



Effets des activités musicales sur la perception de la parole dans le bruit

Mémoire

Elisabeth Maillard

Maîtrise en sciences cliniques et biomédicales - avec mémoire
Maître ès sciences (M. Sc.)

Québec, Canada

Effets des activités musicales sur la perception de la parole dans le bruit

Mémoire

Elisabeth Maillard

Sous la direction de :

Pascale Tremblay, directrice de recherche

Résumé

La perception de la parole dans le bruit (PPB) décline avec l'âge, entraînant des difficultés à communiquer dans des lieux bruyants. À ce jour, il n'existe aucune stratégie pour limiter efficacement les effets de l'âge sur la PPB. Toutefois, la pratique musicale, par son entraînement auditif et cognitif, représente un outil de prévention potentiel. L'objectif de ce mémoire était d'étudier l'effet de la pratique musicale (instrument de musique, chant) sur la PPB. La première étude présente une revue systématique et une méta-analyse de la littérature existante sur les effets de la pratique musicale sur la PPB. Dans la deuxième étude, des musiciens amateurs jeunes (18-59 ans) et âgés (>60 ans) ont été comparés à des individus pratiquant une activité psychomotrice non musicale de niveau amateur. La méta-analyse a montré que l'effet de la pratique musicale varie en fonction des conditions expérimentales de PPB, apportant des avantages en présence de bruits masquants (BM) énergétiques, de BM de parole à deux et quatre locuteurs et lorsque le BM est de même intensité ou plus fort que la parole. Toutefois, la revue systématique a mis en lumière le manque d'études sur des adultes âgés et sur des musiciens amateurs. La deuxième étude a montré qu'une pratique musicale de niveau amateur était associée à une meilleure performance en PPB chez les musiciens jeunes comme âgés, comparée à la pratique d'une activité psychomotrice, indépendamment de l'âge du début et du nombre d'années de pratique. En somme, ce mémoire a montré des bénéfices de la pratique musicale sur la PPB, qui se retrouvent chez des musiciens amateurs jeunes comme âgés. Ces résultats sont encourageants du point de vue de la prévention de la PPB, mais davantage d'études sont nécessaires pour confirmer ces résultats et éclaircir les différents paramètres de pratique qui mènent à des avantages en PPB.

Table des matières

RÉSUMÉ	II
TABLE DES MATIÈRES	III
LISTE DES FIGURES	VI
LISTE DES TABLEAUX	VII
LISTE DES ABRÉVIATIONS, SIGLES, ACRONYMES	VIII
REMERCIEMENTS	IX
AVANT-PROPOS	XI
INTRODUCTION	1
CHAPITRE 1 LA PERCEPTION DE LA PAROLE DANS LE BRUIT	4
1.1 BASES PHYSIOLOGIQUES DES ETAPES DE LA PERCEPTION DE LA PAROLE DANS LE BRUIT	4
1.1.1 <i>Le signal sonore</i>	4
1.1.2 <i>L'audition périphérique</i>	7
1.1.3 <i>Les voies auditives ascendante – la voie lemniscale</i>	9
1.1.4 <i>Perception et compréhension de la parole</i>	12
1.1.5 <i>Perception de la parole et fonctions cognitives</i>	14
1.2 ÉVALUATION COMPORTEMENTALE DE LA PPB	15
1.2.1 <i>La parole cible</i>	16
1.2.2 <i>Le bruit masquant</i>	16
1.2.3 <i>La présentation spatiale des stimuli</i>	19
1.2.4 <i>Le niveau de SNR</i>	20
1.3 LA PPB ET LE VIEILLISSEMENT	22
1.3.1 <i>Les effets de l'âge sur la performance en PPB</i>	22
1.3.2 <i>Contribution du vieillissement auditif et cognitif sur la PPB</i>	23
1.3.2.1 <i>Vieillissement structurel du cerveau</i>	23
1.3.2.2 <i>Vieillissement de l'audition périphérique</i>	24
1.3.2.3 <i>Vieillissement cognitif</i>	25
CHAPITRE 2 MUSIQUE ET PLASTICITÉ CÉRÉBRALE	27
2.1 LA PLASTICITÉ CÉRÉBRALE	27
2.1.1 DIFFÉRENTS TYPES DE PLASTICITÉ	27
2.1.2 <i>Etude de la PDE</i>	28
2.1.2.1 <i>Méthodes</i>	28
2.1.2.2 <i>Devis</i>	29
2.1.3 <i>PDE et vieillissement</i>	30
2.1.3.1 <i>Réserve et maintien</i>	31
2.2 LA MUSIQUE, OUTIL DE RÉSERVE/MAINTIENT POUR LA PPB	33
2.2.1 <i>Bénéfices de la pratique musicale sur le traitement auditif central</i>	33
2.2.2 <i>Bénéfices de la pratique musicale sur les fonctions cognitives</i>	35
2.2.3 <i>Un « OPERA » pour la « cocktail party »</i>	35
CHAPITRE 3 ÉTUDE 1	38
3.1 RÉSUMÉ	39

3.2	ABSTRACT	40
3.3	INTRODUCTION	41
3.4	METHODS	44
3.4.1	<i>Search strategy</i>	44
3.4.2	<i>Study selection</i>	45
3.4.2.1	Selection process.....	45
3.4.2.2	Selection criteria.....	45
3.4.3	<i>Study Designs and Risk of Bias Assessment</i>	46
3.4.4	<i>Data extraction</i>	46
3.4.5	<i>Statistical analysis</i>	47
3.5	RESULTS	49
3.5.1	<i>Study selection</i>	49
3.5.2	<i>Study quality</i>	49
3.5.3	<i>Focus(es) of the selected studies</i>	50
3.5.4	<i>Participant description</i>	51
3.5.5	<i>Musicianship</i>	52
3.5.6	<i>SPiN tasks</i>	53
3.5.6.1	Targets: characteristics and experimental manipulations.....	54
3.5.6.2	Maskers: characteristics and manipulations.....	54
3.5.6.3	Stimulus presentation.....	54
3.5.6.4	SNR Manipulation.....	55
3.5.6.5	Spatial separation between stimuli.....	55
3.5.7	<i>SPiN performance</i>	55
3.5.7.1	Speech in energetic masker.....	55
3.5.7.2	Speech in one-talker maskers.....	57
3.5.7.3	Speech in two-talker maskers.....	59
3.5.7.4	Speech in four-talker maskers.....	61
3.5.7.5	Speech in spatially separated noise.....	63
3.5.7.6	Speech in noise, SNR < 0 dB.....	64
3.5.7.7	Speech in noise, SNR = 0 dB.....	66
3.5.7.8	Speech in noise, SNR > 0 dB.....	68
3.6	DISCUSSION	70
3.6.1	<i>Speech in energetic masker, speech in speech masker</i>	71
3.6.2	<i>Speech in spatially separated masker</i>	72
3.6.3	<i>Speech in noise at various SNR levels</i>	72
3.6.4	<i>Effect of age</i>	73
3.6.5	<i>Critical appraisal of the literature and future considerations</i>	74
3.6.6	<i>Limitations</i>	76
3.7	CONCLUSION	76
3.8	ACKNOWLEDGMENTS	77
3.9	CREDIT AUTHOR STATEMENT	77
3.10	REFERENCES	77
3.11	SUPPLEMENTARY MATERIAL	85
CHAPITRE 4 ÉTUDE 2		131
EFFET DES ACTIVITÉS MUSICALES ET NON MUSICALES SUR LA PPB		131
4.1	INTRODUCTION	131
4.2	OBJECTIFS ET HYPOTHÈSES	135
4.3	MATÉRIEL ET MÉTHODES	136
4.3.1	<i>Participants</i>	136

4.3.1.1	Critères de sélection	138
4.3.1.2	Recrutement des participants	139
4.3.2	<i>Procédure expérimentale</i>	140
4.3.2.1	Evaluation de l'audition : audiométrie tonale	141
4.3.2.2	Tâche de perception de la parole dans le bruit	142
4.3.3	<i>Analyses statistiques</i>	143
4.4	RÉSULTATS	144
4.4.1	<i>ANOVA mixte : Pourcentage de bonnes réponses</i>	144
4.4.2	<i>ANOVA mixte : Temps de réaction</i>	146
4.4.3	<i>Régression linéaires mixtes : pourcentage de bonnes réponses et temps de réaction</i>	149
4.5	DISCUSSION	150
4.5.1	<i>Performance en PPB</i>	150
4.5.2	<i>Prédicteurs de la performance en PPB</i>	154
4.6	CONCLUSION	156
CHAPITRE 5	DISCUSSION	157
5.1	RAPPEL DES RÉSULTATS	157
5.1.1	<i>Etude 1</i>	157
5.1.2	<i>Etude 2</i>	159
5.2	TRANSFERT D'APPRENTISSAGE	161
CONCLUSION	170
BIBLIOGRAPHIE	171

Liste des figures

FIGURE 1. SCHÉMA DE LA FRÉQUENCE D'UNE ONDE. LOW FREQUENCY = BASSE FRÉQUENCE; HIGH FREQUENCY = HAUTE FRÉQUENCE. IMAGE LIBRE DE DROIT NUMÉRO 1854443230, TIRÉE DE HTTPS://WWW.SHUTTERSTOCK.COM/FR/G/PETROUDNY43	5
FIGURE 2. SCHÉMA DE L'AMPLITUDE D'UNE ONDE. LOW AMPLITUDE = FRÉQUENCE BASSE; HIGH AMPLITUDE = AMPLITUDE ÉLEVÉE. IMAGE LIBRE DE DROIT NUMÉRO 2080264282, TIRÉE DE HTTPS://WWW.SHUTTERSTOCK.COM/FR/G/PETROUDNY43	5
FIGURE 3. SCHÉMA D'UNE ONDE COMPLEXE PÉRIODIQUE ET DES ONDES HARMONIQUES QUI LA COMPOSENT. HARMONIC WAVEFORMS = ONDES HARMONIQUES; FUNDAMENTAL 1ST HARMONIC = FRÉQUENCE FONDAMENTALE (F ₀); 2 ND , 3 RD , 4 TH HARMONIC = HARMONIQUES ; COMPLEX WAVEFORM = ONDE COMPLEXE. IMAGE LIBRE DE DROIT NUMÉRO 719150923, TIRÉE DE HTTPS://WWW.SHUTTERSTOCK.COM/FR/G/FOUAD+A.+SAAD	6
FIGURE 4. SCHÉMA DU PASSAGE DU SON DANS LE SYSTÈME AUDITIF PÉRIPHÉRIQUE, DEPUIS LA CAPTATION DE L'ONDE SONORE JUSQU'AU CHEMINEMENT DE L'INFLUX NERVEUX AU CORTEX AUDITIF PRIMAIRE. IMAGE LIBRE DE DROIT NUMÉRO 1960818781 TIRÉE DE HTTPS://WWW.SHUTTERSTOCK.COM/FR/G/AXEL_KOCK . 7	
FIGURE 5. SCHÉMA DE L'APPAREIL AUDITIF PÉRIPHÉRIQUE. EXTERNAL EAR = OREILLE EXTERNE ; EAR CANAL = CONDUIT AUDITIF EXTERNE ; EARDRUM = TYMPAN ; MIDDLE EAR= OREILLE MOYENNE ; MALLEUS = MARTEAU ; INCUS = ENCLUME ; STAPES = ÉTRIER ; OVAL WINDOW= FENÊTRE OVALE ; INNER EAR = OREILLE INTERNE ; SEMICIRCULAR CANALS = CANAUX SEMI-CIRCULAIRE, VESTIBULE = VESTIBULE ; COCHLEA = COCHLÉE ; VESTIBULAR NERVE = NERF VESTIBULAIRE ; COCHLEAR NERVE = NERF COCHLÉAIRE. IMAGE LIBRE DE DROIT NUMÉRO 401539120 TIRÉE DE HTTPS://WWW.SHUTTERSTOCK.COM/FR/G/ALEXPOKUSAY	9
FIGURE 6. VOIES LEMNISCALES SCHÉMATISÉES AVEC LES DIFFÉRENTS NOYAUX-RELAIS. COCHLEAR NUCLEUS = NOYAU COCHLÉAIRE ; SUPERIOR OLIVE = NOYAU OLIVAIRE SUPÉRIEURE ; INFERIOR COLLICULUS = COLLICULE INFÉRIEUR ; MEDIAL GENICULATE = CORPS GENOUILLÉ MÉDIAN DU THALAMUS ; PRIMARY AUDITORY CORTEX = CORTEX AUDITIF PRIMAIRE. IL EST À NOTER QUE LES PROJECTIONS BILATÉRALES NE SONT PAS SCHÉMATISÉES.	11
FIGURE 7. POURCENTAGE DE BONNES RÉPONSES POUR LES QUATRE GROUPES DE PARTICIPANTS DANS LA CONDITION DE SNR 2 (A) ET DANS LA CONDITION DE SNR 1 (B).....	146
FIGURE 8. TEMPS DE RÉACTION POUR LES QUATRE GROUPES DE PARTICIPANTS DANS LA CONDITION DE SNR 2 (A) ET DANS LA CONDITION DE SNR 1 (B).....	148

Liste des tableaux

TABLEAU 1. DONNÉES DÉMOGRAPHIQUES	137
TABLEAU 2. DONNÉES DESCRIPTIVES, POURCENTAGE DE BONNES RÉPONSES	145
TABLEAU 3. ANOVA MIXTE À MESURE RÉPÉTÉE, POURCENTAGE DE BONNES RÉPONSES	145
TABLEAU 4. DONNÉES DESCRIPTIVES, TEMPS DE RÉACTION.....	147
TABLEAU 5. ANOVA MIXTE À MESURE RÉPÉTÉE, TEMPS DE RÉACTION	147
TABLEAU 6. RÉGRESSION LINÉAIRE MIXTE, TEMPS DE RÉACTION	149
TABLEAU 7. RÉGRESSION LINÉAIRE MIXTE, POURCENTAGE BONNES RÉPONSES	149

Liste des abréviations, sigles, acronymes

ABRÉVIATION	SIGNIFICATION FRANÇAISE	SIGNIFICATION ANGLAISE
BM	Bruit masquant	Masking noise
d'	Sensibilité	Sensibility
CV	Consonne-voyelle	Consonnant-vowel
CVC	Consonne-voyelle-consonne	Consonnant-vowel-consonnant
dB	Décibel	Decibel
EEG	Electroencéphalographie	Electroencephalography
ELU		Ease of Language Understanding
F0	Fréquence fondamentale	Fundamental frequency
HINT		Hearing In Noise Test
IRM	Imagerie par résonance magnétique	
M	Musicien	Musician
MoCA		Montreal Cognitive Assessment
NM	Non-musiciens	Non-musician
OPERA		Overlap, Precision, Emotion, Repetition, Attention
PDE	Plasticité dépendante de l'expérience	
PEE		Plasticity experience-expectant
PIE	Plasticité indépendante de l'expérience	
QI / IQ	Quotient intellectuel	Intellectual quotient
PPB / SPiN	Perception de la parole dans le bruit	Speech perception in noise
SNR	Ratio signal sur bruit	Signal to noise ratio
SRM		Spatial release from masking
SRT		Speech reception threshold
SSN		Speech shaped noise
WIN		Words In Noise Test

Remerciements

Je tiens à remercier chaleureusement tous ceux qui, de loin ou de près, ont participé à la réalisation de ce mémoire.

Tout d'abord, la Prof. Pascale Tremblay, qui m'a accueilli dans son laboratoire et qui a dirigé ce travail de maîtrise.

Pascale, je te remercie infiniment pour ton soutien, ta motivation, ton enthousiasme, ta compréhension, ton sens de l'humour et même... ton empathie (oui oui !). La route fut un peu plus longue que prévu, mais tellement enrichissante. Sous ta supervision, je n'ai cessé d'apprendre, que ce soit sur le plan scientifique ou sur le plan organisationnel. J'ai énormément apprécié côtoyer une chercheuse hors pair, honnête et intègre dans ses démarches scientifiques, rigoureuse dans sa méthodologie et toujours passionnée par ses projets.

Je tiens à remercier chaleureusement Marilynne Joyal.

Marilynne, ton intelligence et ton souci du détail ont été d'une aide immense dans la construction de ce mémoire. Ta bienveillance, ta gentillesse et ton caractère si pétillant ont été un vrai cadeau pour la vie au labo.

Je tiens également à remercier du fond du cœur les amis que j'ai eu la chance de rencontrer lors de ma maîtrise au « Labobouffe » : Valérie Brisson, Maxime Perron, Émilie Belley, Julie Poulin, Chloé Gingras, Gabriel Frazer-McKee et plus récemment, Xiyue Zhang et Marjorie Villeneuve.

En plus de votre bonne humeur, j'ai adoré découvrir vos intérêts scientifiques et vos préférences gastronomiques (!). Vos présences ont été d'un soutien précieux et un immense bonheur pendant ces années de maîtrise. Je vous aime !

Un merci tout particulier aux assistantes qui ont travaillé sur le projet « FRQNT », soit Alison Arseneault, Catherine Savard, Alexandra Lavoie, Émilie Belley, Lisa-Marie

Deschenes, Julie Poulin et Lydia Gagnon. Votre super travail a permis à ce projet d'être mené à bien et de fonctionner à plein régime.

Je tiens à remercier également tous les participants qui ont accepté de nous donner un peu de leur temps pour permettre à notre projet d'exister.

Je remercie ma famille du fond du cœur et en particulier mes deux Grand-Mamans pour leur tendresse infinie et leur soutien inconditionnel. Un immense merci également à mes amis en Suisse, dont les encouragements transatlantiques m'ont toujours donné de la force et du courage.

Et finalement, je tiens à remercier infiniment mon amoureux qui m'a écoutée et soutenue tout au long de ce travail.

Avant-propos

Le présent mémoire réunit les projets sur lesquels j'ai travaillé pendant ma maîtrise en sciences cliniques et biomédicales sous la supervision de la Prof. Pascale Tremblay (Centre de Recherche CERVO, Québec). Les deux projets présentés dans ce mémoire portent sur l'effet des activités musicales sur la capacité à percevoir la parole en milieu bruyant chez l'adulte. Ces projets couvrent chacun un versant différent de la recherche scientifique : un versant d'analyse de la littérature existante (Chapitre 3) et un versant expérimental (Chapitre 4).

Les premiers chapitres du mémoire posent les bases théoriques nécessaires afin de contextualiser la démarche dans laquelle s'inscrivent les travaux de ma maîtrise. Le chapitre 1 est consacré à la perception de la parole dans le bruit (PPB). Dans ce chapitre, les bases physiologiques de la perception de la parole et les paradigmes utilisés pour évaluer la PPB sont d'abord présentés. Puis, les effets connus de l'âge sur la PPB sont exposés. Le chapitre 2 introduit la notion de plasticité dépendante de l'expérience et présente des bénéfices comportementaux sur le traitement auditif et la cognition issus de la pratique musicale.

Les chapitres 3 et 4 constituent le cœur de ce travail. Le chapitre 3 présente un article pour lequel je suis première auteure et qui a été soumis au journal *Current Research in Behavioral Sciences* date du 22 novembre 2021. Cet article est une revue systématique et une méta-analyse sur l'effet de la pratique musicale sur la PPB. J'ai travaillé sur toutes les étapes de ce projet, c'est-à-dire que j'ai formulé la question de recherche, procédé à la recherche d'articles, sélectionné ceux qui respectaient les critères d'inclusion, extrait les informations d'intérêt sur leur méthodologie, réalisé les analyses statistiques de méta-analyses et rédigé le manuscrit. Pour ce projet, j'ai travaillé avec ma directrice, la Prof. Pascale Tremblay, ainsi qu'avec Marilyne Joyal, PhD, professionnelle de recherche au laboratoire, Julia Picard, une stagiaire d'été et le Prof. Micah Murray à l'Université de Lausanne.

Le chapitre 4 rapporte des résultats préliminaires d'une tâche de PPB provenant du projet "Effet des activités musicales et non musicales dans le cerveau adulte sur le langage, la cognition et les processus émotionnels" (projet #2019-1733, approuvé par le comité d'éthique de la recherche sectorielle en neurosciences et santé mentale du CIUSSS-CN). C'est un projet sur lequel j'ai travaillé pendant plus d'un an. J'ai contribué à l'élaboration du protocole expérimental et à sa soumission au comité éthique (ainsi que géré plusieurs amendements), développé la tâche de perception de la parole dans le bruit, mené les études pilotes, puis participé au recrutement des 40 premiers participants ainsi qu'à la récolte des données. À cause de la COVID-19, ce projet a été mis en pause. L'incertitude quant à la reprise des activités de recherche sur des participants humains nous a menés à réorienter ma maîtrise vers une revue systématique et une méta-analyse sur cette même thématique. Depuis, le projet a repris, mais n'est toujours pas terminé à cause des difficultés liées au recrutement des participants en situation de crise sanitaire. J'ai tout de même effectué l'analyse des résultats préliminaires de la tâche de PPB. Dans le cadre de ce projet, j'ai travaillé avec une équipe incluant ma directrice, Pascale Tremblay, Catherine Fontaine-Lavallée, ancienne professionnelle de recherche au laboratoire, Gabriel Frazer-McKee, un étudiant à la maîtrise du laboratoire, et plusieurs assistantes de recherche incluant Alison Arseneault, Catherine Savard, Alexandra Lavoie, Émilie Belley, Julie Poulin et Lisa-Marie Deschenes.

Introduction

“ Life’s just a cocktail party – on the streets ”

Mick Jagger.

L’effet “cocktail party » fait référence à la capacité à suivre le discours d’un interlocuteur spécifique au milieu d’un brouhaha ambiant, comme les conversations animées et autres toasts d’un cocktail party. L’audition périphérique, l’audition centrale, les fonctions exécutives et les fonctions langagières agissent de concert pour extraire efficacement le discours d’intérêt du paysage sonore. Mais voilà que la capacité à percevoir la parole au milieu du bruit n’est pas qu’indispensable dans un party, puisque le bruit est omniprésent dans notre société. Que ce soit dans un magasin, dans un transport public, dans un restaurant ou même chez soi, il y a toujours du bruit, organique ou inorganique, qui vient dégrader la parole que l’on cherche à percevoir. *So*, oui, Mick, *life’s just a cocktail party*, mais pas qu’*on the streets*, partout !

La capacité à percevoir la parole dans le bruit est donc essentielle à la communication et aux interactions sociales. Imaginez-vous, à table dans un restaurant populaire, entouré de vos amis. L’un d’eux se lance dans le récit d’une anecdote pleine de rebondissements. Mais les allées et venues des serveurs, les pleurs de l’enfant à la table voisine, les clics de la vaisselle, les clacs des talons, les débats des uns et les rires des autres viennent se mélanger au récit de votre ami dans un imbroglio acoustique. Résultat : seules certaines bribes de l’histoire vous sont perceptibles. Malheureusement, c’est trop peu pour saisir les subtilités du récit et y rebondir habilement. Vous n’aviez pourtant pas de problème à suivre les conversations auparavant et vous n’êtes assurément pas devenu complètement malentendant. Vous réalisez alors que vous n’avez tout simplement ... plus vingt ans.

En effet, la capacité à percevoir la parole dans le bruit diminue avec l'âge et porte atteinte à la qualité de vie des adultes âgés (Working Group on Speech & Aging, 1988). Les relations sociales contribuent au maintien cognitif des personnes âgées ainsi qu'à leur bien-être physique et psychique (Douglas et al., 2017; Kelly et al., 2017; Krueger et al., 2009). Or, les environnements bruyants, dont font partie les lieux de rencontres sociales, sont parmi les contextes de communication les plus difficiles pour les personnes âgées (Palmer et al., 2019). Ne pas percevoir adéquatement son ou ses interlocuteurs encourage le retrait d'activités dans de tels environnements et réduit ainsi les interactions sociales des aînés aînés (Heine & Browning, 2002; Palmer et al., 2016).

À cause des incidences négatives du déclin de la capacité à percevoir la parole dans le bruit, il est primordial de trouver des stratégies de prévention afin de conserver des interactions sociales agréables et une qualité de vie satisfaisante en vieillissant.

Le cerveau adulte demeure malléable à l'effet d'expériences externes, ce phénomène se nomme « plasticité cérébrale dépendante de l'expérience ». La plasticité réfère aux réarrangements structurels et fonctionnels dans le système nerveux, des mécanismes pouvant être mis à profit pour conférer une certaine protection contre le déclin cérébral normal lié à l'âge. La pratique musicale est une activité complexe et un puissant promoteur de plasticité, faisant appel à des traitements multisensoriels, cognitifs et moteurs. Par son implication de différentes fonctions, en particulier auditives et cognitives, la pratique musicale suscite un intérêt particulier pour la prévention du déclin de la perception de la parole dans le bruit.

L'objectif principal de ce mémoire est d'évaluer les bénéfices de la pratique musicale sur la capacité à percevoir la parole dans le bruit. Le premier sous-objectif est de faire un état des lieux de la littérature existante sur l'effet des activités musicales sur la capacité à percevoir la parole dans le bruit, avec la première revue systématique et méta-analyse sur le sujet. Le deuxième objectif est de comparer l'effet de la pratique musicale (chant, instrument

de musique) à des activités non musicales psychomotrices sur la capacité à percevoir la parole dans le bruit chez de jeunes adultes et des adultes âgés, afin d'en évaluer l'influence en fonction de l'âge. Dans le dernier chapitre, les résultats des deux études seront discutés et des pistes de réflexion sont proposées pour l'élaboration de futures études.

Chapitre 1 La perception de la parole dans le bruit

La perception de la parole dans le bruit est la capacité à percevoir un flux de parole d'intérêt au milieu d'un paysage de flux sonores différents. Par exemple, percevoir ce que votre ami vous dit au milieu du brouhaha d'une salle à manger de restaurant. Percevoir la parole dans le bruit repose sur des mécanismes périphériques et centraux, impliquant l'audition périphérique, l'audition centrale, les fonctions exécutives et les fonctions langagières.

Dans la première section de ce chapitre, les bases physiologiques des différentes étapes de la PPB seront abordées. Dans la deuxième section, les méthodes d'évaluation de la PPB seront présentées. Finalement, la troisième section traitera des effets du vieillissement sur la PPB.

1.1 Bases physiologiques des étapes de la perception de la parole dans le bruit

1.1.1 Le signal sonore

Le signal sonore est un signal vibratoire, c'est-à-dire une onde, qui se propage en provoquant des variations de pression sur les molécules du milieu dans lequel il se trouve. Comme toutes ondes, les ondes sonores sont caractérisées, entre autres, par leur fréquence et leur amplitude. La fréquence d'une onde représente le nombre d'oscillations effectuées par unité de temps (voir Figure 1) et correspond à la hauteur perçue : une fréquence élevée donne un son aigu tandis qu'une fréquence basse donne un son grave. L'amplitude d'une onde fait référence à la différence entre la hauteur maximale/minimale qu'atteignent les molécules en mouvement depuis leur position de repos (voir Figure 2). L'amplitude du son correspond à l'intensité perçue : une forte amplitude donne un son fort tandis qu'une faible amplitude un son moins fort.

Frequency

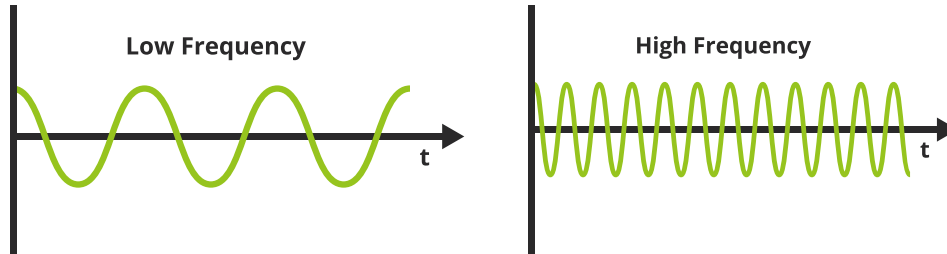


Figure 1. Schéma de la fréquence d'une onde. Low frequency = basse fréquence; high fréquence = haute fréquence. Image libre de droit numéro 1854443230, tirée de <https://www.shutterstock.com/fr/g/petrroudney43>.

Amplitude

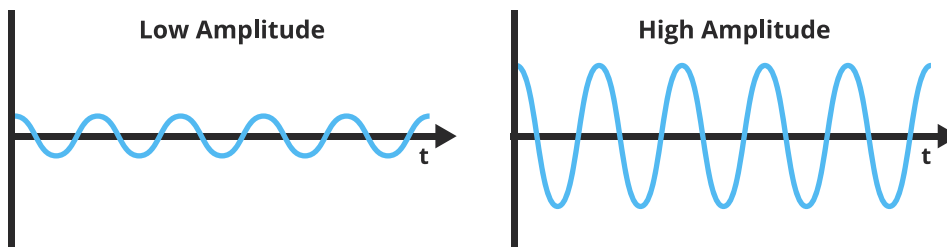


Figure 2. Schéma de l'amplitude d'une onde. Low amplitude = fréquence basse; high amplitude = amplitude élevée. Image libre de droit numéro 2080264282, tirée de <https://www.shutterstock.com/fr/g/petrroudney43>.

Deux types de sons se distinguent, les sons purs et les sons complexes. Les sons purs sont composés d'une onde unique, c'est-à-dire qu'ils ne possèdent qu'une seule fréquence. Les sons purs ne sont pas naturels, ils sont créés par les vibrations d'un diapason ou sont synthétisés artificiellement, comme les sons présentés lors d'un test d'audiométrie tonale. Les sons complexes, eux, sont des sons naturels et sont composés de la superposition de plusieurs ondes de fréquence différente. Lorsque les fréquences qui composent un son

complexe sont des multiples de la fréquence la plus basse, il s'agit de sons complexes périodiques (voir Figure 3). La fréquence la plus basse est appelée la fréquence fondamentale (F0), c'est elle qui va donner la hauteur perçue du son, et ses multiples sont appelés des harmoniques. Les sons périodiques comptent les sons musicaux et les voyelles parlées. En revanche, dans un son apériodique, les fréquences qui le composent ne sont pas des multiples de la F0. Les sons apériodiques comptent ce que l'on appelle communément le « bruit », donc le « bruit » de la ventilation, des appareils électroménagers, de l'eau qui coule, du vent, etc.

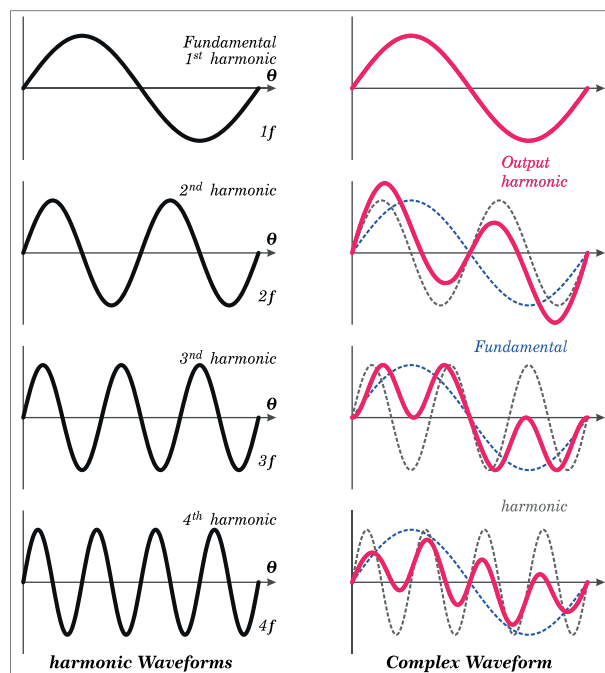


Figure 3. Schéma d'une onde complexe périodique et des ondes harmoniques qui la composent. Harmonic waveforms = ondes harmoniques; fundamental 1st harmonic = fréquence fondamentale (F0); 2nd, 3rd, 4th harmonic = harmoniques ; complex waveform = onde complexe. Image libre de droit numéro 719150923, tirée de <https://www.shutterstock.com/fr/g/Fouad+A.+Saad>.

Pour que la parole soit perçue par notre système nerveux, ses ondes sonores doivent être captées et transformées en influx nerveux par le système auditif périphérique. Puis, l'influx nerveux doit être cheminé jusqu'au cortex (voir Figure 4), où il sera traité par les aires auditives et dans les aires responsables du langage.

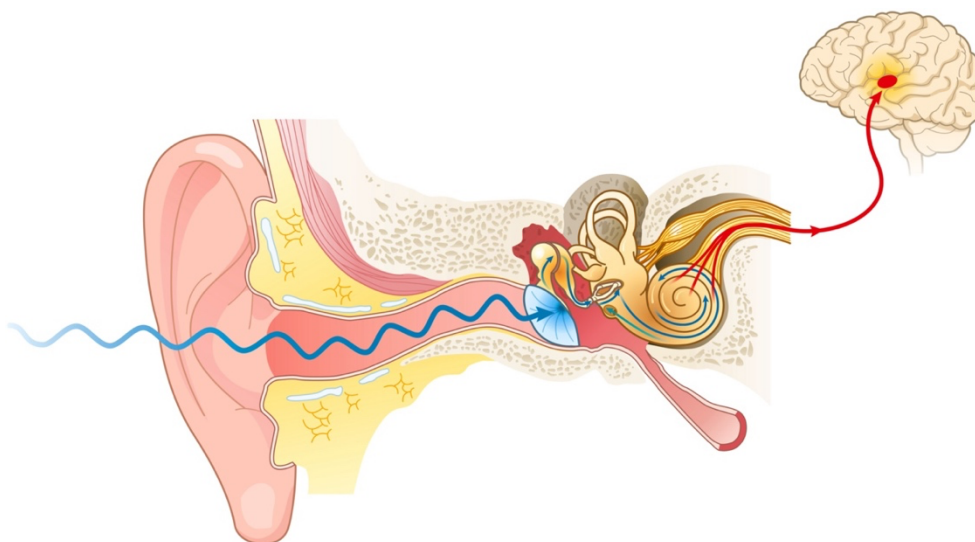


Figure 4. Schéma du passage du son dans le système auditif périphérique, depuis la captation de l'onde sonore jusqu'au cheminement de l'influx nerveux au cortex auditif primaire. Image libre de droit numéro 1960818781 tirée de https://www.shutterstock.com/fr/g/Axel_Kock

1.1.2 L'audition périphérique

Le système auditif périphérique est composé de trois entités interreliées : l'oreille externe, l'oreille moyenne et l'oreille interne (voir Figure 5). Les ondes sonores sont tout d'abord captées par l'oreille externe, constituée du pavillon de l'oreille, du conduit auditif externe et du tympan (voir Figure 5). Le pavillon capte les ondes sonores et, grâce à la conque, il amplifie sélectivement les fréquences essentielles pour le langage. Les ondes sonores sont ensuite dirigées et concentrées dans le conduit auditif externe, qui aboutit sur la membrane tympanique. Sous l'effet des ondes sonores qui arrivent au fond du conduit auditif, la membrane tympanique entrera en vibration et transmettra mécaniquement les ondes sonores à l'oreille moyenne.

