Tapping Tasks

As an indication of sensorimotor synchronization abilities, participants were asked to tap to a metronome. Participants heard seven metronome tones at 120 bpm, and then had to synchronize taps to the metronome for sixty additional metronome tones. The tones were made of a 440 Hz sine wave with a duration of 50 ms. This task was either performed in a different session from the singing task or before the singing task. Participants were asked to tap with the index finger of their dominant hand. Taps were recorded on a square force-sensitive resistor (3.81 cm, Interlink FSR 406) connected to an Arduino UNO (R3; arduino.cc) transmitting timing information to a PC (HP ProDesk 600 G1, Windows 7) via the serial USB port. The metronome stimulus was delivered at a comfortable level through closed headphones (DT 770 PRO, Beyerdynamic).

2.3 Data Analysis

2.3.1 Acoustic Analyses

For each of the 8 singing rendition per participant (amounting to 72 with lyrics and 72 on /la/), the onset and offset of the syllable and vowel associated with each note of the song were marked by hand using Praat software (Boersma & Weenink, 2017). The segmentation was based on visual inspection of the spectrogram (formants), f0 pitch track, intensity changes, and listening to the file. Each song segmentation was revised by another person. The F0 pitch track from the entire production was extracted using Praat with the autocorrelation method (sampling rate at 100 Hz [10 ms], pitch range: 75 Hz to 600 Hz). Before extracting pitch information for further analysis, the F0 track was visually inspected and octave jump errors from the pitch detection algorithm were corrected manually. The marked onsets and offsets time and the F0 pitch track of each file were exported to MATLAB for further analyses.

Using a MATLAB code, an automated procedure extracted the median F0 value from the middle 50% of each vowel based on the marked onset and offset, which was used for further analysis of pitch. Notes having less than 10% of valid F0 information in that time window were excluded from analysis. Less than 12% of the notes per condition and participant were removed on average. There was no difference between groups on the percentage of notes excluded (Welch's test, $p>.15$). For the analysis on the temporal dimension the vowels' onset time served as markers. Note that the first pitch and pitch interval were not included in the acoustic analysis of the *Synchronous* condition, as many participants missed it.

2.3.1.1 Timing

Time interval ratios were based on the duration of two consecutive notes, excluding the last one because it tends to be lengthened. Time interval ratios that deviated from the expected rhythm by more than 25%, based on the songs notation, were considered as rhythm errors and the percentage of rhythm errors was calculated (as in Tremblay-Champoux et al., 2010).

Tempo was based on the inter-onset interval (IOI) in ms between notes positioned at quarter notes, as indicated in Figure 1, expressed in beats per minute (bpm). Tempo variability was measured as a coefficient of variation (CV) between the IOIs by considering the *SD* IOIs/mean IOIs.

Tempo matching was further assessed with circular statistics, using the Circular Statistics Toolbox for MATLAB (Berens, 2009). This technique enabled to determine if the sung IOIs consistently matched the tempo of the model or the metronome. With this technique, quarter note onsets sang by the participant were represented as angles on a circle, with one circle cycle (360°) corresponding to the expected period between two quarter notes of the model/metronome, or in simpler terms, the tempo. The position of each note on the circle is used to compute a mean resultant vector. The length of the mean resultant vector (VL) indicates how clustered are the data points around the circle. Values of VL range from 0 to 1; the larger the value, the more clustered are the points on the circle, indicating that the time interval between quarter notes tends to match the expected tempo with higher consistency. The Rayleigh z test of periodicity is used to test if the mean resultant vector is strong enough to indicate a constant period matching between the sung quarter notes and the expected tempo (Wilkie, 1983). A Rayleigh test with a p-value <.05 was considered as successful period matching. When the VL was used for comparison between groups, a logit transform was applied on the value with the following formula: logVL= -1*ln(1-VL) to normalize the distribution.

In the *Synchronous* and *Metronome* conditions, provided that the participant successfully matched the tempo of the song, the accuracy (or phase matching) of synchronization was further assessed with the mean asynchrony. This measure is the signed difference in milliseconds between participants' note onset time and the corresponding note onset time of the song model or closest metronome tick, using linear statistics. Since participants' singing record was in the same file as the song model and the metronome, timing required no correction in the alignment between tracks to compute the asynchrony.

2.3.1.2 Pitch

Pitch intervals were calculated in cents, where 100 cents equal a semitone, based on the equal tempered scale. The following formula was used to transform F0 ratio in cents: 1200*log2 (median $F0_{[x+1]}$/median $F0_{[x]}$), where *x* was the current note.

Pitch contour corresponds to the direction of the pitch between two successive notes, which can stay the same, goes up or down. Pitch intervals between -50 and 50 cents where consider as no change in pitch contour; pitch intervals larger than 50 cents were considered as ascending and pitch intervals smaller than -50 cents as descending. Pitch contour was then compared to the expected contour of the song according to written notation, and the percentage of contour errors (i.e. pitch intervals that were sung in the wrong direction) was calculated.

Pitch-interval deviation was calculated as the absolute difference in cents, between the signed pitch intervals produced and the corresponding expected pitch interval from the musical notation of the song (see Equation 1 below, *PI* is the produced pitch interval in cents and *MI* the model interval in cents; *x* is the current interval). This means that the direction of the interval was taken into consideration when computing pitch-interval deviation. Pitch-interval deviations of more than 50 cents were considered as pitch-interval errors, and the percentage of pitch-interval errors was calculated.

Equation 1: Pitch-Interval Deviation = $|PI_x - MI_x|$

The mean pitch-interval deviation was averaged across all produced pitch intervals (see below Equation 2, where *N* is the number of pitch intervals in the song). This measure is different from the pitch-interval accuracy measure described by Berkowska & Dalla Bella, (2013) and Pfordresher et al., (2010). Contrary to their measure of pitch-interval accuracy, the direction of the interval was taken into consideration here.

Equation 2: Mean Pitch-Interval Deviation = $\frac{\sum_{x=1}^{N} |PI_x - MI_x|}{N}$

Pitch-interval consistency is indicative of stability in one's production of pitch intervals. It measures how consistent is the interval deviation produced every time by the participant for a given interval. This is measured by the standard deviation of all pitch-interval deviations produced in each pitch-interval class and then averaging across interval classes. *Gens du Pays* has eight interval classes (pitch intervals of -700, -300, -200, 0, 300, 400, 700, 900, cents) and *Happy Birthday* has five (-400, -200, -100, 0, 200). Pitch intervals that only occurred once in a song were excluded from this analysis, which meant removing six intervals from *Happy Birthday* renditions. Smaller values indicate better pitch-interval consistency. This measure was computed as described by Berkowska & Dalla Bella (2013) and Pfordresher et al. (2010) when referring to pitch-interval precision. The following formula only differs by expressing the equation based on the interval in cents rather than F0 (see Equation 3, where *M* is the average produced interval in cents for a given pitch-interval class and *PI* is each produced interval; *x* is the current interval, *PC* the current pitch-interval class, and *N* the number of intervals in *PC*; *Y* is the number of pitch-interval classes). A higher value indicates less consistency in pitch-interval production.

Equation 3: Pitch-interval consistency = $$\frac{\sum_{PC=1}^{Y}\sqrt{\frac{\sum_{x=1}^{N\,PC}(PI_x - M_{PC})^2}{N_{PC}}}}{Y}$$

Since results obtained for the percentage of pitch-interval errors, mean pitch-interval deviation, and pitch-interval consistency were highly correlated (with $r_{(16)}$ being .65 - .97), and that preliminary analysis did not indicate any significant difference between these measures, we report all three measures in the *Memory* singing condition only and focus on the percentage of pitch-interval errors in the other singing conditions.

After the *Memory* condition, participants heard the singing model before singing (*Solo* singing) or while singing (*Synchronous* singing). To measure how participants would adapt to the model's pitch, we examined absolute-pitch deviation errors in these two conditions. Absolute-pitch deviation was measured as the interval in cents between the sang pitch and the corresponding pitch of the model, regardless of octave difference (see Equation 4, where P is the produced pitch, I is the model ideal pitch, and x the current pitch in the melody). Pitch deviations by more than 50 cents were considered as errors and the percentage of absolute-pitch errors was calculated.

Equation 4: Absolute-Pitch Deviation = $$\left|12 * log2\left(\frac{P_x}{I_x}\right)\right| mod12 * 100$$

2.3.2 Tapping data

Recorded taps were first pre-processed to remove inter-tap intervals (ITIs) smaller than 150 ms. Taps were also considered outliers and were removed if they were more than 50% smaller or larger than the median ITI produced by each participant (median ITI ± [median ITI*0.5]). 6.5% of taps were removed on average (*range*: 3.3-20.6%), leaving between 54 and

61 taps per trial and participant for analysis. The first 50 taps produced by each participant were used for the analysis.

The inter-tap variability (CV; *SD* of ITI/mean ITI) was calculated. Period-matching between beat and taps was further assessed with circular statistics, with the same procedure described for the singing task (Rayleigh z test p-value <.05). Group comparisons were computed on the log transform of the mean resultant vector (logVL) from the circular statistics calculations. Using linear statistics, the accuracy of synchronization was measured by the mean relative asynchrony between taps and beats' onset time in milliseconds.

2.3.3 Statistic Analyses

Statistical analyses were performed in SPSS (IBM SPSS Statistics, version 24, 2016). A mixed repeated-measures ANOVA, with Group as the between-subjects factor, were used for group comparison. Because of the small sample sizes, a statistical approach based on sensitivity analysis was applied. This was done to ensure that the significant effects were reliable when assumptions on residuals' normality distribution and homogeneity of variance were violated (Thabane et al., 2013). When these assumptions were violated, the approach employed was as follows: 1) inspect residuals to identify outliers (identified using Q-Q plot and boxplot), 2) re-run the mixed-design ANOVA without the outliers and assess the consistency of the previous significant results, 3) confirm the results with a non-parametric test of the significant comparisons (Thabane et al., 2013). When unequal variance between groups was not resolved by removing the outliers, Welch's test was performed. If the effect was robust to this procedure, the original ANOVA is reported, otherwise the ANOVA without the outliers or the Welch's test results are reported. Greenhouse-Geiser correction was applied whenever the sphericity assumption was violated. A Bonferroni correction was used for post-hoc comparisons. To indicate the estimated effect sizes, partial eta-squared values are reported for repeated-measures ANOVA and Hedge's *g* statistic was computed for the other comparisons.

3. Results and Comments

3.1 Singing from Memory

Because singing abilities have never been studied in beat-deaf individuals, we first examined singing from memory with and without lyrics. Both acoustic and perceptual measures were used to qualify singing proficiency in beat deafness.

3.1.1 Timing

More rhythm errors were produced on average in the *Lyrics* condition than in the */la/* condition, especially by the beat-deaf participants. This was supported by a main effect of Version $F(1,15) = 6.9$, $p = .02$, $\eta^2 = .32$. However, the interaction with Group was no significant, $F(1,15) = 0.4$, $p = .54$, when removing one outlier from the control group (C8, Figure 2A). There was no main effect of group either, $F(1,15) = 1.3$, $p = .28$, $\eta^2 = .08$. There was no group difference in tempo nor in tempo variability (CV; Table 2). Tempo variability was larger in the sung version with lyrics than on /la/ across groups, $F(1,16) = 4.5$, $p = .049$, $\eta^2 = .22$.

3.1.2. Pitch

For the percentage of contour errors, both groups produced more errors when singing with lyrics, $F(1,16) = 5.7$, $p = .03$, $\eta^2 = .26$ (Table2). The groups were similar on the production of contour, $F(1,16) = 1.5$, $p = .24$, $\eta^2 = .09$, and the interaction between Version and Group was not significant, $F(1,16) = 1.6$, $p = .22$, $\eta^2 = .09$. Overall, both groups were generally more accurate to produce the correct pitch intervals when singing with /la/. Note that the version on /la/ was produced after having heard the model twice with lyrics and may explain slightly better performance on /la/ than with lyrics. Nevertheless, beat-deaf participants tended to be worse to produce the expected pitch interval, producing a higher percentage of pitch-interval errors, based on a 50 cents deviation criterion (Figure 2B), $F(1,16) = 2.0$, $p = .17$, $\eta^2 = .11$. Beat-deaf songs deviated more in pitch than controls' ones in both accuracy, $F(1,16) = 4.4$, $p = .052$, $\eta^2 = .22$ (Figure 2A), and consistency, $F(1,16) = 4.4$, $p = .053$, $\eta^2 = .21$ (Figure 3). Three beat-deaf participants (B1, B5 and B6) participants sung very poorly relative to the other participants, which can explain the close to significant group effects. Two control participants, C5 and C8, could also be considered poor singers having a mean pitch-interval error larger than 100 cents (Figure 3A). We should mention, however, that B1, B5, and B6 sang at the fastest rate in the beat-deaf group, as did C5 and C8 in the control group when singing on /la/. Singing at a faster tempo might have worsened their singing performance by reducing pitch precision (Dalla Bella et al., 2007). At the same time, the fact that these participants sang faster in the first place might reflect their uneasiness in singing.

[Insert Table 2 here]

[Insert Figure 2 here]

[Insert Figure 3 here]

3.1.3 Perceptual Judgments

Each rendition of singing from memory was presented to a group of unselected judges via online assessment (Survey Guizmo platform). The final sample included 59 judges (37 females, mean age = 27.9 years, *SD* age = 10.9, age range: 18-61-years-old). On average, judges reported 4.6 years of musical training (*SD* = 5.7, range: 0–20 years). Judges were asked to evaluate on a scale from *1 (Out of tune)* to *10 (In tune)* each song. The instructions specified to focus on pitch alone when making the judgment and ignore other aspects of the performance (such as voice quality and timing). Before starting, the judges heard a perfectly sang version of each song. These examples were the models used in the singing task. Each singing file received at least 15 judgments (19 on average). Ratings were normalized for each judge to reduce the influence of response style when averaging the judgments from all participants. Each judge's responses were transformed to a z-score according to his own mean and *SD* rating across files. A positive value indicates that the singing production was rated toward a good performance and a negative value toward a poor performance.

Participants' singing, based on accuracy of pitch, were hard to rate according to the independent judges. Their normalized ratings did not differentiate the groups, $F(1,16) = 2.4$, $p = .14$, $\eta^2 = .13$, but were more variable for the beat-deaf songs (Figure 4). This was supported by a significant effect of Group, $F(1,16) = 11.2$, $p = .005$, $\eta^2 = .44$ on the variability of ratings only when correcting for the control outlier (C2; Figure 4). Concordant with the acoustic measures of pitch accuracy, poor singers (B1, B5, B6, C5, and C8) were also the one who received the lowest ratings from perceptual judgments.

[Insert Figure 4 here]

3.2 Singing to a Model

As found above, beat-deaf participants do not seem to suffer from a noticeable difference in singing compared to other occasional singers. The main objective of the study was to test if they would show a deficit when synchronizing their songs to others, as spontaneously done for birthdays. This was done by testing how participants adapted their singing productions while singing with a pre-recorded model. For comparison, singing alone after the model and with a metronome were also assessed.

3.2.1 Timing

There was a clear difference between groups on the percentage of rhythm errors produced in *Synchronous singing with lyrics*, Welch's test $t(13.0) = 6.3$, $p = .03$, $g = 1.17$, with only one beat-deaf participant's score overlapping with those of control participants, excluding outlier C7 (Figure 5A). No other comparison between groups or conditions was statistically significant for this measure ($ps > .09$). In general, variability in tempo was higher in beat-deaf than control participants, even when considering the influence of outliers (C4 and C8), $F(1,16) = 4.8$, $p = .04$, $\eta^2 = .23$ (Figure 5B). The difference was not significant between singing conditions, $F(2,32) = 1.9$, $p = .17$, $\eta^2 = .10$, nor the interaction with Group. Both groups were overall more variable when singing with lyrics than on /la/, $F(1,16) = 5.2$, $p = .04$, $\eta^2 = .24$.