L'oreille moyenne est une cavité aérienne positionnée entre l'oreille externe et l'oreille interne (voir Figure 5). Cette cavité contient une chaîne de trois osselets : le marteau, l'enclume et l'étrier. Le marteau est en contact avec la membrane tympanique et l'étrier avec la fenêtre ovale, point de contact avec l'oreille interne. La vibration du tympan

engendrée par les variations de pression de l'onde sonore passe par la chaîne ossiculaire et est ainsi transmise à l'oreille interne. Afin d'être efficacement transmises, les ondes sont amplifiées par leur passage dans l'oreille moyenne, grâce à la différence de taille entre la membrane tympanique, plus grande, et la fenêtre ovale, plus petite. Cette amplification est nécessaire pour mettre en mouvement le liquide contenu dans l'oreille interne et ainsi poursuivre la transmission de l'onde sonore.

L'oreille interne est composée de deux organes neurosensoriels : la cochlée et le vestibule (voir Figure 5). Le vestibule, avec les canaux semi-circulaires, constitue l'organe de l'équilibre et la cochlée représente l'organe de l'audition. La cochlée est un cylindre allongé et enroulé sur lui-même, telle une bobine, qui est divisé en trois chambres liquidiennes : une chambre interne et deux chambres externes qui servent d'amortisseurs. La chambre interne est remplie d'un liquide appelé endolymphe et contient la machinerie nécessaire à la transduction de l'onde sonore en message chimique. La membrane de la chambre interne, la membrane basilaire, porte l'organe de Corti. C'est dans cette structure qu'aura lieu la transduction. L'organe de Corti est recouvert de cellules ciliées, qui sont chacune connectée à une cellule ganglionnaire du nerf auditif. Les cils des cellules ciliées se plient sous les mouvements engendrés par les vibrations de l'onde sonore sur la membrane basilaire. Le mouvement des cils active les mécanorécepteurs des cellules ciliées, ce qui déclenche le relargage des neurotransmetteurs et active ainsi les neurones ganglionnaires du nerf auditif. Ceci transforme l'onde sonore en influx nerveux.

Un premier traitement des fréquences du son est effectué à la hauteur de la cochlée, grâce à la structure de la membrane basilaire. En effet, la membrane basilaire est organisée de façon *tonotopique*, c'est-à-dire que les fréquences de hauteurs différentes sont traitées sur des portions différentes de la membrane basilaire. Celle-ci est plus rigide à sa base et plus molle à son apex. Les sons de haute fréquence, donc plus rapides, parviennent à faire bouger les portions plus rigides de la membrane tandis que les sons plus bas, plus lents, mettent en mouvement les parties plus molles de la membrane.

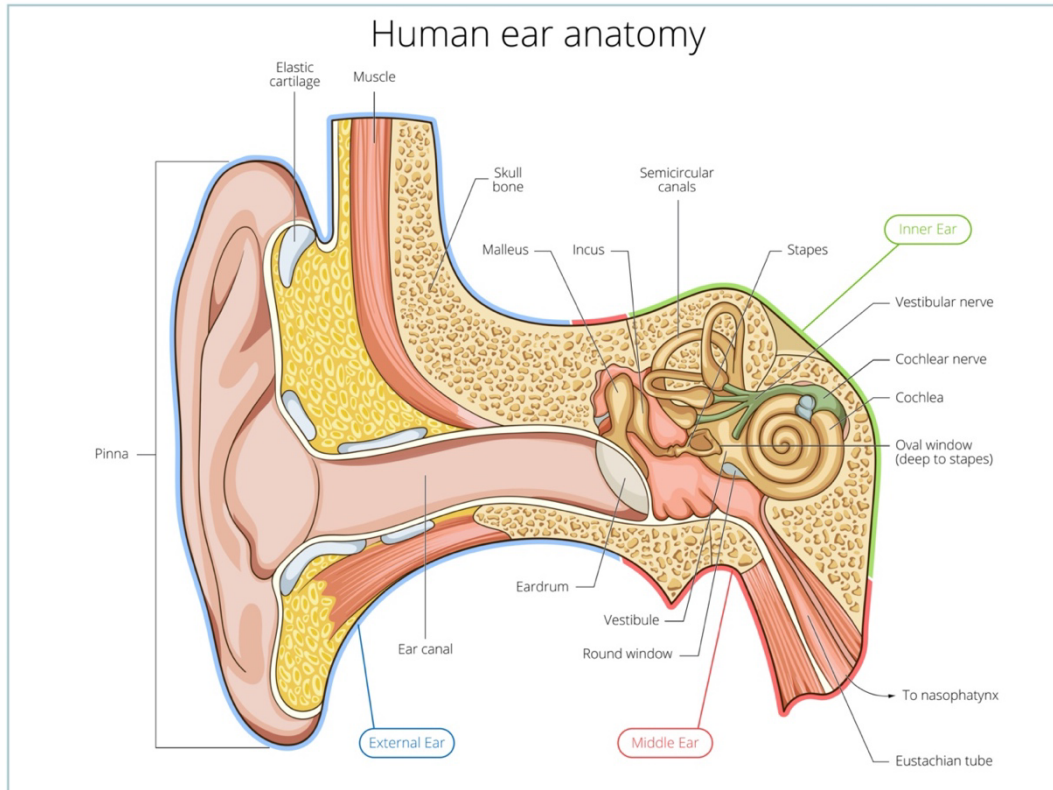


Figure 5. Schéma de l'appareil auditif périphérique. External ear = oreille externe ; ear canal = conduit auditif externe ; eardrum = tympan ; middle ear= oreille moyenne ; malleus = marteau ; incus = enclume ; stapes = étrier ; oval window= fenêtre ovale ; inner ear = oreille interne ; semicircular canals = canaux semi-circulaire, vestibule = vestibule ; cochlea = cochlée ; vestibular nerve = nerf vestibulaire ; cochlear nerve = nerf cochléaire. Image libre de droit numéro 401539120 tirée de <https://www.shutterstock.com/fr/g/alexpokusay>.

1.1.3 Les voies auditives ascendantes – la voie lemniscale

L'influx nerveux chemine depuis l'émergence du nerf auditif (une partie du nerf crânien n° VIII) jusqu'au cortex auditif primaire dans la voie lemniscale. Celle-ci est composée de plusieurs noyaux de relais synaptiques dans le tronc cérébral et dans le thalamus, ainsi que par les axones qui en émergent. La voie lemniscale, tout comme la cochlée, est organisée de façon tonotopique, c'est-à-dire que les fréquences différentes cheminent dans des axones distincts, qui font ensuite synapse dans des portions différentes des noyaux-relais.

À la sortie de l'oreille interne, le nerf auditif chemine dans le nerf crânien VIII en direction du tronc cérébral. Dans ce dernier, les axones font relais dans 3 noyaux différents, le noyau cochléaire, puis le complexe olivaire supérieur et finalement le collicule inférieur. Après le premier relais synaptique dans le noyau cochléaire ipsilatéral, une partie des fibres progresse en direction du complexe olivaire supérieur ipsilatéral, et l'autre partie des axones croise la ligne médiane pour atteindre le complexe olivaire supérieur du côté controlatéral. Recevant des axones ipsi — et controlatéraux, le complexe olivaire supérieur est par conséquent le premier noyau qui reçoit l'influx nerveux en provenance des deux oreilles. Les différences entre les deux oreilles sont codées sous forme de différence interaurale d'intensité et de temps par des neurones qui y répondent spécifiquement. À partir du complexe olivaire supérieur, à chaque relais du tronc cérébral, une portion des axones qui émergent décussent vers le noyau controlatéral du même niveau et l'autre portion chemine en amont vers le prochain relais ipsilatéral.

À la sortie du tronc cérébral, le dernier relais de la voie lemniscale avant le cortex est le corps genouillé médian du thalamus, lequel projette uniquement ses axones du côté ipsilatéral. Par les radiations auditives, l'information issue de la voie lemniscale est projetée au cortex auditif primaire.

Le cortex auditif primaire siège dans le gyrus de Heschl. Dans celui-ci, le signal entrant est traité sur le plan de ses propriétés spectrales (fréquence), temporelles (variations dans le temps), de son intensité et de sa localisation. Cette analyse permettra notamment de séparer le signal entrant en différents objets acoustiques. La théorie de l'analyse de la scène auditive proposée par Bregman (Bregman, 1990) reprend les principes gestaltistes employés en perception visuelle : la proximité, la similarité, la continuité et le destin commun (pour une revue, voir Wagemans et al., 2012). Concrètement, un groupement primitif au sein de la scène auditive s'opère de façon automatique selon ces différents principes. C'est-à-dire que des objets cohérents sont formés en regroupant les composantes acoustiques qui proviennent de la même source sonore et/ou qui partagent des relations harmoniques et/ou qui présentent

des fluctuations d'amplitude similaires dans le temps (Alain & Bernstein, 2015; De Cheveigné, 2002; Sussman, 2017). C'est ainsi que les différentes voix et les autres événements sonores de la scène acoustique pourront être différenciés les uns des autres.

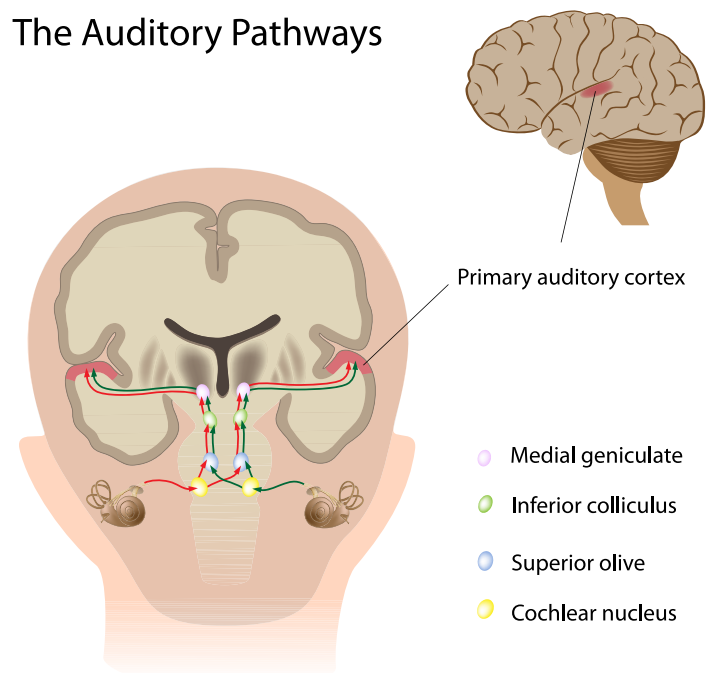


Figure 6. Voies lemniscales schématisées avec les différents noyaux-relais. Cochlear nucleus = noyau cochléaire ; superior olive = noyau olivaire supérieure ; inferior colliculus = collicule inférieur ; medial geniculare = corps genouillé médian du thalamus ; primary auditory cortex = cortex auditif primaire. Il est à noter que les projections bilatérales ne sont pas schématisées.

Image libre de droit numéro 113164525 tirée de <https://www.shutterstock.com/fr/g/alila>

1.1.4 Perception et compréhension de la parole

Après le premier traitement cortical du signal acoustique dans le cortex auditif, les objets auditifs correspondants à de la parole seront traités sur le plan sous-lexical, lexical, puis sémantique pour accéder à la signification des mots entendus.

L'analyse acoustique représente la première étape ; elle permet la transformation d'un signal sonore continu en des représentations phonologiques discrètes. À partir de l'assemblage de ces dernières, la représentation sonore des mots entendus pourra être extraite du lexique phonologique, stocké en mémoire à long terme. Une fois les représentations lexicales des mots sélectionnées, elles seront mises en lien avec leur contenu sémantique, ce qui permettra en dernier lieu d'accéder au sens des mots, et ultimement au sens de la phrase. La compréhension du langage ne peut émerger que lorsque toutes ces analyses ont été effectuées.

Les régions corticales qui contribuent à ces différentes étapes de traitements ne sont pas encore totalement élucidées à ce jour. Néanmoins, différents modèles neuroanatomiques à deux voies, inspirés du modèle à voie ventrale (*What?*) et voie dorsale (*Where?*) de la perception visuelle (Mishkin et al., 1983), ont été proposés pour caractériser les étapes du traitement de la parole (Fridriksson et al., 2016; Hickok & Poeppel, 2007; Rauschecker & Scott, 2009).

Le modèle à deux voies de Hickok et Poeppel (Hickok & Poeppel, 2007, 2016) propose que le traitement du signal de la parole commence par une analyse spectrotemporelle au sein des gyri temporaux supérieurs gauche et droit, suivi d'un traitement phonologique dans le sillon temporal supérieur moyen/postérieur. La suite du traitement se déroulerait ensuite dans deux voies différentes, agissant en parallèle. Une voie ventrale bilatérale serait responsable de mettre en lien les représentations phonologiques de la parole à leurs représentations sémantiques. Une deuxième voie latéralisée à gauche, la voie dorsale, serait

responsable de mettre en relation les représentations phonologiques à des représentations articulatoires en vue de la production de la parole. La voie ventrale comprendrait une interface lexicale, située dans la partie postérieure du gyrus temporal moyen (GTM) et du sillon temporal inférieur (STI), et un réseau combinatoire localisé dans les portions antérieures du GTM et du STI. La voie dorsale comprendrait une interface sensorimotrice, « Spt », logée dans la scissure de Sylvius entre le lobe temporal et le lobe pariétal, et un réseau articulatoire, qui impliquerait le gyrus frontal inférieur et le cortex prémoteur.

Dans le modèle de Hickok & Poeppel, la relation entre les représentations phonologiques et les représentations motrices est unidirectionnelle ; il n'y a pas d'influence des représentations articulatoires sur la perception de la parole. Or, différentes études ayant fait appel à l'IRM fonctionnelle ou la stimulation magnétique transcrânienne (SMT) montrent le recrutement de régions impliquées dans la production de la parole pour la perception, notamment en présence de bruit (pour une revue, voir Skipper et al., 2017).

L'intégration du signal visuel dans la perception de la parole n'a par ailleurs pas de place non plus dans le modèle de Hickok & Poeppel. En effet, bien que l'influence de la modalité visuelle sur la perception de la parole soit bien établie, l'effet McGurk étant l'exemple par excellence (McGurk & Macdonald, 1976), aucun lien avec le cortex visuel n'est proposé. De plus, différentes études ont montré que l'apport visuel facilite non seulement la reconnaissance de la parole en condition d'écoute difficile, comme en présence de bruit (Schubotz et al., 2021; Schwartz et al., 2004), mais également en condition d'écoute normale (Arnold & Hill, 2001; Tremblay, Basirat, et al., 2021). Toutefois, il n'existe pas encore de modèle multimodal et dynamique de la perception de la parole; c'est-à-dire un modèle qui prenne en compte les différentes modalités qui participent à la perception de la parole, ainsi que les voies d'adaptation du système face aux conditions d'écoute difficiles, comme le sont la PPB (Heald & Nusbaum, 2014), la présence de plusieurs locuteurs, de locuteurs avec un débit de parole rapide ou encore de locuteurs avec une articulation modifiée, que ce soit par une pathologie ou tout simplement par un accent particulier (Mattys et al., 2012; Peng & Wang, 2019).

1.1.5 Perception de la parole et fonctions cognitives

La PPB repose également sur les fonctions cognitives. Deux méta-analyses ont mis en évidence l'importance de la mémoire de travail, de l'attention sélective, ainsi que de la vitesse de traitement dans la PPB (Akeroyd, 2008; Dryden et al., 2017).

L'attention permet de moduler le processus automatique de la ségrégation de la scène auditive présenté plus haut (Sussman, 2017). Lorsque des objets similaires, tels que des voix, se retrouvent à la fois dans le flux sonore d'intérêt et dans le bruit de fond, il est possible que la ségrégation des différents locuteurs ne s'opère pas de façon automatique. La redirection de l'attention sur des indices particuliers, comme la F0, permettra alors d'affiner la ségrégation primitive et de séparer les objets acoustiques de F0 proches. Une fois les objets séparés, le rôle de l'attention est tout aussi nécessaire pour percevoir efficacement le flux d'intérêt au sein du bruit de fond. En effet, porter une attention sélective sur le flux d'intérêt permet de renforcer sa représentation et de réduire les interférences du bruit de fond.

La mémoire de travail est définie comme « un système à la fois de stockage temporaire et de manipulation d'information, nécessaire à la réalisation de tâches cognitives complexes, telles que la compréhension du langage, l'apprentissage et le raisonnement » (Baddeley, 1992). Le modèle *Ease of Language Understanding* ou *ELU* (Rönnberg, 2003; Rönnberg et al., 2019; Rönnberg et al., 2013; Rönnberg et al., 2008) propose un rôle central pour la mémoire de travail dans la perception de la parole en condition d'écoute dégradée, comme en présence de bruit. Lorsque le signal acoustique est dégradé, certaines séquences peuvent ne pas être automatiquement mises en lien avec des représentations phonologiques internes. Le modèle *ELU* propose alors que ces séquences sont gardées en mémoire de travail, afin que des processus de compensation, tels que la relocalisation de l'attention, l'inhibition d'interférences ou encore l'intégration d'indices sémantiques, puissent pallier l'inadéquation entre le signal auditif entrant et les représentations internes.

1.2 Évaluation comportementale de la PPB

Dans cette section, les paradigmes employés pour évaluer la PPB et leurs différentes composantes seront présentés. Les tâches utilisées pour évaluer la PPB, que ce soit en contexte clinique ou expérimental, suivent toutes le même principe. Dans ces tâches, des extraits de parole sont présentés dans un bruit masquant (BM) à un ou plusieurs niveaux de ratio signal sur bruit (SNR; “signal-to-noise ratio” (voir la section 2.4)). Les stimuli, autrement dit les paroles cibles (voir la section 2.1) et le BM (voir la section 2.2), peuvent être soit colocalisés, c’est-à-dire provenir de la même source sonore (p. ex. du même haut-parleur), soit séparés spatialement en étant présentés depuis des sources sonores séparées dans l’espace (p. ex. des haut-parleurs différents).

Les tâches de PPB peuvent être des tâches de discrimination ou d’identification (incluant la répétition) (McGuire, 2010). Dans le cas de tâches de discrimination, deux sons (parfois 3) sont présentés et un jugement de similitude doit être effectué. Dans le cas de tâches d’identification, la parole cible doit être soit répétée, oralement ou par écrit, soit associée avec le texte ou le dessin lui correspondant.

La performance à une tâche de PPB peut être évaluée de plusieurs façons. Elle peut être évaluée en calculant le pourcentage de bonnes réponses ou le SNR qui aura engendré un pourcentage prédéfini de bonnes réponses, c’est-à-dire le seuil d’intelligibilité de la parole (SRT ; « *speech reception threshold* »). Dans une tâche de discrimination, les paramètres de la théorie de la détection du signal (Egan, 1971) peuvent être également utilisés ; il s’agit de la sensibilité (d') — une mesure de la capacité sensorielle à discriminer — et du critère de décision — une mesure de la prise de décision en situation de perception ambiguë. Le temps de réaction est aussi fréquemment mesuré.

En résumé, une tâche de PPB se compose de paroles cibles, masquées par un bruit de fond, tous deux présentés depuis la même source sonore ou depuis des sources séparées, à un ou plusieurs niveaux de SNR différents. Ce sont des tâches de discriminations ou d’identification et la performance peut être mesurée à l’aide de différents indices. La parole cible, le BM, le SNR et la présentation spatiale des stimuli sont plus amplement présentés dans les prochaines sous-sections.

1.2.1 La parole cible

La parole utilisée comme cible dans une tâche de PPB peut varier dans son niveau linguistique, tout comme dans ses propriétés spectrotemporales.

La parole cible peut être sous forme de syllabes, de mots ou encore de phrases. Utiliser des syllabes permet d'évaluer les étapes de la perception de la parole de niveau sous lexical tandis que présenter des mots ou des phrases permet d'évaluer la perception de la parole au plan sous-lexical et au plan lexical.

La parole cible peut être soit de la parole naturelle, mais il peut aussi s'agir de parole chuchotée, chantée ou encore modifiée expérimentalement. Ces changements acoustiques dans la cible permettent de tester la robustesse de la perception de la parole en présence notamment de variations spectrotemporales. Parmi les modifications expérimentales, il est possible de faire varier la F0 de la voix du locuteur cible et/ou de locuteurs dans le BM. La modification expérimentale de la F0 permet, par exemple, de comparer la performance en PPB lorsque les F0 des locuteurs de la cible et du BM sont proches et lorsqu'elles sont éloignées (Deroche et al., 2017). Un autre type de modification expérimentale est le « vocoding ». Ce dernier consiste en la création de paroles à partir d'un nombre fini de bandes fréquentielles issues de la décomposition d'un extrait de parole naturelle (Roberts et al., 2011). L'attrait de la parole vocodée est l'utilisation de parole dont le contenu spectral est amoindri, tout en respectant son contenu temporel. Ceci permet d'évaluer l'importance des indices spectraux sur la capacité à percevoir la parole dans le bruit.

En résumé, la parole utilisée comme cible dans une tâche de PPB peut varier dans son niveau linguistique tout comme dans ses propriétés acoustiques. Ces différences permettront de cibler différents niveaux de traitements de la PPB.

1.2.2 Le bruit masquant

Les BMs dans les tâches de PPB sont habituellement de deux types. Il peut s'agir d'un bruit composé uniquement d'extrait de parole, comme du babillage multilocuteurs, ou d'un bruit sans paroles, comme du bruit blanc. Deux types de propriétés masquantes permettent

de caractériser les BMs : les propriétés dites *énergétiques* et les propriétés dites *informationnelles* (Freyman et al., 1999).

Le masquage énergétique est un masquage physique, qui entre en compétition avec la parole cible à la hauteur de la cochlée. Ce masquage entraîne une compétition entre les propriétés spectrotemporales de la parole cible et du BM, rendant une partie de la parole cible inaudible. De ceci résultera la transmission d'un signal acoustique dégradé de la parole cible. Dans la vie de tous les jours, de tels bruits peuvent être représentés par le bruit de la ventilation, de l'eau qui coule d'un robinet, de la machine à laver ou encore d'autres appareils électroménagers. Dans les tâches de PPB, les BMs énergétiques sont des bruits ne comportant que du masquage énergétique. Les exemples les plus courants dans la littérature sont le bruit blanc ou des bruits modulés de parole (« speech shaped noise » (SSN)). Ces derniers sont des bruits composés des fréquences qui possèdent le même spectre moyen vocal à long terme que les extraits de paroles à partir desquels ils sont créés.

Le masquage informationnel est quant à lui un masquage perceptuel, dont les effets agissent en aval de la transduction du signal acoustique. Le masquage informationnel est issu de similarités perceptuelles entre le BM et la parole cible, qui perturbent la PPB à deux niveaux (pour une revue, voir Kidd et al., 2008; Shinn-Cunningham, 2013). Ces similarités peuvent entraver la ségrégation de la scène acoustique, autrement dit, il sera plus difficile de séparer la cible de certains éléments du BM en objets distincts. De plus, la sélection de l'objet auditif d'intérêt peut également être perturbée, c'est-à-dire que l'attention sera détournée de la parole cible vers des éléments du BM. Dans la littérature, les BM informationnels sont généralement représentés par des bruits composés de parole, bien que ceux-ci soient à la fois énergétiques et informationnels.

En effet, les BMs de parole comportent à la fois des éléments énergétiques et des éléments informationnels. La composante énergétique réside dans les propriétés spectrotemporales de la/des voix du BM qui entrent en compétition avec les propriétés spectrotemporales de la voix de la parole cible. Les propriétés informationnelles d'un BM de parole résident dans différentes similarités perceptuelles, telles que celles qui peuvent exister entre différentes voix et celles qui découlent de la présence d'un contenu linguistique à la fois dans la cible et dans le BM (Brouwer et al., 2012; Brungart, 2001; Hoen et al., 2007).

Lorsque les locuteurs de la parole cible et du bruit sont de même sexe, la difficulté de perception est plus élevée que lorsque les locuteurs sont de sexes différents, car il est plus difficile de différencier les voix de locuteurs de même sexe que de sexes différents. Par exemple, dans l'étude de Brungart (2001), la performance en PPB a été évaluée avec des BMs de parole comprenant 2, 3, et 4 locuteurs. Les locuteurs des BMs étaient soit le même locuteur que pour la cible, soit un/des locuteur/s différents de la cible, mais du même sexe, soit un/des locuteur/s de sexe différent. Les meilleures performances ont été obtenues avec des locuteurs de sexes différents, tandis que de moins bonnes performances étaient obtenues avec des locuteurs du même sexe et particulièrement, lorsqu'il s'agissait du même locuteur entre le BM et la cible. Les résultats de cette étude illustrent que peu importe le nombre de locuteurs, la performance est toujours meilleure plus les voix du BM et des cibles diffèrent. Par ailleurs, le pouvoir informationnel d'un BM de parole est aussi modulé en variant l'accès à son contenu sémantique, que ce soit en maniant la langue ou le nombre de locuteurs présents dans le BM. En effet, lorsque la langue des locuteurs du BM diffère de celle du locuteur cible, la performance est facilitée. Par exemple, l'étude de Brouwer et al. (2012) a démontré que la performance de jeunes adultes monolingues de langue maternelle anglaise était significativement meilleure dans un bruit composé de phrases en hollandais (donc non comprises par les participants) que de phrases en anglais. Cette étude a également montré que la performance de ces mêmes participants était significativement améliorée lorsque le BM était composé de phrases en anglais n'ayant pas de sens, comparé à des phrases sémantiquement correctes en anglais. Le nombre de locuteurs influence également l'accès au contenu sémantique du BM et donc son pouvoir masquant informationnel. La situation la plus aisée se rencontre en présence d'un BM composé d'un unique locuteur. Malgré le contenu informationnel de celui-ci, les variations énergétiques d'un discours unique donnent lieu à des fenêtres, appelées « glimpses », pendant lesquelles l'énergie du locuteur cible excédera celle du masqueur d'au moins 3 dB (Cooke, 2006). Ceci donne ainsi l'opportunité au participant d'avoir accès à la parole cible de façon moins dégradée. En augmentant le nombre de locuteurs dans le BM jusqu'à deux ou trois, il a été démontré que la difficulté est à son comble due à l'augmentation du contenu informationnel (Freyman et al., 2001; Hoen et al., 2007; Rosen et al., 2013; Tun & Wingfield, 1999). En revanche, augmenter le nombre de locuteurs au-delà de trois ou quatre diminue l'intelligibilité du

contenu sémantique et rend la perception de la parole cible plus aisée, et ce, malgré l'augmentation du contenu énergétique produit par les voix additionnelles (Freyman et al., 2004; Hoen et al., 2007; Rosen et al., 2013; Tun & Wingfield, 1999). L'étude de Rosen et al. (2013) illustre bien ce phénomène en ayant montré une chute de la performance en PPB chez de jeunes adultes en passant de 1 à 2 locuteurs dans le BM, et une remontée graduelle de la performance entre 4, 8 et 16 locuteurs.

En somme, ces études démontrent l'importance du contenu sémantique sur le pouvoir informationnel du masqueur et ses répercussions négatives sur la PPB.

1.2.3 La présentation spatiale des stimuli

Dans une tâche de PPB, les stimuli (BM et cibles) peuvent être présentés selon plusieurs configurations spatiales : soit de façon colocalisée soit de façon à être séparés spatialement avec $> 0^\circ$ azimut. Des stimuli colocalisés peuvent être présentés depuis un haut-parleur unique, généralement situé face au participant, ou depuis un casque d'écoute. Des stimuli séparés spatialement peuvent être présentés depuis des haut-parleurs différents, séparés de façon symétrique ou asymétrique par rapport au participant, ou également depuis un casque d'écoute lorsque la séparation spatiale est créée artificiellement. Il a été démontré que de séparer spatialement la parole cible du BM augmente la performance en PPB (Hawley et al., 2004; Plomb & Mimpen, 1980). En effet, la séparation spatiale des cibles et du BM ajoute des indices pour la ségrégation des objets acoustiques, sous la forme de différences interaurales de temps et d'intensité. Effectivement, selon la position des sources sonores dans l'espace, les sons n'arrivent ni au même moment ni avec la même intensité dans les deux oreilles (Glyde et al., 2011).

Le bénéfice obtenu par la séparation spatiale est plus important dans un BM avec une forte composante informationnelle (Arbogast et al., 2005; Freyman et al., 2001, 2004; Freyman et al., 1999; Yost, 2017). En effet, dans les BMs de parole, le bénéfice spatial varie en fonction du nombre de locuteurs présent dans le bruit, donc, de la balance entre les composantes informationnelles et énergétiques. L'étude de Freyman et al. (2004) a mesuré la performance de la PPB en présence de bruits de paroles composés de 2, 3, 4, 6 et

10 locuteurs. Ces bruits étaient soit colocalisés aux cibles, soit séparés spatialement. L'étude a montré que la différence de performance entre les deux conditions spatiales était maximale en présence de deux locuteurs dans le BM et que cette différence se resserrait avec l'augmentation du nombre de locuteurs dans le bruit. Cela implique que la séparation spatiale était particulièrement bénéfique lorsque le contenu sémantique du bruit était saillant (2 locuteurs), et que l'avantage de la séparation spatiale se réduisait lorsque le contenu sémantique du bruit devenait moins intelligible.

En conclusion, séparer spatialement les cibles du BM améliore la performance en PPB de façon différentielle en fonction du type de BM et de ses caractéristiques.

1.2.4 Le niveau de SNR

Le niveau de SNR exprime le ratio de la puissance du signal (la parole cible) sur celle du bruit de fond (le BM). La puissance du son étant exprimée en décibel (dB), le niveau de SNR se calcule en soustrayant la puissance du bruit à celle du signal. Le SNR peut donc être négatif, lorsque la puissance du BM excède celle de la cible, il peut être nul, lorsque les puissances sont équivalentes, ou il peut être positif, lorsque la puissance de la cible excède celle du BM.

Dans les tâches de PPB, faire varier le SNR permet de moduler la difficulté de la tâche. Plus le SNR est positif, plus la cible est audible et plus la tâche est aisée. En revanche, plus le SNR est négatif, plus la cible est masquée par le BM ce qui rend la tâche plus ardue.

Le SNR est varié de deux façons différentes dans les tâches de PPB. Les stimuli peuvent être présentés à des niveaux prédéfinis de SNR, comme dans l'étude de Perron et al. (2021), dans laquelle les stimuli sont présentés à deux niveaux de SNR différents, -3 dB et + 3 dB. Dans ces cas, la performance est généralement évaluée en % de bonnes réponses par niveau de SNR. Le niveau de SNR peut également être modulé de façon adaptative en fonction de la performance des participants. Dans de telles tâches, la performance des participants est exprimée à l'aide du niveau de SNR ayant donné un pourcentage fixé à l'avance de bonnes réponses (généralement 50 %), donc un SRT. Par exemple, dans la tâche

de PPB de l'étude de Madsen et al. (2019), la puissance du BM est fixe et c'est la puissance de la cible qui est variée de façon adaptative, pour finalement donner une valeur de SNR pour laquelle 50 % des mots cibles sont rapportés correctement.

En conclusion, modifier le niveau de SNR permet de faire varier la difficulté d'une tâche de PPB en rendant la parole cible plus ou moins audible. Les différents niveaux de SNR d'une tâche de PPB peuvent être prédéfinis ou varier en fonction de la performance des participants.

1.3 La PPB et le vieillissement

La capacité à percevoir la parole dans le bruit décline avec l'âge et les difficultés qui en découlent nuisent à la qualité de vie des personnes âgées, pour lesquelles suivre des conversations dans des environnements bruyants, tels un restaurant ou un café, devient laborieux (Working Group on Speech & Aging, 1988).

Les difficultés des personnes âgées en PPB seront illustrées dans cette section à l'aide des résultats de diverses études. Puis, les changements dans la structure du cerveau, dans la cognition et dans l'audition qui résultent du vieillissement et leur implication dans la PPB seront abordés.

1.3.1 Les effets de l'âge sur la performance en PPB.

Les études comparant les performances en PPB d'adultes jeunes et âgés montrent des différences de performance qui varient en fonction des conditions expérimentales utilisées, en particulier en matière de bruit de fond et de niveau de SNR.

Tout d'abord s'agissant du bruit de fond, les études montrent que le vieillissement affecte la PPB différenciellement selon le type de BM. L'étude de Rajan and Cainer (2008) a comparé les performances en PPB d'adultes entre 20 et 69 ans, divisés en 5 tranches d'âges, à l'aide de phrases présentées dans un BM de parole (4 locuteurs) et un BM énergétique. Cette étude a montré que seule la cohorte la plus âgée (60-69 ans) présentait des performances significativement moins bonnes, et ce, uniquement dans le bruit de parole. Dans l'étude de Buss et al. (2019), qui a comparé la performance en PPB de jeunes adultes (19-30 ans) et d'adultes âgés (60-81 ans) à l'aide de phrases présentées dans du SSN et du babillage à deux locuteurs, une différence significative entre les jeunes et les adultes âgés a été trouvée uniquement dans le babillage à deux locuteurs. Néanmoins, un effet de l'âge sur la performance en PPB a été trouvé dans deux études qui ont présenté des syllabes dans un BM énergétique à des adultes de tout âge (19-86) (Tremblay, Brisson, et al., 2021) ou des adultes jeunes (21-32) et âgés (61-74) (Bilodeau-Mercure et al., 2015). Ces différentes études suggèrent que les adultes âgés sont significativement moins performants que les

adultes jeunes dans des BM de parole lors de la perception de *phrases*. Tandis qu'en présence de syllabes, l'âge affecte la PPB également dans un bruit énergétique.

S'agissant du niveau de SNR, il a été montré que les personnes âgées sont moins performantes que les jeunes pour un même SNR (Billings & Madsen, 2018). De plus, selon le BM utilisé, les personnes âgées requièrent un niveau de SNR plus haut que les jeunes pour obtenir la même performance. L'étude de Taitelbaum-Swead and Fostick (2016) a comparé la performance en PPB de jeunes (20-30 ans) et d'adultes âgés (60-75 ans) dans un bruit de parole (4 locuteurs) et un bruit purement énergétique à deux niveaux de SNR différents (0 dB, 5 dB). Cette étude a montré qu'au niveau de SNR = 0 dB, les performances de jeunes adultes était significativement meilleures que celles d'adultes âgés, peu importe le bruit. De plus, cette étude a montré que la réduction de performance engendrée par la baisse du niveau de SNR était significativement plus grande chez les adultes âgés uniquement dans le bruit de parole. Tandis que dans l'étude de Tun and Wingfield (1999), la baisse du niveau de SNR entre 0 dB et -6 dB s'est révélée aussi délétère en présence d'un BM de parole (1, 2, 20 locuteurs) qu'en présence d'un BM énergétique (bruit blanc) chez des adultes âgés entre 61-79 ans.