[Insert Figure 5 here]

Circular statistics were used to test if participants could synchronize their singing with the model, that is, could match the tempo of the model while singing along. In synchronous singing, nine of 12 control participants successfully matched the tempo of the model. Unexpectedly, six of the eight beat-deaf participants also succeeded to match the tempo of the model with lyrics and on /la/. Consistency of successful synchronization was measured using the logVL (Figure 6). As can be seen in Figure 6, the beat-deaf participants were less consistent than controls in singing with both lyrics and on /la/, Welch's test, $t(10.8) = 5.2$, $p = .04$, $g = 0.96$ and $t(6.7) = 6.1$, $p = .04$, $g = 1.40$, respectively. However, they were as accurate as control participants in their synchronization, with a mean relative asynchrony between sung and model beat onsets of 19 ms ($SE = 9$, range: -15 ms to 49 ms) and 36 ms ($SE = 12$, range: 2 ms to 81 ms) with the lyrics and /la/ version, respectively (controls: 24 ms and 19 ms; Welch's test $t(12.9) = 0.1$, $p = .75$; $t(11.4) = 0.9$, $p = .37$, $g = 0.46$).

[Insert Figure 6 here]

Singing after the model did not seem to influence much the tempo of singing performance. The tempo of songs was highly correlated between the *Memory* and the *Solo* singing conditions, for both singing versions and groups ($r = .80$-$.93$, $ps = .005$-$.001$). Only when singing with the lyrics the difference between participant's singing tempo and the model was smaller in the *Solo* singing condition ($M = 15$ bpm, $SE = 5$ bpm) than in the *Memory* condition ($M = 42$ bpm, $SE = 12$ bpm), paired t-test $t(17) = 3.4$, $p = .004$, $g = 0.67$. However, only two control participants and one beat-impaired participant matched the tempo of the model in the *Solo* singing condition when they sang with the lyrics (according to the results of the Rayleigh test).

Singing to a metronome appears to be a harder task than singing to a pre-recorded voice for occasional singers. When singing with lyrics, only half the control participants matched the tempo of the metronome and two beat-deaf participants did. The task was easier for participants when they sang on /la/; nine control and three beat-deaf participants could match the tempo of the metronome. The same measures of consistency and accuracy as examined above were computed in the best condition, that is when singing on /la/ with the metronome. For consistency, the nine control participants who could period match the metronome had a mean logVL of 0.97 ($SE = 0.13$, range: 0.43-1.39) and the 3 beat-deaf participants obtained a mean logVL of 1.1, 1.08 and .76, within the range of the control group. Accuracy of synchronization was of -7 ms ($SE = 15$, range: -81 ms to 55 ms) for the control participants and of 36 ms, -21 ms, and -1 ms for the beat-deaf participants, again being similar to most control participants. One possible reason for the success of these three beat-deaf participants to match the metronome might be that these participants had the closest tempo to the metronome in singing from memory (118, 119, 123 bpm) and therefore did not have to adjust it much to the metronome of 118 bpm or 120 bpm.

3.2.2 Pitch

For pitch measures, we compared the influence of singing with a model between groups. *Synchronous* and *Solo* singing conditions were compared on the percentage of pitch-interval and absolute-pitch errors produced (Figure 7). In these conditions there was no clear advantage of singing with the lyrics or on /la/. For the percentage of pitch-interval errors, we found no significant difference between groups, $F(1,16) = 2.5$, $p = .13$, $\eta^2 = .14$. In both groups, fewer pitch-interval errors were produced in the *Solo* condition than in the *Synchronous* singing condition, $F(1,16) = 5.7$, $p = .03$, $\eta^2 = .26$. Thus, for pitch intervals we found similar results to those reported for the *Memory* condition. There were further strong positive correlations between the *Memory*, *Solo*, *Synchronous,* and even the *Metronome* singing conditions on this measure ($r_{(16)}$ = .64-.92, *ps* <.005). In contrast to pitch-interval errors beat-deaf participants produced more absolute-pitch errors than control overall, $F(1,16) = 9.4$, $p = .007$, $\eta^2 = .37$ (Figure 7). Both groups also made less absolute-pitch errors in the *Synchronous* condition than in the *Solo* condition, $F(1,16) = 7.8$, $p = .013$, $\eta^2 = .33$. Still, the percentage of absolute-pitch errors produced correlated strongly between the *Solo* and *Synchronous* conditions ($r_{(16)}$ = .66-.68, *ps* <.003). Thus, it seems that beat-deaf participants had more difficulty to match the exact pitch of the model, whether it was after hearing the model or while singing with the model.

[Insert Figure 7 here]

3.3 Sensorimotor synchronization in Relation to Singing Abilities

In this section, we aimed to determine if sensorimotor synchronization abilities were related to synchronization in tapping and in singing, as suggested by Dalla Bella et al. (2015) and Patel (2006).

In tapping to a metronome, variability (CV) did not differ between beat-deaf and control participants, $t(10.8) = 1.6$, $p = .23$, $g = 0.61$. All participants could synchronize their taps to the tones of the metronome, by the Rayleigh z test of periodicity. There was no significant difference between the groups for the relative mean asynchrony between taps and metronome beat, $t(15.9) = 0.7$, $p = .41$, $g = 0.38$. However, the synchronization of beat-deaf participants was less consistent (logVL: $M = 1.7$, $SE = 0.3$) than the control participants (logVL: $M = 2.7$, $SE = 0.2$), $t(13.6) = 9.5$, $p = .01$, $g = 1.43$, as was found in synchronous singing. Tempo variability in tapping (CV) was associated with singing tempo variability in the *Solo* ($r_{(16)} = .49\text{-}.80$, $ps = .03\text{-}.04$), but not the *Synchronous* singing condition ($r_{(16)} = .04\text{-}.11$, $ps = .65\text{-}.87$). The correlation with singing to a metronome was significant only when singing with lyrics ($r_{(16)} = .49$, $ps = .04$), and mostly driven by the control group ($r_{(10)} = .58$; beat-deaf group: $r_{(10)} = .25$). The correlation between tapping and singing to a metronome on /la/ was also strong in the control group, although not significant ($r_{(10)} = .62$, $p = .058$). The consistency of synchronization when tapping (logVL) did not correlate with synchronization in either the *Synchronous* or *Metronome* singing conditions ($r_{(16)} = .07\text{-}.39$, $ps = .11\text{-}.77$).

For the association between singing pitch accuracy and tapping synchronization, less consistent tapping (smaller logVL) was associated with a higher percentage of pitch-interval errors produced in the *Synchronous* condition when singing with lyrics ($r_{(16)} = -.48$, $p = .046$). However, one control participant was influencing significantly this correlation which was no longer significant when he was removed ($r_{(15)} = -.38$, $ps = .13$). The correlation was also not significant when considering solely the beat-deaf group ($r_{(8)} = -.22$, $p = .61$). The accuracy of tapping (asynchrony) was significantly correlated with the percentage of pitch-interval errors only in the control group for the synchronous singing with lyrics ($r_{(10)} = -.66$, $p = .036$). Other correlations between pitch-interval errors and tapping consistency or accuracy were not significant. For the percentage of absolute-pitch errors produced, there was one significant correlation with tapping performance. In the beat-deaf group, absolute-pitch errors produced in the *Solo* condition when singing on /la/ correlated positively with the relative asynchrony when tapping ($r_{(8)} = .79$, $p = .02$). This was not the case for the control group ($r_{(10)} = .04$, $p = .92$).

4. General Discussion

In this study, we compared individuals with a beat-finding impairment and matched neurotypical adults on their singing abilities with a particular interest in vocal imitation and synchronous singing. The idea was to test if a common auditory-motor coupling mechanism may be involved in beat synchronization and singing accuracy. The results show that globally beat-deaf participants were normal at singing from memory. When singing alone or along with a model, the beat-deaf participants produced a higher percentage of absolute-pitch errors than control participants. Beat-deaf participants also showed more difficulty to synchronize with the model and with the metronome when singing. However, the association of deficits between singing accuracy and sensorimotor synchronization to metronome was not systematic. The variability in singing performance found in the beat-deaf group provide mitigated support to the *vocal learning and synchronization hypothesis,* which suggests that synchronization to musical beat is a by-product of vocal imitation abilities (Patel, 2006; Patel et al., 2009).

One main question of this study was whether poor pitch-matching could be associated with poor beat synchronization. When singing a familiar song from memory, beat-deaf participants were generally similar to control participants. Perceptual judgments made on the productions did also not show a significant difference between groups. However, judgments made on singing accuracy was more variable for the beat-deaf than the control group. In both groups, the variability in performance was quite high, with some participants who could more clearly be classified as “poor-pitch singers” (Dalla Bella, 2015; Dalla Bella & Berkowska, 2009a; Pfordresher & Brown, 2007). Considering that five participants could be considered as “poor-pitch singers” in our sample of 18 participants, this corresponds to the proportion of 28% of “poor-pitch singers” reported in the study of Berkowska & Dalla Bella (2013). More interestingly, these five participants were not amongst the poorest to synchronize with a model or a metronome when singing or when tapping. Overall, the variability found in the beat-deaf group might only reflect the normal variability of singing abilities found in the general population, suggesting no direct association between pitch accuracy and beat synchronization in this group.

The beat-deaf group showed clearer distinctions with the control group when singing after a model or with a model. A robust finding of the current study is that the beat impairment of participants extended to the consistency of timing, both in singing with a model and in tapping to a metronome; reduced synchronization consistency was found in tapping and in singing. While most beat-deaf participants could match the tempo of the model in the *Synchronous* singing condition, their synchronization was less consistent than control participants. The fact that most beat-deaf participants could match the tempo of the model might seem surprising at first, but perhaps participants could match the model interval by interval instead of having internalized the tempo of the song. The reduced number of beat-deaf participants who could time their singing to a metronome, which removed the possibility to simply copy the model interval by interval, showed more apparently this group impairment in timing. There was no difference between groups when considering tempo variability of singing from memory and spontaneous tapping, suggesting that the deficit may be primarily in entrainment. These results points toward a deficit in beat deafness to synchronize with a beat, whether it is in a tapping task, as in the M-BAT, or when singing. It also suggests that the same auditory-motor coupling mechanism may be involved for timed behaviors across effectors and tasks.

Auditory-vocal coupling involves a vast brain network, including the posterior superior temporal gyrus (pSTG), the dorsal auditory pathway (dorsal parietal and premotor cortices), the inferior frontal gyrus (IFG), the inferior parietal sulcus (IPS), the anterior cingulate cortex (ACC), the supplementary motor area (SMA) and pre-supplementary motor area (preSMA), the cerebellum, the anterior insula (aINS), and basal ganglia (Belyk, Pfordresher, Liotti, & Brown, 2016; Dalla Bella & Berkowska, 2009b; Kleber & Zarate, 2014; Pfordresher, Demorest, et al., 2015; Zarate, 2013). Interestingly, a recent study comparing singing and cello playing found that activation during a production task overlapped (Segado, Hollinger, Thibodeau, Penhune, & Zatorre, 2018), suggesting that activities requiring the coupling and monitoring of auditory pitch and motor action, even if the effector is not vocal, may recruit the auditory-vocal pathway. Beat-based synchronization seems to use a similar network of regions, including the auditory cortices, the SMA and preSMA, the cerebellum and basal ganglia (Chapin et al., 2010; Grahn & McAuley, 2009; Grahn & Rowe, 2013; Kung, Chen, Zatorre, & Penhune, 2013; Rao, Mayer, & Harrington, 2001). Thus, there seems to be support to the hypothesis of overlap between vocal imitation and beat-based entrainment in human brain networks. However, this remains to be tested directly, as overlap in brain networks does not imply shared neural resources (Peretz, Vuvan, Lagrois, & Armony, 2015).

In both groups, singing along with the model caused more pitch-interval errors, compared to singing solo after the model. In contrast, both groups also produced less absolute-pitch errors when singing along than singing alone. Although the effect of condition on measures of pitch accuracy was found in both groups, beat-deaf participants made more absolute-pitch errors than control participants, while there was no significant difference between groups for pitch-interval errors. This result is interesting because it could be argued that absolute-pitch matching is more representative of vocal-pitch imitation than pitch-interval matching. Indeed, in the context where participants were singing a familiar song, the measure of pitch-interval accuracy is quite independent of the model and more probably based on one's internal representation of the melody. On the other hand, the absolute-pitch measure is based entirely on the participants' ability to reproduce the exact pitch of the model. Prior research suggests that poor singing ability results in most cases from inaccurate sensorimotor translation or inefficient error correction (Dalla Bella et al., 2007; Hutchins, Zarate, Zatorre, & Peretz, 2010; Hutchins & Peretz, 2012; Pfordresher & Brown, 2007; Pfordresher et al., 2010). Sensorimotor translation and error correction mechanisms are viewed as the essential components of the sensorimotor loop supporting auditory-motor coupling, and are relevant to both pitch-matching (Berkowska & Dalla Bella, 2009a; Dalla Bella et al., 2011; Hutchins & Peretz, 2012; Pfordresher & Brown, 2007; Pfordresher, Demorest, et al., 2015; Pfordresher, Halpern, et al., 2015; Pfordresher & Mantell, 2014) and synchronization to the beat (Maes, Leman, Palmer, & Wanderley, 2014; Patel, 2006; Phillips-Silver & Keller, 2012; Ross, Iversen, & Balasubramaniam., 2016). Poor singing found in beat-deaf cases could result from reduced efficiency in pitch perception, sensorimotor translation, motor control, and error correction mechanisms.

We cannot exclude the influence of pitch perception on our results, since beat-deaf participants had lower scores than control participants on the pitch perception tests of amusia. However, choral singing improved pitch-matching in both groups, suggesting sensitivity to pitch in beat-deaf participants. Moreover, singing from memory was correct. Tremblay-Champoux et al. (2010) used the same experimental protocol as here in individuals with pitch deafness. In contrast to our beat-deaf group, pitch-deaf individuals are generally impaired on pitch-interval accuracy, which was not the case in our group of beat-deaf individuals. This suggests that the singing profile in beat deafness is different from pitch-deaf amusics and reflects mostly their timing deficit. Since we did not find clear differences between control and beat-deaf participants when singing a familiar song from memory, it also suggests that most beat-deaf individuals are capable of learning the associated motor plan corresponding to the accurate production of a vocal melodic sequence and are able to accurately translate this representation to a motor action (Berkowska & Dalla Bella, 2009b; Pfordresher, Demorest, et al., 2015; Pfordresher, Halpern, et al., 2015). Further, given that groups were similar when singing from memory and on spontaneous tapping, it seems reasonable to discard the possibility that basic motor skills or motor control differs between groups.

In regard of the correction mechanisms, the additional feedback available in the singing task compared to the tapping task seems to have had little impact on the synchronization performance. When singing with the model or the metronome, participants could monitor their own performance through hearing their production and the model, producing self-generated feedback. Tapping did not produce any sound here and thus only provided tactile feedback and maybe visual feedback if participants looked at their taps. Nevertheless, the same synchronization impairment was observed. The possibility for error correction mechanisms to be inefficient in individuals with a beat-finding impairment has already been suggested by a previous study using a tapping task (Palmer et al., 2014). In their study they found that beat-deaf participants had more difficulty to correct their synchronization to a metronome after a perturbation occurred in the sequence. Poor pitch-matching in beat-deaf participants may reflect either poor error correction mechanism in regards of pitch production or it may be the cost of having to deploy more effort to adjust with the model's timing. Supporting the first hypothesis, in the beat-deaf group only, absolute-pitch errors produced in the *Solo* singing condition when singing on /la/ correlated positively with the relative asynchrony when tapping (i.e. reduced tapping accuracy). This correlation was, however, no significant for synchronous singing. In the future, one should look at single pitch matching task as previously done with pitch-deaf amusics

(Hutchins et al., 2010). This would remove the possible confound or interference caused by the timing implied with singing a melody.

In summary, findings from the current study suggests that the auditory-motor coupling mechanisms involved in beat synchronization are not effector or task specific. However, evidence for a common sensorimotor loop for pitch matching and beat synchronization is not strong and awaits further investigation. Studying the association of both behaviors in other populations like individuals who stutter might give further insight on the question. Indeed, comparing vocal imitation of speech and singing in individuals who stutter could allow to also test for shared sensorimotor translation mechanisms between speech and song.

Acknowledgments

We would like to thank Mihaela Felezeu for her help with data collection. We would also like to thank Michael W. Weiss and Naeem Komeilipoor for advice regarding data analysis and help with collecting the ratings from the perceptual task, as Chanel Marion St-Onge for her contribution to the making of the perceptual task as well. Thanks should be made also to the research assistants (Madeleine Borgeat, Alejandro Hernandez, and Charline Lavigne), for their help with the tedious task of segmenting audio files for acoustic analyses.

References

Bégel, V., Benoit, C.-E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia, 94*, 129–138. doi:https://doi.org/10.1016/j.neuropsychologia.2016.11.022

Belyk, M., Pfordresher, P. Q., Liotti, M., & Brown, S. (2016). The Neural Basis of Vocal Pitch Imitation in Humans. *Journal of Cognitive Neuroscience, 28*(4), 621–635. doi:10.1162/jocn_a_00914

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *J Stat Softw, 31*(10), 1–21.

Berkowska, M., & Dalla Bella, S. (2009a). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology, 5*, 69–83. doi:10.2478/v10053-008-0068-2

Berkowska, M., & Dalla Bella, S. (2009b). Reducing Linguistic Information Enhances Singing Proficiency in Occasional Singers. *Annals of the New York Academy of Sciences, 1169*(1), 108–111. doi:doi:10.1111/j.1749-6632.2009.04774.x

Berkowska, M., & Dalla Bella, S. (2013). Uncovering phenotypes of poor-pitch singing: the Sung Performance Battery (SPB). *Frontiers in Psychology, 4*(714). doi:10.3389/fpsyg.2013.00714

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer (Version 6.0.2). Retrieved from http://www.praat.org

Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: the role of attending. *Frontiers in Psychology, 1*, 224.

Dalla Bella, S. (2015). Defining Poor-Pitch Singing: A Problem of Measurement and Sensitivity. *Music Perception: An Interdisciplinary Journal, 32*(3), 272–282. doi:10.1525/mp.2015.32.3.272

Dalla Bella, S., & Berkowska, M. (2009b). Singing and Its Neuronal Substrates: Evidence from the General Population. *Contemporary Music Review, 28*(3), 279–291. doi:10.1080/07494460903404386

Dalla Bella, S., & Berkowska, M. (2009a). Singing Proficiency in the Majority: Normality and "Phenotypes" of Poor Singing. *Annals of the New York Academy of Sciences, 1169*(1), 99–107. doi:10.1111/j.1749-6632.2009.04558.x

Dalla Bella, S., Berkowska, M., & Sowinski, J. (2011). Disorders of Pitch Production in Tone Deafness. *Frontiers in Psychology, 2*(164). doi:10.3389/fpsyg.2011.00164

Dalla Bella, S., Berkowska, M., & Sowiński, J. (2015). Moving to the Beat and Singing are Linked in Humans. *Frontiers in Human Neuroscience, 9*(663). doi:10.3389/fnhum.2015.00663

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods, 49*(3), 1128–1145. doi:10.3758/s13428-016-0773-6

Dalla Bella, S., Giguere, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *J Acoust Soc Am, 121*(2), 1182–1189. doi:10.1121/1.2427111

Dalla Bella, S., Giguere, J. F., & Peretz, I. (2009). Singing in congenital amusia. *J Acoust Soc Am, 126*(1), 414–424. doi:10.1121/1.3132504

Dalla Bella, S., & Sowiński, J. (2015). Uncovering Beat Deafness: Detecting Rhythm Disorders with Synchronized Finger Tapping and Perceptual Timing Tasks. *Journal of Visualized Experiments: JoVE*(97), 51761. doi:10.3791/51761

Dowling, W. J. (1999). 15—The Development of Music Perception and Cognition. In D. Deutsch (Ed.), *The Psychology of Music (Second Edition)* (pp. 603–625). San Diego: Academic Press.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage, 47*(4), 1894–1903. doi:https://doi.org/10.1016/j.neuroimage.2009.04.039

Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex, 23*(4), 913–921. doi:10.1093/cercor/bhs083

Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General, 141*(1), 76.

Hutchins, S., Zarate, J. M., Zatorre, R. J., & Peretz, I. (2010). An acoustical study of vocal pitch matching in congenital amusia. *J Acoust Soc Am, 127*(1), 504–512. doi:10.1121/1.3270391

Kleber, B. A., & Zarate, J. M. (2014). The neuroscience of singing *The Oxford Handbook of Singing*.

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience, 25*(3), 401–420. doi:10.1162/jocn_a_00325

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. (2014). Action-based effects on music perception. *Frontiers in Psychology, 4*(1008). doi:10.3389/fpsyg.2013.01008

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. doi: https://doi.org/10.1016/j.cortex.2008.06.011

Mithen, S. J. (2006). *The Singing Neanderthals: the Origins of Music, Language, Mind and Body* Cambridge, Mass.: Harvard University Press.

Palmer, C., Lidji, P., & Peretz, I. (2014). Losing the beat: deficits in temporal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi:10.1098/rstb.2013.0405

Patel, A. D. (2006). Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music Perception: An Interdisciplinary Journal, 24*(1), 99–104. doi:10.1525/mp.2006.24.1.99

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology, 19*(10), 827–830. doi:https://doi.org/10.1016/j.cub.2009.03.038

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664). doi: 10.1098/rstb.2014.0090

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal Of Human Genetics, 25*, 625. doi:10.1038/ejhg.2017.15

Pfordresher, P. Q., & Brown, S. (2007). Poor-Pitch Singing in the Absence of "Tone Deafness". *Music Perception: An Interdisciplinary Journal, 25*(2), 95–115. doi:10.1525/mp.2007.25.2.95

Pfordresher, P. Q., & Brown, S. (2017). Vocal mistuning reveals the origin of musical scales. *Journal of Cognitive Psychology, 29*(1), 35–52. doi:10.1080/20445911.2015.1132024

Pfordresher, P. Q., Brown, S., Meier, K. M., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *J Acoust Soc Am, 128*(4), 2182–2190. doi:10.1121/1.3478782

Pfordresher, P. Q., Demorest, S. M., Bella, S. D., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical Perspectives on Singing Accuracy. *An Introduction to the Special Issue on Singing Accuracy (Part 1), 32*(3), 227–231. doi:10.1525/mp.2015.32.3.227

Pfordresher, P. Q., Halpern, A. R., & Greenspon, E. B. (2015). A Mechanism for Sensorimotor Translation in Singing. *The Multi-Modal Imagery Association (MMIA) Model, 32*(3), 242–253. doi:10.1525/mp.2015.32.3.242

Pfordresher, P. Q., & Mantell, J. T. (2014). Singing with yourself: Evidence for an inverse modeling account of poor-pitch singing. *Cognitive Psychology, 70*, 31–57. doi:https://doi.org/10.1016/j.cogpsych.2013.12.005

Phillips-Silver, J., & Keller, P. (2012). Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience, 6*(26). doi:10.3389/fnhum.2012.00026

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia, 49*(5), 961–969. doi:https://doi.org/10.1016/j.neuropsychologia.2011.02.002

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature neuroscience, 4*(3), 317–323. doi:10.1038/85191

Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase, 22*(6), 558–565. doi:10.1080/13554794.2016.1242756

Segado, M., Hollinger, A., Thibodeau, J., Penhune, V., & Zatorre, R. J. (2018). Partially Overlapping Brain Networks for Singing and Cello Playing. *Frontiers in Neuroscience, 12*(351). doi:10.3389/fnins.2018.00351

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51*(10), 1952–1963. doi:https://doi.org/10.1016/j.neuropsychologia.2013.06.027

Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., ... & Debono, V. B. (2013). A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. BMC medical research methodology, 13(1), 92.

Tranchant, P., Lagrois, M., Bellemare Pépin, A., G. Schultz, B., & Peretz, I. (2018). *Beat alignment test of the motor origin of musical entrainment deficits*. Manuscript in preparation

Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the Beat: A Large Sample Study of Bouncing and Clapping to Music. *PLoS ONE, 11*(7), e0160178. doi:10.1371/journal.pone.0160178

Tremblay-Champoux, A., Dalla Bella, S., Phillips-Silver, J., Lebrun, M.-A., & Peretz, I. (2010). Singing proficiency in congenital amusia: Imitation helps. *Cognitive Neuropsychology, 27*(6), 463–476. doi:10.1080/02643294.2011.567258

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion, 11*, 262–266.). doi: 10.1037/a0021302.

Vigneault, G., & Rochon, G. (1978). Gens du pays (chant d'anniversaire). Québec, Canada: Longueil.

Wechsler, D., Coalson, D. L., & Raiford, S. E. (1997). *WAIS-III: Wechsler adult intelligence scale*: Psychological Corporation San Antonio, TX.

Wilkie, D. (1983). Rayleigh test for randomness of circular data. *Applied statistics, 32*(3), 311–312.

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience, 7*(237). doi:10.3389/fnhum.2013.00237

Tables

Table 1. Groups' Characteristics

Variables	Beat-deaf (*SD*) *n* = 8	Control (*SD*) *n* = 10	Group Comparison p-value [a]
Age	27.1 (2.2)	25.8 (4.1)	.39
Education	18.5 (2.1)	18.0 (2.6)	.66
Musical Training	0.8 (1.2)	0.2 (0.6)	.26
WAIS-III Digit Span (ss)	8.4 (1.6)	10.0 (4.3)	.29
WAIS-III Matrix Reasoning (ss) [b]	10.7 (2.4)	12.6 (1.7)	.13
M-BAT			
Production Task—			
% Tempo-matched Trials	40.6 (22.7)	95.4 (5.6)	.001
[range]	[10–65]	[85–100]	
Online Test			
Scale	25.3 (1.6)	27.8 (2.3)	.01
Off-key	16.1 (2.8)	20.1 (2.0)	.01
Off-beat	18.9 (2.4)	19.9 (1.1)	.28

Note: ss = standard score. [a] Results from Welch's tests. [b] This measure includes six beat-deaf and nine control participants, since the other participants were too familiar with the test as students in a clinical neuropsychology doctorate program.

Table 2. Groups' Descriptive Results on Acoustic Measures When Singing from Memory

		Lyrics		/la/	
Variable	Group	Mean (*SE*)	Range	Mean (*SE*)	Range
Tempo (bpm)	Control	131 (*14*)	87 - 205	134 (*7*)	114 - 186
	Beat-deaf	163 (*27*)	93 - 324*	138 (*10*)	107 - 178
Tempo CV	Control	.12 (*.02*)	.05 - .24	.09 (*.02*)	.05 - .24
	Beat-deaf	.16 (*.03*)	.07 - .27	.11 (*.02*)	.05 - .25
% Contour errors	Control	10.3 (*2.1*)	0.0 - 18.2	5.0 (*1.7*)	0.0 - 16.7
	Beat-deaf	12.2 (*3.4*)	0.0 - 25.8	10.6 (*2.5*)	0.0 - 19.2

Note: Mean and standard error in parentheses. * One beat-deaf participant sang at a particularly fast tempo of 324 bpm. The next highest tempo in the beat-deaf group was of 221 bpm, being similar to control participant.

Figure Legends

Figure 1. Notation of the Songs Used in the Experiment. (a) *Gens du pays*, and (b) *Happy Birthday*. The straight bars above the notes indicate the position of beats, or note onsets, considered for measuring tempo and tempo variability.

Figure 2. Boxplot of Rhythm and Pitch-Interval Errors When Singing from Memory. Boxplot where boxes corresponds to a 95% confidence interval from the mean based on standard error mean (SEM). The black line within each box indicates the group mean. The vertical lines show two standard deviations from the mean. Each dot represents a participant (control group in black and beat-deaf group in light-gray). **A** Percentage of rhythm errors produced when singing from memory. Outlier participant C8, who was removed from the ANOVA, is identified in the figure. **B** The percentage of pitch-interval errors produced, based on a 50 cents deviation criterion to count errors.

Figure 3. Boxplot of Mean Pitch-Interval Deviation and Interval Consistency When Singing from Memory. Boxplots with black vertical lines indicating two standard deviations from the mean. **A** Mean pitch-interval deviation, in cents, between the produced intervals and the expected intervals from the musical notation of the song. **B** Consistency to produce each pitch intervals of the song. Higher values indicate larger variability, computed in cents. Participants qualified as poor singers are identified for both measures.

Figure 4. Mean Normalized Perceptual Ratings and Variability of Ratings of Singing from Memory. Boxplots with black vertical lines indicating two standard deviations from the mean. **A** Mean normalized ratings (based on z-score from each online judge's ratings). Higher values indicate a more positive appreciation of the singing performance. **B** Variability between the ratings made for each singing production. Higher values mean more variability in the ratings. * indicates the main effect of Group found for the variability of ratings (p =.005), with more variability in ratings for the beat-deaf group productions.

Figure 5. Rhythm Errors and Coefficient of Variation (CV) of Singing Performances in the *Solo*, *Synchronous* and *Metronome* Conditions. Boxplots with black vertical lines indicating two standard deviations from the mean for **A** the percentage of rhythm errors and **B** coefficient of variation (CV) of singing performances in the *Singing Solo*, *Synchronous* and *Metronome* conditions.

Figure 6. Consistency of Synchronization When Singing with the Model. Boxplots of consistency (logVL) of synchronization with the tempo of the model, for participants who successfully matched the tempo of the model or the metronome, with black vertical lines indicating two standard deviations from the mean.

Figure 7. Boxplot of Pitch-Interval and Absolute-Pitch Errors in the *Solo* and *Synchronous Singing* Conditions. Boxplots with black vertical lines indicating two standard deviations from the mean for **A** the percentage of pitch-interval errors and **B** the percentage of absolute-pitch errors of singing performances in the *Solo* and *Synchronous* singing conditions.