1.3.2 Contribution du vieillissement auditif et cognitif sur la PPB

Malgré les données existantes sur la PPB dans le vieillissement, les mécanismes exacts expliquant le déclin de la PPB ne sont pas encore entièrement élucidés. L'avancée en âge s'accompagne de changements structuraux et fonctionnels qui affectent l'audition périphérique et les fonctions centrales, susceptibles de contribuer au déclin de la perception de la PPB.

1.3.2.1 Vieillesse structurel du cerveau

Le vieillissement s'accompagne de modifications cérébrales, tant sur le plan structurel que fonctionnel. Les changements macroscopiques les plus flagrants dans la structure sont une atrophie de la matière grise, de la matière blanche et un élargissement des ventricules cérébraux. La diminution de volume de la matière grise commence dès la deuxième décennie

de vie et se poursuit ensuite selon une trajectoire linéaire (p. ex., voir Courchesne et al., 2000). Les structures cérébrales les plus affectées sont d'abord les lobes frontaux, surtout au niveau du cortex préfrontal latéral, puis les lobes temporaux (pour une revue, voir Hedden & Gabrieli, 2004), tandis que les lobes occipitaux sont moins affectés. Sur le plan de la matière blanche, structure qui assure le transfert de l'information, la baisse de volume devient apparente dès la cinquième décennie et se poursuit avec l'avancée en âge (p. ex., voir Giorgio et al., 2010). L'imagerie par tenseur de diffusion, qui mesure indirectement l'intégrité de la microstructure de la matière blanche, permet de montrer une détérioration de la microstructure bien plus précoce que la diminution volumique. Celle-ci est en effet détectable dès la troisième décennie, suit une décroissance linéaire jusqu'à la sixième décennie pour ensuite subir un déclin accéléré (p. ex., voir Westlye et al., 2010).

1.3.2.2 Vieillessement de l'audition périphérique

L'avancée en âge s'accompagne de la *presbyacousie*, une perte d'audition sensorineurale qui déforme le signal acoustique entrant. La presbyacousie correspond à une perte auditive progressive liée à l'âge, affectant les deux oreilles de façon relativement symétrique dans la perception des sons de hautes fréquences, pour ensuite toucher les fréquences de la parole. La presbyacousie résulte d'une combinaison de facteurs intrinsèques (susceptibilités génétiques, troubles métaboliques) et extrinsèques (exposition au bruit, agents ototoxiques) qui portent atteinte à l'intégrité de l'oreille interne, avec une perte de cellules ciliées à la base de la cochlée et/ou de cellules ganglionnaires formant le nerf auditif (pour une revue, voir Guinchard et al., 2017).

La contribution de la baisse de l'audition périphérique au déclin de la PPB a été étudiée en comparant la performance en PPB de participants, jeunes et/ou âgés, avec une surdit  sensorineurale et des participants sans trouble auditif dans du bruit  nerg tique (Souza & Turner, 1994) comme dans du bruit de parole (Humes & Roberts, 1990). Il a  galement  t  d montr  que la diminution de l'audition p riph rique   elle seule n'explique pas le d clin en PPB dans le vieillissement. En effet, de jeunes adultes pr sentant un trouble auditif sensorineural obtiennent une meilleure performance en PPB dans un bruit de fond

multilocuteurs que des adultes âgés avec une presbyacousie (Dubno et al., 1984; Gordon-Salant & Fitzgibbons, 1997). De plus, lorsque l'audition périphérique est contrôlée statistiquement, des adultes jeunes obtiennent de meilleures performances en perception de syllabes dans du bruit blanc que des adultes âgés (Bilodeau-Mercure et al., 2015). Par ailleurs, même en l'absence de trouble auditif, des adultes âgés présentent de moins bons scores en PPB que des adultes jeunes, et ce, peu importe le type de bruit (Desjardins & Doherty, 2013; Frisina & Frisina, 1997; Kim et al., 2006; Tun, 1998). Ces études suggèrent qu'au-delà de l'audition périphérique, le vieillissement des structures/fonctions nerveuses centrales pourrait contribuer au déclin de la PPB.

1.3.2.3 Vieillessement cognitif

Le vieillissement s'accompagne d'une diminution de fonctionnement dans différentes sphères cognitives. Ce sont surtout la vitesse de traitement, l'encodage de nouvelles informations et les capacités attentionnelles qui sont touchées. La vitesse de traitement diminue graduellement à partir de la deuxième décennie et affecte le fonctionnement global; ce déclin est apparent lors de tout test neuropsychologique minuté, même les tests ne mesurant pas directement la vitesse de traitement, comme les tests fluence verbale dans lesquels les personnes âgées sont plus lentes que les jeunes (Harada et al., 2013; Salthouse, 1996). La diminution de la vitesse de traitement est associée à la détérioration de la microstructure de la matière blanche (Kerchner et al., 2012). Sur le plan attentionnel, c'est particulièrement la capacité à se concentrer sélectivement sur une information au milieu de distracteurs - c.-à-d. l'attention sélective - qui diminue avec l'âge (Murman, 2015). Soutenir son attention sur une période prolongée devient également plus demandant. Au niveau de la mémoire de travail, les adultes âgés montrent plus de difficultés à manipuler des informations retenues dans la mémoire à court terme que des adultes jeunes. En ce qui concerne la mémoire à long terme, la capacité à encoder de nouvelles informations diminue et le retrait d'information déjà encodée se fait moins rapidement (Harada et al., 2013).

Sur le plan cognitif, l'importance de la vitesse de traitement et de la mémoire de travail ont été mises en évidence dans la capacité à percevoir la parole dans le bruit chez les adultes

âgés. Dans l'étude de Helfer and Freyman (2014), parmi les différentes évaluations cognitives, la vitesse de traitement montre l'association la plus forte avec différentes évaluations de la PPB, comparées à des mesures de mémoire de travail et d'inhibition. Dans l'étude de Pronk et al. (2013), la vitesse de traitement est le seul facteur expliquant la dégradation de la performance en PPB dans une cohorte de personnes âgées suivies sur une période de 3 à 7 ans, comparés aux seuils auditifs, à une mesure générale de la cognition (Mini Mental State Examination) et une mesure de l'intelligence fluide (les matrices de Raven). Dans l'étude de Tun and Wingfield (1999), ce sont à la fois les seuils auditifs et la vitesse de traitement qui expliquent les différences en PPB entre adultes jeunes et âgés. Quant à la mémoire de travail, son rôle est également démontré seul ou en association avec d'autres mesures. L'étude de (Gordon-Salant & Cole, 2016) a montré que des adultes âgés avec une meilleure mémoire de travail performaient mieux que leurs pairs avec une mémoire de travail moins bonne en perception de phrases dans le bruit, mais pas en perception de mots dans le bruit. L'étude de (Vermeire et al., 2019) a montré que la performance en PPB était significativement corrélée à la mémoire de travail et aux seuils auditifs chez des adultes âgés, mais pas chez des adultes jeunes.

Plusieurs facteurs semblent donc contribuer au déclin de la PPB éléments et dont le poids semblerait varier en fonction de facteurs intrinsèques (audition, cognition, âge) et extrinsèques (propriété expérimentale des tâches de PPB). Certains sont des éléments sur lesquels il est déjà possible d'intervenir et d'autres pas encore. En effet, la contribution de l'audition périphérique sur la PPB peut être améliorée grâce au port d'appareils auditifs. Néanmoins, ceux-ci ne suffisent pas, puisque des difficultés à percevoir la parole dans le bruit subsistent malgré le rétablissement des seuils auditifs (El-Assal et al., 2020; Humes et al., 2002). Il est par conséquent nécessaire de développer de nouvelles stratégies qui agissent en amont de l'audition périphérique, afin de préserver ou de rétablir des fonctions centrales nécessaires à la PPB.

Chapitre 2 Musique et plasticité cérébrale

Comme vu au chapitre précédent, le vieillissement s'accompagne de déclin structurels et fonctionnels qui affectent la capacité à percevoir la parole dans le bruit chez les adultes âgés. Toutefois, notre système nerveux est dynamique, capable de modifier sa structure et son fonctionnement en réponse à des stimuli internes et externes : c'est le phénomène de plasticité cérébrale. Cette dernière peut-elle être induite pour tenter de prévenir ou de limiter le déclin de la capacité à percevoir la parole dans le bruit ?

Ce chapitre commence par une présentation de la notion de plasticité cérébrale et des devis d'étude utilisés pour évaluer ses effets comportementaux, fonctionnels et structurels. Le chapitre se poursuit par la présentation des concepts de réserve ainsi que de maintien cognitif et cérébral et se termine par une section sur les effets comportementaux de la plasticité induite par la pratique d'activités musicales.

2.1 La plasticité cérébrale

2.1.1 Différents types de plasticité

Il y a trois types de plasticité : la plasticité indépendante de l'expérience (PIE), la plasticité « *experience expectant* » (PEE) et la plasticité dépendante de l'expérience (PDE) (Kolb, 2018). Les deux premiers types de plasticité sont l'apanage du cerveau en maturation. La PIE orchestre le développement du système nerveux in utero sous le contrôle de l'expression génique. À partir de la naissance, la PEE sous-tend l'installation et la consolidation des circuits ainsi que de leur fonction grâce aux expériences sensorielles déclenchées au contact de l'environnement (Kolb et al., 2013). La PEE comprend les périodes de développement dites « critiques ». Ces dernières correspondent à des fenêtres de plasticité limitées dans le temps, lors desquelles un contact sensoriel particulier avec l'environnement est nécessaire au développement de fonctions données (Cisneros-Franco et al., 2020). Le troisième type de plasticité - la PDE - représente la capacité que conserve le système nerveux à moduler, en réponse à des expériences, le fonctionnement et la structure de circuits déjà installés. C'est la forme de plasticité qui constitue le cœur de ce mémoire.

Longtemps il a été pensé que la PDE était l'exclusivité du système en développement. Toutefois, la PDE est conservée tout au long de la vie (Holman & de Villers-Sidani, 2014). C'est ainsi qu'au travers d'expériences et d'entraînements, comme la pratique d'activités musicales, nous sommes en mesure d'acquérir de nouvelles compétences, de développer des expertises et de récupérer des fonctions cérébrales lésées.

Au niveau cérébral, les changements fonctionnels induits par la PDE peuvent se traduire par une augmentation ou une diminution de l'activité neuronale dans un circuit ou une région donnée (Jonides, 2004). Sur le plan structurel, des changements dans la microstructure se rencontrent dans la matière grise comme dans la matière blanche. Les changements dans la matière grise englobent des remaniements synaptiques dans le nombre et la taille des synapses, et dans la morphologie des épines dendritiques (Draganski & May, 2008). En ce qui concerne la matière blanche, des modifications sont trouvées dans la morphologie des astrocytes et de leur surface de contact avec les neurones adjacents, ainsi que dans la quantité de myéline présente dans les faisceaux axonaux (Markham & Greenough, 2004).

2.1.2 Etude de la PDE

2.1.2.1 Méthodes

Les changements cérébraux résultant de la PDE issus d'un apprentissage peuvent être mesurés *in vivo* grâce aux techniques de neuroimagerie fonctionnelle et structurelle. Sur le plan structurel, le volume de la matière grise et de la matière blanche sont étudiés à l'aide de l'imagerie par résonance magnétique (IRM) (pour une revue, voir Tardif et al., 2016), grâce à des indices d'épaisseur corticale ou des mesures de densité dans une région donnée. La microstructure de la matière blanche est étudiée à l'aide de l'IRM par tenseur de diffusion. Cette technique permet de mesurer l'organisation au sein des fibres de matière blanche, en mesurant la direction du déplacement des molécules d'eau dans les fibres pendant un temps donné (Baumann et al., 2010; Le Bihan, 2003). Sur le plan fonctionnel, l'activité neuronale peut être mesurée pendant une tâche spécifique ou au repos, grâce à différentes techniques, comme l'IRM fonctionnelle ou l'électroencéphalographie (EEG). L'IRM fonctionnelle

mesure les changements dans l'oxygénation de l'hémoglobine induits par l'activité neuronale. L'électroencéphalographie (EEG) mesure l'activité électrique des réseaux neuronaux à l'aide d'électrodes placées sur le scalp. L'IRM fonctionnelle est plus précise sur le plan spatial, tandis que l'EEG est plus précise sur le plan temporel.

2.1.2.2 Devis

Concrètement, l'évaluation de la plasticité induite par un apprentissage peut être réalisée par des études en neuro-imagerie, dans lesquelles la structure et/ou l'activité neuronale sont mesurées. L'étude de la plasticité induite peut également se faire au moyen d'études comportementales qui s'intéressent au gain de performance issu d'un apprentissage. Le comportemental et la neuro-imagerie sont souvent couplés à l'intérieur d'une même étude. Coupler ces deux approches permet à la fois d'observer le substrat anatomique/fonctionnel sollicité par l'entraînement et de comparer les effets comportementaux avec les effets anatomiques/fonctionnels de l'entraînement. Il arrive toutefois parfois de trouver des changements structurels sans la présence de changements dans la performance comportementale (p. ex., voir Hennessy et al., 2021).

Le devis des études en neuro-imagerie et/ou comportementales peut être transversal ou longitudinal (May, 2011). Dans le premier cas, des personnes ayant été exposées à un apprentissage sont comparées à des personnes n'y ayant pas été exposées. Prenons comme exemple l'étude de Maguire et al. (2006) qui s'est intéressée aux conséquences structurelles de la mémorisation du réseau routier londonien par les chauffeurs de taxi. Cette étude a montré que les chauffeurs de taxi londoniens, qui naviguent dans l'ensemble du réseau routier, présentent un volume hippocampique plus élevé que celui des chauffeurs de bus, dont les trajets sont limités géographiquement, rectilignes et redondants, et qui, par conséquent, n'ont pas eu à mémoriser la carte du réseau routier.

Dans le cas des devis longitudinaux, les participants sont répartis, la plupart du temps aléatoirement, soit dans un groupe expérimental et dans un groupe contrôle. Le groupe

expérimental sera exposé à un apprentissage d'intérêt pendant une période déterminée. Le ou les groupes contrôles seront exposés à d'autres apprentissages ou ne seront exposés à aucun apprentissage. Des mesures de performance et/ou de neuro-imagerie seront prises avant et après l'exposition à l'apprentissage, pour y évaluer les potentiels changements. Une comparaison est donc effectuée au sein du même groupe entre pré-et post-exposition, et entre le groupe expérimental et le/les groupe/s contrôle/s en pré- comme en post-exposition pour évaluer la présence de différences structurelles, fonctionnelles ou comportementales entre les groupes. Par exemple, les études de Draganski et al. (2004) et de Boyke et al. (2008) ont montré qu'exposer des adultes jeunes et des adultes âgés à 3 mois d'entraînement au jonglage, comparé à des contrôles passifs, augmentait le volume cortical dans la région du cortex extra-strié (hMT/V5s), région responsable des mouvements guidés par la vision.

La plasticité cérébrale peut donc être évaluée au niveau structurel, fonctionnel et/ou comportemental. Les études peuvent utiliser un devis transversal en recrutant des participants déjà exposés à un entraînement donné, ou bien recruter des participants naïfs, qui seront alors exposés à un entraînement selon un régime de pratique contrôlé et pendant un temps prédéfini.

2.1.3 PDE et vieillissement

Comme discuté au début de ce présent chapitre, la PDE, bien qu'amoindrie, est conservée dans le cerveau vieillissant. En effet, même à un âge avancé, il est possible de développer de nouvelles habiletés et d'observer des changements structurels et fonctionnels à la suite de l'exposition à un entraînement particulier, même si celui-ci est commencé à un âge avancé. C'est notamment le cas de l'étude de Boyke et al. (2008) citée plus haut. Dans celle-ci, des participants âgés de 50 à 67 ans ont été entraînés au jonglage pendant 3 mois. Au terme de l'entraînement, des changements corticaux ont été observés chez les jongleurs jeunes et âgés, bien que la performance comportementale des jongleurs âgés ait été moins élevée que celle des jeunes jongleurs.

Des entraînements de nature variée sont associés à des changements comportementaux et/ou neurobiologiques chez les adultes âgés. C'est le cas d'entraînements cardiovasculaires, comme l'aérobique (Colcombe et al., 2006) ou la danse (Rehfeld et al., 2018), d'entraînements ciblant les fonctions exécutives (Dahlin et al., 2008; Nguyen et al., 2019) ou d'activités plus complexes comme l'apprentissage d'une nouvelle langue (Nilsson et al., 2021) ou la pratique d'un instrument de musique (Bugos et al., 2007). Et si les résultats comportementaux et/ou neurobiologiques de certains entraînements conféraient une protection contre certaines conséquences du vieillissement ?

2.1.3.1 Réserve et maintien

Deux concepts différents sont proposés pour expliquer les variations retrouvées chez des adultes âgés face aux effets du vieillissement ou de processus pathologiques : la réserve cognitive/cérébrale et la maintenance cérébrale. La réserve cognitive fait référence à « l'adaptabilité des processus cognitifs qui explique les différences de susceptibilité des habiletés cognitives ou du fonctionnement dans la vie de tous les jours au vieillissement cérébral ou aux pathologies » (Stern, 2002, 2006, 2009; Stern et al., 2020). Cette différence de susceptibilité est conférée par des facteurs innés, mais également par l'exposition à différentes expériences au cours de la vie. Parmi ces dernières, l'on retrouve l'éducation, l'activité physique, les activités sociales et les occupations complexes, comme la pratique musicale. La notion de « réserve cérébrale » est le pendant neurobiologique de la réserve cognitive, qui fait référence aux variations structurelles (nombre de neurones, nombre de synapses) qui confèrent une certaine protection contre le vieillissement cérébral ou les pathologies (Stern et al., 2020). La maintenance cérébrale illustre la capacité du système nerveux à rétablir un certain niveau de ressources cérébrales en réponse au déclin cellulaire et/ou moléculaire. Comme pour le concept de réserve, la maintenance cérébrale résulte d'une combinaison de facteurs innés et acquis (Barulli & Stern, 2013; Cabeza et al., 2018). La réserve cognitive et la réserve cérébrale impliquent donc qu'à un âge avancé, les personnes ayant accumulées certaines expériences possèdent un « capital » cognitif/structurel augmenté et que celles-ci seront moins rapidement ou moins fortement affectés par le vieillissement normal ou par un processus pathologique que des personnes ayant accumulé

moins d'expériences. La maintenance cérébrale signifie que la pratique de certaines activités pourrait maintenir un certain niveau de fonctionnement pour des habiletés déjà touchées par le vieillissement cérébral, comme certaines fonctions cognitives. En résumé, la réserve et la maintenance supposent un ralentissement des effets néfastes du vieillissement cérébral.

La possibilité de ralentir le déclin cérébral lié à l'âge ouvre la voie au développement de programmes de prévention et/ou de réhabilitation de certaines fonctions touchées par le vieillissement. Comme vu au chapitre précédent, la capacité à percevoir la parole décline et la baisse de l'audition périphérique n'en est pas l'unique cause. La PPB dépend notamment du traitement auditif central, de fonctions exécutives et du traitement linguistique. L'on pourrait alors imaginer un entraînement de mémoire de travail ou d'attention combiné à des exercices ciblant le traitement auditif central ou linguistique. Toutefois, il a été montré que le résultat d'entraînements focalisés sur une faculté, comme sur la mémoire de travail, n'augmentait la performance que dans la faculté entraînée, voire même uniquement dans l'exercice pratiqué (Dahlin et al., 2008; Nguyen et al., 2019; Noack et al., 2009; Park & Bischof, 2013; Shipstead et al., 2012; Teixeira-Santos et al., 2019). Un effet de transfert, c'est-à-dire la généralisation d'un apprentissage à d'autres habiletés non-entraînées directement par cet apprentissage, n'est que rarement observé dans les études utilisant des entraînements focalisés (Dahlin et al., 2008; Nguyen et al., 2019; Noack et al., 2009; Park & Bischof, 2013; Shipstead et al., 2012; Teixeira-Santos et al., 2019). Il se pourrait par conséquent que pour entraîner des facultés complexes, les entraînements le soient aussi. Il est possible que pour ralentir le déclin de la PPB, une activité qui puisse solliciter un large éventail de fonction soit nécessaire, afin de couvrir et entraîner le plus de processus impliqués dans la PPB.

2.2 La musique, outil de réserve/maintien pour la PPB

La pratique musicale, grâce à son implication de différentes fonctions en parallèle, suscite un intérêt particulier pour la prévention de la PPB.

Pratiquer un instrument de musique est un entraînement cérébral complet, faisant appel à l'intégration et la coordination de fonctions exécutives, de la mémoire sémantique, de systèmes sensoriels (visuel, auditif, somatosensoriel) et d'habiletés motrices (Schlaug, 2001, 2015; Wan & Schlaug, 2010; Zatorre et al., 2007). En effet, jouer d'un instrument implique, dans la plupart des cas, de réaliser avec succès la lecture et la reconnaissance de notes sur une partition, leur traduction en schémas moteurs et en représentations auditives, l'exécution de séquences précises et rapides de mouvements, le monitoring constant des notes jouées, le maintien des notes jouées en mémoire de travail et, dans le cas de la pratique en groupe, l'allocation de son attention sur les voies à suivre tout en résistant à la distraction des autres. De surcroît, la pratique d'activités musicales est une activité agréable, motivante, grâce à laquelle il est possible de vivre et de transmettre des émotions (Ferreri et al., 2019; Juslin & Laukka, 2004; Zatorre & Salimpoor, 2013). Des études comparant des musiciens et des non-musiciens ont montré des différences cérébrales entre les groupes sur le plan structurel et fonctionnel, ainsi que des différences comportementales chez des participants de tout âge, démontrant ainsi le potentiel plastique des activités musicales.

2.2.1 Bénéfices de la pratique musicale sur le traitement auditif central

Plusieurs études se sont intéressées aux effets de l'entraînement musical sur différents traitements auditifs centraux, comme le traitement des fréquences, incluant les relations harmoniques dans un son périodique (Zendel & Alain, 2009; B. R. Zendel & C. Alain, 2013), et le traitement du décours temporel.

Les musiciens démontrent de meilleures performances que les non-musiciens pour la détection de fréquence, comme le montrent de nombreuses études dans lesquelles des musiciens de tout âge ont un seuil de discrimination fréquentiel beaucoup plus bas que les non-musiciens ; c'est-à-dire qu'ils sont capables de différencier deux sons purs de

fréquences très proches, ou deux sons complexes de F0 très proche (Deguchi et al., 2012; Grassi et al., 2017; Hutka et al., 2015; Kishon-Rabin et al., 2001; Kuhnis et al., 2013; Liang et al., 2016; Strait et al., 2010). Cet avantage pour le traitement de la fréquence se retrouve également dans leur capacité supérieure à détecter des violations de fréquences dans une phrase musicale (Sares et al., 2018; Schon et al., 2004). Lorsque la fréquence fondamentale d'une des notes est modifiée, les musiciens sont plus performants que les non-musiciens pour repérer de petites comme de grandes variations.

Il a été montré que les musiciens ont des capacités supérieures comparées aux non-musiciens pour séparer l'information sonore entrante en objets acoustiques distincts en se basant sur les relations harmoniques entre les fréquences (Zendel & Alain, 2009; Benjamin Rich Zendel & Claude Alain, 2013). Les jeunes adultes comme les adultes âgés pratiquant la musique sont capables de percevoir deux objets acoustiques différents, basés sur la détection de violation dans les relations harmoniques (Zendel & Alain, 2009; Benjamin Rich Zendel & Claude Alain, 2013).

Des études démontrent également des performances supérieures chez les musiciens comparés à des non-musiciens dans des tâches ciblant le traitement temporel des stimuli auditifs. En effet, des performances supérieures pour les musiciens comparés aux non-musiciens sont trouvées dans des tâches de détection de silence dans un bruit (« Gap detection ») (Donai & Jennings, 2016; Grassi et al., 2017; Kuhnis et al., 2013; Kumar et al., 2014) et de discrimination de durée entre les différents sons présentés (« duration discrimination ») (Grassi et al., 2017; Kumar et al., 2014).

Certaines études ont démontré que ces avantages dans le traitement auditif central se retrouvaient également lors de l'utilisation de stimuli linguistiques. Par exemple, la capacité à détecter un changement de F0 sur une note au sein d'une phrase musicale se répercute sur la détection de changements de la F0 d'un mot dans la prosodie d'une phrase (Deguchi et al., 2012; Sares et al., 2018; Schon et al., 2004). Ces avantages comportementaux sont appuyés par des données en électroencéphalographie qui montrent que les musiciens présentent un encodage sous-cortical plus précis de la F0 dans des syllabes que des non-musiciens, précision corrélée au nombre d'années de pratique musicale (Bidelman et al., 2014; Musacchia et al., 2007). Une étude effectuée sur des musiciens et des non-musiciens

âgés n'a pas montré de différence dans la précision de l'encodage de la F0, mais plutôt que celui-ci se faisait plus rapidement chez les musiciens âgés que chez les non-musiciens (Bidelman & Alain, 2015), démontrant tout de même un bénéfice de la pratique musicale sur le cerveau vieillissant.

2.2.2 Bénéfices de la pratique musicale sur les fonctions cognitives

Dans le domaine cognitif, des études ont démontré des avantages pour des musiciens de tout âge dans différentes fonctions cognitives. Les musiciens démontrent une mémoire de travail supérieure pour des séquences de tons comparés aux non-musiciens (Bowe et al., 2021; Ding et al., 2018; Fennell et al., 2020). Les musiciens démontrent également une meilleure mémoire de travail dans le domaine verbal que les non-musiciens, par exemple dans des tâches auditives, telles que le « Listening Span » (Cools, 2010), le « Digit Span » (Wambach et al., 2011) ou le « Letter-number sequencing test » (Saklofske & Schoenberg, 2011) (D'Souza et al., 2018; Grassi et al., 2017; Groussard et al., 2020; Hanna-Pladdy & Gajewski, 2012; Mansens et al., 2018; Zuk et al., 2014). Des performances supérieures des musiciens dans une tâche mesurant la vitesse de traitement, le sous-test « Code » du WAIS, (Saklofske & Schoenberg, 2011), ont également été montrées (Groussard et al., 2020; Hanna-Pladdy & Gajewski, 2012; Mansens et al., 2018; Zuk et al., 2014). De plus, les musiciens démontrent de meilleures performances en fluence verbale phonologique, mais pas en fluence verbale sémantique (Hanna-Pladdy & Gajewski, 2012; Mansens et al., 2018). De meilleures performances dans des tâches testant certaines habiletés visuospatiales ont également été trouvées, mais de façon moins consistante que pour les autres fonctions (Grassi et al., 2017; Hanna-Pladdy & Gajewski, 2012).

2.2.3 Un « OPERA » pour la « cocktail party »

Des bénéfices de la pratique musicale sont donc retrouvés dans des tâches ciblant le traitement auditif central avec des stimuli auditifs simples et des stimuli linguistiques, ainsi

que dans des tâches ciblant différentes fonctions cognitives. Sous quelles conditions est-ce que ces différentes fonctions pourraient améliorer la perception de la parole ?

Le modèle “OPERA”, élaboré par Anirudh Patel (Patel, 2011, 2012, 2014), propose que les bénéfices engendrés par la pratique musicale améliorent la perception de la parole sous cinq conditions : le chevauchement (« Overlap »), la précision (« Precision »), l’émotion (« Emotion »), la répétition (« Repetition ») et l’attention (« Attention »). Pour illustrer ces conditions, Patel propose de prendre le traitement de la fréquence fondamentale (F0) des sons comme exemple. En musique comme en parole, la F0 des sons est traitée afin de percevoir la hauteur d’une note ou la hauteur d’une voix. De plus, les structures nerveuses mobilisées lors du traitement de la fréquence sont les mêmes, peu importe l’origine sonore. Il y a donc un chevauchement entre les structures neurales nécessaires à la perception d’une composante musicale et d’une composante vocale, première condition pour un transfert d’apprentissage selon le modèle OPERA. La deuxième condition, la précision, implique que la pratique musicale impose une sollicitation plus importante sur le traitement commun que la perception de la parole. Lors de la pratique musicale, le traitement de la F0 est essentiel. En effet, une fausse note entrave totalement la mélodie, tandis qu’en perception de la parole, le sens d’un message véhiculé oralement peut être compris malgré un changement dans la hauteur de la voix. La pratique musicale nécessite donc un traitement et un contrôle plus précis de la F0. La troisième condition est la répétition, c’est-à-dire qu’un transfert nécessite un entraînement répété du traitement commun à la musique et à la perception de la parole. La quatrième condition est l’attention ; le traitement commun doit être l’objet d’une attention soutenue par la pratique musicale. Finalement, la pratique musicale impliquant le traitement commun doit être associée à la récompense et à des émotions positives, cinquième condition pour un effet de transfert entre la pratique musicale et la perception de la parole. L’exemple utilisé ci-dessus concerne un traitement auditif, mais Patel étend son modèle également à tout processus cognitif qui serait partagé par la pratique musicale et la parole, par exemple la mémoire de travail ou l’attention.

Un certain nombre d’études ont été réalisées à ce jour sur la perception de la PPB chez des musiciens et des non-musiciens, afin d’évaluer les potentiels bénéfices de la pratique musicale sur la PPB. Le prochain chapitre fait un résumé qualitatif et quantitatif de la

littérature parue sur l'effet de la pratique musicale sur la PPB, à l'aide d'une revue systématique d'une méta-analyse.

Chapitre 3 Étude 1

Do musical activities enhance the perception of speech in noise? A systematic review and meta-analysis.

Elisabeth Maillard^{1,2,3}, Marilyne Joyal^{1,2}, Micah M. Murray³⁻⁵, Pascale Tremblay^{1,2}

1. CERVO Brain Research Center, Quebec City, Canada, G1J 2G3

2. Université Laval, Faculté de Médecine, Département de Réadaptation, Quebec City, Canada, G1V 0A6

3. The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland.

4. CIBM Center for Biomedical Imaging, Lausanne, Switzerland

5. Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

3.1 Résumé

La capacité à percevoir la parole dans le bruit diminue avec l'âge et nuit à la qualité de vie des personnes âgées. La pratique d'activités musicales, par leur entraînement cognitif et auditif, suscite l'intérêt comme potentiel outil de prévention pour la PPB. Toutefois, les résultats obtenus dans la littérature existante sont mitigés. L'objectif de cette étude était alors d'évaluer l'effet de la pratique musicale en fonction des conditions expérimentales de PPB les plus utilisées et en fonction de l'âge. Quarante-neuf (49) études ont été incluses dans la revue systématique et 32 dans la méta-analyse. Les résultats de cette étude ont montré que l'effet de la pratique musicale varie en fonction des conditions expérimentales de PPB et qu'il est plus important dans les conditions d'écoute les plus difficiles. Dû aux nombres limités d'études sur les adultes âgés, l'effet de l'âge n'a pas pu être évalué de façon adéquate.

3.2 Abstract

The ability to process speech in noise (SPiN) declines with age and negatively impacts quality of life. Musical activities have raised interests as a potential prevention strategy for SPiN perception due to their auditory and cognitive components. However, the literature on musicianship and SPiN performance has yielded mixed results. By assessing the existing literature with a systematic review and a meta-analysis, we aim to assess the strength of a musician effect in function of the most frequent experimental conditions. A total of 32 articles, the majority on young adults, were included in the quantitative analysis. The results show that the effect of musicianship varies with the experimental SPiN condition and is the strongest in the most challenging conditions. This pattern of results supports the notion of an advantage for musicians on SPiN performance. However, further studies, especially on older adults, are needed to extend the present conclusions and assess the potential for musical activities to be used as a rehabilitation strategy.

3.3 Introduction

The capacity to comprehend speech in noise (SPiN) is an essential tool for everyday life. From busy streets to supermarkets and restaurants, we are constantly faced with the challenge of untangling speech from competing sounds. SPiN capacity develops slowly, reaching adult-like performances in adolescence (e.g. for review see Leibold, 2017), and beginning to decline around the age of 50 (Moore et al., 2014) to become a common complaint among aging adults (Working Group on Speech & Aging, 1988). Difficulties comprehending speech in noisy environments can lead to reduced social participation (Heine & Browning, 2002). Social disconnectedness and perceived loneliness have in turn been linked to depressive symptoms and an increased risk of dementia in elderly populations (Taylor et al., 2018; Sundström et al., 2020; Santini et al., 2020). Finding strategies to mitigate the decline of SPiN perception is therefore crucial to maintain a satisfying quality of life and wellbeing in older adults. The exact etiology of SPiN difficulties remains elusive, however, factors such as hearing impairment, cognitive decline and brain aging appear to contribute with varying degrees to these difficulties (Souza & Turner, 1994; Tun, 1998; Bilodeau-Mercure et al., 2015; Presacco et al., 2016; Gordon-Salant & Cole, 2016; Vermeire et al., 2016; Humes, 2021).

Musical practice is a powerful promoter of brain plasticity (Wan & Schlaug, 2010; Herholz & Zatorre, 2012). It is thought to solicit cognitive, multisensory and motor processes as well as their integration (Zatorre et al., 2007). Musical practice has gained widespread interests as a potential prevention/rehabilitation tool for SPiN perception in the elderly, largely due to demonstrated enhancement of auditory and cognitive processes in musicians compared to non-musicians. Musicians exhibit better performance in auditory temporal processing (Kumar et al., 2014; Donai & Jennings, 2016; Grassi et al., 2017), enhanced spectral processing with better frequency and pitch discrimination for both musical and speech stimuli (Schon et al., 2004; Micheyl et al., 2006; Deguchi et al., 2012; Kuhnis et al., 2013) and better enhanced concurrent sound segregation (Zendel & Alain, 2009; Zendel & Alain, 2013). Musicians also exhibit better performance on verbal working memory, especially in the auditory modality (Hanna-Pladdy & Gajewski, 2012; Zuk et al., 2014; Mansens et al., 2018; D'Souza et al., 2018; Groussard et al., 2020). However, it is still unclear

whether the skills trained with musical activities benefit speech perception abilities. The expanded OPERA hypothesis (Patel, 2014) proposes that cross-domain plasticity between music and speech processing could occur when the following five conditions are satisfied: 1) musical training and speech solicit overlapping brain networks in sensory or cognitive processes, 2) musical training places higher demands on these networks, 3) the engagement of these networks by musical training is associated with strong positive emotions, 4) the engagement of these networks by musical training is frequently repeated and 5) the engagement of these networks by musical training engages focused attention.

The study of Parbery-Clark et al. (2009)—which showed that young adult musicians performed better than age-matched non-musicians on the HINT, a clinical SPiN task (see Suppl. Material 1)—opened the path for the numerous studies that have explored the potential benefits of musicianship on SPiN perception under various experimental conditions. A recent review of 20 published articles (Coffey, Mogilever, et al., 2017) on SPiN performance in musicians, while leaning towards a musician’s advantage on SPiN perception, underlined the heterogeneity of the results.