Figures

Figure 1. Notation of the Songs Used in the Experiment

(a)

(b)

Figure 2. Boxplot of Rhythm and Pitch-Interval Errors When Singing from Memory

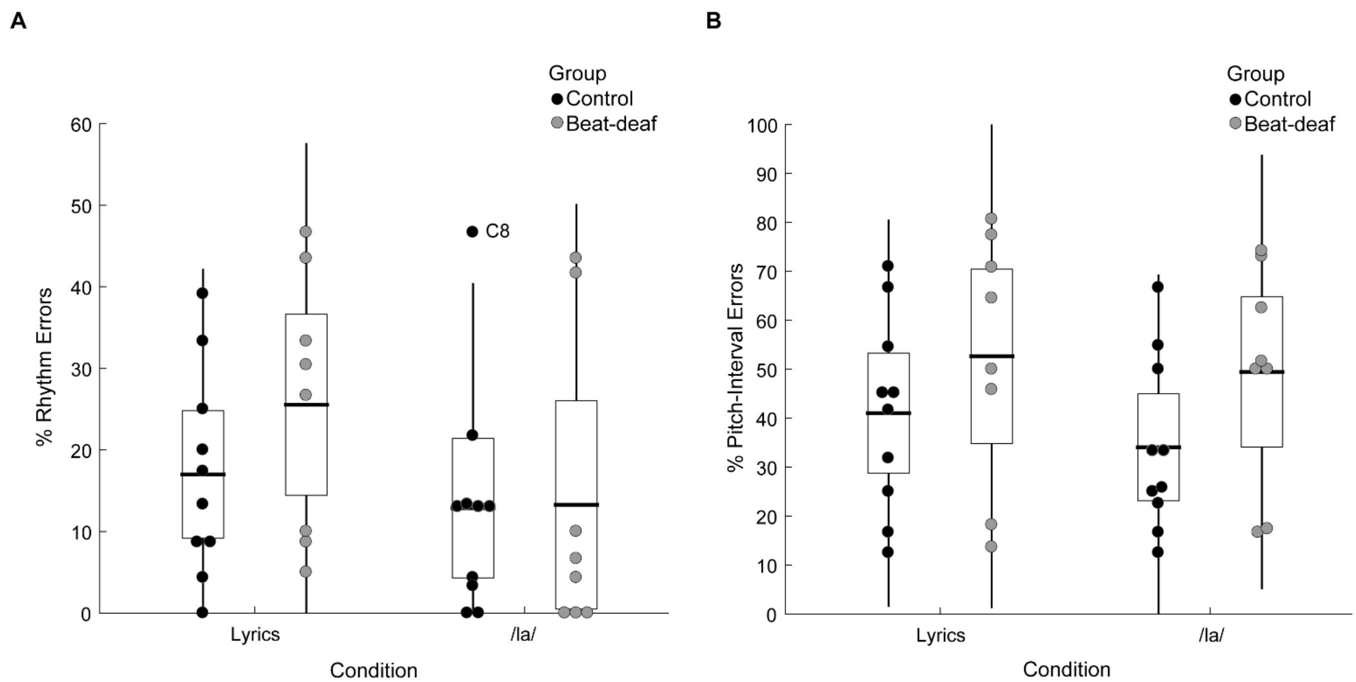


Figure 3. Boxplot of Mean Pitch-Interval Deviation and Interval Consistency When Singing from Memory

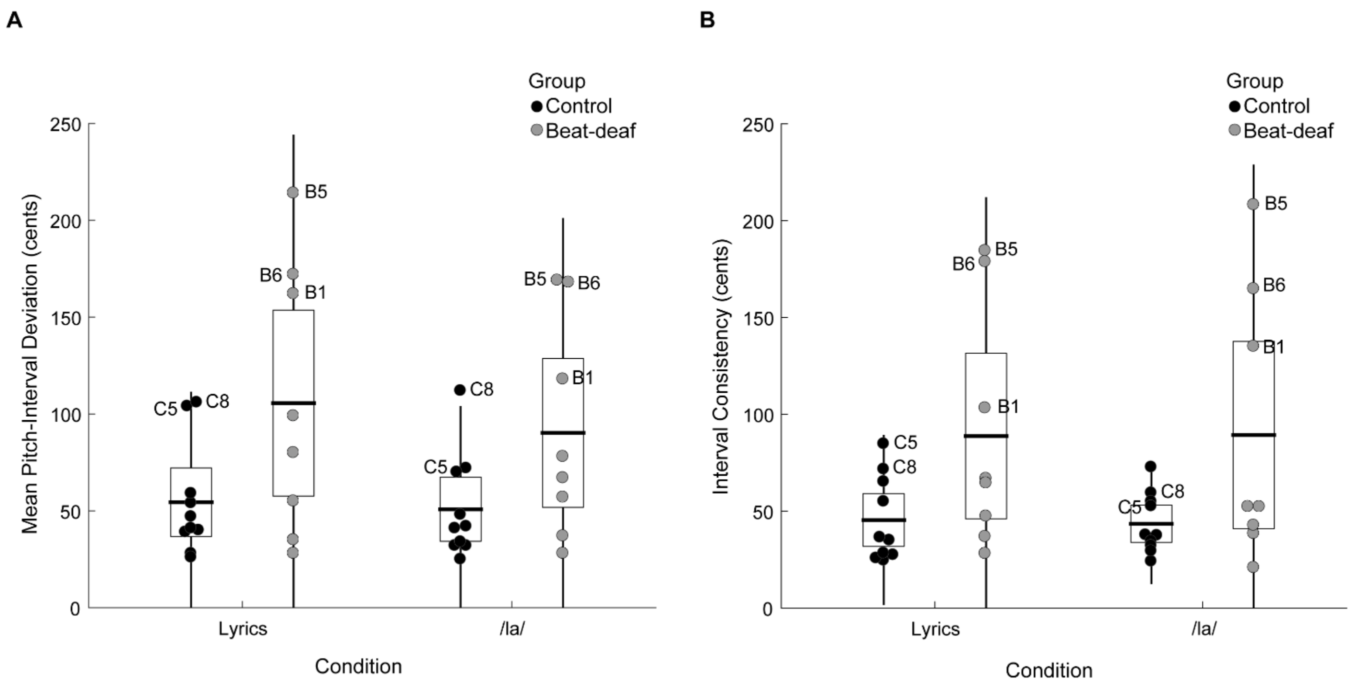


Figure 4. Mean Normalized Perceptual Ratings and Variability of Ratings of Singing from Memory

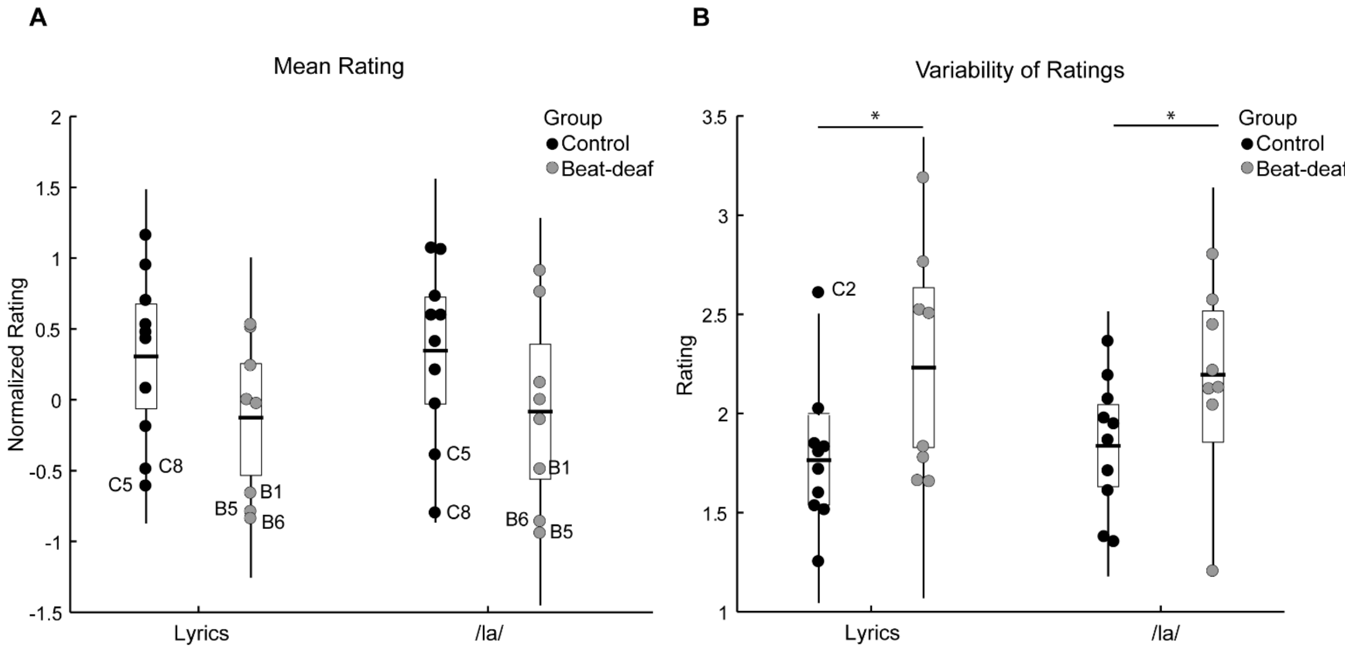


Figure 5. Rhythm Errors and Coefficient of Variation (CV) of Singing Performances in the *Solo*, *Synchronous* and *Metronome* Singing Conditions

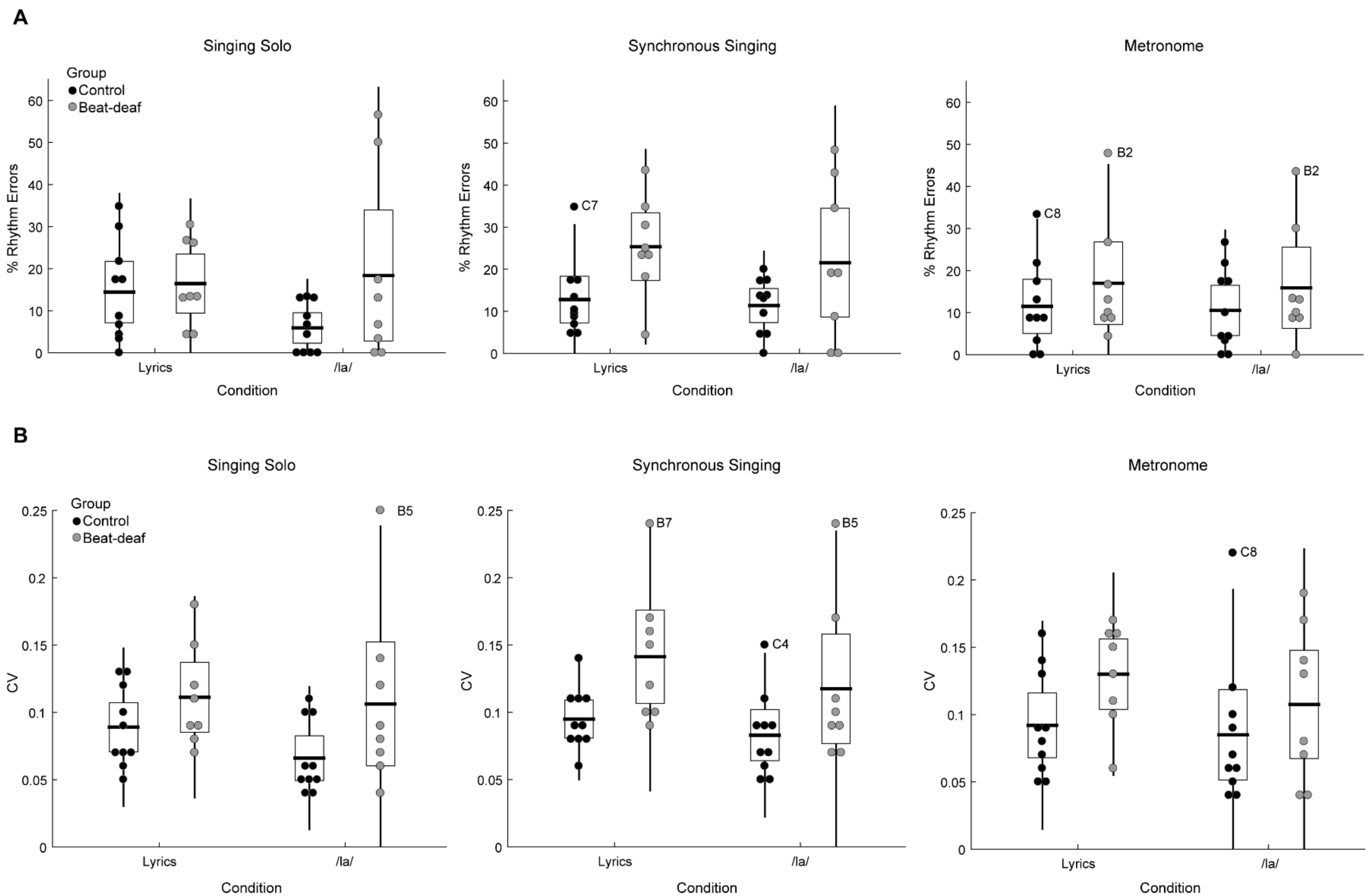


Figure 6. Consistency of Synchronization When Singing with the Model

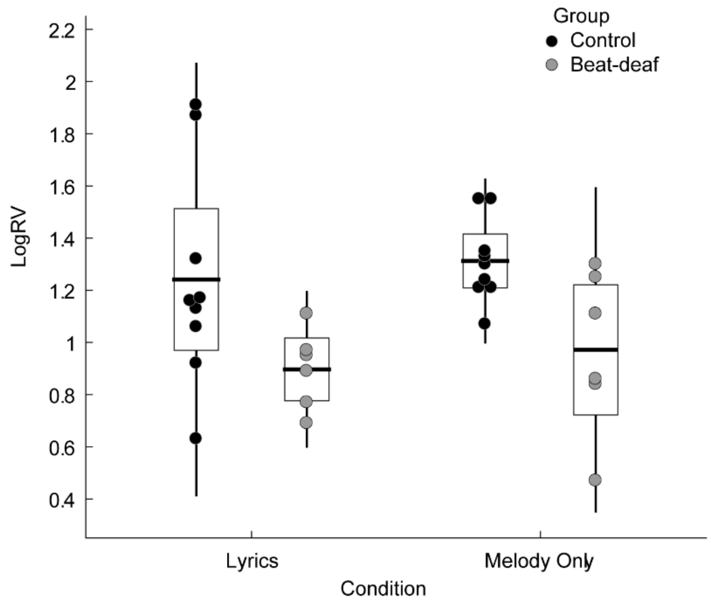
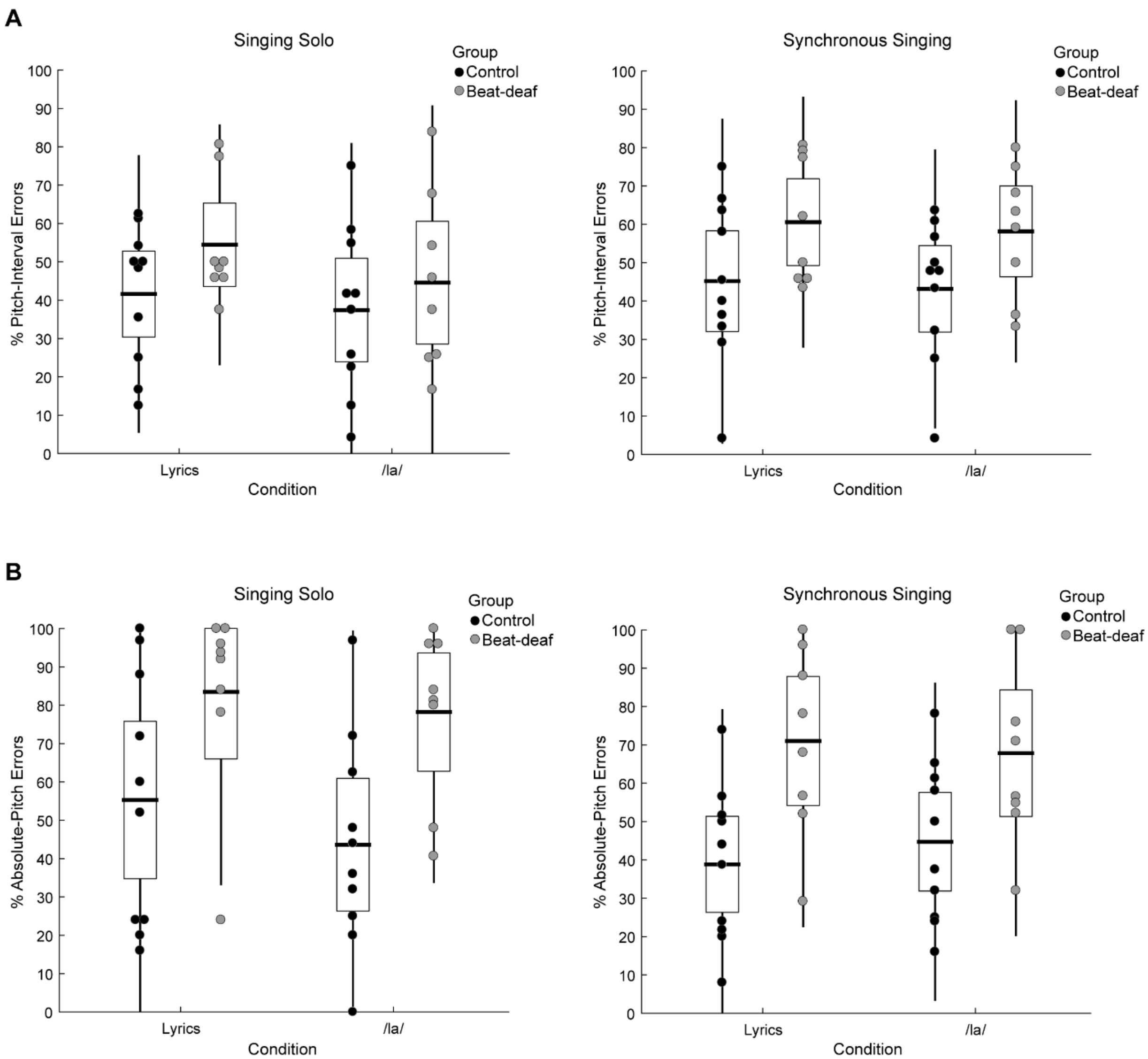


Figure 7. Boxplot of Pitch-Interval and Absolute-Pitch Errors in the *Solo* and *Synchronous* Singing Conditions



Chapitre III : Discussion générale

Rappel des objectifs et sommaire des résultats

L'objectif général de la thèse était de mieux cerner la spécificité des mécanismes d'entrainement au beat musical par l'étude de la synchronisation sensorimotrice à la musique, au chant et à la parole dans l'amusie congénitale. Trois études ont été menées à cette fin.

Le but de la première étude (Article 1) était de déterminer l'impact d'un trouble de la perception des hauteurs de notes sur la perception et la synchronisation au beat musical. Dans cette étude, un groupe de dix individus présentant un trouble de la perception des hauteurs a pris part à deux expériences. Dans la première expérience, la perception et la synchronisation au beat ont été évaluées à l'aide d'extraits musicaux variés en genre et en tempo. Les résultats ont montré que, parmi le groupe d'amusiques, deux étaient comparables aux contrôles à la fois lors de la perception et de la synchronisation au beat, et deux autres étaient également similaires aux participants du groupe contrôle, mais moins réguliers, ou constants, dans leur synchronisation au beat. La performance aux tâches de perception et de synchronisation au beat n'était pas associée à la sévérité du trouble de perception des hauteurs.

Dans la deuxième expérience, l'hypothèse selon laquelle la difficulté à traiter le rythme proviendrait d'un effet 'd'interférence', chez les participants ayant un trouble de perception des hauteurs, a été ciblée plus directement. Les mêmes participants ont accompli une tâche de synchronisation au beat, mais cette fois sur des extraits percussifs, éliminant donc l'influence possible des variations de hauteurs. Les résultats de cette seconde expérience se sont révélés très similaires aux résultats de la première expérience, les mêmes quatre participants amusiques sur dix ayant réussi la tâche de façon comparable aux participants contrôles. Les participants amusiques de cette étude n'ont donc pas montré d'amélioration à la tâche de synchronisation au beat par le retrait de la composante mélodique. Ainsi, les résultats de cette étude semblent aller à l'encontre de l'hypothèse d'interférence et confirmer la dissociation entre le trouble de perception des hauteurs et de synchronisation au beat.

La deuxième étude et la troisième étude de la thèse visaient à évaluer la généralisation du trouble de la synchronisation au beat de la musique à d'autres contextes, soit la parole dans l'étude 2 et le chant dans l'étude 3.

Dans l'étude 2 (Article 2), un groupe de treize amusiques ayant un trouble de synchronisation au beat, dont huit ayant un trouble de perception des hauteurs en concomitance, ont participé à l'étude. La tâche était ici de synchroniser des tapes au beat perçu dans des extraits de chant, de parole régulière et de parole naturelle. Dans l'étude, les amusiques, comparativement aux participants contrôles neurotypiques, ont produit des intervalles de temps plus variables entre leurs tapes (coefficient de variation [CV]) lors de la synchronisation, tant aux extraits de chant que de parole. Les participants amusiques ont, en revanche, également produit des tapes avec un coefficient de variation plus élevé dans des tâches contrôles nécessitant de taper de façon régulière en l'absence d'un stimulus (tempo spontané) et lors de la synchronisation à un métronome. Par ailleurs, les individus amusiques ont montré le même profil de synchronisation que les participants contrôles, demeurant plus réguliers lors de la synchronisation avec des stimuli réguliers (chant et parole régulière) qu'avec la parole naturelle. Ainsi, le trouble de synchronisation au beat de la musique semble affecter également la synchronisation sensorimotrice à la parole dans une tâche de *tapping*. Par ailleurs, en considérant les résultats aux tâches contrôles, il a été proposé que le trouble de la synchronisation chez les participants inclus à l'étude pourrait découler d'un défaut des mécanismes de *timing* interne, lié à la génération et au maintien d'intervalles de temps périodiques.

Dans la troisième étude de cette thèse (Article 3), l'association possible entre les habiletés de synchronisation au beat musical et les habiletés de chant a été explorée. Au-delà de l'intérêt pour le *timing* du chant des participants ayant un trouble de la synchronisation, une attention a également été portée à la justesse du chant, en regard de l'hypothèse de Patel (2006) suggérant un mécanisme commun de couplage sensorimoteur entre l'imitation vocale et la synchronisation au beat. Cette étude incluait huit participants présentant un trouble de la synchronisation au beat, sans atteinte de la perception des hauteurs. Les participants ont chanté une mélodie familière dans quatre contextes : de mémoire, après avoir entendu un modèle (imitation implicite), avec un modèle (chant choral) et avec un métronome. D'emblée, le résultat principal de cette étude est que le trouble de la synchronisation au beat, évalué par une tâche de *tapping*, se manifeste également dans la production du chant choral. Les participants amusiques ont en effet montré plus de difficulté à chanter de manière synchronisée avec le modèle et ont également échoué, pour la plupart, à chanter en synchronie avec un métronome. Lors du chant produit seul, de mémoire, les deux groupes se sont avérés très similaires, et ce tant pour les mesures de justesse que de *timing*, suggérant l'absence de trouble de contrôle vocal et une capacité préservée à se représenter mentalement une mélodie. En contexte d'imitation et de chant choral, les participants amusiques ont montré une plus grande variabilité dans le tempo de leur production et plus de difficulté à reproduire les notes du modèle (c.-à-d. ajuster la hauteur des notes de leur production). En revanche, nous n'avons pas trouvé d'association directe entre l'ampleur de la difficulté à se synchroniser au beat et le pourcentage d'erreurs de hauteur.

Synchronisation au beat dans l'amusie congénitale

Au travers des diverses études de la thèse, nous nous sommes principalement intéressés à la synchronisation au beat, ce comportement reflétant le plus directement l'entrainement (*entrainment*). En introduction, il a été indiqué que la capacité à pouvoir se synchroniser au beat dépendrait principalement de trois processus, soit 1) percevoir et prédire l'occurrence des beats, 2) synchroniser une action motrice au beat et 3) ajuster la coordination temporelle de la réponse motrice pour corriger les erreurs de synchronisation pouvant survenir. Il a été avancé que le dysfonctionnement de l'un ou l'autre de ces processus pourrait causer un trouble de la synchronisation au beat (Van der Steen & Keller, 2013). Dans les prochaines sections, par l'intégration des divers résultats obtenus dans le cadre de la thèse, nous tenterons de qualifier le trouble de synchronisation chez les participants amusiques étudiés, et par la même occasion la spécificité des mécanismes d'entrainement au beat de la musique.

Dans la thèse, une seule mesure directe de la perception du beat a été utilisée. La tâche de la M-BAT implique de déterminer si une piste de métronome entendue en simultané avec un extrait de musique correspond ou non au beat de la pièce. Les participants présentant un trouble de la synchronisation, car ayant été sélectionnés sur cette base, montrent en grande majorité une performance inférieure à des adultes neurotypiques dans cette tâche de perception. Seulement trois de ces participants se sont avérés comparables aux participants contrôles, mais tout en demeurant faible par rapport à la moyenne de ces derniers. Ainsi, on pourrait croire que la majorité des participants étudiés ici présentent à la base une capacité limitée à extraire le beat dans la musique.

De manière assez intéressante, des études récentes ont indiqué des cas où synchronisation et perception pouvaient être dissociées. Sowiński & Dalla Bella (2013) rapportent deux cas de perception intacte, mais de synchronisation déficiente. Peut-être plus intrigant encore, une étude récente de Bégel et collaborateurs (2017) rapporte deux cas d'individus démontrant une faible performance dans une tâche de perception du beat, malgré une performance dans la norme en synchronisation. En étudiant davantage ces derniers, ils ont trouvé que ceux-ci démontraient tout de même un avantage à entendre une séquence de sons régulière versus non-régulière sur le temps de réaction à un stimulus cible, ce que les auteurs ont interprété comme une perception implicite du beat intacte. Ainsi, les auteurs proposent qu'un traitement implicite des régularités rythmiques puisse suffire à la synchronisation au beat, alors qu'un accès à la représentation du beat serait nécessaire pour accomplir la tâche de perception. Cet accès conscient au beat serait donc altéré chez ces participants, au même titre que ce qui a été décrit dans l'amusie portant sur la discrimination des hauteurs (Dalla Bella et al., 2011 ; Dalla Bella, Giguère, & Peretz, 2009 ; Peretz et al., 2009). Étant donné le trouble généralisé des participants inclus dans cette thèse aux tâches de perception et de synchronisation, on pourrait penser que ceux-ci présentent un trouble déjà au niveau des mécanismes de *timing* implicite.

La tâche de *tapping* à un métronome, utilisée au travers des diverses études, avait pour objectif de servir d'indicateur de synchronisation ou de *timing* en enlevant la charge de devoir 'extraire' le beat. Dans cette tâche, les participants amusiques avec un trouble de la synchronisation au beat ont été capable de se synchroniser au tempo du métronome, mais en montrant un plus grand coefficient de variation. Une asynchronie négative a tout de même été retrouvée entre les tapes et les beats du métronome, n'indiquant pas une simple réaction aux sons du métronome. Ainsi, la performance à la tâche de métronome pourrait refléter une sous-estimation de la période du beat et/ou, considérant le plus haut coefficient de variation, une difficulté à ajuster la réponse en fonction de l'erreur de synchronisation.

Par ailleurs, les résultats obtenus à la tâche de *tapping* spontané se sont avérés plus mitigés. Dans l'étude 2, le coefficient de variation du *tapping* spontané des participants amusiques étaient plus élevé que celui des contrôles. Par contre, cette différence n'est pas ressortie à l'étude 3. Un tempo spontané plus variable impliquerait également un trouble au niveau des mécanismes de *timing* internes, mais pourrait également refléter une variabilité dans la planification du geste moteur. Dans les études précédentes auprès de participants ayant un trouble de synchronisation au beat, les résultats sont aussi mitigés concernant le tempo spontané. En effet, Sowiński & Dalla Bella (2013), Bégel et collaborateurs (2017), ainsi que Phillips-Silver et collaborateurs (2011), n'ont pas rapporté de différence dans le tapping spontané de leurs participants. Le tempo spontané d'individus ayant un trouble de synchronisation au beat a été mesuré plus systématiquement dans des travaux de recherche récents menés dans le cadre de la thèse doctorale de madame Pauline Tranchant. Madame Tranchant a obtenu que huit participants beat-deaf avaient un tempo spontané plus variable qu'un groupe contrôle apparié, sur la base de six productions de tempo spontané. Ce résultat a été interprété comme une déficience possible de la régularité des oscillateurs internes chez ces participants. Il est possible que la variabilité des résultats d'une étude à l'autre soit due à des effets de puissance statistique ou reflète une variabilité dans l'origine du trouble de synchronisation à la musique entre les participants.

Ainsi, dans l'ensemble, il est encore difficile à l'heure actuelle de cibler l'origine du trouble de la synchronisation au beat musical étudié ici. Par ailleurs, le trouble ne semble pas découler d'une simple perception appauvrie du beat, considérant l'augmentation du coefficient de variation des productions motrices observée au travers des différents contextes à l'étude dans la thèse. Si l'on assume que le trouble vient de la génération d'un modèle interne du beat, ou de mécanismes de *timing* implicite, les aires cérébrales impliquées dans le trouble pourraient être soit le cortex prémoteur, le SMA et/ou les ganglions de la base. Il se pourrait également que le couplage des aires auditives et motrices via la voie auditive dorsale soit atteinte. Les ganglions de la base pourraient être cependant de bons candidats, étant impliqués dans le maintien de la régularité d'une action répétitive et l'extraction des périodicités dans le signal, jouant ainsi un rôle à la fois dans le tempo spontané, l'extraction et la prédiction du beat (Grahn & McAuley, 2009 ; Grahn & Rowe, 2013 ; Schwartze et al., 2011 ; Schwartze & Kotz, 2013 ; Zatorre et al., 2007). Si par ailleurs, on présume que le déficit vient plutôt des mécanismes de correction et de planification motrice, alors un circuit impliquant davantage le cervelet serait probablement à remettre en cause (Schwartze et al., 2016 ; Schwartze & Kotz, 2016 ; Steen et al., 2015). L'étude de patients atteints de la maladie de Parkinson ou avec des troubles cérébelleux suggère que des réseaux neuroanatomiques distincts pourraient être impliqués dans le traitement de rythmes périodiques et non-périodiques, particulièrement au niveau de l'implication des ganglions de la base et du cervelet (Iversen & Balasubramaniam, 2016 ; Leow & Grahn, 2014 ; Schwartze & Kotz, 2013 ; Steen, Schwartze, Kotz, & Keller, 2015 ; Teki et al., 2011). Les ganglions de la base seraient plus particulièrement impliqués dans l'extraction du beat par l'encodage des régularités temporelles dans le signal (Grahn, 2009 ; Grahn & Brett, 2009 ; Grahn & Rowe, 2013 ; Schwartze, Keller, Patel, & Kotz, 2011 ; Teki et al., 2011). Le cervelet, quant à lui, jouerait un rôle dans l'encodage d'intervalles de temps plus complexes et l'ajustement de la réponse de synchronisation (Rao, Mayer, & Harrington, 2001 ; Schwartze, Keller, & Kotz, 2016 ; Teki et al., 2011 ; Zatorre et al., 2007).

Entrainement au beat : mécanisme spécifique ou multi-domaine ?

Dissociation des troubles du traitement des hauteurs et du beat

Dans un premier temps, nous avons pu montrer dans l'étude 1 que la perception du beat pouvait être dissociée de la perception des hauteurs dans l'amusie congénitale. En effet, nous avons identifié quatre cas d'amusiques ayant un trouble de la perception des hauteurs avec des capacités préservées à se synchroniser au beat de la musique. Dans l'autre sens, les participants ayant été identifiés pour leur trouble de synchronisation ont des scores dans la norme aux tâches mélodiques de la MBEA. Ce résultat concorde avec les données existantes, découlant de l'étude de patients cérébrolésés, démontrant que la perception des hauteurs et la perception du beat peuvent être dissociées (Liégeois-Chauvel et al., 1998 ; Peretz & Kolinsky, 1993 ; Peretz et al., 1994). En accord avec ces résultats, l'étude 3 a montré que la production d'intervalles mélodiques dans le chant est normale chez la majorité des participants ayant un trouble de la synchronisation. Ceci confirme que, tant sur le plan de la perception que de la production, le traitement des hauteurs et du beat peuvent être dissociés dans l'amusie congénitale.

Une association a tout de même été trouvée dans l'étude 1 entre la perception des hauteurs et la perception du beat. Il a été proposé dans la discussion de cette étude que l'association des deux domaines en soit probablement une de haut niveau, au plan plutôt des représentations cognitives, et non en lien avec l'entrée sensorielle ou le contrôle moteur. En effet, l'association entre la perception du beat et la discrimination simple de différences de hauteurs n'était pas significative. De plus, les participants amusiques de cette étude ne différaient pas des contrôles à la tâche de synchronisation de tapes au métronome. L'étude 3 a par contre montré que les participants ayant un trouble de la synchronisation avaient plus de difficulté à produire les notes exactes du modèle en condition de chant choral. Un lien n'a pu être établi cependant entre les aptitudes mélodiques et rythmiques dans ce contexte, une importante variabilité dans la performance des amusiques étant présentes. Ainsi, concernant de l'hypothèse de Patel (2006) qui avance une association entre l'imitation vocale et la synchronisation au beat, nos données ne permettent actuellement pas de l'appuyer avec certitude. Des études futures devraient mesurer la capacité à reproduire une note unique, tel que cela a été fait avec des amusiques présentant un trouble de la perception des hauteurs (Hutchins, Zarate, Zatorre, & Peretz, 2010). Un tel protocole éliminerait l'effet d'interférence possible induit par la composante 'temporelle' associée à la production d'une mélodie.