Multiple factors can explain this variability, including the characteristics of the participants, such as their age, their cognitive and education levels. Characteristics related to musical training could also influence SPiN performance: the age of onset, the number of years and the intensity of musical practice could have an influence on SPiN skills, with more experience leading to better SPiN performance. In addition to participants’ personal and music-related characteristics, heterogeneity could also be related to differences in SPiN tests, such as the type of task (e.g., repetition, discrimination, recognition), the response format (closed or open set), the nature of the speech targets (syllables, words, sentences), the use of spatial separation between the targets and the masker, the noise level (either positive signal-to-noise ratios (SNRs) or more challenging negative SNRs) and, importantly, the type of masking noise.

The masking properties of noises can be separated into two broad categories: energetic and informational properties (Freyman et al., 1999). Energetic masking competes with the target speech on a spectro-temporal level, reducing its audibility at the auditory periphery (Rennies et al., 2019). Typical energetic maskers (EM) used in SPiN tasks are speech-shaped

or white noises. In contrast, informational masking interferes with the target speech beyond the auditory periphery, due to perceptual similarities between the target and the masker. Informational masking impacts speech perception by affecting the segregation of the speech from the masker (object formation) and/or by drawing attention away from the target to the masker (object selection) (for a review, see Kidd et al., 2008; Shinn-Cunningham, 2013). Competing speech contains both energetic and informational masking components. The energetic masking is produced by the overlap of spectro-temporal energy between the target and the masker (Rennies et al., 2019). The informational masking of speech maskers derives from similarities between the target and masker talker's voice (e.g., same-sex speakers in both the target and the masker (Brungart, 2001)), as well as by the linguistic properties of the masker, such as the phonetic and semantic content (Hoen et al., 2007; Brouwer et al., 2012). Varying the number of talkers in a speech masker modulates both its informational and energetic masking properties. The informational power peaks with two-three talkers and then gradually decreases with the addition of talkers, whereas the energetic power increases proportionally with the number of talkers (Tun & Wingfield, 1999; Hoen et al., 2007; Rosen et al., 2013; Helfer & Freyman, 2014). Studies on non-musicians have shown that both energetic and speech maskers produce significant masking when presented collocated to the target speech (Taitelbaum-Swead & Fostick, 2016; Goossens et al., 2017). For speech maskers, it has been shown that task difficulty varies with the number of talkers. A one-talker masker represents the least challenging condition, whereas a 2–3 talker maker represents the most challenging condition (Freyman et al., 2004; Rosen et al., 2013; Helfer & Freyman, 2014). Spatially separating the target from the masker can alleviate the masking power of maskers in general, but the effect of spatial separation is particularly salient in speech maskers with high informational content (Freyman et al., 1999; Freyman et al., 2001; Freyman et al., 2004; Arbogast et al., 2005; Yost, 2017).

The effect of age on SPiN perception has been studied in non-musicians and has been shown to negatively impact performance under various experimental conditions. The negative impact of age is systematically shown in speech maskers, whereas SPiN performance in energetic maskers seems to be less affected by age (Tun & Wingfield, 1999; Rajan & Cainer, 2008; Taitelbaum-Swead & Fostick, 2016; Goossens et al., 2017; Buss et al., 2019). Compared to younger adults, older adults require a more positive SNR than

younger adults to reach the same performance (Desjardins & Doherty, 2013; Buss et al., 2019) and at the same SNR levels, older adults have worse performance than younger adults (Helfer & Freyman, 2014; Taitelbaum-Swead & Fostick, 2016).

The main objective of the present study was to provide a comprehensive, qualitative and quantitative analysis of the literature on SPiN performance in musicians compared to non-musicians through a systematic review of the literature and a meta-analysis. We aimed to evaluate the strength of the evidence for a musician advantage in SPiN performance as a function of different experimental conditions, namely speech in energetic maskers, speech in speech maskers, speech in spatially separated maskers and different SNR levels (<0 dB; 0 dB; > 0 dB). Because musical activities are both sensory and cognitive, we expected to find a musician advantage with both energetic and speech maskers. However, we expected the musician advantage to vary according to the difficulty of the SPiN task condition. We expected the musician advantage to be stronger in collocated conditions (VS in spatially separated conditions), with two-talker maskers and with negative SNR levels. The second objective was to evaluate whether the impact of musicianship changes as a function of the participants' age. We expected to find a musician advantage at all ages, but we expected the impact of musicianship to be more salient in older musicians especially in speech maskers, with the hypothesis of a potential cognitive reserve (Stern, 2002, 2009; Stern et al., 2020) mitigating SPiN perception decline.

3.4 Methods

3.4.1 Search strategy

A systematic literature review was conducted in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Liberati et al., 2009). The electronic literature search was performed using PubMed and PsycNet on May 9, 2020, to identify studies comparing the speech-in-noise performance of musicians compared to non-musicians. The following keywords in these specific Boolean combinations were used: (music*[Title] OR sing [Title] OR singing [Title] OR singer [Title] OR instrument*[Title]) AND (noise OR intelligibility OR comprehensibility OR clear) AND

(speech OR word OR signal OR sentence OR verbal) AND (perception OR comprehension OR intelligibility OR discrimination OR recognition OR understanding OR listening).

3.4.2 Study selection

3.4.2.1 Selection process

A preselection on the title and abstract of all retrieved articles was performed independently by two team members. The preselected articles were then independently assessed thoroughly on the full text by the same two reviewers. After full text assessment, the bibliographies of the selected articles were screened for additional articles, which were then independently assessed via the same preselection and selection processes by the two team members.

3.4.2.2 Selection criteria

To be included, studies had to be quantitative cross-sectional or longitudinal group studies published in French or English. There were no restrictions on publication dates. Studies had to include healthy male and female participants of any age, with both musician and non-musician participants. In cross-sectional studies, musicianship was defined as the active solo/group practice of one/more musical instrument(s) and/or of singing—both practised at any level (amateur, professional). In longitudinal training studies, “musicians” were those enrolled in a musical instrument and/or singing training programs (solo and/or group practice). Finally, studies had to report the behavioural performance of musicians and non-musicians in at least one SPiN task. SPiN tasks had to present speech targets of any linguistic level (syllables, words, sentences) with an energetic (speech-shaped noise, white noise) or a speech masker.

Single case, case series and multiple case studies were excluded. Studies with participants affected by pathologies of any kind, including clinically diagnosed hearing impairments and/or wearing hearing aids were also excluded.

Specifically for the meta-analysis, we included in the analyses only SPiN tasks which presented stimuli (i.e., targets and maskers) *binaurally*, either collocated or spatially separated. Furthermore, SPiN tasks had to present natural speech targets, meaning that compressed, cochlear-implant (CI) simulated, vocoded, monotonized, whispered and sung speech tasks were excluded. With regards to energetic maskers, only continuous energetic maskers were included, meaning that conditions with gated energetic maskers were excluded.

3.4.3 Study Designs and Risk of Bias Assessment

The quality of each study—focusing on SPiN tasks—was assessed with the QualSyst (Kmet et al., 2004), which consists of 14 questions and a scoring system. As shown in supplementary material 1a, the QualSyst being a general tool, subcriteria were added to several questions in order to adapt the assessment to our specific area of interest.

According to the QualSyst’s scoring system, each question was answered by “No” (0 point), “Partial” (1 point), “Yes” (2 points) or “N/A” and a summary score was calculated. Based on previous systematic reviews (Lee et al., 2008; Maharaj & Harding, 2016), the summary score of each study was classified as either strong (score of $> .8$), good (score $.7-.8$), adequate ($.5-.7$) or limited ($<.5$). Two of the authors (EM, PT) independently assessed the quality of each study and their agreement was measured with an intraclass correlation coefficient (ICC) calculated with a two-way mixed model (Koo & Li, 2016). Discrepancies between authors were solved by consensus.

3.4.4 Data extraction

Data extraction was done independently by one author (EM) and a research intern (JP) and verified by a second author (MJ). Discrepancies were solved by consensus. The extracted data are presented in supplementary materials 2 and 3. The following information was extracted for each study: group characteristics (sample size, mean age, sex distribution, education level and language), musicianship, detailed in supplementary material 2 (Table 2a

and 2b) (type of musical activity practised [vocals, instruments], level of practice [amateur, professional], mean age of practice onset, mean number of years of practice, mean hours of weekly practice), and SPiN tasks, detailed in supplementary material 3 (table 3) (task type and characteristics of the target and the masker, target to masker localization, presentation mode). The dependant variables were also extracted, including accuracy (% correct), speech reception threshold (SRT), signal-to-noise ratio (SNR) loss score, reaction time (RT) and variables based on signal detection theory (sensitivity, criterion). Group differences in speech performance and their direction were documented (supplementary material 3, table 3).

Finally, for the meta-analysis, we also extracted the mean and standard deviations for the performance of each group in the selected SPiN tasks.

3.4.5 Statistical analysis

The SPiN tasks of 32 cross-sectional studies were included in the meta-analysis. To test our hypotheses, separate meta-analyses were run on the following conditions:

- a) *speech in energetic maskers;*
- b) *speech in speech masker—one talker;*
- c) *speech in speech masker—two talkers;*
- d) *speech in speech masker—four talkers;*
- e) *speech in a spatially separated masker;*
- f) *speech in noise at fixed SNR levels—negative SNR;*
- g) *speech in noise at fixed SNR levels—null SNR;*
- h) *speech in noise at fixed SNR levels—positive SNR;*

Task conditions with collocated stimuli were used in all meta-analyses except for e). Due to the small number of studies, separate meta-analyses for speech and energetic maskers for the e)-h) conditions were not possible, hence both energetic and speech maskers were included in these analyses.

In several studies, there were multiple (non-independent) outcomes, which were dealt in the following ways. First, for studies reporting SPiN performances in conditions eligible for

different meta-analyses, for example one outcome with an energetic masker and another one with a two-talker speech masker, the issue was automatically resolved by separating them in the different meta-analyses (Borenstein et al., 2009). Second, when studies presented various outcomes eligible for the same meta-analysis, the most represented experimental condition in the meta-analysis was always preferred. For example, when a study presented outcomes with two different energetic maskers such as a white noise and a speech-shaped noise, the speech-shaped noise was always used in meta-analysis a); or when a study presented both the WIN and the QuickSIN—two tasks with a four-talker masker—the QuickSIN was always used in meta-analysis d). Third, when a study presented various outcomes eligible for the same meta-analysis but with an additional experimental manipulation, such as results with the same energetic masker but with variations in the stimuli's $F(0)$ or the SNR levels, an average score of these different results was calculated and used in meta-analysis a) (Borenstein et al., 2009).

For each meta-analysis, a random-effect model was used to take into account the variability in SPiN tasks and participant characteristics (Borenstein et al., 2009). The random effect meta-analyses were conducted using the R *Metafor* package (R 3.5.0; R Development Core Team) running on Mac OS 10.15.16. Due to the small number of studies per analysis, Hedge's g (Borenstein et al., 2009) was used for each analysis as a power estimate (bias-corrected standardized mean differences). Hedge's g of 0.2–0.5, 0.5–0.8, and ≥ 0.8 were considered as small, moderate, and high, respectively (Cohen, 1988). The Cochrane's Q test and I^2 statistics (Higgins et al., 2003) were used to assess the between-study heterogeneity. For the Cochrane's Q test, to indicate the presence of between-study heterogeneity, an alpha level of .05 was chosen for statistical significance. I^2 values of 25%—50%, 50%—75%, and $\geq 75\%$ were considered as low, moderate, and high between-study heterogeneity, respectively (Higgins et al., 2003). To test whether any one study was overly influential on the effect size estimate, we conducted a leave-one-out cross-validation in each meta-analysis (Viechtbauer, 2010). Finally, meta-regressions (Huizenga et al., 2011) were also conducted to assess the potential moderating impact of different factors on SPiN performance. First, the mean of the Ms and NMs mean ages (“mean age”) was computed and used as a continuous moderator in each analysis. Second, to evaluate the effect of study quality, the scores obtained on the QualSyst evaluation tool were used as a continuous moderator, “study

quality”, in each meta-analysis. The little number of available studies prevented the assessment of other potential moderators, such as the target’s linguistic level (syllable, word, sentences) or the masker type (energetic masker, speech masker) in meta-analyses c) and d).

The extracted data and analysis scripts will be publicly available on the Scholar Portal Dataverse upon publication.

3.5 Results

3.5.1 Study selection

Electronic literature searches identified a total of 579 studies matching the search terms: 322 in Pubmed and 257 in PsycNet. 455 records remained after the removal of duplicates. References not meeting the inclusion criteria on the screening of titles and abstracts were discarded (N = 394 records). All disagreements were solved by either consensus (23/28) or by involving the senior author (PT) (5/28). After evaluation of the remaining 61 articles on their full text, 17 articles that did not meet the inclusion criteria for the systematic review were discarded, and 44 articles were retained. All disagreements (3/3) were resolved by involving the senior author (PT), who evaluated the articles independently. A record of both the articles and the reason for the exclusion were kept, as well as the number of disagreements. Five additional records were obtained: 3 from the bibliography screening of all 44 previously selected articles and 2 additional articles were obtained directly from their authors. After assessment of their titles, abstracts and then full texts, all 5/5 articles were retained for the systematic review. In total, the systematic review included 49 studies, including 43 cross-sectional and 6 longitudinal studies, conducted between 2009 and 2020.

15 of the included cross-sectional studies reported the mean and standard deviation for each group on the SPiN task(s). The authors of the remaining 28 studies were contacted. The descriptive data were obtained for 20 additional articles. Thus, 32 articles were included in the quantitative analyses.

3.5.2 Study quality

The mean summary score over all studies was $.72 \pm .09$, corresponding to a good quality score (.7-.8). Specifically, 21 studies exhibited an adequate quality score (.5-.7), 18 studies exhibited a good quality score (.7-.8) and 10 studies exhibited a strong ($> .8$) quality score. The detailed results are presented in supplementary material 1, Table 1b.

The single measure ICC between both judges was high (.952, with a 95% confidence interval from .917 to .973, $F_{(48, 48)} = 40.048, p < .001$).

The most frequent issues were the following. Few studies specified in their introduction the age group of interest (item 1). The sample size for the groups in most studies was small, with 20 or fewer participants, and not justified using power calculation (item 9). Specifically, there were between 15 and 20 participants in 17 studies and fewer than 15 participants in 9 studies. Regarding the recruitment of participants, the selection criteria were generally well stated but the recruitment setting (e.g., place of recruitment such as a university or a music school or the general community), and especially the recruitment methods (e.g., emails, flyers) were lacking in most studies (item 3). While participants demographic information was generally reported, 9 studies did not provide any information about the musicians' musical practices, such as a mean number of years of practice and/or a mean age of onset. Musicians were generally matched on age with the control group, but were matched on hearing in only 11 studies, introducing potentially important confounds (item 12). The statistical analyses were in general well described and appropriate (item 10). The reporting of results was, however, less strong (item 13), due to several articles reporting the descriptive data of SPiN performances either graphically only (22 studies) or not at all (6 studies).

3.5.3 Focus(es) of the selected studies

In 46/49 of the articles, one of the main goals was to explore the effect of musical training on SPiN tasks. In 3/46 cross-sectional studies, performance in the SPiN tasks was used as predictor or comparator and was not the focus of the analyses. 38/46 cross-sectional studies reported the performance of musicians and non-musicians, and 5/46 studies consisted of correlation/regression studies.

3.5.4 Participant description

Across all studies, there was 2427 participants in total (47,48% female). The language of the participants was mentioned in 28 studies. The majority (15) of the studies involved North American English speakers. The other languages that were reported are British English (1 study), Australian English (2 studies), French from Québec (3 studies), Dutch (3 studies), Kannada (1 study), Brazilian Portuguese (1 study), Danish (1 study) and Hebrew (1 study).

Cross-sectional studies on participants of all ages were found. Namely, there were 3 studies on children (7–17 years old), 33 studies on young adults (18–40 years old), 2 studies on middle-aged adults (45–65 years) and 1 study on older adults (≥ 60 years old). 4 studies employed groups with a broad age range (18–91, 30–60, 22–59, 20–87 years old). Among the longitudinal studies, 3 involved children aged between 3 and 9 years, while the 3 other studies focused on older adults (≥ 60 years old).

Information on the participants' education was inconsistently available in 21 studies, with musicians and non-musicians being matched on education in only 8/21 studies. Furthermore, the available information revealed that in most studies, a majority of highly educated participants (college students) were recruited.

The hearing status of participants was measured in 33/49 studies. In 13 of those studies, it was reported that M and NM were matched on hearing status. In 18 of those studies, it was reported that participants all had under a given auditory threshold in the tested frequencies (i.e., 15 dB [1 study], 20 dB [10 studies], 25 dB [5 studies] and 35 dB [1 study]), but direct group comparison was reported. In 2/33 studies, no information was provided regarding the hearing status of participants.

The cognitive functioning of participants was verified in only 23/49 studies. The scores at the Montreal Cognitive Assessment were measured in 1 study, for which participants were matched. Specific executive functions (working memory, attention) were measured in 8 other studies, with participants being matched in 6/8 studies. In the other 2 studies, the cognitive scores were used as a factor in linear regression analyses. Participants' IQ was measured in 14 studies; participants were matched across groups in 13 of those

studies. A significant difference in IQ was found in one study, which was then controlled for in the analyses.

3.5.5 Musicianship

Among the cross-sectional studies, two studies involved a single group of participants, with musicianship measured in terms of the number of years of musical training (continuous quantitative variable). For the other 41 cross-sectional studies, the musicianship factor was defined by specific inclusion criteria to differentiate a group of musicians (M) from a group of non-musicians (NM). Across all cross-sectional studies, there were 996 M and 953 NM. Across all longitudinal studies, there were 214 M and 176 NM. Of the 176 NM in longitudinal studies, 31 were enrolled in a non-musical training group and 145 were enrolled in a no-activity group. The average number of participants per cross-sectional study was relatively small with 21.59 participants (SD = 11.51; range: 10–74) in the M group and 22.31 (SD = 13.84; range: 8–89) in the NM groups. The average number of participants per longitudinal study was slightly higher, with 30.57 participants (SD = 27.23; range: 13–90) in the M group and 25.14 (SD=29.43; range: 8–90) in the NM groups.

The inclusion criteria reported for the musician groups were either a maximum onset age for musical practice (26 studies, range: 5–16 years old), a minimum number of years of practice (29 studies, range: 2–10) and a minimum weekly practice (10 studies), either in terms of the number of practice sessions per week (3x/week) or in terms of the number of hours of practice per week (range: 5–15 hours/week). In 17 studies, both onset age and minimum number of years of practice were used as inclusion criteria. In 3 studies, a maximum onset age and a minimal amount of training per week were used as inclusion criteria. In 2 studies, a minimum number of years of practice and a minimal amount of training per week were used as inclusion criteria. Finally, only 2 studies combined all three inclusion criteria (i.e., age of onset, number of sessions per week and session duration).

Among the cross-sectional studies that provided information about the participants' type of musical activities (32/43), 13 involved instrumentalists (any instrument) and 17 involved both instrumentalists (any instrument) and vocalists. In four studies, the M group

was homogeneous: there was one study on choir singers, one study on singers and percussionists, one study on keyboard players (piano or organ) and one study on percussionists. Four studies focused on amateur musicians, 6 on professionals and 8 on both (participants enrolled in a university music program were considered as professionals). Most studies did not provide this information (31/43). The average number of musical training years was available for the musician group in 28 studies, with an average number of 6.85 ± 1.49 years of musical training for children, 15.46 ± 2.7 years for young adults, 49.68 ± 0.45 years for middle-aged adults and 39.7 ± 12.03 years for one study with a broad age group. The age of onset of musical training was available in 27/43 studies with an average starting age of 4.05 ± 2.9 years old for children, 6.43 ± 1.26 years old for young adults, 6.06 ± 0.65 years old for middle-aged adults and 7.2 ± 1.9 years old for one study with a broad age group. This information was not available for older adults. The average hours of weekly practice for the musician group were available in only 4 studies: 1.82 ± 1.02 hours for one study on children, 19.85 ± 5.01 hours for studies on young adults and 24.65 ± 7.12 hours for one study on a broad age group.

All longitudinal studies compared a musical training against another kind of training, a do-nothing control group or both. The different trainings lasted from 10 weeks to 3 years, with a total weekly training duration of 30 minutes up to 5 hours.

3.5.6 SPiN tasks

70 SPiN tasks were used across the 49 studies; the QuickSIN (16 studies) (Killion et al., 2004), the Hearing in Noise Test (HINT, 12 studies) (Nilsson et al., 1994), a modified (computerized) version of the HINT (2 studies) (Coffey, Chepesiuk, et al., 2017), the Listening in spatialized noise sentence test (LiSN-S T, 3 studies) (Cameron & Dillon, 2007), the PRESTO (1 study) (Gilbert et al., 2013), a modified SPiN discrimination subtest from the Auditory Skills Assessment (ASA; Geffner & Goldman, 2010), the Words-in-noise test (WIN, 4 studies) (Wilson, 2003; Wilson et al., 2007), the Revised Speech-in-noise (R-SPIN, 2 studies) and the AB-word test (Hebrew version, n=1) (Boothroyd, 1968; Fostick et al., 2014) and 30 experimental tasks.

3.5.6.1 Targets: characteristics and experimental manipulations

Sentences were the most common targets in the selected studies, being used in 56/70 tasks. Words were used in 11/70 tasks, non-words in 1/70 tasks and syllables in 2/70 tasks. Targets were acoustically modified in 12/70 experimental tasks (7 studies), including variation of the fundamental frequency (F0), whispered speech, monotone speech, sung speech, CI-simulated speech and vocoded speech.

3.5.6.2 Maskers: characteristics and manipulations

The most frequent masker type was a speech masker, which was used in 47/70 tasks (42 studies): four-talker babble was the most represented speech masker (23 tasks), followed by one-talker babble (13 tasks), two-talker babble (8 tasks), six-talker babble (3 tasks), twelve-talker babble (2 tasks), eight-talker babble (1 task), multi-talker babbles not otherwise specified (2 tasks), and time-reversed speech (1 task).

Energetic maskers were used in 30/70 tasks (26 studies). Speech shaped noises were the most frequent (22 tasks) and the other energetic maskers found were white noises (2 tasks), buzzes (4 tasks), Gaussian noise (1 task) and time-reversed speech (1 task).

The masker's properties were modified in 13/70 experimental tasks. Of these, 9/13 tasks used both speech and non-speech maskers. In 2/13 tasks the speech masker's amount of informational content was varied by increasing the number of talkers or by playing the one-talker babble in reverse. In 2/13 tasks, the vocal properties of the masker's talker were modified by varying the fundamental frequency (F0) of the talker(s) (1 task) or by varying both the F0 and the vocal tract length (VTL) of the talker (1 task). In 1/13 task, the masking power of the non-speech masker was varied by using a continuous and a gated noise.

3.5.6.3 Stimulus presentation

The most common task presentation was via headphones (35/70 tasks). Loudspeakers were used in 17/70 tasks and insert earphones in 15/70 tasks. The presentation was not specified for 3 tasks

3.5.6.4 SNR Manipulation

A SNR manipulation was present in 61/70 tasks. In 31 tasks, the SNR was varied along predefined SNR levels and in 30 tasks, the SNR varied in an adaptive procedure, according to the participants' performance. In 9 tasks, the stimuli were presented at a single SNR level.

3.5.6.5 Spatial separation between stimuli

A spatial separation between targets and maskers ($^{\circ}$ azimuth) was present in only 10/70 tasks. Of these, there were 5 bilateral masker presentations ($\pm 90^{\circ}$, $\pm 15^{\circ}$ azimuth), 4 unilateral masker presentations ($+90^{\circ}$, -90°) and one presentation via a circular array of loudspeakers with azimuths between 20° and 180° .

3.5.7 SPiN performance

3.5.7.1 Speech in energetic masker

21 studies comprising 26 tasks included a speech in non-speech masker condition. In the 2 studies involving children (2 tasks), the M did not outperform the NM. In the 16 studies involving adults, M outperformed NM in 10/21 tasks (10/16 studies). In the 2 studies involving middle-aged adults, M outperformed NM in 2/2 tasks. In the one study involving older adults, M outperformed NM in 1/1 task. Overall, M outperformed NM in 13/21 studies (13/26 tasks).

Meta-analysis: There were 17 studies comprising 18 tasks in this meta-analysis (Table 3). Note that the study of Deroche et al. (2017) reported the outcome for two different (independent) sets of participants, labelled as “exp.1” and “exp.4”. As shown in Figure 1, musicianship had a moderate but significant effect on SPiN perception (Hedge's $g = .61$; $p < .0001$). The confidence interval was narrow and did not include 0. The between-study heterogeneity across studies was moderate (Q 's $p = .0002$; $I^2 = 57.78\%$). Cross-validation indicated that the effect of musicianship was moderate and significant in each iteration of the analysis (all $p < .0001$; Hedge's g between .56 and .66). The moderator analysis (Table 4)

showed a significant effect of mean age ($p = .027$). The effect of study quality was not significant ($p = .90$).

Table 3. Tasks and descriptive data used in the speech in energetic masker analysis				
Article	Tasks	Score	Musicians M (SD)	Non-Musicians M (SD)
Baskent et al., 2018	Words in speech-shaped noise — mean score of all SNR conditions (0 dB, 5 dB, 10 dB)	% correct	84.44 (6.44)	81.82 (8.27)
Anaya et al., 2016	HINT at -3 dB	% correct	50.74 (8.59)	47.22 (8.11)
Puschmann et al., 2019	HINT	SRT	-3.02 (1.06)	-2.24 (0.95)
Madsen et al., 2017	Sentences in Gaussian noise	% correct	73.53 (9.59)	72.91 (6.84)
Du & Zatorre, 2017	Syllables in white noise — mean score of SNR conditions (-8 dB, -4 dB, 0 dB, 8 dB)	% correct	76.73 (3.57)	69 (6.03)
Escobar et al., 2019	HINT	SRT	-0.57 (0.95)	-0.69 (0.93)
Ruggles et al., 2014	HINT	SRT	-2.34 (0.59)	-2.21 (0.58)
Deroche et al., 2017 (exp.1)	Sentences in buzz – mean score for buzz and diotic conditions	SRT	-8.67 (0.50)	-7.96 (1.44)
Deroche et al., 2017 (exp.4)	Sentences in buzz – mean score for buzz and fixed masker F0 conditions	SRT	-8.95 (1.57)	-7.52 (1.70)
Madsen et al., 2019	Sentences in Gaussian noise (anechoic condition)	SRT	-0.52 (1.02)	-0.58 (0.82)
Fuller et al., 2014	Sentences in steady speech shaped noise	SRT	-6.25 (1.20)	-5.77 (1.38)
Varnet et al., 2015	Syllables in white noise	Mean SNR threshold	-13.37 (1.19)	-11.91 (1.01)
Yoo & Bidelman, 2019	HINT	SRT	-6.5 (2.00)	-6.13 (2.07)
Coffey et al., 2019	modified HINT	% correct	-0.66 (0.06)	-0.58 (0.08)
Boebinger et al., 2015	Sentences in steady speech shaped noise	SRT	-3 (1.00)	-2.7 (0.90)
Parbery-Clark et al., 2012b	HINT	SRT	-3.16 (0.61)	-2.34 (0.63)
Parbery-Clark et al., 2011b	HINT	SRT	-3.37 (0.52)	-2.24 (0.87)
Fostick, 2019	AB word test: words in speech shaped noise	% correct	72 (15)	55 (15)

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

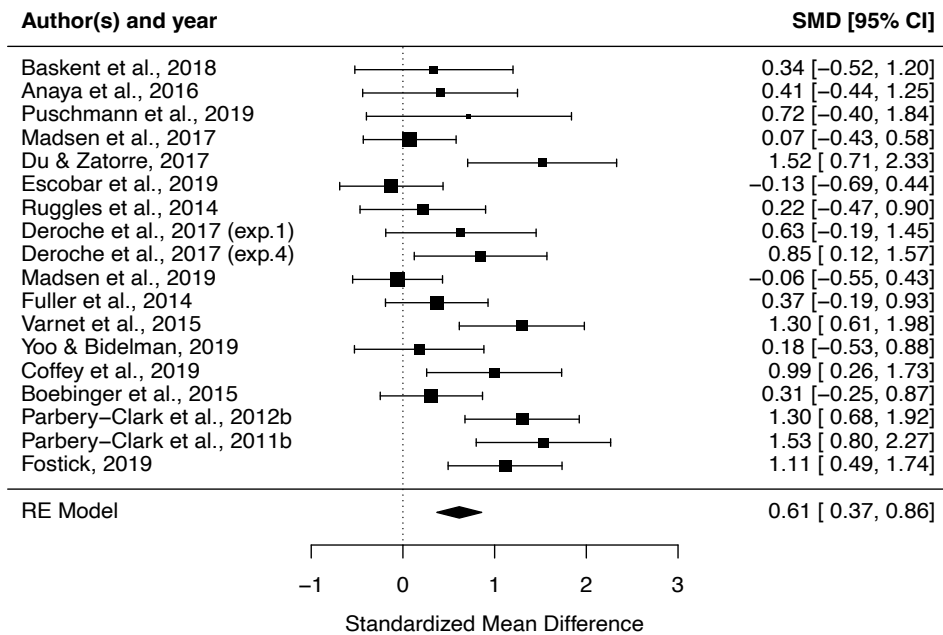


Figure 1. Effect sizes for the speech in energetic masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI= 95% confidence interval.

Table 4. Moderator analysis for the speech in energetic masker condition							
Moderator	β	β 's <i>p-value</i>	β 's $CI_{95\%}$		SE	Q's <i>p-value</i>	I^2 (%)
			LL	UL			
Mean age	0.0216	0.0041	0.0069	0.0364	0.2427	0.0182	42.05
Study quality	0.2366	0.1257	-3.4514	3.9245	1.8817	0.0001	60.45

Note. β = regression coefficient, $CI_{95\%}$ =95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity.

3.5.7.2 Speech in one-talker maskers

There were 8 studies comprising 8 experimental tasks in this analysis. Of those, 1 study focused on children, 6 on younger adults and 1 involved a broad age group. In the only study on children (1 task) and on the broad age group (1 task), no musician advantage was

found. In the younger adult group, M outperformed NM in only 1/6 studies (1 task), Overall, M outperformed NM in only 1/8 studies (1/8 task).

Meta-analysis: There were 8 studies in this analysis (Table 5). As shown in Figure 2, the effect of musicianship was small and not significant (*Hedge's g* = .17; *p* = .569; *Q's p-value* = .005; *I*² = 82.36%). Cross-validation indicated that the effect of musicianship became significant but remained small when removing the study of Puschmann et al. (2019) (*Hedge's g* = .34; *p* = .044), which showed a disadvantage for M compared to NM. The effect of musicianship remained nonsignificant in every other iteration (all *p-values* > .573). The moderator analysis indicated that the effect of both mean age and study quality were nonsignificant (Table 6).

Article	Task	Score	Musicians M (SD)	Non- Musicians M (SD)
Baskent et al., 2018	Sentences in sentences	SRT	-14.42 (1.54)	-12.00 (3.63)
Puschmann et al., 2019	Sentences in sentences	SRT	6.77 (0.95)	4.96 (1.25)
Baskent & Gaudrain, 2016	Sentences in sentences —composite score of all F0 and VTL modification	% correct	0.81 (0.05)	0.70 (0.12)
Clayton et al., 2016	Sentences in sentences	SRT	3.8 (2.24)	4.4 (2.05)
Madsen et al., 2019	Sentences in sentences	SRT	2.03 (1.49)	2.22 (1.08)
Boebinger et al., 2015	Sentences in clear speech	SRT	-11.9 (2.40)	-11.5 (3.20)
Mandikal-Vasuki et al., 2016	LiSN-S test: natural speech (talker1) in natural speech (talker 2)	SRT	-7.8 (3.50)	-7.9 (3.10)
Meha-Bettison et al., 2018	LiSN-S test: natural speech (talker1) in natural speech (talker 2)	SRT	11.9 (3.47)	11.64 (2.81)

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

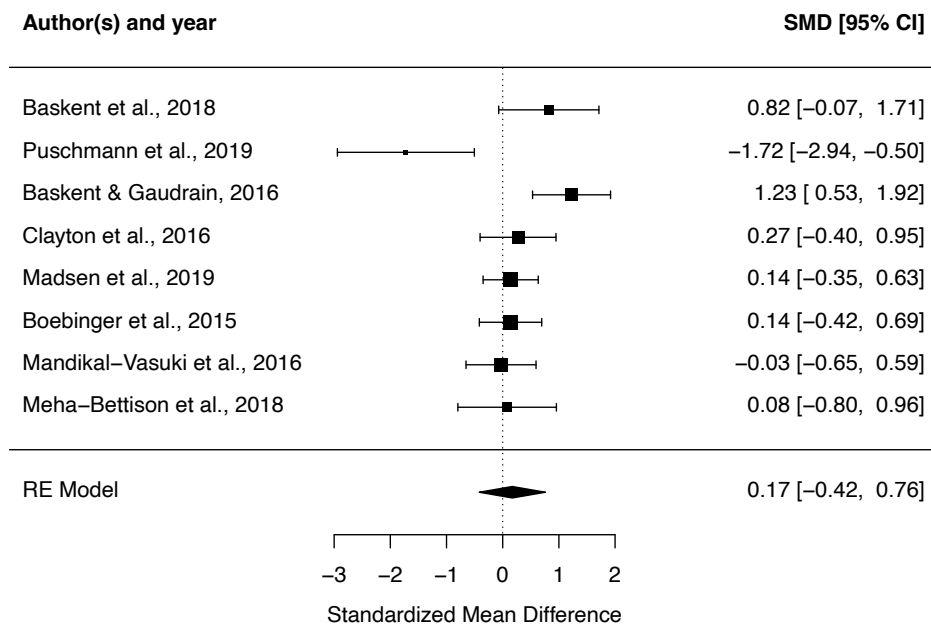


Figure 2. Effect sizes for the speech in one-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p</i> -value	β 's CI _{95%}		SE	Q's <i>p</i> -value	I ² (%)
			LL	UL			
Mean age	-.0111	.7855	-.0907	.0686	.0406	.7855	87.60
Study quality	5.0040	0.2879	-4.2253	14.2333	4.7089	.0027	79.54

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity.

3.5.7.3 Speech in two-talker maskers

There were 4 studies comprising 7 experimental tasks in this analysis. All 4 studies involved younger adults. M outperformed NM in 2 studies (3/7 tasks).

Meta-analysis: There were 4 studies in this meta-analysis (Table 7). As shown in Figure 3, the effect of musicianship on this analysis was large and significant (*Hedge's* = .79;

$p = .0007$). The confidence interval was broad but did not include 0. Between study heterogeneity was not significant (Q 's $p = .122$; $I^2 = 45.75\%$). Cross-validation indicated that the effect of musicianship was significant in every iteration (*Hedge's g* between .56 and .98; all p -values $< .013$). The moderating effect of mean age and study quality were not significant (all $p > .437$) (Table 8).

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Morse-Fortier et al., 2017	CVC words in nonsense sentences	SNR threshold	-7.75 (5.19)	-4.56 (3.04)
Deroche et al., 2017	Exp.1: sentences in sentences— mean score for diotic presentation	SRT	-2.76 (1.14)	-1.81 (1.20)
Deroche et al., 2017	Exp.4: sentences in sentences— mean score for fixed masker presentation	SRT	-2.26 (1.06)	0.02 (1.79)
Madsen et al., 2019	Sentences in sentences	SRT	2.41 (0.94)	2.83 (1.27)

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

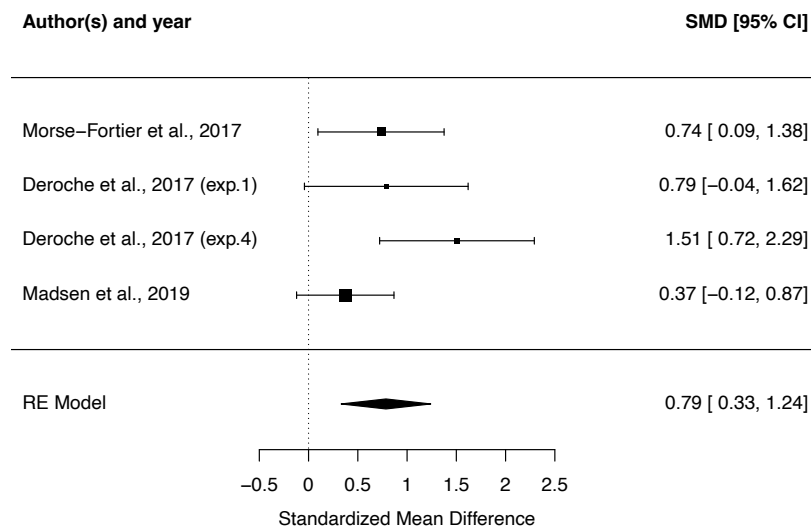


Figure 3. Effect sizes for the speech in two-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (*Hedge's G*), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p-value</i>	β 's CI _{95%}		SE	Q's <i>p-value</i>	I ² (%)
			LL	UL			
Mean age	.1509	.4374	-.4337	.7354	.2982	.6130	58.72
Study quality	-2.2030	.6112	-7.7629	3.3569	2.8367	.4374	46.91

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity.

3.5.7.4 Speech in four-talker maskers

There were 19 tasks in this analysis from 16 studies; 1 study focused on children (1 task), 13 studies on younger adults (15 tasks), 1 study on middle-aged adults (2 tasks) and 1 study on a broad age group (1 task). In the studies on children and the broad age groups, M did not outperform NM. In studies on younger adults, M outperformed NM in 8/15 tasks (8 studies). Overall, M outperformed NM in 8/16 studies (8/19 tasks).

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Escobar et al., 2019	QuickSIN	SNR loss	0.19 (0.85)	0.15 (0.90)
Ruggles et al., 2014	QuickSIN	SNR loss	0.8 (1.03)	1.19 (0.72)
Skoe et al., 2019	QuickSIN	SNR loss	0.61 (0.75)	0.86 (0.58)
Zendel et al., 2015	Words in four talker babble—composite score for 0 and 15 dB	% correct	91.67 (3.34)	88.72 (3.42)
Zhang et al., 2019	QuickSIN	SNR loss	1.48 (1.42)	1.35 (1.01)
Bidelman & Yoo, 2020	QuickSIN	SNR loss	-1 (1.29)	-0.21 (1.14)
Yoo & Bidelman, 2019	QuickSIN	SNR loss	-1.56 (1.81)	0.7 (1.32)
Slater et al., 2018	QuickSIN	SNR loss	-1.04 (0.70)	-0.36 (1.00)
Jain & Nataraja, 2019	QuickSIN	SNR loss	-7.55 (1.75)	-6.58 (1.90)
Perron et al., 2021	CVC syllables in four talker babble composite score for -3 and +3 dB	Sensitivity	2 (0.58)	1.92 (0.50)

Zendel & Alain, 2012	QuickSIN	SNR loss	0.84 (1.57)	1.65 (2.50)
Parbery-Clark et al., 2011b	QuickSIN	SNR loss	0.22 (0.39)	-0.51 (0.38)

Meta-analysis: There were 12 studies in this meta-analysis (Table 9). As shown in figure 4, the effect of musicianship was moderate and significant (*Hedge's g* = .56; *p* = .0003). The confidence interval was broad but did not include 0. Between study heterogeneity was significant (*Q's p* = .005; *I*² = 68.54%). Cross-validation indicated that the effect of musicianship was moderate in each iteration (*Hedge's g* between .49 and .62; *p* < .001). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all *p* > .592) (Table 10).

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

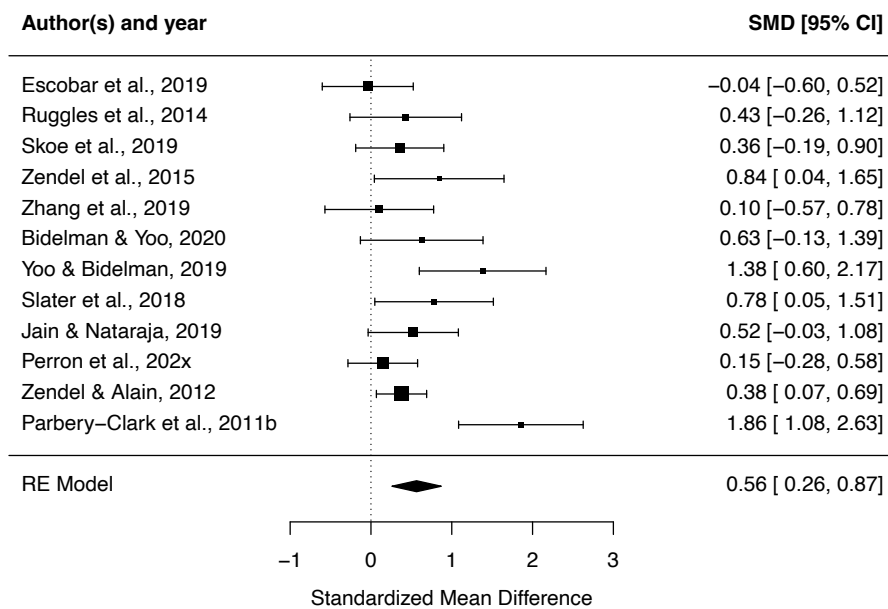


Figure 4. Effect sizes for the speech in four-talker masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (*Hedge's G*), 95% CI= 95% confidence interval.

Moderator	β	β 's	β 's CI _{95%}	SE	Q's	I ² (%)
-----------	---------	------------	------------------------------	----	-----	--------------------

		<i>p-value</i>	LL	UL		<i>p-value</i>	
Mean age	.0078	.5920	-.0207	.0362	.0145	.0011	75.09
Study quality	-.2684	.8980	-4.3710	3.8341	2.0932	.8980	71.76

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity

3.5.7.5 Speech in spatially separated noise

There were 8 studies in this analysis comprising 9 tasks, 1 study on children (1 task), 6 studies on younger adults (7 tasks) and 1 study on a sample with a broad age range (1 task). In the study on children and the study with the broad age sample, M did not outperform NM. For the younger adults, M outperformed NM in 4/6 tasks (4 studies). Overall, M outperformed NM in 4/8 studies (4/9 tasks).

Meta-analysis: There were 5 studies in this analysis (Table 11). As seen on Figure 5, the effect of musicianship was nonsignificant (*Hedge's g* = .25; *p* = .538). The confidence interval included 0. Between study heterogeneity was nonsignificant (*Q's p* = .342; *I*² = 1.06%). Cross-validation indicated that the effect of musicianship was significant when removing the study of Yeend et al. (2017) (*Hedge's g* = .39; *p* = .014) and was nonsignificant in all other iterations (all *p-values* > .097). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all *p-values* > .118) (Table 12).

Article	Task	Score	Musicians M (SD)	Non- Musicians M (SD)
Morse-Fortier et al., 2017	CVC words in sentences (2-talker babble), +60° azimuth	SNR threshold	-25,18 (1.65)	-24,56 (1.70)
Clayton et al., 2016	Sentences in sentences (1-talker-babble), ± 15° azimuth	SRT	-11,7 (5.14)	-6,6 (6.34)
Madsen et al., 2019 (1)	Sentences in sentences (1-talker-babble), ±15° azimuth	SRT	-10,27 (3.55)	-9,33 (3.89)
Meha-Bettison et al., 2018	LiSN-S test: sentences (talker1) in natural speech (talker 2), ±90° azimuth	SRT	-18,6 (2.6)	-18,1 (2.93)

Yeend et al., 2017	LiSN-S test: sentences (talker1) in natural speech (talker 2), $\pm 90^\circ$ azimuth	SRT	-19,59 (2.31)	-19,65 (2.74)
--------------------	---	-----	---------------	---------------

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

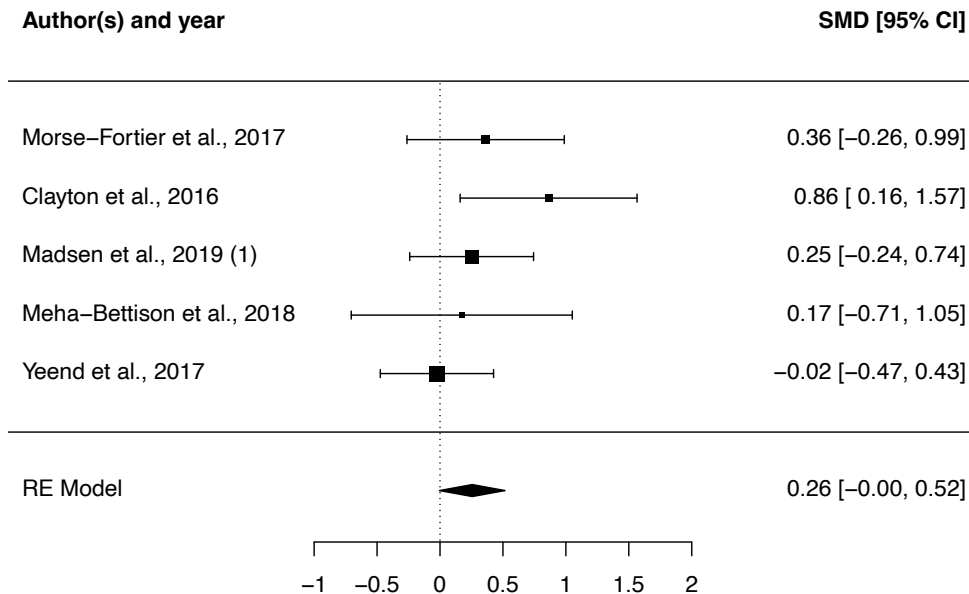


Figure 5. Effect sizes for the speech in spatially separated masker condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p-value</i>	β 's CI _{95%}		SE	Q's <i>p-value</i>	I ² (%)
			LL	UL			
Mean age	-.0182	.1184	-.0411	.0047	.0117	.1184	0.00
Study quality	.9117	.5421	-2.0194	3.8427	1.4955	.5421	18.08

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity

3.5.7.6 Speech in noise, SNR < 0 dB

There were 5 studies comprising 5 experimental tasks in this analysis, all including younger adults. Overall, M outperformed NM in 2 studies (2/5 tasks).

Meta-analysis. There were 6 studies in this meta-analysis (Table 13). As shown in Figure 6, the effect of musicianship was moderate and significant (*Hedge's g* = .56; *p* = .027). The confidence interval was broad but did not include 0. Between study heterogeneity was moderate (*Q*'*p-value*=.005; *I*² = 72.66%). Cross-validation indicated that the effect size was significant in 3/6 iterations (*Hedge's g* between .64-.67; all *p-values* <.029) but was non-significant in the other 3/6 iterations (all *p-values*>.052). Moderators indicated that the effect of mean age and study quality were nonsignificant (all *p-values*>.479) (see table 14).

Article	Task	Score	Musicians M (SD)	Non- Musicians M (SD)
Madsen et al., 2017	Sentences in Gaussian noise, SNR = -3 dB	SRT	73.53 (9.59)	72.91 (6.84)
Baskent & Gaudrain, 2016	Sentences in sentences, SNR = -6 dB	% correct	0.82 (0.05)	0.70 (0.12)
Anaya et al., 2016	HINT, -3 dB	% correct	50.74 (8.59)	47.22 (8.11)
Perron et al., 2021	Syllables in four-talker babble, SNR = -3 dB	Sensitivity	1.26 (0.55)	1.17 (0.56)
Ruggles et al., 2014	Sentences in continuous noise, composite score for -3 and -6 dB	% correct	64.84 (5.82)	63.30 (10.20)
Du & Zatorre, 2017	CV syllables in white noise, composite score for -4, -8 and -12 dB	% correct	75.21 (6.35)	65.04 (7.37)

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

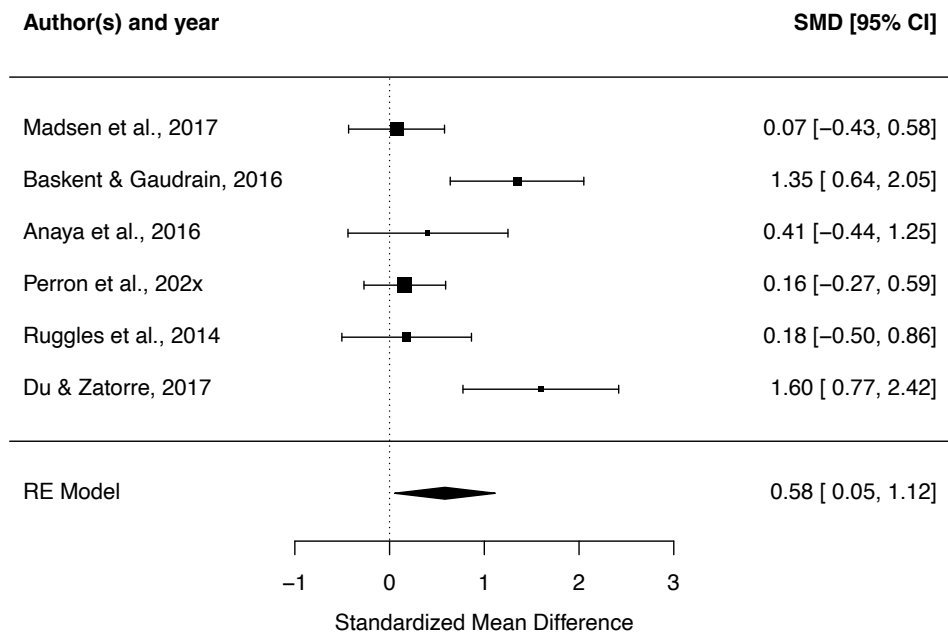


Figure 6. Effect sizes for the speech in SNR < 0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study's weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge's G), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p-value</i>	β 's CI _{95%}		SE	Q's <i>p-value</i>	I ² (%)
			LL	UL			
Mean age	-0.0151	.4878	-0.0579	.0276	.0176	.0218	74.48
Study quality	2.0720	.6613	-7.1961	11.3401	4.7287	.6613	79.86

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity

3.5.7.7 Speech in noise, SNR = 0 dB

There were 6 studies comprising 6 tasks in this condition, 1 study on children, 4 studies on younger adults and 1 study on older adults. M did not outperform NM in the children's group. M outperformed NM in 2/4 tasks in 2/4 studies on younger adults. M

outperformed NM in 1/1 tasks (1 study) on older adults. Overall, M outperformed NM in 3/6 studies (3/6 tasks).

Meta-analysis: There were 7 studies in this analysis (Table 15). As shown in Figure 7, the effect of musicianship was moderate but significant ($G = .56$; $p = .00001$). The confidence interval was narrow and did not include 0. Between study heterogeneity was nonsignificant (Q 's $p = .2430$; $I^2 = 9.77\%$). Cross-validation indicated that the effect was moderate in 6/7 iterations (G between .50 and .66; all p -values $< .00014$) and was small in 1/7 iterations ($G = .42$; $p = .004$). Moderator analysis indicated that the effect of mean age was significant ($p = .041$), but the effect of study quality was non-significant ($p = .94$).

Article	Task	Score	Musicians M (SD)	Non- Musicians M (SD)
Baskent et al., 2018	Words in speech shaped noise	% correct	73.33(12.30)	71.21(13.62)
Anaya et al., 2016	PRESTO	% correct	68.13(6.11)	63.5(7.48)
Du & Zatorre, 2017	CV syllables in white noise	% correct	92.33(2.91)	88(7.56)
Ruggles et al., 2014	Sentences in continuous noise	% correct	86.97(6.01)	85.17(11.73)
Fuller et al., 2014	Words in speech shaped noise	% correct	89.55(6.72)	88(9.14)
Zendel et al., 2015	Words in multitalker babble noise	% correct	85.28(5.51)	79.54(6.10)
Fostick et al., 2019	Words in speech shaped noise	% correct	0.72(0.15)	0.55(0.15)

Note. M= mean, SD = standard deviation, SRT = speech reception threshold

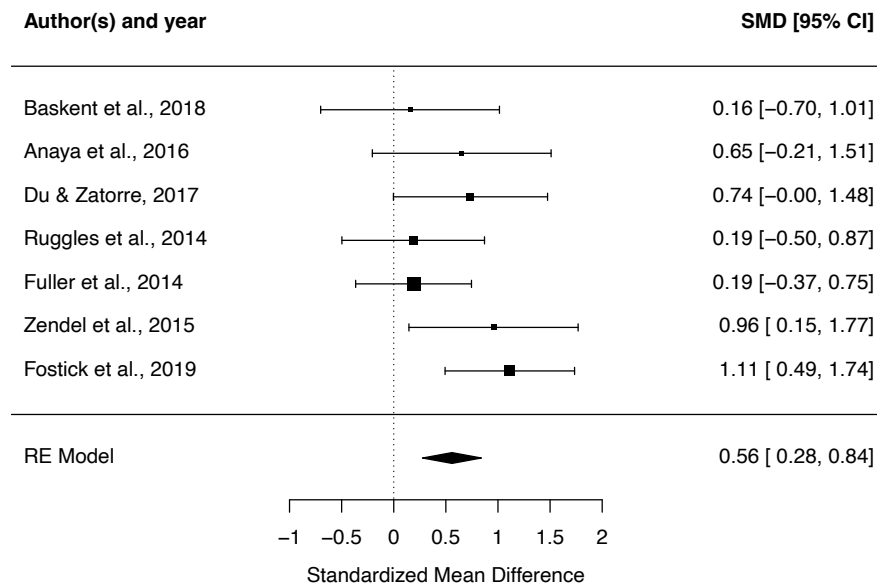


Figure 7. Effect sizes for the speech in SNR = 0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p-value</i>	β 's $CI_{95\%}$		SE	Q's <i>p-value</i>	I^2 (%)
			LL	UL			
Mean age	.0158	.0415	.0006	.0310	.0078	.5819	0.00
Study quality	.1826	.9367	-4.3266	4.6919	2.3007	.1610	25.62

Note. β = regression coefficient, $CI_{95\%}$ =95% confidence interval, LL = lower limit of the $CI_{95\%}$, UL = upper limit of the $CI_{95\%}$, SE = standard error, I^2 = residual heterogeneity

3.5.7.8 Speech in noise, SNR > 0 dB

There were four studies comprising 4 experimental tasks in this analysis, one focusing on children and three on younger adults. For the children, M did not outperform NM. For the younger adults, M outperformed NM in 2/3 tasks in the young adult group, namely one task of syllables in white noise and one task of words in speech-shaped noise. Overall, M outperformed NM in 2/4 tasks.

Meta-analysis: There were 5 studies in this meta-analysis (Table 17). As seen in Figure 8, the effect of musicianship was not significant ($G = .32$; $p = .084$). The confidence interval included 0. Between study heterogeneity was low ($Q, p = .237$; $I^2 = 37.41\%$). Cross-validation indicated that the effect of musicianship was nonsignificant in every iteration (all p -values $> .073$). Moderator analysis indicated that the effect of mean age and study quality were nonsignificant (all p -values $> .19$) (Table 18).

Article	Task	Score	Musicians M (SD)	Non-Musicians M (SD)
Baskent et al., 2018	Words in speech shaped noise, composite score for +5 and +10 dB SNR	% correct	90(5.96)	87.12(9.58)
Du & Zatorre, 2017	CV syllables in white noise, SNR = +8 dB	% correct	97.33(2.31)	91.5(6.72)
Fuller et al., 2014	Words in speech shaped noise, composite score for +5 and +10 dB SNR	% correct	96.10(2.54)	95.88(2.67)
Zendel et al., 2015	Words in multitalker babble, SNR = +15 dB	% correct	98.05(1.35)	97.90(1.21)
Perron et al., 2021	CVC syllables in four-talker babble, SNR = +3dB	Sensitivity	2.75(0.69)	2.63(0.58)

Note. M= mean, SD = standard deviation.

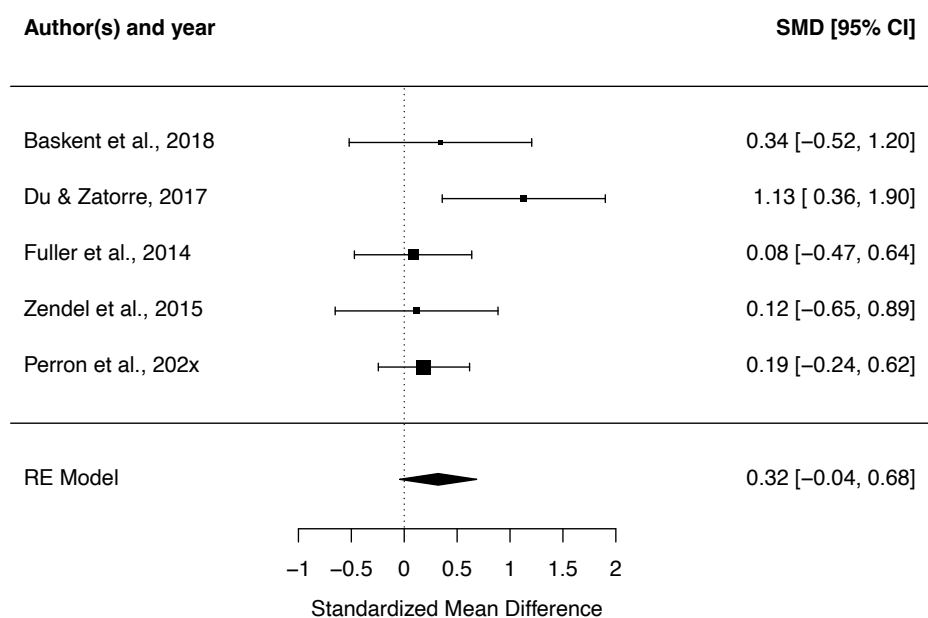


Figure 8. Effect sizes for the speech in SNR > 0 dB condition. Each box represents the effect estimate of an individual study and its size is proportional to that study’s weight on the pooled result. The horizontal lines through the boxes represent the 95% confidence interval. The diamond represents the pooled result from the random effect meta-analysis. RE Model = Random effect model, SMD = Standardized mean difference (Hedge’s G), 95% CI= 95% confidence interval.

Moderator	β	β 's <i>p-value</i>	β 's CI _{95%}		SE	Q's <i>p-value</i>	I ² (%)
			LL	UL			
Mean age	-.0058	.6556	.0314	.0197	.0130	.1614	42.76
Study quality	2.4628	.1871	-1.1961	6.1217	1.8668	.2856	0.00

Note. β = regression coefficient, CI_{95%}=95% confidence interval, LL = lower limit of the CI_{95%}, UL = upper limit of the CI_{95%}, SE = standard error, I²= residual heterogeneity

A summary of all analyses is presented in Table 19.

TABLE 19. SUMMARY STATISTICS FOR ALL META-ANALYSES

Experimental condition	N	G	G's <i>p-value</i>	SE	G's CI _{95%}		Q's <i>p-value</i>	I ² (%)	I ² 's CI _{95%}	
					LL	UL			LL	UL
Energetic masker	17	.6145	<.0001	.1249	.3697	.8593	.0002	57.78	31.9200	82.4727
1-talker masker	8	.1713	.5696	.3012	-.4190	.7615	.0046	82.38	31.5652	95.7093
2-talker masker	4	.7855	.0007	.2320	.3308	1.2402	.1228	45.75	0.0000	96.2904
4-talker masker	12	.5645	.0003	.1558	.2592	.8698	.0058	68.54	21.3576	89.4320
Spatially separated noise	5	.2562	.0538	.2660	-.0042	.5165	.3423	1.06	0.0000	89.7405
SNR < 0 db	6	.5845	.0308	.2707	.0539	1.1150	.0028	76.17	32.0310	96.1697
SNR = 0 db	7	.5586	.0001	.1442	.2760	.8411	.2430	9.77	0.0000	82.7345
SNR > 0 db	5	.3208	.0841	.1857	-.0432	.6848	.2375	37.41	-0.0432	0.6848

3.6 Discussion

The present study is the first systematic review and meta-analysis on the effect of musicianship on SPiN performance. Our systematic review (49 studies) documents the state of knowledge on this topic, describing the participants and experimental approaches that have been used, evaluating their methodological quality and examining their results. With the meta-analyses (31 studies), we quantitatively evaluated the effect of musicianship in different experimental conditions, namely speech in energetic masker, speech in different speech maskers, speech in spatially separated noise and speech in noise at different SNR

levels. We hypothesized that there would be a musician advantage in both energetic and speech maskers, but that this advantage would vary according to the difficulty of the task. Furthermore, we expected to observe a greater musician advantage in older adult musicians, especially in speech maskers. Our results show that musicianship was associated with better SPiN performance in both energetic and speech masker conditions, as well as in SNR <0 dB and SNR=0 dB conditions. The present review also highlighted that very few studies included participants aged ≥ 55 years old, limiting the possibility to adequately assess the effect of musicianship in an ageing population. Overall, our results argue in favour of a musicianship effect on SPiN perception.

3.6.1 Speech in energetic masker, speech in speech masker

Our main hypothesis was that there would be a musician advantage in both energetic and speech maskers and that the effect would be stronger with two-talker maskers. The meta-analyses for energetic maskers revealed a moderate effect size of .61 and the meta-analyses for speech in different speech maskers revealed the largest effect size in the two-talker condition, with an effect of .79 (though with the broadest confidence interval), followed by a moderate effect in the four-talker condition (.56) and a small non-significant effect in the 1-talker condition (.17). These observations suggest that musicianship impacts both masking conditions and that the effect of musicianship varies according to the number of talkers in the masker. Speech maskers display both energetic and informational properties and their masking power varies according to the characteristics and the number of talkers in the masker (Hoen et al., 2007; Rosen et al., 2013; Helfer & Freyman, 2014). In the literature, perceiving speech in a one-talker masker has been shown to be less challenging than in maskers with additional talkers (Rosen et al., 2013). Despite the informational load present in a one-talker masker, its lower masking power could be explained by the energy modulation present in the speech of a single talker, allowing more access to the target speech due to glimpses, defined as “spectrotemporal regions where the energy of the target speech exceeds that of the masker by at least 3 dB” (Cooke, 2006). The addition of a second talker in the masker reduces glimpsing opportunities and raises its informational load, leading to a dramatic decrease in SPiN performance (Tun & Wingfield, 1999; Rosen et al., 2013).

However, increasing further the number of talkers in the masker to ≥ 4 gradually improves SPiN performance by decreasing the informational masking power, due to a reduced access to the background's semantic content (Hoen et al., 2007). Thus, our results suggest that musicianship has the greatest impact on SPiN performance in the most challenging condition in terms of informational masking, the two-talker masker, and a slightly smaller impact on four-talker maskers and energetic maskers, while having no significant effect on the less challenging condition, the one-talker masker.

3.6.2 Speech in spatially separated masker

We hypothesized to find a smaller musician advantage in the spatially separated conditions compared to all collocated conditions. The results of the meta-analysis showed that the effect of musicianship on speech in spatially separated masker was small and non-significant.

Previous studies on non-musicians have shown that spatially separating the target speech from the masker improves SPiN perception, a phenomenon referred to as “spatial release from masking” (SRM) (for a review, see Wang and Xu (2021)). Spatially separating stimuli improves perception by allowing binaural processes of interaural time and level differences (Glyde et al., 2011; Yost, 2017).

SRM has been shown to be less effective on energetic maskers (Jones & Freyman, 2012) compared to speech maskers, for which SRM is inversely proportional to the number of talkers in the masker (Freyman et al., 2004; Helfer & Freyman, 2014). The result of our meta-analysis on speech in spatially separated speech maskers suggests that musicianship provides little to no significant advantages for the perception of spatially separated stimuli. Non-musicians benefitting already from SRM, it is conceivable that the present situation reflects a ceiling effect, that mitigates the advantages of musicianship when masker and targets are collocated.

3.6.3 Speech in noise at various SNR levels

We hypothesized to find a greater musician advantage in the most challenging listening condition, which is the SNR < 0 dB condition. The meta-analysis revealed an effect of musicianship in the SNR < 0 dB condition, as was expected, but also in the 0 dB SNR condition, with a moderate effect size (.56) in both analyses.

Lowering the SNR level raises the loudness of the background noise compared to the target speech, rendering it less audible. The performance in SPiN tasks drops with a decrease in the SNR in young adults; and the effect of SNR is more detrimental in older adults, especially at SNRs \leq 0 dB (Taitelbaum-Swead & Fostick, 2016; Heidari et al., 2018). It would thus appear that the perception of musicians is more resistant to the disruptive effect of low SNR than non-musicians, but not when the target is louder than the masker.

Studies having reported real life SNRs calculated from recordings made from a listener's position have shown that SNRs at 0 dB or below are scarce in real life situations (Smeds et al., 2015; Wu et al., 2018; Brungart et al., 2020). However, SNRs calculated from real life situations are probably overestimated, and corresponding experimental SNRs would be lower (possibly negative), especially for bars and restaurants (Brungart et al., 2020). Additional data is needed to determine whether experimental SPiN tasks mimic to real life SPiN situations, and to understand the extent to which a musician advantage may facilitate communication in day-to-day situations.

3.6.4 Effect of age

We hypothesized that there would be a greater effect of musicianship with advancing age, especially in speech maskers, due to a potential reserve resulting from years of musical training, which would mitigate the normal age-related decline of speech perception. In the systematic review, 2/2 of the studies on middle-aged and 1/1 older adults showed a musician advantage in energetic (3 tasks) and in speech maskers (2 tasks). In the meta-analysis, a significant positive association between the mean age of study participants and the effect of musicianship was found in the analyses containing the only study on older adults (\geq 60 years old), hence in the energetic masker and the SNR = 0 dB conditions. No positive association between mean age and the effect of musicianship was observed in the other analyses, despite

the presence of a study on middle-aged adults in the four-talker masker analysis. However, the limited number of studies on middle-aged (55–65 years old, n=2) and older adults (> 60 years old, n=1) in the meta-analyses does not warrant the interpretation of any age effect, or lack thereof, in the different experimental conditions.

Difficulties perceiving speech in noisy situations appear gradually in adulthood, beginning to emerge in midlife, in the absence of abnormal audiograms (Demeester et al., 2012). These difficulties appear predominantly in conditions of speech maskers, rather than in energetic masking (for a review, see Helfer and Jesse (2021)). Echoing the literature on middle-aged adults, studies comparing young and older adults have also shown that the effect of increasing age systematically decreases the performance of speech perception in speech maskers as opposed to energetic maskers for which there is little effect (Tun & Wingfield, 1999; Rajan & Cainer, 2008; Taitelbaum-Swead & Fostick, 2016; Goossens et al., 2017). It is plausible that this discrepancy in the age effect between energetic and speech maskers could be partially explained by the age-related decline in cognitive faculties (Harada et al., 2013), on which speech maskers place higher demands than energetic maskers (Brungart, 2001; Freyman et al., 2004; Freyman et al., 2007; Meister et al., 2013; Fitzhugh et al., 2021). Hence, to properly assess whether musicianship mitigates the decline in speech perception, and whether this mitigation impacts the perception in energetic and/or informational masker, additional studies on middle-aged and older adult comparing the performance of M and NM in both masking conditions are required.

3.6.5 Critical appraisal of the literature and future considerations

The present review, through the QualSyst evaluation and the extraction of the studies' characteristics, highlights a few limitations in the selection and the characterization of the musician participants in this literature.

First, regarding the selection of participants. The musician participants were selected mainly based on two criteria: the age of onset of musical practice and the number of years of practice. Very few studies selected musician participants based on additional criteria such

as their level (amateur/professional), a measure of their weekly/daily practice frequency and duration, or the type of instrument and music played, leading to heterogeneous samples.

Second, echoing the selection criteria, even if lacking in some studies, the characterization of musician participants was homogeneous in terms of reporting the age of onset and the years of musical practice. However, practice level (amateur, professional), frequency and duration of practice were widely unreported. Additionally, the type of music played by the musicians was not reported. Hence, it is difficult to determine which aspect of the musical practice is associated with a musicianship advantage and to outline a profile of successful musical practice for SPiN perception. On a rehabilitation standpoint, one of the main concerns is the lack of selection and information on the musicians' level (amateurs vs. professionals), which complicates the disentangling of the potential benefits stemming from intense musical practices that may be associated with innate advantages in professional musicians, compared to the accessible recreative practices of amateur musicians.

Finally, the control for potential confounders was scant and uneven between studies. Since musicianship impacts both auditory and cognitive functions, controlling for these aspects appears essential. Musicians and non-musicians were matched on their hearing status in very few studies (14/49). In other studies, either the distribution of hearing statuses between musicians and non-musicians was missing or no information was provided at all. Because peripheral hearing affects speech perception, and given that the practice of a musical instrument, especially in a band/orchestra, can cause hearing loss (Teie, 1998; Ramrattan & Gurevich, 2020) this raises the question of whether potential cognitive benefits of musical practices could be lessened, especially in older adults, by music-induced hearing impairments. Regarding cognition, when participants were matched, it was mostly on their IQ and very few studies measured specific cognitive functions. Measuring cognitive functions that have been associated with SPiN perception such as auditory working memory and speed of processing (Dryden et al., 2017) and evaluating their impact on SPiN performance would contribute to revealing the extent to which these domains are impacted by musicianship, if at all, and their contribution to SPiN perception.