Synchronisation au chant et à la parole

Un des résultats intéressants de la thèse est la généralisation du trouble de la synchronisation au beat de la musique aux tâches de synchronisation sur la parole. Dans la deuxième étude de la thèse, les participants présentant un trouble de la synchronisation au beat ont montré non seulement une difficulté à se synchroniser au beat de phrases chantées et de parole régulière, mais se sont également avérés moins précis que les participants contrôles lors de la synchronisation à la parole naturelle. En fonction de ces résultats, il semblerait qu'une boucle de couplage auditivo-moteur serait conjointement impliquée lors de la synchronisation à la musique et à la parole, malgré la distinction fondamentale entre la présence d'un beat dans le premier cas et non dans le second. Une réponse d'entrainement aux caractéristiques rythmiques de la parole et de la musique pourrait effectivement être partagée entre les deux domaines. Ding

& Simon (2014) proposent que ce mécanisme d'entrainement pourrait se faire dans la bande de fréquence delta (1 à 4 Hz), qui englobe la majorité des tempi retrouvés préférentiellement dans la musique (Ding et al., 2017 ; van Noorden & Moelants, 1999), ainsi que la fenêtre temporelle de production des accents prosodiques (*stress*), un marqueur important du rythme dans la parole (Cummins & Port, 1998 ; Kotz & Schwartze, 2010 ; Port, 2003 ; Turk & Shattuck-Hufnagel, 2013).

Il est possible que le mécanisme commun de synchronisation entre les deux domaines passe par la voie auditive dorsale. La voie auditive dorsale est reconnue comme une voie de couplage sensorimoteur (*sensorimotor transformation*) impliquée dans la mise en place d'actions motrices coordonnées et la production de modèles prédictifs moteurs régulant la synchronisation (Patel & Iversen, 2014 ; Zatorre et al., 2007). Des modèles de perception et de production de la parole impliquent également cette voie (Hickok, 2012 ; Hickok, Houde, & Rong, 2011 ; Hickok & Poeppel, 2004). Le modèle de Hickok & Poeppel (2004), entre autres, suggère un rôle important de l'interaction auditivo-motrice dans le traitement de la parole. Ces auteurs proposent qu'un réseau ventral est utilisé pour déterminer le sens des sons, alors que la voie dorsale servirait à la représentation motrice du son (*articulatory-based representation*). Cette voie dorsale serait recrutée non seulement au moment de la production de la parole, mais aussi dans la perception, au même titre que le propose l'hypothèse ASAP pour la musique (Patel & Iversen, 2014). La voie auditive dorsale pourrait donc jouer un rôle dans l'extraction des caractéristiques rythmiques à la fois de la musique et de la parole, sans égard à la périodicité du son. Il demeure, néanmoins, que la présence d'une périodicité dans le rythme semble faciliter le couplage entre les aires auditives et motrices (Chen et al., 2008 b ; Chen, Penhune, & Zatorre, 2009 ; Chen et al., 2006).

On pourrait argumenter, en revanche, que l'effet trouvé ici soit un artéfact de la tâche qui impliquait de synchroniser des tapes sur le 'beat' perçu dans les extraits présentés, les participants ayant reçu spécifiquement cette instruction. Cependant, nous avons également obtenu comme résultat à l'étude 3 que les participants présentant un trouble de synchronisation au beat parvenaient difficilement à se synchroniser par le chant, qui fait appel au même système moteur que la parole. Ainsi, il semblerait que le trouble de synchronisation, présent chez les participants étudiés dans le cadre des travaux de cette thèse, soit indifférencié de l'effecteur utilisé ou de la nature de la tâche de synchronisation.

Il demeure cependant que nous n'avons pas mesuré directement l'imitation vocale ou la synchronisation à la parole, mais bien le chant. De plus, un recoupement demeure entre la performance des amusiques et des contrôles lors de la synchronisation à la parole naturelle, comparativement à la tâche de synchronisation au beat musical, qui permet de dissocier complètement les deux groupes. Il serait donc important que l'effet trouvé ici fasse l'objet d'études futures, en s'attardant davantage à la synchronisation ou à l'entrainement tel que typiquement mesuré dans le domaine du langage. Un protocole classique serait de demander à deux participants d'alterner à la lecture d'un texte ou de lire un texte en synchronie (Cummins, 2001, 2002a, 2002b, 2009b).

Limites et directions futures

Bien que les résultats obtenus au travers des trois études soient dans l'ensemble cohérents et permettent de soulever de nouvelles hypothèses concernant la spécificité de l'entrainement au beat, certaines limites méritent d'être adressées.

Une première limite des résultats de ces travaux de thèse est que dans l'ensemble des études, les stimuli utilisés environnaient la période de 500 ms (2 Hz). Cette période est communément utilisée dans les études de synchronisation, correspondant à la fenêtre de tempo préférentiel de synchronisation de 500 à 600 ms décrit dans la littérature, soit le tempo où la synchronisation semble pouvoir être la plus optimale (Levitin et al., 2018 ; McAuley, 2010 ; Repp & Su, 2013). Il se pourrait que des profils de performance différents soient ainsi observés si d'autres tempi venaient à être utilisés, demandant une plus grande adaptation de la réponse de synchronisation par rapport à la période de tempo spontané des participants. Une étudiante du laboratoire de recherche (Pauline Tranchant) a d'ailleurs récemment montré dans ces travaux de thèse que les participants ayant un trouble de la synchronisation auraient plus de difficulté à se synchroniser à des tempi de métronome très lents (> 1139 ms) ou très rapide (225 ms), suggérant donc, en effet, une possible difficulté chez ces individus à adapter leur tempo. Nous avons par ailleurs pu constater dans la deuxième étude (Article 2) que les participants amusiques étaient en mesure d'ajuster la période de leurs tapes entre des essais ayant une période de 500 ms versus 625 ms. Ceci avait également été constaté chez Mathieu, le premier cas de *beat deafness* étudié, qui avait montré être en mesure de s'adapter à un changement de tempo de 20 % (Phillips-Silver et al., 2011). Il reste néanmoins que ces deux périodes se rapprochent de la fourchette de tempi préférentiels.

L'amusie congénitale touchant la synchronisation au beat n'est étudiée que depuis peu et offre une occasion importante de tester la spécificité des mécanismes d'entrainement au beat musical. Nous avons, entre autres, pu montrer clairement que le trouble ne se limite pas, en fait, à la synchronisation au beat de la musique. En effet, les participants ont produit une synchronisation moins précise à l'ensemble des tâches présentées. Par contre, la taille réduite de l'échantillon étudié et la variabilité au sein de celui-ci amène à préconiser la prudence quant à la généralisation des résultats. La variabilité dans les profils de performance fait en sorte qu'il est difficile de proposer une explication unique pour expliquer les résultats. Par exemple, certains participants ayant un trouble de la synchronisation se sont également avérés être de mauvais chanteurs, alors qu'un autre était parmi les meilleurs chanteurs, incluant les contrôles. On peut donc se demander encore si le trouble de la synchronisation pourrait provenir de différentes sources selon le cas. Pour répondre à cette question, plusieurs avenues de recherche semblent intéressantes.

Il serait pertinent de déterminer si les participants étudiés ici présentent des difficultés avec d'autres aspects du *timing*, soit, entre autres, l'intégration des durées d'intervalles. Il est supposé que les mécanismes de timing *beat-based* et *duration-based* pourraient être en partie différenciés (Teki et al., 2011). Jusqu'à maintenant, nos données semblent indiquer que les mécanismes d'intégration de la durée pourraient être préservés chez nos participants, considérant que la majorité de ceux-ci ont bien réussi à la tâche de rythme de la MBEA. Toutefois, cela demeure à être exploré systématiquement.

Un autre aspect qui serait intéressant à évaluer chez les participants ayant un trouble de la synchronisation serait l'intégration de la rétroaction dans la synchronisation. Deux études suggèrent que les mécanismes de correction pourraient être atteints chez ces individus (Palmer et al., 2014 ; Sowiński & Dalla Bella, 2013). Nous avons observé ici que la rétroaction présente dans la tâche de chant, comparativement à une tâche de tapping, semble avoir été peu profitable. L'étude de l'effet de la rétroaction sur la synchronisation pourrait donc être pertinente pour déterminer si, effectivement, les mécanismes de correction de l'erreur dans la coordination temporelle sont déficients dans cette population. Un protocole de perturbation de la rétroaction pourrait être facilement combiné à une tâche de reproduction de hauteurs de notes, évaluant donc par la même occasion l'hypothèse de Patel (2006) concernant l'association entre la synchronisation au beat et l'imitation vocale.

Comme mentionné plus tôt, une limite méthodologique de la thèse est de ne pas avoir mesuré directement la synchronisation vocale parlée des participants ayant un trouble de la synchronisation au beat. En effet, il a été décidé de s'intéresser d'abord à la synchronisation vocale chantée. Ce choix a été motivé par le fait que la seule étude chez l'humain ayant testé l'association entre l'imitation vocale et la synchronisation au beat a utilisé le chant, offrant ainsi un comparatif pour la présente étude (Dalla Bella et al., 2015). De plus, le chant n'avait jamais été étudié chez ces participants et l'utilisation d'une tâche de chant choral apparaissait plus écologique qu'une tâche de parole synchronisée. Par ailleurs, il serait primordial qu'une étude se penche sur la synchronisation vocale, et non chantée, afin de s'assurer de tester avec plus de certitude la généralisation du trouble à la synchronisation à la parole. L'adaptation du rythme de parole en contexte conversationnel permettrait également un test direct de l'hypothèse de Wilson & Wilson (2005) ayant mis de l'avant un modèle d'entrainement dans ce contexte. Si effectivement, un mécanisme d'entrainement multi-domaine sous-tend la coordination temporelle pour la musique et la parole, on pourrait s'attendre à ce que nos participants présentent peu d'adaptation spontanée de leur rythme de parole ou produisent plus de 'bris' dans le rythme d'une conversation. Par ailleurs, l'intégration de signaux non-verbaux (la gestuelle) pourrait peut-être leur permettre de compenser. Ainsi, il faudrait s'assurer de bien contrôler pour ces facteurs.

Enfin, dans une autre perspective, bien que l'amusie congénitale offre un cadre intéressant pour étudier la spécificité des mécanismes d'entrainement, l'étude de personnes avec un trouble du langage d'origine neurodéveloppementale pourrait également être une approche pertinente à adopter. Bon nombre d'études menées auprès d'enfants dyslexiques, dysphasiques et de personnes bègues (personnes avec un bégaiement) montrent chez ceux-ci certains troubles associés au traitement du beat. Certains enfants dyslexiques présenteraient de la difficulté à taper en synchronie avec un métronome (Flaugnacco et al., 2014 ; Thomson & Goswami, 2008) et cette difficulté perdurerait à l'âge adulte (Thomson, Fryer, Maltby, & Goswami, 2006). Des résultats similaires ont été rapportés auprès d'enfants dysphasiques (Corriveau & Goswami, 2009 ; Goswami, 2012). Les personnes bègues seraient également moins précises lors de la synchronisation à la musique (Falk, Müller, & Dalla Bella, 2015; van de Vorst & Gracco, 2017). Selon les auteurs de ces études, le lien se ferait, chez les enfants dyslexiques, via le trouble de la conscience phonologique (Corriveau, Pasquini, & Goswami, 2007 ; Goswami, 2011, 2012 ; Leong & Goswami, 2014). Le modèle PATH (*Precise Auditory Timing Hypothesis*) de Tierney & Kraus (2014a) relie les habiletés phonologiques et l'entrainement auditivo-moteur par le recours commun à une représentation neurale précise de la signature temporelle de la stimulation auditive dans le système auditif, et l'intégration de cette information via les aires motrices, entre autres. Ainsi, en se basant avant tout sur l'étude du langage, ceux-ci proposent une hypothèse qui semble se rapprocher en partie de l'hypothèse ASAP (Patel & Iversen, 2014). Considérant ces données de la littérature, on pourrait s'attendre à ce que les personnes avec un trouble de la synchronisation au beat présentent des particularités sur le plan du langage. Les participants de notre échantillon ne rapportent pas, cependant, de trouble de langage ou de dyslexie. En l'absence d'évaluation formelle, on ne peut exclure, en revanche, que ces participants puissent montrer certaines fragilités sur le plan des habiletés phonologiques.

Conclusion

En somme, les résultats de la thèse se sont révélés plutôt constants d'une étude à l'autre et permettent de mettre de l'avant trois conclusions principales en regard de la spécificité des mécanismes d'entrainement au beat. D'abord, le traitement du pitch et du beat semblent pouvoir se dissocier dans l'amusie congénitale, tant sur le plan des habiletés de perception que de production. Nous avons montré que les difficultés relevées dans la synchronisation au beat musical chez les personnes ayant un trouble de la perception des variations de hauteurs ne peuvent s'expliquer que par un effet d'interférence, et que les deux habiletés peuvent être dissociées. Ensuite, les résultats de la thèse ont permis de montrer que le trouble de synchronisation à la pulsation musicale n'est pas spécifique à un effecteur ou à la tâche, et serait de plus non exclusif à des stimuli mélodiques. Les personnes présentant un trouble de synchronisation au beat ont montré une difficulté marquée à se synchroniser à des extraits de parole régulière et à chanter en synchronie avec un métronome. Enfin, il est possible qu'un mécanisme d'entrainement commun soit impliqué dans le traitement de la pulsation musicale (beat) et de la parole, les amusiques étant moins précis pour synchroniser une réponse motrice dans ce contexte également. Cette dernière conclusion demande toutefois à être investiguée plus en profondeur dans des recherches futures, considérant la nature écologique limitée de cette tâche pour le domaine langagier. Le trouble de la synchronisation à la pulsation musicale d'origine congénitale fait l'objet d'études empiriques depuis moins de dix ans. Ainsi, beaucoup reste encore à apprendre sur ce trouble et son origine. Par ailleurs, l'étude de l'amusie congénitale offre une occasion unique de mieux cerner les mécanismes d'entrainement au beat musical, une habileté qui semble fondamentalement propre à l'humain.

Bibliographie (Introduction & discussion générale)

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238-251. doi: 10.1093/brain/awf028

Bégel, V., Benoit, C.-E., Correa, A., Cutanda, D., Kotz, S. A., & Dalla Bella, S. (2017). "Lost in time" but still moving to the beat. *Neuropsychologia, 94*, 129-138. doi: https://doi.org/10.1016/j.neuropsychologia.2016.11.022

Benichov, J. I., Globerson, E., & Tchernichovski, O. (2016). Finding the Beat: From Socially Coordinated Vocalizations in Songbirds to Rhythmic Entrainment in Humans. *Frontiers in Human Neuroscience, 10*, 255. doi: 10.3389/fnhum.2016.00255

Berkowska, M., & Dalla Bella, S. (2009a). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology, 5*, 69-83. doi: 10.2478/v10053-008-0068-2

Berkowska, M., & Dalla Bella, S. (2009b). Reducing Linguistic Information Enhances Singing Proficiency in Occasional Singers. *Annals of the New York Academy of Sciences, 1169*(1), 108-111. doi: doi:10.1111/j.1749-6632.2009.04774.x

Berkowska, M., & Dalla Bella, S. (2013). Uncovering phenotypes of poor-pitch singing: the Sung Performance Battery (SPB). *Frontiers in Psychology, 4*(714). doi: 10.3389/fpsyg.2013.00714

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a Coordinating Device: Entrainment With Disordered Speech. *Journal of Speech, Language, and Hearing Research, 57*(3), 815-824. doi: 10.1044/2014_JSLHR-S-13-0149

Brown, S., Pfordresher, P. Q., & Chow, I. (2017). A musical model of speech rhythm. *Psychomusicology: Music, Mind, and Brain, 27*(2), 95-112. doi: 10.1037/pmu0000175

Chapin, H. L., Zanto, T., Jantzen, K. J., Kelso, S., Steinberg, F., & Large, E. W. (2010). Neural responses to complex auditory rhythms: the role of attending. *Frontiers in psychology, 1*, 224.

Chauvigné, L. A. S., Gitau, K. M., & Brown, S. (2014). The neural basis of audiomotor entrainment: an ALE meta-analysis. *Frontiers in Human Neuroscience, 8*, 776. doi: 10.3389/fnhum.2014.00776

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008a). Listening to musical rhythms recruits motor regions of the brain. *Cerebral cortex, 18*(12), 2844-2854.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008b). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of cognitive neuroscience, 20*(2), 226-239.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences, 1169*(1), 15-34.