3.6.6 Limitations

There are a few limitations to our study. Our analyses and conclusions are limited by the available literature: a restricted number of studies, and descriptive statistics (mean and standard deviation) not available for every study. Due to the limited number of publications, the different meta-analyses include an uneven number of studies, varying between 4 and 17. This impacts, first, the reliability of the effect sizes, which should be interpreted with some caution. Because of the paucity of studies on middle-aged and older adults, the moderating impact of age could not be consistently assessed in every analysis. In addition, the effect of masker type (energetic VS. speech) in analyses of different SNR levels ((f) to h)) could not be analyzed because of a lack of studies. Furthermore, the assessment of other potentially moderating variables, such as the targets' linguistic level (e.g., words vs. sentences) were also impossible, limiting our ability to explore the between-study heterogeneity.

Finally, we could not obtain the descriptive data for all studies. Specifically, 9 studies that were eligible for our meta-analyses could not be included. As a result, the effect sizes do not reflect the entirety of the literature selected for the systematic review.

3.7 Conclusion

The study of a potential musician advantage on SPiN performance has garnered a wide interest for the prevention/rehabilitation of SPiN perception in the aging population. Despite the little number of studies on the topic, our results show a significant effect of musicianship on SPiN performance in challenging conditions, while having no significant effect in easier ones. This pattern of results is compelling and supports the notion of a musician advantage. Further studies, especially on older adults, are needed to confirm and refine the present results. Furthermore, a more careful selection of musicians, e.g. on their practice level and intensity, as well as a thorough characterization of their practice would hone our understanding on the practices associated with benefits in SPiN performance and contribute to the elaboration of prevention/rehabilitation strategies.

3.8 Acknowledgments

The project was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC grant # RGPIN-2019-06534) to PT. PT holds a Career award from the Fonds de la Recherche en Santé du Québec (FRQ-S, #35016). EM is funded by a graduate scholarship from the Fonds Theodor and Gabriela Kummer (Université de Lausanne, Switzerland).

3.9 Credit author statement

Elisabeth Maillard: Conceptualization, Methodology, Investigation, Validation, Project administration, Formal analysis, Visualization, Writing- Original draft preparation.

Marilyne Joyal: Methodology, Validation, Writing—Reviewing and Editing. **Micah Murray.** Writing—Reviewing and Editing. **Pascale Tremblay:** Conceptualization,

Methodology, Investigation, Validation, Supervision, Resources, Project administration, Writing- Reviewing and Editing, Data Curation.

3.10 References

- Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2005, Apr). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, *117*(4 Pt 1), 2169-2180. <https://doi.org/10.1121/1.1861598>
- Bilodeau-Mercure, M., Lortie, C. L., Sato, M., Guitton, M. J., & Tremblay, P. (2015, Mar). The neurobiology of speech perception decline in aging. *Brain Struct Funct*, *220*(2), 979-997. <https://doi.org/10.1007/s00429-013-0695-3>
- Boothroyd, A. (1968). Statistical Theory of the Speech Discrimination Score. *The Journal of the Acoustical Society of America*, *43*(2), 362-367. <https://doi.org/10.1121/1.1910787>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009, 03/01). An Introduction to Meta-Analysis. *Introduction to Meta-Analysis*, *19*. <https://doi.org/10.1002/9780470743386>
- Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012, Feb). Linguistic contributions to speech-on-speech masking for native and non-native listeners: language

- familiarity and semantic content. *J Acoust Soc Am*, 131(2), 1449-1464. <https://doi.org/10.1121/1.3675943>
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101-1109. <https://doi.org/10.1121/1.1345696>
- Brungart, D. S., Barrett, M. E., Cohen, J. I., Fodor, C., Yancey, C. M., & Gordon-Salant, S. (2020, Nov/Dec). Objective Assessment of Speech Intelligibility in Crowded Public Spaces. *Ear Hear*, 41 Suppl 1(Suppl 1), 68s-78s. <https://doi.org/10.1097/aud.0000000000000943>
- Buss, E., Hodge, S. E., Calandruccio, L., Leibold, L. J., & Grose, J. H. (2019, Sep/Oct). Masked Sentence Recognition in Children, Young Adults, and Older Adults: Age-Dependent Effects of Semantic Context and Masker Type. *Ear Hear*, 40(5), 1117-1126. <https://doi.org/10.1097/AUD.0000000000000692>
- Cameron, S., & Dillon, H. (2007, Apr). Development of the Listening in Spatialized Noise-Sentences Test (LISN-S). *Ear Hear*, 28(2), 196-211. <https://doi.org/10.1097/AUD.0b013e318031267f>
- Coffey, E. B. J., Chepesiuk, A. M. P., Herholz, S. C., Baillet, S., & Zatorre, R. J. (2017). Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. *Front Neurosci*, 11, 479. <https://doi.org/10.3389/fnins.2017.00479>
- Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017, Sep). Speech-in-noise perception in musicians: A review. *Hear Res*, 352, 49-69. <https://doi.org/10.1016/j.heares.2017.02.006>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Routledge. <https://doi.org/https://doi.org/10.4324/9780203771587>
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, 119(3), 1562-1573. <https://doi.org/10.1121/1.2166600>
- D'Souza, A. A., Moradzadeh, L., & Wiseheart, M. (2018). Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cogn Res Princ Implic*, 3(1), 11. <https://doi.org/10.1186/s41235-018-0095-6>
- Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schon, D., & Colombo, L. (2012, May 21). Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: behavioral, electrophysiological and psychoacoustic study. *Brain Res*, 1455, 75-89. <https://doi.org/10.1016/j.brainres.2012.03.034>
- Demeester, K., Topsakal, V., Hendrickx, J.-J., Franssen, E., van Laer, L., Van Camp, G., Van de Heyning, P., & van Wieringen, A. (2012). Hearing Disability Measured by the Speech, Spatial, and Qualities of Hearing Scale in Clinically Normal-Hearing and Hearing-Impaired Middle-Aged Persons, and Disability Screening by Means of a Reduced SSQ (the SSQ5). *Ear and Hearing*, 33(5), 615-616. <https://doi.org/10.1097/AUD.0b013e31824e0ba7>
- Deroche, M. L. D., Limb, C. J., Chatterjee, M., & Gracco, V. L. (2017). Similar abilities of musicians and non-musicians to segregate voices by fundamental frequency. *The Journal of the Acoustical Society of America*, 142(4), 1739-1755. <https://doi.org/10.1121/1.5005496>
- Desjardins, J. L., & Doherty, K. A. (2013, May-Jun). Age-related changes in listening effort for various types of masker noises. *Ear Hear*, 34(3), 261-272. <https://doi.org/10.1097/AUD.0b013e31826d0ba4>
- Donai, J. J., & Jennings, M. B. (2016). Gaps-in-noise detection and gender identification from noise-vocoded vowel segments: Comparing performance of active musicians to non-musicians. *The Journal of the Acoustical Society of America*, 139(5), EL128-EL134. <https://doi.org/10.1121/1.4947070>
- Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017, Jan-Dec). The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic

- Literature Review and Meta-Analysis. *Trends Hear*, 21, 2331216517744675. <https://doi.org/10.1177/2331216517744675>
- Fitzhugh, M. C., Schaefer, S. Y., Baxter, L. C., & Rogalsky, C. (2021). Cognitive and neural predictors of speech comprehension in noisy backgrounds in older adults. *Lang Cogn Neurosci*, 36(3), 269-287. <https://doi.org/10.1080/23273798.2020.1828946>
- Fostick, L., Babkoff, H., & Zukerman, G. (2014, Jun 1). Effect of 24 hours of sleep deprivation on auditory and linguistic perception: a comparison among young controls, sleep-deprived participants, dyslexic readers, and aging adults. *J Speech Lang Hear Res*, 57(3), 1078-1088. [https://doi.org/10.1044/1092-4388\(2013/13-0031\)](https://doi.org/10.1044/1092-4388(2013/13-0031))
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001, May). Spatial release from informational masking in speech recognition. *J Acoust Soc Am*, 109(5 Pt 1), 2112-2122. <https://doi.org/10.1121/1.1354984>
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*, 115(5), 2246-2256. <https://doi.org/10.1121/1.1689343>
- Freyman, R. L., Helfer, K. S., & Balakrishnan, U. (2007, Feb). Variability and uncertainty in masking by competing speech. *J Acoust Soc Am*, 121(2), 1040-1046. <https://doi.org/10.1121/1.2427117>
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999, Dec). The role of perceived spatial separation in the unmasking of speech. *J Acoust Soc Am*, 106(6), 3578-3588. <https://doi.org/10.1121/1.428211>
- Gilbert, J. L., Tamati, T. N., & Pisoni, D. B. (2013, Jan). Development, reliability, and validity of PRESTO: a new high-variability sentence recognition test. *J Am Acad Audiol*, 24(1), 26-36. <https://doi.org/10.3766/jaaa.24.1.4>
- Glyde, H., Hickson, L., Cameron, S., & Dillon, H. (2011, Sep). Problems hearing in noise in older adults: a review of spatial processing disorder. *Trends Amplif*, 15(3), 116-126. <https://doi.org/10.1177/1084713811424885>
- Goossens, T., Vercammen, C., Wouters, J., & van Wieringen, A. (2017, Feb). Masked speech perception across the adult lifespan: Impact of age and hearing impairment. *Hear Res*, 344, 109-124. <https://doi.org/10.1016/j.heares.2016.11.004>
- Gordon-Salant, S., & Cole, S. S. (2016, Sep-Oct). Effects of Age and Working Memory Capacity on Speech Recognition Performance in Noise Among Listeners With Normal Hearing. *Ear Hear*, 37(5), 593-602. <https://doi.org/10.1097/AUD.0000000000000316>
- Grassi, M., Meneghetti, C., Toffalini, E., & Borella, E. (2017). Auditory and cognitive performance in elderly musicians and nonmusicians. *PLoS One*, 12(11), e0187881. <https://doi.org/10.1371/journal.pone.0187881>
- Groussard, M., Coppalle, R., Hinault, T., & Platel, H. (2020). Do Musicians Have Better Mnemonic and Executive Performance Than Actors? Influence of Regular Musical or Theater Practice in Adults and in the Elderly. *Front Hum Neurosci*, 14, 557642. <https://doi.org/10.3389/fnhum.2020.557642>
- Hanna-Pladdy, B., & Gajewski, B. (2012). Recent and past musical activity predicts cognitive aging variability: direct comparison with general lifestyle activities. *Front Hum Neurosci*, 6, 198. <https://doi.org/10.3389/fnhum.2012.00198>
- Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013, Nov). Normal cognitive aging. *Clin Geriatr Med*, 29(4), 737-752. <https://doi.org/10.1016/j.cger.2013.07.002>
- Heidari, A., Moossavi, A., Yadegari, F., Bakhshi, E., & Ahadi, M. (2018, Jun). Effects of Age on Speech-in-Noise Identification: Subjective Ratings of Hearing Difficulties and Encoding of

- Fundamental Frequency in Older Adults. *J Audiol Otol*, 22(3), 134-139. <https://doi.org/10.7874/jao.2017.00304>
- Heine, C., & Browning, C. J. (2002, Oct 15). Communication and psychosocial consequences of sensory loss in older adults: overview and rehabilitation directions. *Disabil Rehabil*, 24(15), 763-773. <https://doi.org/10.1080/09638280210129162>
- Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, 136(2), 748-759. <https://doi.org/10.1121/1.4887463>
- Helfer, K. S., & Jesse, A. (2021). Hearing and speech processing in midlife. *Hearing Research*, 402. <https://doi.org/10.1016/j.heares.2020.108097>
- Herholz, Sibylle C., & Zatorre, Robert J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, 76(3), 486-502. <https://doi.org/10.1016/j.neuron.2012.10.011>
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003, Sep 6). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557-560. <https://doi.org/10.1136/bmj.327.7414.557>
- Hoen, M., Meunier, F., Grataloup, C.-L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., & Collet, L. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication*, 49(12), 905-916. <https://doi.org/10.1016/j.specom.2007.05.008>
- Huizenga, H. M., Visser, I., & Dolan, C. V. (2011, Feb). Testing overall and moderator effects in random effects meta-regression. *Br J Math Stat Psychol*, 64(Pt 1), 1-19. <https://doi.org/10.1348/000711010X522687>
- Humes, L. E. (2021). Factors Underlying Individual Differences in Speech-Recognition Threshold (SRT) in Noise Among Older Adults. *Front Aging Neurosci*, 13, 702739. <https://doi.org/10.3389/fnagi.2021.702739>
- Jones, J. A., & Freyman, R. L. (2012, Jan-Feb). Effect of priming on energetic and informational masking in a same-different task. *Ear Hear*, 33(1), 124-133. <https://doi.org/10.1097/AUD.0b013e31822b5bee>
- Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., & Durlach, N. I. (2008). Informational Masking. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 143-189). Springer US. https://doi.org/10.1007/978-0-387-71305-2_6
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S. (2004, Oct). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 116(4 Pt 1), 2395-2405. <https://doi.org/10.1121/1.1784440>
- Kmet, L., Lee, R., & Cook, L. (2004). Standard quality assessment criteria for evaluating primary research papers for a variety of fields. (AHFMR - HTA Initiative #13). <https://doi.org/http://www.ahfmr.ab.ca/download.php/>
- Koo, T. K., & Li, M. Y. (2016, Jun). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kuhnis, J., Elmer, S., Meyer, M., & Jancke, L. (2013, Jul). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study. *Neuropsychologia*, 51(8), 1608-1618. <https://doi.org/10.1016/j.neuropsychologia.2013.04.007>
- Kumar, P. V., Rana, B., & Krishna, R. (2014). Temporal processing in musicians and non-musicians [journal article]. *Journal of Hearing Science*, 4(3), 35-42.

- <https://www.journalofhearingscience.com/TEMPORAL-PROCESSING-IN-MUSICIANS-AND-NON-MUSICIANS,120632,0,2.html>
- Lee, L., Packer, T. L., Tang, S. H., & Girdler, S. (2008, Dec). Self-management education programs for age-related macular degeneration: a systematic review. *Australas J Ageing*, 27(4), 170-176. <https://doi.org/10.1111/j.1741-6612.2008.00298.x>
- Leibold, L. J. (2017, Oct 17). Speech Perception in Complex Acoustic Environments: Developmental Effects. *J Speech Lang Hear Res*, 60(10), 3001-3008. https://doi.org/10.1044/2017_JSLHR-H-17-0070
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gotzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009, Jul 21). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*, 6(7), e1000100. <https://doi.org/10.1371/journal.pmed.1000100>
- Maharaj, S., & Harding, R. (2016, Jan 22). The needs, models of care, interventions and outcomes of palliative care in the Caribbean: a systematic review of the evidence. *BMC Palliat Care*, 15, 9. <https://doi.org/10.1186/s12904-016-0079-6>
- Mansens, D., Deeg, D. J. H., & Comijs, H. C. (2018, Aug). The association between singing and/or playing a musical instrument and cognitive functions in older adults. *Aging Ment Health*, 22(8), 964-971. <https://doi.org/10.1080/13607863.2017.1328481>
- Meister, H., Schreitmuller, S., Grugel, L., Ortmann, M., Beutner, D., Walger, M., & Meister, I. G. (2013, Mar 1). Cognitive resources related to speech recognition with a competing talker in young and older listeners. *Neuroscience*, 232, 74-82. <https://doi.org/10.1016/j.neuroscience.2012.12.006>
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006, Sep). Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res*, 219(1-2), 36-47. <https://doi.org/10.1016/j.heares.2006.05.004>
- Moore, D. R., Edmondson-Jones, M., Dawes, P., Fortnum, H., McCormack, A., Pierzycki, R. H., & Munro, K. J. (2014). Relation between speech-in-noise threshold, hearing loss and cognition

- from 40-69 years of age. *PLoS One*, 9(9), e107720. <https://doi.org/10.1371/journal.pone.0107720>
- Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994, Feb). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *J Acoust Soc Am*, 95(2), 1085-1099. <https://doi.org/10.1121/1.408469>
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician Enhancement for Speech-In-Noise. *Ear and Hearing*, 30(6), 653-661. <https://doi.org/10.1097/AUD.0b013e3181b412e9>
- Patel, A. D. (2014, Feb). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hear Res*, 308, 98-108. <https://doi.org/10.1016/j.heares.2013.08.011>
- Presacco, A., Simon, J. Z., & Anderson, S. (2016, Nov 1). Effect of informational content of noise on speech representation in the aging midbrain and cortex. *J Neurophysiol*, 116(5), 2356-2367. <https://doi.org/10.1152/jn.00373.2016>
- Puschmann, S., Baillet, S., & Zatorre, R. J. (2019, Jul 22). Musicians at the Cocktail Party: Neural Substrates of Musical Training During Selective Listening in Multispeaker Situations. *Cereb Cortex*, 29(8), 3253-3265. <https://doi.org/10.1093/cercor/bhy193>
- Rajan, R., & Cainer, K. E. (2008, Jun 23). Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. *Neuroscience*, 154(2), 784-795. <https://doi.org/10.1016/j.neuroscience.2008.03.067>
- Ramrattan, H., & Gurevich, N. (2020). Prevalence of Noise-Induced Hearing Loss in Middle and High School Band Members: A Preliminary Study. *Folia Phoniatr Logop*, 72(4), 302-308. <https://doi.org/10.1159/000501154>
- Rennies, J., Best, V., Roverud, E., & Kidd, G., Jr. (2019, Jan-Dec). Energetic and Informational Components of Speech-on-Speech Masking in Binaural Speech Intelligibility and Perceived Listening Effort. *Trends Hear*, 23, 2331216519854597. <https://doi.org/10.1177/2331216519854597>
- Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013, Apr). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J Acoust Soc Am*, 133(4), 2431-2443. <https://doi.org/10.1121/1.4794379>
- Santini, Z. I., Jose, P. E., York Cornwell, E., Koyanagi, A., Nielsen, L., Hinrichsen, C., Meilstrup, C., Madsen, K. R., & Koushede, V. (2020). Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *The Lancet Public Health*, 5(1), e62-e70. [https://doi.org/10.1016/s2468-2667\(19\)30230-0](https://doi.org/10.1016/s2468-2667(19)30230-0)
- Schon, D., Magne, C., & Besson, M. (2004, May). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341-349. <https://doi.org/10.1111/1469-8986.00172.x>
- Shinn-Cunningham, B. (2013). *Understanding informational masking from a neural perspective*
- Smeds, K., Wolters, F., & Rung, M. (2015, Feb). Estimation of Signal-to-Noise Ratios in Realistic Sound Scenarios. *J Am Acad Audiol*, 26(2), 183-196. <https://doi.org/10.3766/jaaa.26.2.7>
- Souza, P. E., & Turner, C. W. (1994, Jun). Masking of speech in young and elderly listeners with hearing loss. *J Speech Hear Res*, 37(3), 655-661. <https://doi.org/10.1044/jshr.3703.655>
- Stern, Y. (2002, Mar). What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc*, 8(3), 448-460.
- Stern, Y. (2009, Aug). Cognitive reserve. *Neuropsychologia*, 47(10), 2015-2028. <https://doi.org/10.1016/j.neuropsychologia.2009.03.004>
- Stern, Y., Arenaza-Urquijo, E. M., Bartres-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., Vuoksima, E., the Reserve, R., Protective Factors, P. I. A. E. D., & Conceptual Frameworks, W. (2020, Sep). Whitepaper:

- Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimers Dement*, 16(9), 1305-1311. <https://doi.org/10.1016/j.jalz.2018.07.219>
- Sundström, A., Adolfsson, A. N., Nordin, M., & Adolfsson, R. (2020, Apr 16). Loneliness Increases the Risk of All-Cause Dementia and Alzheimer's Disease. *J Gerontol B Psychol Sci Soc Sci*, 75(5), 919-926. <https://doi.org/10.1093/geronb/gbz139>
- Taitelbaum-Swead, R., & Fostick, L. (2016). The Effect of Age and Type of Noise on Speech Perception under Conditions of Changing Context and Noise Levels. *Folia Phoniatr Logop*, 68(1), 16-21. <https://doi.org/10.1159/000444749>
- Taylor, H. O., Taylor, R. J., Nguyen, A. W., & Chatters, L. (2018, Feb). Social Isolation, Depression, and Psychological Distress Among Older Adults. *J Aging Health*, 30(2), 229-246. <https://doi.org/10.1177/0898264316673511>
- Teie, P. U. (1998, Jan). Noise-induced hearing loss and symphony orchestra musicians: risk factors, effects, and management. *Md Med J*, 47(1), 13-18.
- Tun, P. A. (1998, Sep). Fast noisy speech: age differences in processing rapid speech with background noise. *Psychol Aging*, 13(3), 424-434. <https://doi.org/10.1037//0882-7974.13.3.424>
- Tun, P. A., & Wingfield, A. (1999, Sep). One voice too many: adult age differences in language processing with different types of distracting sounds. *J Gerontol B Psychol Sci Soc Sci*, 54(5), P317-327.
- Vermeire, K., Knoop, A., Boel, C., Auwers, S., Schenus, L., Talaveron-Rodriguez, M., De Boom, C., & De Sloovere, M. (2016, Apr). Speech Recognition in Noise by Younger and Older Adults: Effects of Age, Hearing Loss, and Temporal Resolution. *Ann Otol Rhinol Laryngol*, 125(4), 297-302. <https://doi.org/10.1177/0003489415611424>
- Viechtbauer, W. (2010, 2010-08-05). Conducting Meta-Analyses in R with the metafor Package. 2010, 36(3), 48. <https://doi.org/10.18637/jss.v036.i03>
- Wan, C. Y., & Schlaug, G. (2010). Music Making as a Tool for Promoting Brain Plasticity across the Life Span. *The Neuroscientist*, 16(5), 566-577. <https://doi.org/10.1177/1073858410377805>
- Wang, X., & Xu, L. (2021, Apr). Speech perception in noise: Masking and unmasking. *J Otol*, 16(2), 109-119. <https://doi.org/10.1016/j.joto.2020.12.001>
- Wilson, R. H. (2003, Nov). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *J Am Acad Audiol*, 14(9), 453-470.
- Wilson, R. H., McArdle, R. A., & Smith, S. L. (2007). An Evaluation of the BKB-SIN, HINT, QuickSIN, and WIN Materials on Listeners With Normal Hearing and Listeners With Hearing Loss. *Journal of Speech, Language, and Hearing Research*, 50(4), 844-856. [https://doi.org/doi:10.1044/1092-4388\(2007/059\)](https://doi.org/doi:10.1044/1092-4388(2007/059))
- Working Group on Speech, U., & Aging. (1988, 1988/03/01). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83(3), 859-895. <https://doi.org/10.1121/1.395965>
- Wu, Y. H., Stangl, E., Chipara, O., Hasan, S. S., Welhaven, A., & Oleson, J. (2018, Mar/Apr). Characteristics of Real-World Signal to Noise Ratios and Speech Listening Situations of

- Older Adults With Mild to Moderate Hearing Loss. *Ear Hear*, 39(2), 293-304. <https://doi.org/10.1097/aud.0000000000000486>
- Yeend, I., Beach, E. F., Sharma, M., & Dillon, H. (2017, Sep). The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hear Res*, 353, 224-236. <https://doi.org/10.1016/j.heares.2017.07.006>
- Yost, W. A. (2017, Mar). Spatial release from masking based on binaural processing for up to six maskers. *J Acoust Soc Am*, 141(3), 2093. <https://doi.org/10.1121/1.4978614>
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007, 2007/07/01). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547-558. <https://doi.org/10.1038/nrn2152>
- Zendel, B. R., & Alain, C. (2009). Concurrent Sound Segregation Is Enhanced in Musicians. *Journal of Cognitive Neuroscience*, 21(8), 1488-1498. <https://doi.org/10.1162/jocn.2009.21140>
- Zendel, B. R., & Alain, C. (2013, Apr). The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. *J Cogn Neurosci*, 25(4), 503-516. https://doi.org/10.1162/jocn_a_00329
- Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One*, 9(6), e99868. <https://doi.org/10.1371/journal.pone.0099868>

3.11 Supplementary material

Supplementary Material 1: Quality Evaluation

Nr	Checklist	Items used to answer the question				Scoring method
		A	B	C	D	
Q1	Question/objectives sufficiently described?	Purpose of the study	Subject/target population	Age of interest	Parameter(s) under investigation	All 4 criteria = Yes; 2-3 criteria = Partial; 0 criteria = No
Q2	Study design evident and appropriate?	No additional criteria				
Q3	Method of subject/comparison group selection or source of information/input variables described and appropriate?	Recruitment method	Recruitment setting	Sampling method	Inclusion/exclusion criteria for the different groups	All 4 criteria = Yes; 2-3 criteria = Partial; 0 criteria = No
Q4	Subject (and comparison group, if applicable) characteristics sufficiently described?	Mean age and SD per group	Sex distribution in both groups	Mean onset age of musical practice and SD	Mean total years/hours of musical practice and SD	All 4 criteria = Yes; 2-3 criteria = Partial; 0 criteria = No
Q5	If interventional and random allocation was possible, was it described?	No additional criteria				
Q6*	If interventional and blinding of investigators was possible,	No additional criteria				

	was it reported?					
Q7*	If interventional and blinding of subjects was possible, was it reported?	No additional criteria				
Q8*	Outcomes and (if applicable) exposure measure(s) well defined and robust to measurement / misclassification bias? Means of assessment reported?	Testing environment	Test(s) performed	Presentation mode (loudspeaker, headphones or earphones)	Dependant variable	All 4 criteria = Yes; 2-3 criteria = Partial; 0 criteria = No
Q9*	Sample size appropriate?	Sample size calculation	Effect size	Power estimate	Over 10 participants per group	A/B/C = Yes; Only D = Partial; 0 criteria = No
Q10*	Analytic methods described/justified and appropriate?	No additional criteria				
Q11*	Some estimate of variance is reported for the main results?	Confidence interval - exact value or plot	Standard error of measure - exact value or plot	Group mean and standard deviation - exact value or plot	Individual data	A/B/C (exact value) = Yes; D + A/B/C (plot) = Yes; Only A/B/C (plot) = Partial; Only D = Partial; 0 criteria = No
Q12	Controlled for confounding?	Groups matched in age	Groups matched in cognition/intelligence/ socio-economic status	Groups matched in audition	N/A	All 3 criteria = Yes; 1-2 criteria = Partial; 0 criteria = No
Q13*	Results reported in	Descriptive data	Inferential statistics	Effect size	N/A	All 3 criteria = Yes; 1-2 criteria = Partial; 0 criteria = No

	sufficient details?			
Q14	Conclusion supported by the results?	No additional criteria		

Note. We focused the evaluation of the items with an asterisk* on the SPiN tasks. When present, other tasks were not evaluated.

Table 1b. *Qualsyst scores*

Authors, year	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Total sum	Possible total sum	Summary score
Zhang (2019)	1	2	1	0	N/A	N/A	N/A	1	1	1	1	1	1	1	11	22	0,50
Puschmann (2019)	1	2	1	1	N/A	N/A	N/A	1	1	2	0	0	1	2	12	22	0,55
Soncini (2006)	1	2	1	1	N/A	N/A	N/A	2	1	1	1	0	1	2	13	22	0,59
Parbery-Clark (2012)	1	2	1	1	N/A	N/A	N/A	1	1	1	0	2	1	2	13	22	0,59
Morse-Fortier (2017)	1	2	1	1	N/A	N/A	N/A	2	1	1	2	0	1	1	13	22	0,59
Slater (2016)	2	2	1	1	N/A	N/A	N/A	1	1	1	1	1	1	1	13	22	0,59
Mankel & Bidelman (2018)	1	2	1	1	N/A	N/A	N/A	1	1	2	1	1	1	2	14	22	0,64
Yoo & Bidelman (2019)	1	2	1	1	N/A	N/A	N/A	1	1	1	2	1	1	2	14	22	0,64
Strait (2012)	1	2	1	1	N/A	N/A	N/A	2	1	2	1	1	1	1	14	22	0,64
Parbery-Clark (2013)	1	2	1	1	N/A	N/A	N/A	1	1	2	0	2	1	2	14	22	0,64
Başkent et al., (2018)	2	2	1	1	N/A	N/A	N/A	2	1	1	1	0	1	2	14	22	0,64
Deroche (2017)	1	2	1	1	N/A	N/A	N/A	2	1	2	1	0	1	2	14	22	0,64
Yeend (2017)	2	2	1	1	2	N/A	N/A	1	1	1	0	N/A	1	2	14	22	0,64
Coffey (2019)	2	2	1	1	N/A	N/A	N/A	2	1	1	1	0	1	2	14	22	0,64
Prabery-Clark (2011)	1	2	1	2	N/A	N/A	N/A	1	1	1	1	2	1	2	15	22	0,68
Anaya (2016)	1	2	1	1	N/A	N/A	N/A	2	1	1	2	1	2	1	15	22	0,68
Swaminathan (2015)	1	2	1	1	N/A	N/A	N/A	2	1	1	2	0	2	2	15	22	0,68
Başkent & Gaudrain (2016)	1	2	1	1	N/A	N/A	N/A	2	1	2	1	1	1	2	15	22	0,68
Parbery-Clark (2009)	1	2	1	1	N/A	N/A	N/A	1	1	2	1	2	1	2	15	22	0,68
Zendel (2015)	1	2	2	2	N/A	N/A	N/A	1	1	1	1	1	1	2	15	22	0,68
Fuller (2014)	1	2	1	2	N/A	N/A	N/A	2	1	1	1	1	1	2	15	22	0,68
Coffey (2017)	1	2	1	1	N/A	N/A	N/A	2	1	2	2	N/A	1	1	14	20	0,70
Fostick (2019)	2	2	1	1	N/A	N/A	N/A	1	1	2	1	1	2	2	16	22	0,73
Varnet (2015)	2	2	1	2	N/A	N/A	N/A	1	1	1	2	1	1	2	16	22	0,73

Table 1b. Qualsyst scores

Nie (2018)	2	2	1	2	N/A	N/A	N/A	1	1	2	1	0	2	2	16	22	0,73
Skoe (2019)	1	2	1	1	N/A	N/A	N/A	2	1	2	2	1	1	2	16	22	0,73
Slater (2018)	1	2	2	1	N/A	N/A	N/A	1	1	1	1	2	2	2	16	22	0,73
Jain & Nataraja, 2019	1	2	1	1	N/A	N/A	N/A	2	1	2	2	0	2	2	16	22	0,73
Clayton (2016)	1	2	1	1	N/A	N/A	N/A	2	1	1	2	1	2	2	16	22	0,73
Meha-Bettison (2018)	2	2	1	2	N/A	N/A	N/A	1	1	2	1	2	1	2	17	22	0,77
Ruggles (2014)	1	2	1	1	N/A	N/A	N/A	2	1	2	2	1	2	2	17	22	0,77
Parbery-Clark (2011b)	2	2	1	1	N/A	N/A	N/A	1	1	2	2	2	1	2	17	22	0,77
Parbery-Clark (2012b)	2	2	1	1	N/A	N/A	N/A	1	1	2	2	2	1	2	17	22	0,77
Boebinger (2015)	1	2	1	1	N/A	N/A	N/A	1	1	2	2	1	2	2	17	22	0,77
Zendel & Alain, 2012	2	2	1	1	N/A	N/A	N/A	2	1	2	2	1	1	2	17	22	0,77
Mandikal-Vasuki (2016)	1	2	2	2	N/A	N/A	N/A	1	1	2	2	1	1	2	17	22	0,77
Bidelman & Yoo (2020)	1	2	1	2	N/A	N/A	N/A	2	1	2	2	1	1	2	17	22	0,77
Slater (2015)	2	2	1	2	1	N/A	N/A	2	1	1	1	2	2	2	19	24	0,79
Zendel (2019)	2	2	2	2	2	N/A	N/A	2	1	1	1	1	2	1	19	24	0,79
Hutchins (2018)	2	2	1	2	N/A	N/A	N/A	2	1	2	2	N/A	2	2	18	22	0,82
MacCutcheon (2020)	2	2	1	2	N/A	N/A	N/A	2	1	1	2	1	2	2	18	22	0,82
Madsen (2017)	1	2	1	2	N/A	N/A	N/A	2	2	2	2	1	1	2	18	22	0,82
Madsen (2019)	1	2	1	2	N/A	N/A	N/A	2	2	1	2	1	2	2	18	22	0,82
Escobar (2019)	1	2	1	2	N/A	N/A	N/A	2	1	2	2	1	2	2	18	22	0,82
Perron (2021)	1	2	2	1	N/A	N/A	N/A	2	1	2	2	2	1	2	18	22	0,82
Fleming (2019)	2	2	2	2	2	N/A	N/A	2	1	1	1	1	2	2	20	24	0,83
Vanden Bosch der Nederlanden (2018)	1	2	1	2	N/A	N/A	N/A	2	2	2	2	1	2	2	19	22	0,86
Du & Zatorre (2017)	2	2	1	2	N/A	N/A	N/A	1	1	2	2	2	2	2	19	22	0,86
Dubinsky (2019)	2	2	2	2	N/A	N/A	N/A	2	1	1	2	1	2	2	19	22	0,86

Supplementary material 2: participants characteristics

Table 2a. Cross-sectional studies											
Article	Demographics					Participants education	Musical Experience				
	Group	N	Age (Mean and SD/range)	% women	Participants language		Age at onset (Mean and SD/range)^a	Total years of practice (Mean and SD/range)^a	Hours of training per week (Mean and SD)^a	Activity type	Proficiency Level^b
Strait (2012)	Musicians	15	10.3 (1.6)	-	-	-	2.0 (1.4)	7.9 (2.2)	≥ 1,66	Mixed instruments	Amateur
	Non-musicians	16	10.1 (1.9)	-			N/A	N/A	N/A	N/A	N/A
Nie (2018)	Musicians	15	12.1 (2.8)	80	-	-	6.1 (1.6)	5.8 (2.1)	1.82(1.02)	Mixed instruments	Amateur
	Non-musicians	15	10.5 (2.1)	33			N/A	N/A	N/A	N/A	N/A
Başkent (2018)	Musicians	10	12.4 (11-13)	70	Dutch	-	≤ 7	> 5	-	-	-
	Non-musicians	11	12.3 (11-14)	27			N/A	N/A	N/A	N/A	N/A
Vanden Bosch der Nederlanden (2018)	Musicians	30	21.60 (4.22)	43	-	-	9.10 (2.38)	11.63 (4.87)	-	-	-
	Non-musicians	30	20.27 (2.32)	47			11.09 (2.72)	1.24 (1.47)	N/A	N/A	N/A
Zhang (2019)	Musicians	17	24.27 (1.58)	-		College students	≤ 7	≥ 10	-	Mixed instruments	Professional

	Non-musicians	17	23.65 (2.83)*	-	North American English		N/A	≤ 3	N/A	N/A	N/A
Parbery-Clark (2011a)	Musicians	16	22.4 (3.4)	69	North American English	-	5.1 (1.2)	16.4 (3.4)	-	Mixed instruments	-
	Non-musicians	15		73			N/A	< 3	N/A	N/A	N/A
Varnet (2015)	Musicians	20	23 (2.89)	ratio M/F= 9/5	-	-	6.03 (1.99)	15.84 (3.95)	-	Mixed instruments, vocals	-
	Non-musicians	20	22.68 (4.39)	ratio M/F= 6/13			N/A	N/A	N/A	N/A	N/A
Ruggles (2014)	Musicians	16	21,8	50	North American English	University of Minnesota students.	6.02 (1.98)	(10 – 22)	≥ 5	Mixed instruments, vocals	Mixed
	Non-musicians	17	20,7	59			N/A	< 2	N/A	N/A	N/A
Mankel (2018)	Non-musicians	28	22.2 (3.1)	82	North American English	The groups were matched on education.	N/A	0.7 (0.8)	N/A	N/A	N/A
	Musicians	14	age matched	gender matched			7.14 (2.47)	16.0 (4.9)	-	-	-
Yoo & Bidelman (2019)	Musicians	16	25.4 (4.2)	63	North American English	The groups were matched in formal education.	7.0 (2.37)	15.81 (4.82)	-	Mixed instruments	-
	Non-musicians	15		67			N/A	0.37 (0.79)	N/A	N/A	N/A

Anaya (2016)	Musicians	11	23.42 (3.76)	-	North American English	The groups were matched on education levels. (musicians: recruited from Indiana University's Jacob School of Music, non-musicians: no information)	4.9 (1.44)	15.45 (2.69)	16.31 (10.44)	Piano or organ	Professional
	Non-musicians	11	20.75 (1.85)	-		N/A	1.72 (1.9)	N/A	N/A	N/A	
Parbery-Clark (2013)	Musicians	15	20 (2)	63	North American English	-	5.1 (0.88)	16.2 (3.19)	-	Mixed instruments	-
	Non-musicians	15					N/A	0.73 (1.03)	N/A	N/A	N/A
Du & Zatorre (2017)	Musicians	15	21.4 (2.7)	47	-	Post-secondary education (years): M=2.5, SD=1.7	5.13 (1.60)	16.27 (3.77)	-	Mixed instruments, vocals	-
	Non-musicians	15	22.1 (4.4)	47		Post-secondary education (years): M=2.5, SD=1.5	N/A	< 1	N/A	N/A	N/A

Swaminathan (2015)	Musicians	12	23.0 (2.8)	-	North American English	Mostly students at the School of Music at Boston University	8.50 (3.32)	13.75 (3.11)	≥ 5	Mixed instruments, vocals	Mixed
	Non-musicians	12	20.3 (1.1)	-		-	N/A	N/A	N/A	N/A	N/A
Baskent & Gaudrain (2016)	Musicians	18	21.89 (1.97)	56	Dutch	MBO=0%, HBO=44%, BSc=33%, MSc=22%, Other=0%	≤ 7	≥ 10	-	-	-
	Non-musicians	20	22.75 (2.43)	70		MBO=20%, HBO=30%, BSc=45%, MSc=0%, Other=5%.	N/A	N/A	N/A	N/A	N/A
Parbery-Clark (2009)	Musicians	16	23 (3)	61	North American English	-	4.68 (1.4)	16 (2.49)	-	Mixed instruments, vocals	-
	Non-musicians	15					N/A	N/A	0.75 (1.06)	N/A	N/A
Boebinger (2015)	Musicians	25	27.2 (6.9)	42	British English	The groups were matched on years of post-secondary education.	5.9 (1.0)	22.7 (7.8)	-	Mixed instruments	-
	Non-musicians	25					N/A	< 3	N/A	N/A	N/A

Madsen (2017)	Musicians	30	21.13 (2.47)	70	North American English	21 are music majors	5.43 (0.94)	14.57 (2.74)	≥ 5	Mixed instruments, vocals	Mixed
	Non-musicians	30	20.9 (2.70)	73		-	N/A	< 2	N/A	N/A	N/A
Parbery-Clark (2012a)	Musicians	23	22.0 (3.54)	65	-	-	5.4 (0.89)	17.30 (3.83)	-	Mixed instruments, vocals	-
	Non-musicians	27		55			N/A	< 3	N/A	N/A	N/A
Skoe (2019)	Musicians	35	18 - 24	77	North American English	College students	-	11.31 (2.49)	-	Mixed instruments, vocals	Mixed
	Non-musicians	21					N/A	1.95 (2.31)	N/A	N/A	N/A
Deroche (2017)	Musicians	12	21.9 (2.6)	78	North American English	-	≤ 8	≥ 8	-	Mixed instruments, vocals	-
		12									
		16									
		16									
	Non-musicians	12	25.1 (5.9)	67	North American English	-	N/A	≤ 2	N/A	N/A	N/A
		12									
		16									
		16									

Madsen (2019)	Musicians	32	22.84 (3.48)	50	Danish	-	6.13 (0.91)	15.25 (3.12)	-	Mixed instruments, vocals	-
	Non-musicians	32	22.94 (2.2)	53			N/A	N/A	N/A	N/A	N/A
Slater (2018)	Musicians	17	25.8 (5.9)	29	-	-	-	≥ 5	-	Percussions	-
	Non-musicians	14	23.4 (3.7)	29			N/A	≤ 3	N/A	N/A	N/A
Escobar (2019)	Musicians	27	21.7	59	North American English	Undergraduate and graduate students	8.22 (2.2)	13.41 (3.27)	≥ 5	Mixed instruments, vocals	-
	Non-musicians	22	21.1	68			N/A	< 3	N/A	N/A	N/A
Morse-Fortier (2017)	Musicians	20	20.1 (0.67)	65	-	College students	-	11.4 (3.27)	-	Mixed instruments, vocals	Professional
	Non-musicians	20	22.5	95			N/A	N/A	N/A	N/A	N/A
Zendel (2015)	Musicians	13	23.4 (4.3)	38	French (Québec)	Years of education: M=16.7, SD=1.9	7.8 (3.6)	15.5 (5.1)	23.4 (4.3)	Mixed instruments, vocals	-
	Non-musicians	13	21.9 (2.6)	69			Years of education: M=15.2, SD=1.1	N/A	< 1	N/A	N/A

Coffey (2019)	Musicians	15	26.2 (5.1)	57	North American English	-	6.35 (3–13)	> 10	-	Mixed instruments, vocals	-
	Non-musicians	17	23.9 (5.4)				N/A	< 2	N/A	N/A	N/A
Fuller (2014)	Musicians	25	22.9 (18–27)	72	Dutch	-	5.8 (3–7)	14.6 (10–20)	-	Mixed instruments	-
	Non-musicians	25	22.4 (19–28)	72			9.1 (6–13)	1.6 (0–10)	N/A	N/A	N/A
Slater & Kraus (2016)	Musicians	16	25.4 (5.7)	0	-	N/A	-	16.7 (7.1)	-	Percussions	-
	Musicians	21	23.4 (3.6)	0			-	14.7 (4.4)	-	Vocals	-
	Non-musicians	17	23.2 (3.8)	0			N/A	2.35 (1.5)	N/A	N/A	N/A
Clayton (2016)	Musicians	17	22.5 (2.8)	-	North American English	Nearly all were enrolled in the School of Music at Boston University.	-	14.4 (2.96)	-	Mixed instruments, vocals	Mixed
	Non-musicians	17	20.47 (1.4)	-			-	N/A	< 3 (on average)	N/A	N/A
Mandikal-Vasuki (2016)	Musicians	18	median = 28	72	Australian English	The groups were matched in formal education.	6.33 (1.71)	21.77 (11.30)	-	Mixed instruments, vocals	-
	Non-musicians	22	median = 25	77			N/A	< 3	N/A	N/A	N/A

Bidelman & Yoo (2020)	Musicians	14	range: 19-33	64	-	Groups were matched in education	7.2 (2.49)	15.07 (4.14)	-	Mixed instruments, vocals	-
	Non-musicians	14		71			N/A	0.89 (1.23)	N/A	N/A	N/A
Soncini (2006)	Musicians	45	30,93	0	Brazilian Portuguese	High school degree: 51,11%, University education: 48,89%	-	≥ 5	-	Mixed instruments	Professional
	Non-musicians	55	31,69	0			High school degree: 63,63%, University education: 36,36%	N/A	N/A	N/A	N/A
Parbery-Clark (2012b)	Musicians	23	56 (5)	-	-	-	6.52 (1.5)	49.36 (5.51)	-	Mixed instruments	-
	Non-musicians	25		-			N/A	< 4	N/A	N/A	N/A
Parbery-Clark (2011b)	Musicians	18	55 (4.24)	-	-	-	5.6 (1.2)	50 (4.77)	-	Mixed instruments	-
	Non-musicians	19	54 (6.02)	-			N/A	N/A	N/A	N/A	N/A
Fostick (2019)	Musicians	23	65.57 (5.84)	48	Hebrew	-	-	-	≥ 7	Mixed instruments	Amateur

	Non-musicians	23	65.39 (5.57)				-	-	≥ 3	Card playing	Amateur
	Non-musicians	23	65.61 (6.36)				N/A	N/A	N/A	N/A	N/A
Meha-Bettison (2018)	Musicians	10	46.9 (11.84)	30	-	The groups were matched on education levels.	7.2 (1.9)	39.7 (12.03)	24.65 (7.12)	Mixed instruments, vocals	Professional
	Non-musicians	10	44.8 (12.02)	30			N/A	N/A	N/A	N/A	N/A
Zendel & Alain (2012)	Musicians	74	45,3	47		The groups were matches in years of education (musicians: 16.9 years, non-musicians: 16.2 years).	≤ 16	-	-	-	Mixed
	Non-musicians	89	49,3	57	-		N/A	≤ 2	N/A	N/A	N/A
Perron (2021)	Musicians	42	54.95 (19.25)	66	French (Québec)	Years of education: M=15.14, SD=2.78	-	≥ 2	≥ 1	Choir singing	Amateur
	Non-musicians	41	53.90 (19.75)	50		Years of education: M=15.14, SD=2.78.	N/A	N/A	N/A	N/A	N/A
Coffey (2017)	Musicians	12	25.7 (4.2)	60	-	-	range: 5 - 12	-	-	Mixed instruments, vocals	Mixed

	Non-musicians	8					N/A	N/A	N/A	N/A	N/A
Yeend (2017)	Musicians	22	range: 30-57	52	Australian English	University degree: 68%, trade or technical qualification: 25%, high school degree: 7%.	-	< 8	-	-	Amateur
	Musicians	49					-	≥ 8	-	-	Amateur
	Musicians	20					-	-	-	-	Professional
	Non-musicians	31					N/A	N/A	N/A	N/A	N/A
Puschmann (2019)	Musicians	16	21 (3)	55	-	-	7.81 (2.54)	"intense training years" : 10.5 (5.62)	-	-	-
	Non-musicians	4					N/A	0	N/A	N/A	N/A
Jain (2019)	Musicians	26	33,71	0	Kannada	-	-	≥ 10	(6-8h) x 4-5 days	Mixed instruments	Professional
	Non-musicians	25	32,84	0			N/A	≥ 10	(6-8h) x 4-5 days	Theater	Professional

Table 2b. Longitudinal studies									
Article	Demographics					Participants education	Activity		
	Group	N	Age (Mean and SD /range)	% women	Participants language		Training regimen	Activity type	Level
Hutchins (2018)	Musicians (post-training)	90	5.13 (3.68–7.43)	45,6	-	3-year-old group: 26%, 4-year-old group: 42%, 5-year-old group: 21%, 6-year-old group: 11%.	28 classes for 8 months	Musical theory (basic concepts of rhythm and pitch, musical notation), musical listening, musical practice (singing, practice of pitched and non-pitched percussion instruments).	Amateur
	Controls (same children pre-training)	90	4.55 (3.09 - 6.83)						
MacCutcheon (2020)	Musical training group	26	6.3 (0.5)	0	-	-	30-60 minutes/week for 38 weeks	Group recorder lesson (30 min, all participants), additional individual piano or violin classes (30 min, 12 participants)	Amateur
	Sport training group (controls)	15		0					

Slater (2015)	Musical training - 1st group	19	8,1	63	-	-	1st, 2nd and 3rd year of study : 4h/week (all participants)	Musical theory, singing and instrument classes (Harmony project's musical curriculum)	Amateur
	Musical training - 2nd group	19	8,3	47			1st year of study : no activity 2nd and 3rd year of study : 4h/week (14 participants), 5 h/wk (5 participants)	Musical theory, singing and instrument classes (Harmony project's musical curriculum)	Amateur
Zendel (2019)	Musical training group	13	67.5 (4.2)	77	-	Years of education (M(SD)): 14.5(2.2)	≥30 minutes 5x/week during 6 months	Piano	Amateur
	Video games training group	8	66.9 (3.9)	50		Years of education (M(SD)): 17.5(2.3)		Video game : Super mario 64	Amateur
	No-contact control group	13	69.3 (5.7)	77		Years of education (M(SD)): 15.2(3.1)	N/A	N/A	N/A

Dubinsky (2019)	Choir singing group	34	67.6 (6.1)	91	-	-	2h/week sessions for 10 weeks + 1h vocal exercises for 10 weeks	Choir singing	Amateur
	No-contact control group	29	67.7 (4.9)	90			N/A	N/A	N/A
Fleming (2019)	Piano training group	13	67.7 (4.3)	83,3	French (Québec)	Years of education (M(SD)): 14.7(2.3)	≥30 minutes 5x/week during 6 months	Piano	Amateur
	Video games training group	8	69.3 (5.7)	50		Years of education (M(SD)): 15.2(3.2)		Video game : Super mario 64	Amateur
	No-contact control group	13	66.9 (3.9)	76,9		Years of education (M(SD)): 17.5(2.3)	N/A	N/A	N/A

Supplementary material 3: SPiN task characteristics and results

Table 3. SPiN tasks and results													
Article	Group comparison	SPiN task	Task type	Target: linguistic level	Target: vocal style (talker gender)	Target: language	Presentation type	Masker	Target-masker spatial localization	DV	Experimental manipulation	Condition with sign. better performance from the musician group	Condition without sign. better performance from musicians
Soncini (2006)	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken	Brazilian Portuguese	Headphones	Energetic masker (speech-shaped noise)	Colocated in right ear only (monaural presentation), colocated in left ear only (monaural presentation)	SRT-quiet, SNR level	SNR manipulation (adaptive procedure) No-noise condition	Target sentences in speech masker delivered monorally a) left ear only, b) right ear only (SNR level).	Target sentences in quiet delivered monorally a) left ear only, b) right ear only (SRT-quiet).
Fostick (2019)	Musicians VS. Card players VS. Controls	AB word test	Oral repetition	CVC words	Spoken	Hebrew	Headphones	Energetic masker (speech-shaped noise, white noise)	Colocated	% correct	Target manipulation (natural speech, compressed speech) Masker manipulation (speech-noise frequencies, white noise)	Every condition of the task	None

											SNR = 0 dB No noise condition		
Zhang (2019)	Musicians VS. Non- musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Headphones	Four- talker masker (one male, three females)	Colocated	SNR loss	SNR manipulati on (from 25 dB to 0 dB with 5 dB steps) Presentatio n manipulati on (right ear only, left ear only and binaurally)	None	Every condition of the task
Parbery -Clark (2011a)	Musicians VS. Non- musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeake r	Energetic masker (speech- shaped noise)	Colocated	SRT	SNR manipulati on (adaptive procedure)	HINT	None

Meha-Bettison (2018)	Musicians VS. Non-musicians	LiSN-S test	Oral repetition	Sentences	Spoken (female)	Australian English	Headphones	One-talker masker (continuous discourse; same voice as the target or a different female voice)	Colocated, $\pm 90^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure) Masker manipulation (same voice as the target, different voice) Azimuth manipulation (0° , $\pm 90^\circ$)	Same voice for target and masker sentences at 0° azimuth ("low-cue SRT"; no group effect, result of planned comparison)	Different voices for target and masker sentences at $\pm 90^\circ$ azimuth ("high-cue SRT"); Same voice for target and masker sentences at $\pm 90^\circ$; Different voices for target and masker sentences at 0° azimuth; Same voice for target and masker sentences at 0° azimuth
Varnet (2015)	Musicians VS. Non-musicians	Other	Syllable categorization (da or ga)	VCCV non-words	Spoken (male)	-	Headphones	Energetic masker (white noise)	Colocated	% correct, signal detection theory (sensitivity (d'), reaction time (RT), decision criteria (c)), SNR	SNR manipulation (adaptive procedure)	Target syllables in white noise masker (% correct, d', SNR)	Target syllables in white noise masker (RT, c)

Ruggles (2014)	Musicians VS. Non-musicians	Other	Keywords written repetition	Sentences (nonsense)	Spoken (female), whispered (same female), modified whispered speech (same female)	-	Headphones	Energetic maskers (1) continuous speech-shaped noise, 2) square-wave gated speech-shaped noise)	Colocated	% correct, masking release (between gated and continuous noise)	SNR manipulation (-6 dB, -3 dB, 0 dB) Target manipulation (voiced speech, whispered speech, modified whispered speech) Masker manipulation (continuous VS. gated noise)	None	Every condition of the task
	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Headphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (from 25 dB to 0 dB with 5 dB steps)	None	Quick-Sin
	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	None	HINT

Mankel (2018)	Musicians VS. Non-musicians (high- and low-PROMS)	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	-	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (from 25 dB to 0 dB with 5 dB steps)	Quick-Sin	None
Yoo & Bidelman (2019)	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Headphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None
	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	None	HINT
	Musicians VS. Non-musicians	WIN	Oral repetition	Monosyllabic words	Spoken (female)	North American English	Headphones (binaurally)	Four-talker masker	Colocated	SNR threshold	SNR manipulation (24 dB to 0 dB with 4 dB steps)	None	WIN
Anaya (2016)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (speech-shaped noise)	Colocated	% correct	Fixed SNR = -3dB	None	HINT & PRESTO composite score
	Musicians VS. Non-musicians	PRESTO	Oral repetition	Sentences	Spoken (different talkers)	-	Headphones	Six-talker masker	Colocated	% correct	Fixed SNR = 0 dB		

Nie (2018)	Musicians VS. Non-musicians	Other	Word identification (multiple choice)	Sentences	Spoken, sung (male)	-	Headphones	Energetic masker (steady noise)	Colocated in the right ear only (monaural presentation)	% correct	Target manipulation (spoken speech, sung speech with fixed pitch, sung speech with mixed pitch) Fixed SNR = 0 dB No-noise condition	Spoken words in steady noise; Sung words with mixed pitch in steady noise; Sung words with fixed speech in steady noise.	Spoken words in quiet; Sung words with mixed pitch in quiet; Sung words with fixed speech in quiet.
Parbery-Clark (2011b)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	HINT	None
	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None
	Musicians VS. Non-musicians	WIN	Oral repetition	Words	Spoken (female)	North American English	Insert earphones	Four-talker masker	Colocated	SNR threshold	SNR manipulation (24 dB to 0 dB with 4 dB steps)	WIN	None

Parbery-Clark (2012b)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	HINT	None
Strait (2012)	Musicians VS. Non-musicians (7-13 ans)	WIN	Oral repetition	Words	Spoken (female)	North American English	Loudspeaker	Four-talker masker	Colocated	SNR threshold	SNR manipulation (24 dB to 0 dB with 4 dB steps)	None	WIN
	Musicians VS. Non-musicians (7-13 ans)	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated, +90° azimuth - 90° azimuth	HINTleft/right composite score	SNR manipulation (adaptive procedure) Azimuth manipulation (0°, +90°, -90°)	HINTright (target sentences in speech shaped noise separated by +90° azimuth); HINTleft (target sentences in speech shaped noise separated by -90° azimuth)	HINTfront (target sentences in speech shaped noise colocated at 0° azimuth)
Parbery-Clark (2013)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	HINT	None

Du & Zatorre (2017)	Musicians VS. Non-musicians	Other	Identification (multiple choices)	CV syllables	Spoken (female)	North American English	Insert earphones	Energetic masker (white noise)	Colocated	% correct, RT	SNR manipulation (-12 dB, -8 dB, -4 dB, 0 dB, 8 dB) No-noise condition	CV syllables in white noise (%correct)	CV syllables in white noise (RT)
Swaminathan (2015)	Musicians VS. Non-musicians	Other	Word identification (multiple choice)	Sentences (syntactically correct, but semantically anomalous)	Spoken (female)	-	Headphones	One-talker masker (sentences ; female)	Colocated, $\pm 15^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure) Masker manipulation (forward or time-reversed) Azimuth manipulation (0° , -15° , $+15^\circ$)	Target sentences and forward sentence masker spatially separated ($\pm 15^\circ$ azimuth); Target sentences and reversed sentence masker colocated.	Target sentences and forward sentence masker colocated; Target sentences and reversed sentences masker spatially separated ($\pm 15^\circ$ azimuth)
Başkent & Gaudrain (2016)	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken (male)	Dutch	Headphones	One-talker masker (sentence sequences ; same male as the target)	Colocated	% correct	Masker manipulation (VTL shifted down by 0, 0.75, 1.5 semitones and F0 shifted up by 0, 4, 8 semitones) Fixed	Every condition of the task	None

											SNR = -6 dB		
Baskent (2018)	Musicians VS. Non-musicians	Other	Oral repetition	CVC words	Spoken (female)	Dutch	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated	% correct	SNR manipulation (10 dB, 5 dB, 0 dB) Target manipulation (natural VS. vocoded speech)	None	Every condition of the task
	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken (female)	Dutch	Loudspeaker	One-talker masker (sentences ; same female talker as the target or a male talker)	Colocated	SRT	SNR manipulation (adaptive procedure) Target manipulation (natural speech, vocoded speech) Masker manipulation (same female as the target, male)	None	Every condition of the task

Parbery-Clark (2009)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated, +90° azimuth -90° azimuth	SRT	SNR manipulation (adaptive procedure) Azimuth manipulation (0°, +90° azimuth, -90° azimuth)	HINTfront (target sentences in speech shaped noise colocated at 0° azimuth)	HINTright (target sentences in speech shaped noise separated by +90° azimuth); HINTleft (target sentences in speech shaped noise separated by -90° azimuth).
	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None

Boebinger (2015)	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken (female)	British English	Headphones	One talker maskers (1) clear speech, male talker), 2) spectrally rotated speech, male talker); Energetic maskers (1) speech amplitude modulated noise, 2) steady speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure) Masker manipulation (clear speech, spectrally rotated speech, speech amplitude modulated noise, speech spectrum steady noise)	None	Every condition of the task
Madsen (2017)	Musicians VS. Non-musicians	Other	Written word repetition	Sentences	Spoken (male)	British English	Headphones	Two-talker masker (sentences); Speech-shaped noise	Colocated	% correct	Target manipulation (natural speech, monotone speech) Masker manipulation (natural speech with F0 equal or 1/2/4/8 semitones lower than the target), monotone	None	Every condition of the task

											speech, speech- shaped noise) Fixed SNR = -3 dB		
Parbery -Clark (2012a)	Musicians VS. Non- musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four- talker masker (one male, three females)	Colocated	SNR loss	SNR manipulati on (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None
Skoe (2019)	Musicians VS. Non- musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four- talker masker (one male, three females)	Colocated	SNR loss	SNR manipulati on (25 dB to 0 dB with 5 dB steps)	Quick-Sin (with noise exposure as a covariate)	Quick-Sin (without noise exposure as a covariate)

Deroche (2017)	Musicians VS. Non-musicians	Other	Written repetition	Sentences	Spoken (male)	-	Headphones	Two-talker masker (sentences, same male as the target); Speech-modulated buzzes	Colocated (diotic); spatially separated (dichotic)	SRT	SNR manipulation (adaptive procedure) Manipulation of the difference in fundamental frequency (0, -2, -8 semitones) Masker manipulation (speech-modulated buzzes, two talker sentences) Presentation manipulation (diotic, dichotic)	None	Every condition of the task
-----------------------	-----------------------------	-------	--------------------	-----------	---------------	---	------------	---	--	-----	--	------	-----------------------------

		Other	Written repetition	Sentences	Spoken (male)		Headphones	Two-talker masker (sentences, same male as the target); Speech-modulated buzzes	Colocated	SRT	SNR manipulation (adaptive procedure) Manipulation of the difference in fundamental frequency between target and masker (0, -2, -8 semitones) Masker manipulation (speech-modulated buzzes, two talker sentences) Presentation manipulation (with priming, without priming)	None	Every condition of the task
--	--	-------	--------------------	-----------	---------------	--	------------	---	-----------	-----	---	------	-----------------------------

		Other	Written repetition	Sentences	Spoken (male)	Headphones	Two-talker masker (sentences, same male as the target); Speech-modulated buzzes	Colocated	SRT	SNR manipulation (adaptive procedure) Manipulation of the difference in fundamental frequency (-2, 0, +2, ± 2 semitones) Masker manipulation (masker type: speech-modulated buzzes, two talker sentences; masker roving: fundamental frequency fixed at 125 Hz, variable over ten logarithmic steps between 100 and 150 Hz)	Target sentences in two-voice babble with fixed and variable F0; Target sentences in buzzes with variable F0.	Target sentences in buzzes with fixed F0.
--	--	-------	--------------------	-----------	---------------	------------	---	-----------	-----	---	---	---

		Other	Written repetition	Sentences	Spoken (male)	Headphones	Two-talker masker (sentences, same male as the target); Speech-modulated buzzes	Colocated	SRT	SNR manipulation (adaptive procedure) Manipulation of the difference in fundamental frequency (-8, 0, +8, ±8 semitones) Masker manipulation (masker type: speech-modulated buzzes, two talker sentences; masker roving: fundamental frequency fixed at 125 Hz, variable over ten logarithmic steps between 100 and 150 Hz)	Every condition of the task	None
--	--	-------	--------------------	-----------	---------------	------------	---	-----------	-----	--	-----------------------------	------

Madsen (2019)	Musicians VS. Non-musicians	Other	Word identification (multiple choice)	Sentences	Spoken (3 different females, one at a time)	Danish	Loudspeaker	One-talker masker (sentences ; same 3 females as the target, one at a time)	Colocated, $\pm 15^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure) Azimuth manipulation (0° , $\pm 15^\circ$)	None	Every condition of the task
	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken (male)		Loudspeaker	Two-talker babble (two males); Energetic masker (speech-shaped noise)	Colocated, $\pm 15^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure) Masker manipulation (two-talker babble; Gaussian noise) Azimuth manipulation (0° , $\pm 15^\circ$) Reverberation manipulation (anechoic, reverberant)	None	Every condition of the task
Slater (2018)	Percussionists VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	American English	Loudspeaker	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None

Jain (2019)	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	Kannada	Headphones	Four-talker masker (one male, three females)	Colocated	SRT	SNR manipulation (+8 dB to -10 dB with 3 dB steps)	None	Quick-Sin
Escobar (2019)	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	None	Quick-Sin
	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Loudspeaker	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)	None	HINT
	Musicians VS. Non-musicians	Spin-R	Oral repetition	Sentences	Spoken (male)	-	Insert earphones	Twelve-talker masker	Colocated	% correct for final word	Target manipulation (high-predictability final word, low-predictability final words); Fixed SNR level = +2 dB	None	SPIN-R test
Zendel (2015)	Musicians VS. Non-musicians	Other	Oral repetition	CVC words	Spoken (male)	French (Québec)	Insert earphones	Four-talker masker (two women, two men)	Colocated	% correct	SNR manipulation (0 dB, 15 dB) No-noise condition	Target words in four-talker babble at SNR = 0 dB	Target words in four-talker babble at SNR = 15 dB; Target words in quiet

Morse-Fortier (2017)	Musicians VS. Non-musicians	Other	Recall of target words	CVC words	Spoken (female)	North American English	Loudspeaker	Two-talker babble (nonsense sentences; female)	Colocated, colocated + (+60°) azimuth	SNR threshold	SNR manipulation (adaptive procedure) Target and masker manipulation (natural speech, vocoded speech) Azimuth manipulation (masker and speech colocated at 0°, masker and speech colocated at 0° + masker at +60° azimuth)	Every condition of the task	None
Coffey (2019)	Musicians VS. Non-musicians	Modified HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (Speech-shaped noise)	Colocated	% correct	SNR manipulation (-2 dB to -7 dB with 1 dB steps)	Modified HINT	None
Fuller (2014)	Musicians VS. Non-musicians	Other	Oral repetition	CVC Words	Spoken (female)	Dutch	Loudspeaker	Energetic masker (Speech-shaped noise)	Colocated	% correct (phones)	SNR manipulation (0 dB, 5 dB, 10 dB) Target manipulation	Every condition of the task	None

											on (natural speech, CI simulation speech) No-noise condition		
	Musicians VS. Non-musicians	Other	Oral repetition	Sentences	Spoken (female)	Dutch	Loudspeaker	Energetic masker (Steady speech-shaped noise; fluctuating speech-shaped noise); Six-talker babble	Colocated	SRT	SNR manipulation (adaptive procedure) Target manipulation (natural speech, CI simulation speech) Masker manipulation (steady speech shaped noise, fluctuating speech shaped noise, 6 talker babble)	None	Every condition of the task
Slater (2016)	Percussionist and Vocalists VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	North American English	Insert earphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None

	Percussionist and Vocalists VS. Non-musicians	WIN	Oral repetition	Words	Spoken (female)	North American English	Insert earphones	Four-talker babble	Colocated	SNR threshold	SNR manipulation (24 dB to 0 dB with 4 dB steps)	None	WIN
Clayton (2016)	Musicians VS. Non-musicians	Other	Word identification (multiple choice)	Sentences (syntactically correct, but semantically anomalous)	Spoken (female)	-	Headphones	One-talker masker (sentences, different female as the target)	Colocated, $\pm 15^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure) Azimuth manipulation ($0^\circ, \pm 15^\circ$)	Target sentences spatially separated from the masker.	Target sentences colocated with the masker.
Mandikal-Vasuki (2016)	Musicians VS. Non-musicians	LiSN-S test (2 conditions)	Oral repetition	Sentences	Spoken (female)	Australian English	Headphones	One-talker masker (continuous discourse; same voice as the target or a different female voice)	Colocated	SRT	SNR manipulation (adaptive procedure) Masker manipulation (same female, different female as the target)	None	Every condition of the task
Bidelman & Yoo (2020)	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	American English	Headphones	Four-talker masker (one male, three females)	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	Quick-Sin	None

	Musicians VS. Non-musicians	Other	Recall of callsigns (colour-number combination)	Sentences	Spoken (female (50%) or male (50%))	-	Circular array of 16 loudspeakers	Two- to eight-talker masker	Spatially separated (from +/- 20° to +/- 180°)	Accuracy, RT, Localization error	Masker manipulation (number of different maskers from 1 to 8 talkers, with each talker in a different loudspeaker) Azimuth manipulation (random spatial separation of targets and maskers from +/- 20° to +/- 180°) SNR = 0dB	Target sentences with rising masker numbers (accuracy, reaction time and localization error).	N/A
Perron (2021)	Singers VS. Non-singers	Other	Discrimination	CVC syllables	Spoken (male)	French (Québec)	Headphones	Multi-talker masker	Colocated	Sensitivity (d') (signal detection theory)	SNR manipulation (-3 dB, +3 dB)	None	Target CVC syllables in multi-talker babble noise (average d' between -3dB and +3dB)

Studies with correlation/regression :													
Vanden Bosch der Nederlanden (2018)	Musicians VS. Non-musicians	Revised speech-in-noise (R-Spin)	Oral repetition of the last word in a carrier sentence	Sentences	Spoken (male)	-	Headphones	Multi-talker masker	Colocated (monaural presentation, left ear)	% correct (at -1 dB SNR)	SNR manipulation (-1 dB to 23 dB with 3 dB steps) Target manipulation (high predictability carrier sentences, low predictability carrier sentences)		

Puschmann (2019)	Continuum	Other	Word detection	Sentences	Spoken (male)	American English	Insert earphones	One-talker masker (sentences ; male)	Colocated	F1 score (= harmonic mean of recall (i.e., number of detected target words divided by the total number of targets words) and precision (number of detected target words divided by the total number of button presses))	Fixed SNR = 0 dB No-noise condition		
	Continuum	HINT	Oral repetition	Sentences	Spoken (male)	British English	-	Energetic masker (speech-shaped noise)	Colocated	SRT	SNR manipulation (adaptive procedure)		

	Continuum	Other	Oral repetition	Sentences	Spoken (3 different males)	-	-	One-talker masker (sentences ; 3 same males as the target)	Colocated	SRT	SNR manipulation (adaptive procedure)		
Yeend (2017)	Continuum	LiSN-S test (2 conditions)	Oral repetition	Sentences	Spoken (female)	Australian English	Headphones	One-talker masker (continuous discourse; female)	$\pm 90^\circ$ azimuth	SRT	SNR manipulation (adaptive procedure)		
Coffey (2017)	Continuum	Modified HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (speech-shaped noise)	Colocated	% correct	SNR manipulation (2 dB, -2 dB, -6 dB)		
Zendel & Alain (2012)	Musicians VS. Non-musicians	QuickSIN	Oral repetition	Sentences	Spoken (female)	American English	Insert earphones	Four-talker babble	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)		
Longitudinal studies :													

Hutchins (2018)	Children pre-training VS. same children post training.	Modified Speech Discrimination in Noise subtest (from the ASA)	Word-picture matching	Words	Spoken	-	Headphones	Multi-talker babble	Colocated	Number of correct answers at the final SNR level	SNR manipulation (3 different SNR - level not precised)	
Slater (2015)	Musicians VS. Non-musicians	HINT	Oral repetition	Sentences	Spoken (male)	British English	Headphones	Energetic masker (Speech-shaped noise)	Colocated	SNR threshold	SNR manipulation (adaptive procedure)	
MacCutcheon (2020)	Musicians VS. Non-musicians (longitudinal study musical training VS. sport training)	Other	Oral word repetition (numbers and colours)	Sentences	Spoken (male)	British English	Headphones	Energetic masker (speech-shaped noise) One-talker masker (male)	Colocated, +90° azimuth	SRT	SNR manipulation (adaptive procedure) Masker manipulation (speech-shaped noise; one-talker masker) Azimuth manipulation (0°, +90°)	
Zendel (2019)	Musicians VS. Video game players VS. Controls (longitudinal study)	Other	Oral repetition	Monosyllabic words	Spoken (male)	-	Insert earphones	Four-talker babble	Colocated	% correct	SNR manipulation (0 dB, 15 dB) No-noise condition	

Dubinsky (2019)	Choir singers VS. do nothing controls (longitudinal)	QuickSIN	Oral repetition	Sentences	Spoken (female)	American English	Headphones	Four-talker babble	Colocated	SNR loss	SNR manipulation (25 dB to 0 dB with 5 dB steps)	
Fleming (2019)	Piano players VS. Video game players VS. No contact group (longitudinal)	Other	Sentence-picture matching	Sentences	Spoken (male)	French (Québec)	Insert earphones	Multi-talker babble	Colocated	% correct	SNR manipulation (5 dB, 20 dB) No noise condition	

Chapitre 4 Étude 2

Effet des activités musicales et non musicales sur la PPB

Ce chapitre rapporte les résultats préliminaires d'une étude comportementale en PPB issue d'un projet qui s'intéresse aux effets des activités musicales (chant, instrument de musique) et non musicales dans le cerveau adulte, sur la cognition, le langage et les processus émotionnels. Ce projet est composé de deux volets, un volet comportemental, qui inclut la tâche en PPB, et un volet en neuro-imagerie.