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. doi: https://doi.org/10.1016/j.neuroimage.2006.04.207

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology, 127*(4), 412-427. doi: 10.1037/a0032345

Corriveau, K., Pasquini, E., & Goswami, U. (2007). Basic Auditory Processing Skills and Specific Language Impairment: A New Look at an Old Hypothesis. *Journal of Speech, Language, and Hearing Research, 50*(3), 647-666. doi: 10.1044/1092-4388(2007/046)

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex, 45*(1), 119-130. doi: https://doi.org/10.1016/j.cortex.2007.09.008

Cummins, F. (2001). Reducing expressive variation in speech with synchronous speech. *The Journal of the Acoustical Society of America, 109*(5), 2416-2417. doi: 10.1121/1.4744550

Cummins, F. (2002a). Entraining speech with speech and metronomes. *Cad. Est. Ling., 43*, 55-70.

Cummins, F. (2002b). On synchronous speech. *Acoustics Research Letters Online, 3*(1), 7-11.

Cummins, F. (2009a). Rhythm as an Affordance for the Entrainment of Movement. *Phonetica, 66(suppl 1-2)*(1-2), 15-28.

Cummins, F. (2009b). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics, 37*(1), 16-28.

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*(2), 145-171.

Dalla Bella, S., Berkowska, M., & Sowinski, J. (2011). Disorders of Pitch Production in Tone Deafness. *Frontiers in Psychology, 2*(164). doi: 10.3389/fpsyg.2011.00164

Dalla Bella, S., Berkowska, M., & Sowiński, J. (2015). Moving to the Beat and Singing are Linked in Humans. *Frontiers in Human Neuroscience, 9*(663). doi: 10.3389/fnhum.2015.00663

Dalla Bella, S., Białuńska, A., & Sowiński, J. (2013). Why movement is captured by music, but less by speech: Role of temporal regularity. *PloS one, 8*(8), e71945.

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2017). BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behavior Research Methods, 49*(3), 1128-1145. doi: 10.3758/s13428-016-0773-6

Dalla Bella, S., Giguere, J. F., & Peretz, I. (2009). Singing in congenital amusia. *J Acoust Soc Am, 126*(1), 414-424. doi: 10.1121/1.3132504

Dalla Bella, S., & Peretz, I. (2003). Congenital Amusia Interferes with the Ability to Synchronize with Music. *Annals of the New York Academy of Sciences, 999*(1), 166-169. doi: 10.1196/annals.1284.021

Dalla Bella, S., Tremblay-Champoux, A., Berkowska, M., & Peretz, I. (2012). Memory disorders and vocal performance. *Annals of the New York Academy of Sciences, 1252*(1), 338-344. doi: doi:10.1111/j.1749-6632.2011.06424.x

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*(1), 51-62.

Di Pietro, M., Laganaro, M., Leemann, B., & Schnider, A. (2004). Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia, 42*(7), 868-877. doi: https://doi.org/10.1016/j.neuropsychologia.2003.12.004

Di Liberto, Giovanni M., O'Sullivan, James A., & Lalor, Edmund C. (2015). Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing. *Current Biology, 25*(19), 2457-2465. doi: https://doi.org/10.1016/j.cub.2015.08.030

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neurosci Biobehav Rev, 81*(Pt B), 181-187. doi: 10.1016/j.neubiorev.2017.02.011

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience, 8*, 311. doi: 10.3389/fnhum.2014.00311

Dowling, W. J. (1999). 15 - The Development of Music Perception and Cognition. Dans D. Deutsch (dir.), *The Psychology of Music (Second Edition)* (p. 603-625). San Diego: Academic Press.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition, 77*(3), 251-288. doi: https://doi.org/10.1016/S0010-0277(00)00106-2

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in Time with Mechanically and Expressively Performed Music. *Music Perception: An Interdisciplinary Journal, 18*(1), 1-23. doi: 10.2307/40285899

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264-280. doi: 10.1037/a0013482

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*(847). doi: 10.3389/fpsyg.2015.00847

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., . . . Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience, 8*(392). doi: 10.3389/fnhum.2014.00392

Foxton, J. M., Nandy, R. K., & Griffiths, T. D. (2006). Rhythm deficits in "tone deafness". *Brain and Cognition, 62*(1), 24-29. doi: https://doi.org/10.1016/j.bandc.2006.03.005

Fujii, S., & Schlaug, G. (2013). The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. *Frontiers in Human Neuroscience, 7*(771). doi: 10.3389/fnhum.2013.00771

Ghitza, O. (2011). Linking Speech Perception and Neurophysiology: Speech Decoding Guided by Cascaded Oscillators Locked to the Input Rhythm. *Frontiers in Psychology, 2*(130). doi: 10.3389/fpsyg.2011.00130

Ghitza, O. (2013). The theta-syllable: a unit of speech information defined by cortical function. *Frontiers in Psychology, 4*(138). doi: 10.3389/fpsyg.2013.00138

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15*(1), 3-10. doi: https://doi.org/10.1016/j.tics.2010.10.001

Goswami, U. (2012). Entraining the brain: applications to language research and links to musical entrainment.

Grahn, J. A. (2009). The Role of the Basal Ganglia in Beat Perception. *Annals of the New York Academy of Sciences, 1169*(1), 35-45. doi: doi:10.1111/j.1749-6632.2009.04553.x

Grahn, J. A., & Brett, M. (2007). Rhythm and Beat Perception in Motor Areas of the Brain. *Journal of Cognitive Neuroscience, 19*(5), 893-906. doi: 10.1162/jocn.2007.19.5.893

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex, 45*(1), 54-61. doi: https://doi.org/10.1016/j.cortex.2008.01.005

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage, 47*(4), 1894-1903. doi: https://doi.org/10.1016/j.neuroimage.2009.04.039

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*(23), 7540-7548.

Grahn, J. A., & Rowe, J. B. (2013). Finding and Feeling the Musical Beat: Striatal Dissociations between Detection and Prediction of Regularity. *Cerebral Cortex, 23*(4), 913-921. doi: 10.1093/cercor/bhs083

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2014). Speech Rhythms and Multiplexed Oscillatory Sensory Coding in the Human Brain. *PLOS Biology, 11*(12), e1001752. doi: 10.1371/journal.pbio.1001752

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews, 86*, 150-165. doi: https://doi.org/10.1016/j.neubiorev.2017.12.002

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 956.

Hausen, M., Torppa, R., Salmela, V., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: a common rhythm. *Frontiers in Psychology, 4*(566). doi: 10.3389/fpsyg.2013.00566

Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders, 45*(6), 393-402. doi: https://doi.org/10.1016/j.jcomdis.2012.06.004

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor Integration in Speech Processing: Computational Basis and Neural Organization. *Neuron, 69*(3), 407-422. doi: https://doi.org/10.1016/j.neuron.2011.01.019

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(1), 67-99. doi: https://doi.org/10.1016/j.cognition.2003.10.011

Honing, H. (2012). Without it no music: beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*(1), 85-91. doi: doi:10.1111/j.1749-6632.2011.06402.x

Hutchins, S., & Peretz, I. (2012a). Amusics can imitate what they cannot discriminate. *Brain and Language, 123*(3), 234-239. doi: https://doi.org/10.1016/j.bandl.2012.09.011

Hutchins, S., & Peretz, I. (2012b). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General, 141*(1), 76.

Hutchins, S., Zarate, J. M., Zatorre, R. J., & Peretz, I. (2010). An acoustical study of vocal pitch matching in congenital amusia. *J Acoust Soc Am, 127*(1), 504-512. doi: 10.1121/1.3270391

Hyde, K. L., Lerch, J. P., Zatorre, R. J., Griffiths, T. D., Evans, A. C., & Peretz, I. (2007). Cortical Thickness in Congenital Amusia: When Less Is Better Than More. *The Journal of Neuroscience, 27*(47), 13028-13032. doi: 10.1523/jneurosci.3039-07.2007

Hyde, K. L., & Peretz, I. (2003). “Out-of-Pitch” but Still “In-Time”. *Annals of the New York Academy of Sciences, 999*(1), 173-176. doi: 10.1196/annals.1284.023

Hyde, K. L., & Peretz, I. (2004). Brains That Are out of Tune but in Time. *Psychological Science, 15*(5), 356-360. doi: 10.1111/j.0956-7976.2004.00683.x

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. *Brain, 129*(10), 2562-2570. doi: 10.1093/brain/awl204

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI Evidence of an Abnormal Neural Network for Pitch Processing in Congenital Amusia. *Cerebral Cortex, 21*(2), 292-299. doi: 10.1093/cercor/bhq094

Iversen, J. R. (2016). In the beginning was the beat. Dans R. Hartenberger (dir.), *The Cambridge Companion to Percussion* (p. 281-295). Cambridge: Cambridge University Press.

Iversen, J. R., & Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioral Sciences, 8*, 175-180. doi: http://dx.doi.org/10.1016/j.cobeha.2016.02.027

Iversen, J. R., & Patel, A. D. (2008). *The Beat Alignment Test (BAT): surveying beat processing abilities in the general population.* Communication présentée Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10), Sapporo.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *J Acoust Soc Am, 124* (4), 2263-2271. doi: 10.1121/1.2973189

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-Down Control of Rhythm Perception Modulates Early Auditory Responses. *Annals of the New York Academy of Sciences, 1169*(1), 58-73. doi: doi:10.1111/j.1749-6632.2009.04579.x

Jadoul, Y., Ravignani, A., Thompson, B., Filippi, P., & de Boer, B. (2016). Seeking Temporal Predictability in Speech: Comparing Statistical Approaches on 18 World Languages. *Frontiers in Human Neuroscience, 10*(586). doi: 10.3389/fnhum.2016.00586

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological review, 83*(5), 323.

Jones, M. R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics, 41*(6), 621-634. doi: 10.3758/bf03210494

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*(3), 459-491. doi: 10.1037/0033-295X.96.3.459

Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51* (4), 271-291. doi: 10.1037/1196-1961.51.4.271

Jungers, M. K., Palmer, C., & Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 1*(2), 21-35.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392-399. doi: https://doi.org/10.1016/j.tics.2010.06.005

Kung, S.-J., Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2013). Interacting Cortical and Basal Ganglia Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive Neuroscience, 25*(3), 401-420. doi: 10.1162/jocn_a_00325

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *The psychology of time*, 189-232.

Large, E. W., & Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (Pan paniscus). *Journal of Comparative Psychology, 129*(4), 317-328. doi: 10.1037/com0000011

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience, 9*. doi: 10.3389/fnsys.2015.00159

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological review, 106*(1), 119.

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science, 26*(1), 1-37. doi: http://dx.doi.org/10.1016/S0364-0213(01)00057-X

Large, E. W., & Snyder, J. (2009). Pulse and Meter as Neural Resonance. *Annals of the New York Academy of Sciences, 1169*(1), 46-57. doi: doi:10.1111/j.1749-6632.2009.04550.x

Leong, V., & Goswami, U. (2014). Impaired extraction of speech rhythm from temporal modulation patterns in speech in developmental dyslexia. *Frontiers in Human Neuroscience, 8*(96). doi: 10.3389/fnhum.2014.00096

Leong, V., Stone, M. A., Turner, R. E., & Goswami, U. (2014). A role for amplitude modulation phase relationships in speech rhythm perception. *J Acoust Soc Am, 136* (1), 366-381. doi: 10.1121/1.4883366

Leow, L.-A., & Grahn, J. A. (2014). Neural Mechanisms of Rhythm Perception: Present Findings and Future Directions. Dans H. Merchant & V. de Lafuente (dir.), *Neurobiology of Interval Timing* (p. 325-338). New York, NY: Springer New York.

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception: An Interdisciplinary Journal, 1*(2), 229-252.

Levitin, D. J., Grahn, J. A., & London, J. (2018). The Psychology of Music: Rhythm and Movement. *Annual Review of Psychology, 69*(1), 51-75. doi: 10.1146/annurev-psych-122216-011740

Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011a). Entrainment to speech and song. Dans A. Williamon, D. Edwards & L. Bartel (dir.), *Proceedings of the International Symposium on Performance Science* (p. 123-128). Utrecht, The Netherlands: EAC.

Lidji, P., Palmer, C., Peretz, I., & Morningstar, M. (2011b). Listeners feel the beat: Entrainment to English and French speech rhythms. *Psychonomic Bulletin & Review, 18*(6), 1035-1041. doi: 10.3758/s13423-011-0163-0

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121*(10), 1853-1867. doi: 10.1093/brain/121.10.1853

London, J. (2002). Cognitive Constraints on Metric Systems: Some Observations and Hypotheses. *Music Perception: An Interdisciplinary Journal, 19*(4), 529-550. doi: 10.1525/mp.2002.19.4.529

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone Deafness: A New Disconnection Syndrome? *The Journal of Neuroscience, 29*(33), 10215-10220. doi: 10.1523/jneurosci.1701-09.2009

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. (2014). Action-based effects on music perception. *Frontiers in Psychology, 4*(1008). doi: 10.3389/fpsyg.2013.01008

Magne, C., Jordan, D. K., & Gordon, R. L. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain and Language, 153-154*, 13-19. doi: https://doi.org/10.1016/j.bandl.2016.01.001

Mai, G., Minett, J. W., & Wang, W. S. Y. (2016). Delta, theta, beta, and gamma brain oscillations index levels of auditory sentence processing. *NeuroImage, 133*, 516-528. doi: https://doi.org/10.1016/j.neuroimage.2016.02.064

Martens, P. A. (2011). The Ambiguous Tactus: Tempo, Subdivision Benefit, And Three Listener Strategies. *Music Perception: An Interdisciplinary Journal, 28*(5), 433-448. doi: 10.1525/mp.2011.28.5.433

McAuley, J. D. (2010). Tempo and Rhythm. Dans M. Riess Jones, R. R. Fay & A. N. Popper (dir.), *Music Perception* (p. 165-199). New York, NY: Springer New York.

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.

McKinney, M. F., & Moelants, D. (2006). Ambiguity in Tempo Perception: What Draws Listeners to Different Metrical Levels? *Music Perception: An Interdisciplinary Journal, 24*(2), 155-166. doi: 10.1525/mp.2006.24.2.155

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience, 7*(274). doi: 10.3389/fnins.2013.00274

Meyer, L. (2017). The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *Eur J Neurosci, 0*(0). doi: 10.1111/ejn.13748

Mithen, S. J. (2006). *The Singing Neanderthals: the Origins of Music, Language, Mind and Body* Cambridge, Mass.: Harvard University Press.

Moreau, P., Jolicœur, P., & Peretz, I. (2009). Automatic Brain Responses to Pitch Changes in Congenital Amusia. *Annals of the New York Academy of Sciences, 1169*(1), 191-194. doi: doi:10.1111/j.1749-6632.2009.04775.x

Moreau, P., Jolicœur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. *Brain and Cognition, 81*(3), 337-344. doi: https://doi.org/10.1016/j.bandc.2013.01.004

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: a metaphor? *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi: 10.1098/rstb.2013.0396

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi: 10.1098/rstb.2013.0393

Nozaradan, S., Peretz, I., & Keller, P. E. (2016). Individual Differences in Rhythmic Cortical Entrainment Correlate with Predictive Behavior in Sensorimotor Synchronization. *Scientific Reports, 6*, 20612. doi: 10.1038/srep20612

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the Neuronal Entrainment to Beat and Meter. *The Journal of Neuroscience, 31*(28), 10234-10240. doi: 10.1523/jneurosci.0411-11.2011

Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective Neuronal Entrainment to the Beat and Meter Embedded in a Musical Rhythm. *The Journal of Neuroscience, 32*(49), 17572-17581. doi: 10.1523/jneurosci.3203-12.2012

O'Dell, M., & Nieminen, T. (1999, août). *Coupled oscillator model of speech rhythm.* Dans *Proceedings of the XIVth international congress of phonetic sciences* (vol. 2, pp. 1075-1078). Berkeley: University of California.

Palmer, C. (1997). Music Performance. *Annual Review of Psychology, 48*(1), 115-138. doi: 10.1146/annurev.psych.48.1.115

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 728-741. doi: 10.1037/0096-1523.16.4.728

Palmer, C., Lidji, P., & Peretz, I. (2014). Losing the beat: deficits in temporal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1658). doi: 10.1098/rstb.2013.0405

Patel, A. D. (2006). Musical Rhythm, Linguistic Rhythm, and Human Evolution. *Music Perception: An Interdisciplinary Journal, 24*(1), 99-104. doi: 10.1525/mp.2006.24.1.99

Patel, A. D. (2008). *Music, Language, and the Brain*. New York, United States of America: Oxford University Press.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in systems neuroscience, 8*, 57.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. *Current Biology, 19*(10), 827-830. doi: https://doi.org/10.1016/j.cub.2009.03.038

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental brain research, 163*(2), 226-238.

Peelle, J. E., & Davis, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology, 3*, 320. doi: 10.3389/fpsyg.2012.00320

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral Cortex, 23*(6), 1378-1387. doi: 10.1093/cercor/bhs118

Peretz, I. (2016). Neurobiology of Congenital Amusia. *Trends in Cognitive Sciences, 20*(11), 857-867. doi: https://doi.org/10.1016/j.tics.2016.09.002

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain, 132*(5), 1277-1286. doi: 10.1093/brain/awp055

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of Musical Disorders. *Annals of the New York Academy of Sciences, 999*(1), 58-75. doi: 10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688. doi: 10.1038/nn1083

Peretz, I., & Kolinsky, R. (1993). Boundaries of Separability between Melody and Rhythm in Music Discrimination: A Neuropsychological Perspective. *The Quarterly Journal of Experimental Psychology Section A, 46*(2), 301-325. doi: 10.1080/14640749308401048

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117*(6), 1283-1301. doi: 10.1093/brain/117.6.1283

Peretz, I., & Vuvan, D. T. (2017). Prevalence of congenital amusia. *European Journal Of Human Genetics, 25*, 625. doi: 10.1038/ejhg.2017.15

Pfeuty, M., & Peretz, I. (2010). Abnormal pitch-time interference in congenital amusia: Evidence from an implicit test. *Attention, Perception, & Psychophysics, 72*(3), 763-774. doi: 10.3758/app.72.3.763

Pfordresher, P. Q. (2003). The Role of Melodic and Rhythmic Accents in Musical Structure. *Music Perception: An Interdisciplinary Journal, 20*(4), 431-464. doi: 10.1525/mp.2003.20.4.431

Pfordresher, P. Q., & Brown, S. (2007). Poor-Pitch Singing in the Absence of "Tone Deafness". *Music Perception: An Interdisciplinary Journal, 25*(2), 95-115. doi: 10.1525/mp.2007.25.2.95

Pfordresher, P. Q., Demorest, S. M., Bella, S. D., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical Perspectives on Singing Accuracy. *An Introduction to the Special Issue on Singing Accuracy (Part 1), 32*(3), 227-231. doi: 10.1525/mp.2015.32.3.227

Pfordresher, P. Q., Halpern, A. R., & Greenspon, E. B. (2015). A Mechanism for Sensorimotor Translation in Singing. *The Multi-Modal Imagery Association (MMIA) Model, 32*(3), 242-253. doi: 10.1525/mp.2015.32.3.242

Pfordresher, P. Q., & Mantell, J. T. (2014). Singing with yourself: Evidence for an inverse modeling account of poor-pitch singing. *Cognitive Psychology, 70*, 31-57. doi: https://doi.org/10.1016/j.cogpsych.2013.12.005

Phillips-Silver, J., Aktipis, C. A., & A. Bryant, G. (2010). The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement. *Music Perception: An Interdisciplinary Journal, 28*(1), 3-14. doi: 10.1525/mp.2010.28.1.3

Phillips-Silver, J., & Keller, P. (2012). Searching for Roots of Entrainment and Joint Action in Early Musical Interactions. *Frontiers in Human Neuroscience, 6*(26). doi: 10.3389/fnhum.2012.00026

Phillips-Silver, J., Toiviainen, P., Gosselin, N., & Peretz, I. (2013). Amusic does not mean unmusical: Beat perception and synchronization ability despite pitch deafness. *Cognitive Neuropsychology, 30*(5), 311-331. doi: 10.1080/02643294.2013.863183

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piché, O., Nozaradan, S., Palmer, C., & Peretz, I. (2011). Born to dance but beat deaf: A new form of congenital amusia. *Neuropsychologia, 49*(5), 961-969. doi: https://doi.org/10.1016/j.neuropsychologia.2011.02.002

Port, R. F. (2003). Meter and speech. *Journal of phonetics, 31*(3-4), 599-611.

Prince, J. B. (2011). The Integration of Stimulus Dimensions in the Perception of Music. *Quarterly Journal of Experimental Psychology, 64*(11), 2125-2152. doi: 10.1080/17470218.2011.573080

Prince, J. B. (2014). Pitch structure, but not selective attention, affects accent weightings in metrical grouping. *Journal of Experimental Psychology: Human Perception and Performance, 40*(5), 2073-2090. doi: 10.1037/a0037730

Prince, J. B., & Pfordresher, P. Q. (2012). The role of pitch and temporal diversity in the perception and production of musical sequences. *Acta Psychologica, 141*(2), 184-198. doi: https://doi.org/10.1016/j.actpsy.2012.07.013

Rajendran, V. G., Teki, S., & Schnupp, J. W. H. (2017). Temporal Processing in Audition: Insights from Music. *Neuroscience*. doi: https://doi.org/10.1016/j.neuroscience.2017.10.041

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience, 4*(3), 317-323. doi: 10.1038/85191

Ravignani, A., Bowling, D. L., & Fitch, W. T. (2014). Chorusing, synchrony, and the evolutionary functions of rhythm. *Frontiers in Psychology, 5*(1118). doi: 10.3389/fpsyg.2014.01118

Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of motor behavior, 35*(4), 355-370.

Repp, B. H. (2004). On the nature of phase attraction in sensorimotor synchronization with interleaved auditory sequences. *Human Movement Science, 23*(3), 389-413. doi: http://dx.doi.org/10.1016/j.humov.2004.08.014

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic bulletin & review, 12*(6), 969-992.

Repp, B. H., Iversen, J. R., & Patel, A. D. (2008). Tracking an Imposed Beat within a Metrical Grid. *Music Perception: An Interdisciplinary Journal, 26*(1), 1-18. doi: 10.1525/mp.2008.26.1.1

Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: Effects of intention, attention, and awareness. *The Quarterly Journal of Experimental Psychology Section A, 57*(3), 499-521. doi: 10.1080/02724980343000369

Repp, B. H., & Keller, P. E. (2008). Sensorimotor synchronization with adaptively timed sequences. *Human Movement Science, 27*(3), 423-456. doi: https://doi.org/10.1016/j.humov.2008.02.016

Repp, B. H., London, J., & Keller, P. E. (2008). Phase Correction in Sensorimotor Synchronization with Nonisochronous Sequences. *Music Perception: An Interdisciplinary Journal, 26*(2), 171-175. doi: 10.1525/mp.2008.26.2.171

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic bulletin & review, 20*(3), 403-452.

Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase, 22*(6), 558-565. doi: 10.1080/13554794.2016.1242756

Schachner, A. (2010). Auditory-motor entrainment in vocal mimicking species. *Communicative & Integrative Biology, 3*(3), 290-293. doi: 10.4161/cib.3.3.11708

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology, 19*(10), 831-836. doi: https://doi.org/10.1016/j.cub.2009.03.061

Schultz, B. G., O'Brien, I., Phillips, N., McFarland, D. H., Titone, D., & Palmer, C. (2015). Speech rates converge in scripted turn-taking conversations. *Applied Psycholinguistics, 37*(5), 1201-1220. doi: 10.1017/S0142716415000545

Schwartze, M., Keller, P. E., & Kotz, S. A. (2016). Spontaneous, synchronized, and corrective timing behavior in cerebellar lesion patients. *Behavioural Brain Research, 312*, 285-293. doi: https://doi.org/10.1016/j.bbr.2016.06.040

Schwartze, M., Keller, P. E., Patel, A. D., & Kotz, S. A. (2011). The impact of basal ganglia lesions on sensorimotor synchronization, spontaneous motor tempo, and the detection of tempo changes. *Behavioural Brain Research, 216*(2), 685-691. doi: https://doi.org/10.1016/j.bbr.2010.09.015

Schwartze, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience & Biobehavioral Reviews, 37*(10, Part 2), 2587-2596. doi: https://doi.org/10.1016/j.neubiorev.2013.08.005

Schwartze, M., & Kotz, S. A. (2016). Contributions of cerebellar event-based temporal processing and preparatory function to speech perception. *Brain and Language, 161*, 28-32. doi: http://dx.doi.org/10.1016/j.bandl.2015.08.005

Selkirk, E. O. (1986). *Phonology and syntax: the relationship between sound and structure*. London, UK: MIT press.

Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception: An Interdisciplinary Journal, 18*(4), 455-489.

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia, 51*(10), 1952-1963. doi: https://doi.org/10.1016/j.neuropsychologia.2013.06.027

Steen, M. C., Schwartze, M., Kotz, S. A., & Keller, P. E. (2015). Modeling effects of cerebellar and basal ganglia lesions on adaptation and anticipation during sensorimotor synchronization. *Annals of the New York Academy of Sciences, 1337*(1), 101-110. doi: doi:10.1111/nyas.12628

Stupacher, J., Wood, G., & Witte, M. (2017). Neural Entrainment to Polyrhythms: A Comparison of Musicians and Non-musicians. *Frontiers in Neuroscience, 11*(208). doi: 10.3389/fnins.2017.00208

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Golumbic, E. Z. (2017). Neural entrainment to the beat: the “missing-pulse” phenomenon. *Journal of Neuroscience, 37*(26), 6331-6341.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *The Journal of Neuroscience, 31*(10), 3805-3812. doi: 10.1523/jneurosci.5561-10.2011

ten Cate, C., Spierings, M., Hubert, J., & Honing, H. (2016). Can Birds Perceive Rhythmic Patterns? A Review and Experiments on a Songbird and a Parrot Species. *Frontiers in Psychology, 7*(730). doi: 10.3389/fpsyg.2016.00730

Thaut, M. H. (2003). Neural basis of rhythmic timing networks in the human brain. *Ann N Y Acad Sci, 999*(1), 364-373. doi: doi:10.1196/annals.1284.044

Thomson, J. M., Fryer, B., Maltby, J., & Goswami, U. (2006). Auditory and motor rhythm awareness in adults with dyslexia. *Journal of Research in Reading, 29*(3), 334-348. doi: doi:10.1111/j.1467-9817.2006.00312.x

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: Auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris, 102*(1), 120-129. doi: https://doi.org/10.1016/j.jphysparis.2008.03.007

Tierney, A., & Kraus, N. (2014a). Auditory-motor entrainment and phonological skills: precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience, 8*, 949. doi: 10.3389/fnhum.2014.00949

Tierney, A., & Kraus, N. (2014b). Neural Entrainment to the Rhythmic Structure of Music. *Journal of Cognitive Neuroscience, 27*(2), 400-408. doi: 10.1162/jocn_a_00704

Tierney, A., & Kraus, N. (2015). Evidence for Multiple Rhythmic Skills. *PLOS ONE, 10*(9), e0136645. doi: 10.1371/journal.pone .0136645

Tranchant, P., Lagrois, M., Bellemare Pépin, A., G. Schultz, B., & Peretz, I. (2018). *Beat alignment test of the motor origin of musical entrainment deficits*. Article en preparation

Tranchant, P., & Vuvan, D. T. (2015). Current conceptual challenges in the study of rhythm processing deficits. *Frontiers in Neuroscience, 9*(197). doi: 10.3389/fnins.2015.00197

Tranchant, P., Vuvan, D. T., & Peretz, I. (2016). Keeping the Beat: A Large Sample Study of Bouncing and Clapping to Music. *PLOS ONE, 11*(7), e0160178. doi: 10.1371/journal.pone .0160178

Tremblay-Champoux, A., Dalla Bella, S., Phillips-Silver, J., Lebrun, M.-A., & Peretz, I. (2010). Singing proficiency in congenital amusia: Imitation helps. *Cognitive Neuropsychology, 27*(6), 463-476. doi: 10.1080/02643294.2011.567258

Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong *Laboratory Phonology* (Vol. 4, p. 93).

van de Vorst, R., & Gracco, V. L. (2017). Atypical non-verbal sensorimotor synchronization in adults who stutter may be modulated by auditory feedback. *Journal of Fluency Disorders, 53*, 14-25. doi: https://doi.org/10.1016/j.jfludis.2017.05.004

Van Der Steen, M. C., & Keller, P. E. (2013). The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience, 7*, 253.

van Noorden, L., & Moelants, D. (1999). Resonance in the Perception of Musical Pulse. *Journal of New Music Research, 28*(1), 43-66. doi: 10.1076/jnmr.28.1.43.3122

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex, 69*, 186-200. doi: https://doi.org/10.1016/j.cortex.2015.05.002

Vuvan, D. T., Paquette, S., Mignault Goulet, G., Royal, I., Felezeu, M., & Peretz, I. (2017). The Montreal Protocol for Identification of Amusia. *Behav Res Methods*. doi: 10.3758/s13428-017-0892-8

Wilson, M., & Cook, P. F. (2016). Rhythmic entrainment: Why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychonomic Bulletin & Review, 23*(6), 1647-1659. doi: 10.3758/s13423-016-1013-x

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic bulletin & review, 12*(6), 957-968.

Wilson, S. J., Pressing, J. L., & Wales, R. J. (2002). Modelling rhythmic function in a musician post-stroke. *Neuropsychologia, 40*(8), 1494-1505. doi: https://doi.org/10.1016/S0028-3932(01)00198-1

Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: An ALE meta-analysis. *NeuroImage, 42*(1), 343-356. doi: https://doi.org/10.1016/j.neuroimage.2008.04.025

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience, 8*, 547. doi: 10.1038/nrn2152

Zendel, B. R., Lagrois, M.-É., Robitaille, N., & Peretz, I. (2015). Attending to Pitch Information Inhibits Processing of Pitch Information: The Curious Case of Amusia. *The Journal of Neuroscience, 35*(9), 3815-3824. doi: 10.1523/jneurosci.3766-14.2015

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences, 107*(13), 5768-5773. doi: 10.1073/pnas.1000121107

Zhang, W., & Ding, N. (2017). Time-domain analysis of neural tracking of hierarchical linguistic structures. *NeuroImage, 146*, 333-340. doi: https://doi.org/10.1016/j.neuroimage.2016.11.016

Annexes

Annexe I : Caractéristiques des extraits musicaux inclus dans le *Montreal - Beat Alignment Test* (M-BAT)

Titre de la chanson	**Genre**	**Tempo (BPM)**	**Durée (sec)**
Party at your mama's house	Rock	82	22
Superstition	Pop	100	19
Solsbury hill	Rock	103	18
Since you've been gone	Soul	117	17
The flow	Dance lounge	129	16
Suavemente	Merengue	124	16
Brand New Carpet	Pop rock	126	16
What a feeling	Pop dance	132	15
Don't stop me now	Rock	156	14
Take five	Jazz	170	17

Annexe II: Autres articles publiés dans des revues scientifiques au cours du doctorat

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. Philosophical Transactions of the Royal Society B: Biological Sciences, 370(1664). doi: 10.1098/rstb.2014.0090

Zendel, B. R., Lagrois, M.-É., Robitaille, N., & Peretz, I. (2015). Attending to Pitch Information Inhibits Processing of Pitch Information: The Curious Case of Amusia. The Journal of Neuroscience, 35(9), 3815-3824. doi: 10.1523/jneurosci.3766-14.2015

Lagrois, M.-É., Peretz, I., & Zendel, B. R. (2018). Neurophysiological and Behavioral Differences between Older and Younger Adults When Processing Violations of Tonal Structure in Music. Frontiers in Neuroscience, 12(54). doi: 10.3389/fnins.2018.00054