Le recrutement pour ce projet a commencé en juin 2019 et a dû être interrompu cinq mois en raison des mesures liées à la COVID-19. Ce chapitre présente par conséquent les résultats sur 70 participants recrutés en date du 30 avril 2021, ce qui correspond à 58 % de l'échantillon total attendu (70/120 participants). Le recrutement est en cours et se terminera fin décembre 2021.

4.1 Introduction

Le déclin de la PPB qui accompagne le vieillissement nuit à la participation sociale des personnes âgées (Working Group on Speech & Aging, 1988). En effet, les adultes âgés tendent à adopter des comportements d'évitement envers les situations dans lesquelles ils éprouvent de la difficulté à communiquer, telles que celles qui se déroulent en groupe et/ou dans des environnements bruyants (Heine & Browning, 2002). Chez les aînés, la participation à des activités sociales est associée à de meilleurs scores dans des évaluations globales de la cognition (Kelly et al., 2017; Krueger et al., 2009), tandis que des mesures du sentiment de solitude sont associées à des effets néfastes, tant au niveau cognitif qu'au niveau affectif (Santini et al., 2020; Taylor et al., 2018). Trouver des stratégies efficaces de prévention/réadaptation pour la PPB est donc primordial pour conserver une qualité de vie satisfaisante chez les aînés.

La pratique d'activités musicales, par les avantages démontrés chez des musiciens jeunes et âgés dans le traitement auditif central (Deguchi et al., 2012; Donai & Jennings,

2016; Grassi et al., 2017; Hutka et al., 2015; Kishon-Rabin et al., 2001; Kuhn et al., 2013; Kumar et al., 2014; Liang et al., 2016; Sares et al., 2018; Schon et al., 2004; Strait et al., 2010; Zendel & Alain, 2009; B. R. Zendel & C. Alain, 2013) et certaines fonctions cognitives (Grassi et al., 2017; Groussard et al., 2020; Hanna-Pladdy & Gajewski, 2012; Mansens et al., 2018; Zuk et al., 2014), suscite un intérêt particulier en tant qu'outil de prévention/réadaptation potentiel pour la capacité à percevoir la parole dans le bruit.

Tel que présenté dans la méta-analyse du chapitre précédent, des bénéfices de la pratique musicale (chant/instrument de musique) sont retrouvés en condition de BM énergétique, de BM de parole avec 2 et 4 locuteurs, et dans des conditions de SNR à 0 dB ou de SNRs négatifs. Ces avantages sont retrouvés chez des musiciens qui sont pour la plupart de jeunes adultes ayant un minimum de 7 ans de pratique et ayant commencé leur pratique majoritairement entre l'âge de 7 ans et 10 ans. Dans la plupart des études, ces avantages ont été trouvés en comparant des musiciens avec des participants contrôles ne pratiquant pas d'activités précises. Dans la majorité des études, les tâches de PPB présentent des mots ou des phrases, c'est-à-dire que les avantages sont trouvés en présence de cibles lexicales.

Sur le plan de la pratique musicale, malgré la relative homogénéité dans les critères d'inclusion pour l'âge de début et le nombre d'années de pratique, il ne ressort pas de profil complet du régime de pratique qui mène à ces avantages. Par exemple, il n'est pas possible de tirer des conclusions sur le niveau de pratique nécessaire (amateur/professionnel) ainsi que sur la fréquence et la durée des pratiques qui mènent à ces avantages. De plus, les participants étant surtout de jeunes adultes, il n'est de surcroît pas possible d'affirmer que ces avantages se retrouvent chez des adultes âgés. Sur le plan expérimental, il est difficile de cerner quel/s niveau/x de traitement de la PPB est/sont influencés par la pratique musicale, puisque les cibles des tâches de PPB sont essentiellement de niveau lexical. Ces différents points seront abordés plus amplement dans les prochains paragraphes.

Premièrement, la revue systématique fait ressortir le manque d'études incluant des adultes âgés de plus de 60 ans, bien que ce soit cette population qui soit affectée par un déclin de la PPB. Depuis la recherche effectuée pour la revue systématique, il y a ce jour deux études transversales (Mussoi, 2021; Zhang et al., 2020) et trois études longitudinales

(Hennessy et al., 2021; Merten et al., 2021; Worschech et al., 2021) qui ont été publiées sur des participants âgés.

Un autre problème majeur soulevé dans la revue systématique est le manque d'information sur le niveau de pratique des musiciens (amateur VS professionnel). Lorsque le niveau de pratique des participants est toutefois mentionné, les groupes de musiciens sont composés majoritairement d'étudiants dans des programmes musicaux de niveau universitaire ou de musiciens professionnels, donc dans les deux cas de musiciens experts, ou à la fois de musiciens experts et de musiciens amateurs. Dans beaucoup d'études s'étant intéressées à la plasticité structurelle, à la cognition ou à l'audition centrale des musiciens, les données sont également récoltées chez des musiciens experts (Amer et al., 2013; Anaya et al., 2016; Grassi et al., 2017; Pantev et al., 1998; Schlaug, 2001; Schlaug et al., 1995; Zendel & Alain, 2009). Parmi les études sur les musiciens experts, certaines ont montré que les différences entre les musiciens et les non-musiciens dans les variables étudiées étaient corrélées avec des indices de la pratique, comme l'âge de début (Pantev et al., 1998; Schlaug, 2001) ou le nombre d'années de pratique (Anaya et al., 2016). Comparés à des musiciens amateurs, les musiciens experts ont une intensité de pratique beaucoup plus importante, avec des pratiques beaucoup plus longues, fréquentes et spécialisées. En effet, arrivés à l'âge adulte, les musiciens experts peuvent avoir accumulé plus de 10 000 heures de pratique structurée, personnalisée et dédiée entièrement à l'amélioration de techniques spécifiques (Ericsson, 2020; Ericsson & Harwell, 2019; Ericsson et al., 1993; Krampe & Ericsson, 1996). Les avantages trouvés avec de telles pratiques seront-ils présents chez des musiciens amateurs ? Pour le développement de programmes de prévention, cette question est essentielle. Il faut en effet pouvoir établir la présence d'avantages obtenus à partir de pratiques dont les paramètres (années de pratique, fréquence et durée des entraînements) sont réalistes pour la population générale. Par ailleurs, les avantages trouvés chez des musiciens experts pourraient également résulter non seulement de leur intensité de pratique, mais également d'un terrain génétique favorable, par exemple pour le traitement auditif ou certaines fonctions cognitives. Afin de diminuer la probabilité de recruter des participants prédisposés génétiquement pour le traitement auditif et/ou cognitif et ainsi évaluer les effets d'une pratique plus récréative et accessible, le recrutement de musiciens amateurs est nécessaire.

Dans les articles de la revue de littérature, les musiciens sont en général comparés à un groupe contrôle composé de participants sélectionnés sur la base de leur quasi-inexistence de pratique musicale (généralement moins de 2 ans de pratique). Très peu nombreuses sont les études qui ont recruté un groupe contrôle formé de non-musiciens qui pratiquent une activité non musicale donnée. Une étude a comparé les musiciens avec un groupe composé d'acteurs de théâtre (Jain & Nataraja, 2019) et une autre étude a comparé les musiciens avec des joueurs de cartes (Fostick, 2019). Il s'agit de trop peu d'études pour établir avec certitude la supériorité de la pratique musicale sur la PPB comparée à d'autres activités. La pratique d'activités musicales intéresse surtout par son entraînement cognitif, mais surtout par son entraînement auditif. Or, le lien entre ces différents entraînements et un avantage en PPB n'est aujourd'hui pas encore formellement établi. Des corrélations entre la performance en PPB et des tâches ciblant le traitement auditif central, comme la discrimination de fréquence, sont trouvées dans une étude (Meha-Bettison et al., 2018) tandis que d'autres n'en ont pas trouvé (Boebinger et al., 2015; Ruggles et al., 2014). Certaines études ont trouvé des corrélations entre certaines fonctions cognitives chez les musiciens, comme la mémoire de travail, et le score en PPB (Parbery-Clark et al., 2009; Yoo & Bidelman, 2019), tandis que d'autres n'en ont pas trouvé (Du & Zatorre, 2017; Slater & Kraus, 2016). Considérant ces résultats partagés, il est envisageable que l'entraînement auditif ne soit pas essentiel et que seul un entraînement cognitif suffise pour obtenir des bénéfices en PPB. Par conséquent, d'autres activités complexes qui entraînent les fonctions cognitives pourraient améliorer la PPB. Par exemple, le yoga, dont la pratique sur six sessions de 60 minutes améliore le score dans différentes tâches de mémoire de travail (Digit Span, Listening Span) Brunner et al. (2017).

Concernant l'aspect expérimental, la revue systématique a montré que dans les tâches de PPB, les cibles utilisées étaient essentiellement des cibles lexicales comme des mots et surtout des phrases. Néanmoins, très peu d'études utilisaient des cibles sous-lexicales. En effet, seulement trois études se sont intéressées à la perception de syllabes. Or, l'utilisation du contexte linguistique est facilitante dans les tâches de perception de mots ou de phrases dans le bruit, car le contexte permet de compenser une entrée auditive incomplète, par exemple grâce à ses connaissances syntactiques et sémantiques, et tout de même deviner les

mots masqués (Benichov et al., 2012; Buss et al., 2019; Cervera & Rosell, 2015; Hutchinson, 1989). Il est ainsi probable que les scores aux tests de PPB avec des phrases ne reflètent pas uniquement le processus de perception en tant que tel, mais également la capacité à utiliser ses connaissances linguistiques pour restaurer l'information manquante.

Il est par conséquent, à ce stade, encore difficile de tirer des conclusions claires sur les effets de la pratique musicale chez les personnes âgées, sur le niveau et les paramètres de pratique requis pour obtenir des avantages en PPB, sur le/les niveaux/x de traitements de la PPB touchés par la pratique musicale et sur la supériorité de la pratique musicale sur la PPB comparée à d'autres activités complexes.

4.2 Objectifs et hypothèses

Le premier objectif de cette étude était d'explorer les effets d'activités musicales (chant, instrument de musique) et non musicales psychomotrices, toutes deux de niveau amateur, sur la PPB à différents niveaux de SNR chez des adultes jeunes et âgés. Le deuxième objectif était d'évaluer, chez les musiciens, la contribution de l'âge du début et du nombre d'années de pratique sur la performance en PPB.

Recruter des participants ayant une pratique de niveau amateur a permis d'évaluer les effets issus d'une pratique récréative, potentiellement implantable dans le cadre de programmes de prévention. La présence d'un groupe contrôle pratiquant des activités psychomotrices non musicales a permis de comparer la pratique musicale avec une autre activité de loisirs sollicitant les fonctions exécutives et leur coordination avec des programmes moteurs, mais contrairement aux activités musicales, n'incluant pas d'entraînement auditif.

La première hypothèse de cette étude était que l'âge serait associé à une diminution de la performance en PPB, peu importe le type d'activité pratiquée et peu importe la condition de SNR. La deuxième hypothèse était que la pratique musicale serait associée avec une diminution de l'effet de l'âge sur la PPB. La troisième hypothèse était que, chez les

musiciens, l'âge de début ainsi que le nombre d'années de pratique musicale seraient tous deux associés à la performance en PPB.

4.3 Matériel et méthodes

Cette étude est issue d'un projet s'intéressant aux effets des activités musicales et non musicales sur le cerveau adulte. Ce projet, sous la supervision de la Prof. Pascale Tremblay, a été évalué et approuvé par le Comité d'éthique de la recherche sectorielle en neurosciences et santé mentale du CIUSSS-CN (#2019-1733).

4.3.1 Participants

La présente étude inclut 70 participants (recrutés en date du 30.04.2021) en bonne santé auto rapportée et divisés en quatre groupes : un groupe de musiciens jeunes, un groupe psychomoteurs jeunes, un groupe de musiciens âgés et un groupe de psychomoteurs âgés (voir Tableau 1)¹. Les musiciens pratiquaient le chant ou un/plusieurs instruments de musique à niveau amateur. Les psychomoteurs pratiquaient une/plusieurs activités psychomotrices (voir 4.1.1), également au niveau amateur. Les participants définis comme étant « jeunes » avaient entre 20 et 60 ans, tandis que les participants définis comme étant « âgés » avaient 65 ans ou plus. Les 4 groupes étaient appariés sur le score de latéralité (Oldfield, 1971). Toutefois, des différences significatives ont été trouvées entre les groupes pour un score d'audition (PTA oreille droite, voir 4.3.2.1), pour les années d'éducation ainsi que pour le score global au MoCA (voir tableau 1). Ces mesures ont donc été ajoutées en co-variables dans les analyses.

¹ L'objectif général du projet dont cette étude est issue est de comparer des chanteurs, des instrumentistes et des psychomoteurs. Toutefois, en date du 30 avril 2021, les participants recrutés incluaient 35 psychomoteurs (20 jeunes et 15 âgés), 33 chanteurs (8 jeunes et 25 âgés) et 30 instrumentistes (19 jeunes et 11 âgés). Afin d'obtenir le même nombre de participants dans tous les groupes, deux groupes de musiciens (musiciens jeunes, musiciens âgés) a été formé en regroupant des instrumentistes (n=20) et des chanteurs (n=15). Les groupes de psychomoteurs jeunes et de musiciens jeunes sont appariés sur l'âge, tout comme les groupes de psychomoteurs âgés et de musiciens âgés (tous $p < .737$).

Tableau 1. Données démographiques				
Groupes	<u>Musiciens jeunes</u>	<u>Musiciens âgés</u>	<u>Psychomoteurs jeunes</u>	<u>Psychomoteurs âgés</u>
N	20	15	20	15
Âge M(ET)	41.85 (11.38)	73.07 (5.71)	39.25 (12.13)	71.40 (5.99)
PTA oreille droite M(ET)	10.083 (8.76)	31.055 (13.59)	10.208 (7.23)	22.055 (9.77)
PTA oreille gauche M(ET)	10.875 (8.16)	31.166 (16.02)	8.583 (6.76)	24.667 (10.30)
MoCA M(ET)	28.15 (1.60)	26.20 (1.08)	28.35 (1.18)	27.20 (2.18)
Latéralité M(ET)	95.50 (7.59)	94.33 (8.42)	89.75 (13.03)	97.67 (4.58)
Education M(ET)	14.45 (2.30)	15.13 (2.26)	14.65 (2.28)	16.67 (2.19)

Note. M(ET) = moyenne (écart-type), PTA = pure tone average, MoCA = Montreal Cognitive Assessment, Education = années d'éducation. Latéralité mesurée à l'aide du test Oldfield (voir section 4.3.1.1).

4.3.1.1 Critères de sélection

Tous les participants ont pris part aux deux volets du projet (volet comportemental, volet en neuro-imagerie) et devaient par conséquent tous répondre aux critères pour participer à un examen en IRM. Les critères d'exclusion généraux étaient les suivants : ne pas avoir le français québécois comme langue principale², être atteint d'une maladie neurologique, être atteint d'un trouble psychiatrique, présenter un trouble cognitif (exclusion avec le Montréal Cognitive Assessment (Nasreddine et al., 2005)), être traité avec un ou plusieurs médicaments psychotropes, avoir présenté par le passé ou présenter actuellement un trouble diagnostiqué du langage et/ou de la voix, présenter un trouble de l'audition diagnostiqué, porter un appareil auditif, souffrir d'un acouphène. Les critères d'exclusion spécifique pour un examen en IRM sont les suivants : la présence de tout corps étranger métallique dans le corps, le port d'un stimulateur électrique, la claustrophobie et la grossesse. De plus, l'examen en IRM s'intéressant aux zones impliqués dans la parole, les participants devaient tous être gauchers (score au test Oldfield (Oldfield, 1971) ≥ 14).

Sur le plan des activités, les activités musicales incluses étaient le chant et la pratique d'un instrument de musique de niveau amateur. Pour les activités psychomotrices, l'étude s'intéressant spécifiquement aux effets des activités musicales (chant, instrument de musique), des loisirs non musicaux divers et faisant appel aux fonctions exécutives et à la motricité ont été inclus, comme le golf, la pétanque, le curling, les quilles, le yoga, le tricot, les jeux vidéo ou encore le billard.

Les critères d'inclusion pour les musiciens et les psychomoteurs étaient les suivants : pratiquer son activité (musique/activité psychomotrice) depuis au moins 6 ans, dont au moins 5 ans en continu avec une pratique d'au minimum 3 heures/semaine³. Afin de ne pas

² Les participants doivent répondre à l'une de ces conditions:

- avoir le français québécois comme langue maternelle
- avoir le français d'une autre région linguistique francophone comme langue maternelle, mais avoir demeuré au Québec depuis au moins 10 ans et ne pas avoir d'accent autre que l'accent québécois.

³ Une étude effectuée dans notre laboratoire sur des chanteurs avec au minimum 1h de chant par semaine et 2 ans minimum de pratique avait montré des résultats mitigés ; nous avons par conséquent décidé d'augmenter le minimum d'heures/semaine d'activité requis, ainsi que le nombre d'années minimum d'activité.

avoir de chevauchement d'activités entre les musiciens et les psychomoteurs, un participant psychomoteur ne pouvait pas pratiquer d'activité musicale (chant/instrument de musique) à raison de plus de la moitié du nombre d'heures par semaine de son activité psychomotrice, et vice-versa pour une activité psychomotrice chez un participant musicien. De plus, pratiquer une autre activité musicale en plus du chant et d'un instrument, telle que la danse ou le patinage artistique, à raison de plus de la moitié du nombre d'heure par semaine de la pratique principale (musicale ou psychomotrice) était également un critère d'exclusion. Si un musicien potentiel jouait 5h de son instrument par semaine et qu'il pratiquait la danse 3h par semaine, il était par conséquent exclu.

4.3.1.2 Recrutement des participants

Le recrutement s'est fait dans la ville de Québec au moyen d'appels téléphoniques, de courriels de recrutement, de la diffusion d'affiches et de courtes présentations du projet. Les courriels de recrutement étaient destinés à des organismes en lien avec les activités recherchées, comme des écoles de musique ou des chorales, ou à des participants potentiels présents dans une base de données du laboratoire. Cette dernière rassemble les coordonnées de personnes ayant déjà participé à un projet du laboratoire et ayant donné leur accord pour être contactées dans le cadre d'une prochaine étude. La base de données réunit également les coordonnées de personnes n'ayant jamais participé à des projets du laboratoire, mais ayant exprimé leur intérêt à participer à une étude. En plus du contact par courriel, ces participants potentiels pouvaient également être contacté par téléphone, à l'aide d'un script standardisé. La diffusion d'affiches s'est faite au moyen d'affiches papiers et numérique. Les premières ont été posées dans différents lieux publics de la ville de Québec, tels que des commerces, ou alors de façon plus ciblée dans des lieux propres aux activités recherchées pour le projet. Des affiches ont donc été posées, par exemple, dans des écoles de musique, des centres communautaires proposant des activités psychomotrices, des arénas de curlings ou encore des salons de quilles. Les affiches numériques ont été diffusées soit par courriel aux étudiants de l'Université Laval et à des organismes d'intérêts, attachée à un courriel standardisé de recrutement, soit par Facebook à divers groupes/pages ciblant spécifiquement les activités du projet ou alors à des groupes de petites annonces. La dernière méthode de

recrutement était la présentation du projet par un ou deux étudiants de l'équipe dans différents lieux après avoir obtenu l'autorisation des responsables (p.ex. clubs de golfs, centres d'hébergement). La présentation ne dépassait pas 5 minutes et incluait le contexte et le but du projet, ainsi qu'un aperçu de la méthodologie.

Toute personne intéressée à participer à ce projet était invitée à passer une entrevue téléphonique de sélection. Lors de celle-ci, les critères d'inclusion et d'exclusion étaient vérifiés à l'aide de questionnaires. Un premier questionnaire documentait l'état de santé, les potentielles contre-indications à un examen IRM, les années d'éducation, la maîtrise de différentes langues, les activités sociales et les activités psychomotrices en termes d'années et de fréquence de pratique. Un deuxième questionnaire documentait les différentes pratiques musicales, également en termes d'années et de fréquence de pratique, mais également en termes de contexte (formation, pratique de groupe, pratique solo). En plus du chant et de la pratique d'un instrument de musique, la pratique d'autres activités musicales comme la danse ou le patinage artistique était vérifiée. Les habitudes d'écoute musicale des participants étaient également documentées. Finalement, la latéralité était évaluée à l'aide du questionnaire Oldfield (Oldfield, 1971). À la fin de cette entrevue, si le participant potentiel répondait aux critères d'inclusion et ne présentait pas de critères d'exclusion, celui-ci était invité au laboratoire des neurosciences de la parole et de l'audition, situé dans le Centre de Recherche CERVO (2301 avenue d'Estimauville, Québec G1E 1T2) pour le volet comportemental du projet. Cette visite commençait par la passation du MoCA qui terminait la sélection du participant potentiel. Toute personne de 40 ans et moins était exclue avec un score au MoCA inférieur à 26 (Nasreddine et al., 2005) et toute personne de 41 ans et plus était exclue avec un score Z au-dessous de -1.33 (Larouche et al., 2016).

4.3.2 Procédure

Tout participant ayant été inclus après la passation du MoCA était éligible aux deux volets du projet et pouvait alors procéder avec la suite de la visite du volet comportemental. Le projet s'intéressant aux effets des activités sur le langage, la cognition et les émotions, la visite comportementale consistait en différents tests cognitifs, l'évaluation de la perception

de la parole dans le bruit, des évaluations de la production de la parole, un test de reconnaissance des émotions et une évaluation de l'audition. Une partie des tests cognitifs était passée dans une salle d'entrevue et les autres tâches, dont l'évaluation de la perception de la parole dans le bruit et celle de l'audition, étaient effectuées dans la salle insonorisée du laboratoire. La présente étude se concentre sur l'évaluation de la PPB.

4.3.2.1 Évaluation de l'audition : audiométrie tonale

Les seuils auditifs des participants étaient évalués par audiométrie tonale aérienne (AC40, Interacoustic) indépendamment pour chaque oreille dans les fréquences suivantes : .5, 1, 2, 3, 4 et 6 kHz. La moyenne en sons pur était ensuite calculée pour l'oreille gauche (PTA OG) et pour l'oreille droite (PTA OD) en moyennant les seuils obtenus à chaque fréquence (voir Tableau 1).

4.3.2.2 Tâche de perception de la parole dans le bruit

La tâche de PPB était une tâche de discrimination de syllabes consonnes-voyelles-consonnes (CVC), présentées dans un bruit masquant de parole à quatre locuteurs selon trois niveaux de SNR. Des syllabes plutôt que des mots ou des phrases ont été choisies afin d'évaluer spécifiquement la perception de la parole plutôt que l'utilisation du contexte linguistique pour compenser les effets du BM. Un BM à quatre locuteurs a été choisi puisque les personnes âgées ont dans la plupart des cas plus de difficulté dans un BM de parole que dans un BM énergétique.

174 paires de syllabes étaient présentées dans la tâche. Les paires étaient composées de syllabes identiques (86 paires), comme la paire [myz] – [myz], ou de syllabes différentes. Les paires pouvaient différer sur l'attaque (44 paires), comme dans la paire [vɛs] – [zɛs], ou alors sur la coda (44 paires), comme dans la paire [kab] – [kad]. Les syllabes qui composaient la paire étaient soit produites par le même locuteurs (86 paires), soit par deux locuteurs différents (88 paires). 91.3% des 174 paires de syllabes ne contenaient aucun mot unisyllabique. 11 paires incluaient un mot (p. ex. : [vis] - [zis], **vis**) et neuf paires paires étaient formées de deux mots (p. ex : [dus] - [gus], **douce-gousse**).

Les syllabes ont été choisies à partir du corpus SyllabO+, qui réunit plus de 360 000 syllabes du français québécois oral (<http://syllabo.speechneurolab.ca>) (Bédard et al., 2017). Les syllabes ont été produites par deux locuteurs différents, deux hommes de 22 ans et de 21 ans parlant le français québécois. Afin de conserver une prosodie neutre tout au long de l'enregistrement des syllabes, les locuteurs devaient prononcer la phrase « Maintenant, je dis ... » avant chaque syllabe à enregistrer. Les enregistrements ont été réalisés via un micro-casque Shure (Microflex Beta 53) connecté à une interface audio USB Quartet (Apogee Electronics, Santa Monica) connectée à un ordinateur iMac. Les fichiers audios ont été sauvegardés via le logiciel Sound Studio 4.8.14 (Felt Tip Software, New-York). Les fichiers audios ont ensuite été édités sur le logiciel PRAAT (Boersma & Weenink, 2018) à l'aide d'un script pour normaliser la valeur quadratique moyenne de l'intensité des syllabes aux différents niveaux souhaités (67 dB, 70 dB et 73 dB). Les phrases ont été prononcées et enregistrées trois fois et la meilleure production de chaque syllabe a ensuite été choisie indépendamment par deux membres du laboratoire parlant le français québécois.

Le bruit masquant de parole était composé de quatre locuteurs français entre 25 et 45 ans, deux hommes et deux femmes, qui lisent le journal dans une salle insonorisée (Perrin & Grimaud, 2005). La valeur quadratique moyenne de l'intensité du BM a été normalisée sur le logiciel PRAAT à 70 dB.

Les syllabes et le bruit de fond étaient présentés aléatoirement selon 3 niveaux de SNR pour créer trois niveaux de difficulté : -3 dB (condition 1, difficulté élevée), 0 dB (condition 2, difficulté moyenne) et +3 dB (condition 3, difficulté basse). Le niveau du bruit de fond était fixe à 70 dB et le niveau des paires a été varié entre 67 dB, 70 dB et 73 dB pour générer les 3 niveaux de SNR.

La tâche de PPB a été codée et était présentée via le logiciel Presentation (Version 20.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com). Le participant, assis confortablement dans la salle insonorisée et muni d'un casque d'écoute (Beyer, DT 770 Pro) par lequel étaient transmis les stimuli (les paires de syllabes et le bruit de fond), devait fixer un moniteur de 27 pouces (HP EliteDisplay, E272q), sur lequel apparaissait une croix blanche lors de la transmission des stimuli et un point d'interrogation vert lors de la phase de réponse. Lors de cette période, d'une durée de 3 secondes, le participant devait déterminer si la paire était composée de syllabes identiques ou de syllabes différentes en cliquant sur l'un ou l'autre des boutons d'un boîtier de réponse (Cedrus, Model RB- 530).

4.3.3 Analyses statistiques

Toutes les analyses ont été effectuées sur le logiciel IBM SPSS 27 pour Mac OS. Pour évaluer l'effet des activités musicales et non musicales sur la PPB dans les deux groupes d'âge aux différents niveaux de SNR (objectif 1), deux ANOVAS mixtes à trois facteurs ont été réalisées. La première ANOVA mixte a été réalisée avec le pourcentage de bonnes réponses comme variable dépendante et la deuxième, avec le temps de réaction comme variable dépendante. Dans les deux analyses, le facteur intra-sujet était le niveau de SNR (1, 2, 3) et les deux facteurs inter-sujets étaient le groupe d'activité (musiciens, psychomoteurs) et le groupe d'âge (jeunes, âgés). Une mesure d'audition, le PTA oreille droite, le score au MoCA et le nombre d'années d'éducation ont été introduites en

covariables. La normalité des données a été vérifiée avec le test de Shapiro-Wilk et l'homogénéité des variances avec le test de Levene.

Pour évaluer l'effet de l'âge du début et du nombre d'années de pratique musicale sur la performance en PPB (objectif 2), des régressions linéaires mixtes ont été effectuées sur le groupe de musiciens en contrôlant pour l'âge des participants. Pour chaque niveau de SNR, une régression a été effectuée avec le pourcentage de bonne réponse en variable dépendante et une autre avec le temps de réaction en variable dépendante. L'indépendance des résidus a été évaluée avec la statistique de Durbin-Watson. L'homogénéité des variances a été évaluée en inspectant visuellement un graphique des résidus standardisé VS les valeurs prédites pour chaque régression. La normalité des résidus a été évaluée en inspectant visuellement les tracés de probabilité normale.

4.4 Résultats

4.4.1 ANOVA mixte à mesure répétées : Pourcentage de bonnes réponses

Les résultats dans la condition 3 de SNR (la plus facile) ne respectant pas la normalité, seules les conditions de SNR 1 et 2 ont été incluses dans la présente analyse. L'ANOVA mixte à 3 facteurs a montré un effet principal du groupe d'activité ($F[1,62]= 5.717$; $p = 0.020$; $\eta_p^2 = 0.084$) et du groupe d'âge ($F[1,62]= 10.733$; $p = 0.002$; $\eta_p^2 = 0.148$) (voir Tableau 3), avec une meilleure performance des musiciens comparé aux psychomoteurs, et des adultes jeunes comparé aux adultes âgés (voir Tableau 2, voir Figure 7). L'effet principal de la condition n'était pas significatif ($p = .395$), tout comme les interactions à deux et à trois facteurs ($p > .123$) (voir Tableau 3).

Condition de SNR	Groupe d'âge	Groupe d'activité	Moyenne	Ecart-type
1, SNR = - 3 dB	Jeunes	Musiciens	71.052	5.988
		Psychomoteurs	65.689	7.908
	Âgés	Musiciens	61.724	7.933
		Psychomoteurs	58.045	12.341
2, SNR = 0 dB	Jeunes	Musiciens	77.858	5.754
		Psychomoteurs	73.275	6.677
	Âgés	Musiciens	67.011	8.860
		Psychomoteurs	59.310	13.958

Tests des effets intersujets				
Facteur	dl	F	p	partial η^2
Audition (PTA OD)	1	.130	.719	.002
MOCA	1	.575	.451	.009
Education	1	.091	.764	.001
Activité	1	5.717	.020*	.084
Âge	1	10.733	.002**	.148
Âge x Activité	1	.000	.986	.000
Test des effets intrasujets				
Facteur	dl	F	p	partial η^2
Condition	1	.735	.395	.012
Condition x Audition	1	.200	.656	.003
Condition x MOCA	1	.040	.842	.001
Condition x Éducation	1	1.692	.198	.027
Condition x Activité	1	1.388	.243	.022
Condition x Âge	1	2.247	.139	.035
Condition x Activité x Âge	1	2.450	.123	.038

Note. dl = degré de liberté, Condition = condition de SNR 1 de SNR, Audition = PTA OD, MOCA = le score obtenu au MOCA, Éducation = nombre d'années d'éducation, Activité = musiciens ou psychomoteurs, Âge = jeunes ou âgés.

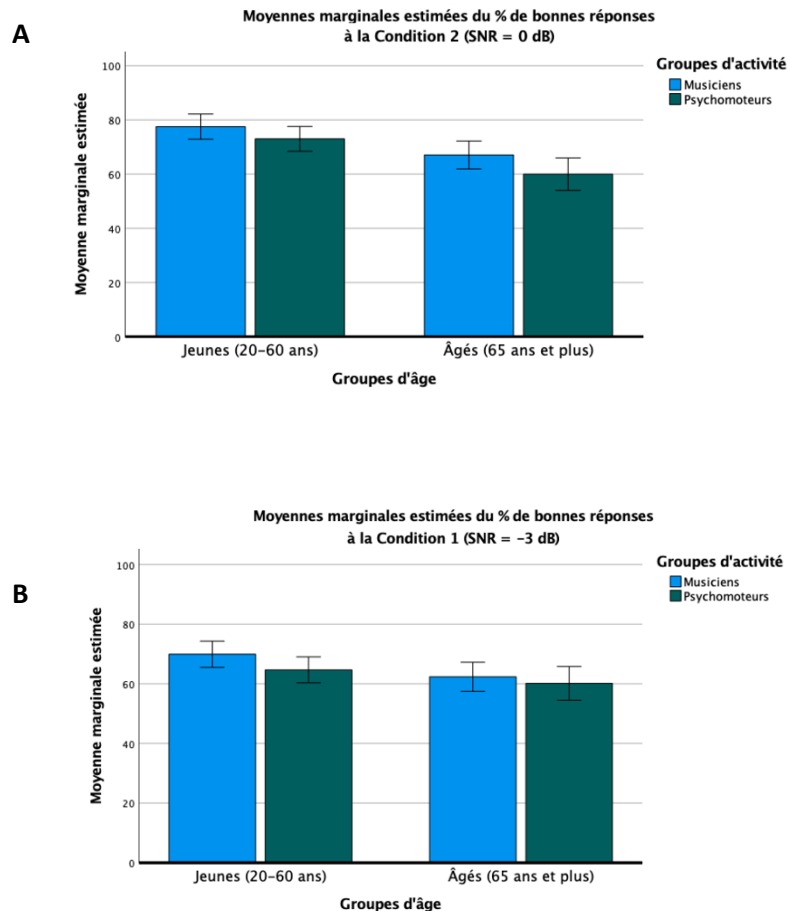


Figure 7. Pourcentage de bonnes réponses pour les quatre groupes de participants dans la condition de SNR 2 (A) et dans la condition de SNR 1 (B).

4.4.2 ANOVA mixte à mesure répétées : Temps de réaction

Pour avoir des résultats comparables au pourcentage de bonnes réponses, seules les conditions de SNR 1 et 2 ont été utilisées pour les temps de réaction. L'ANOVA mixte à 3 facteurs a montré un effet principal du groupe d'activité ($F[1,62]= 6.526; p = 0.013 ; \eta_p^2 = 0.094$) et du groupe d'âge ($F[1,62]= 10.939; p < .002; \eta_p^2 = 0.148$) (voir Tableau 5), avec des temps de réaction plus rapides pour les adultes jeunes comparés aux adultes âgés et pour les musiciens comparés aux psychomoteurs (voir tableau 4, voir Figure 8). Comme pour le pourcentage de bonnes réponses, ni effet principal de la condition ni interaction à trois ou deux facteurs n'ont été trouvés ($p > .475$) (voir Tableau 5).

Condition de SNR	Groupe d'âge	Groupe d'activité	Moyenne	Écart-type
1, SNR = - 3 dB	Jeunes	Musiciens	5220.67015	1579.740
		Psychomoteurs	5564.782	1375.717
	Âgés	Musiciens	5943.986	1880.084
		Psychomoteurs	7090.107	1684.789
2, SNR = 0 dB	Jeunes	Musiciens	4817.877	1300.338
		Psychomoteurs	5277.844	1454.725
	Âgés	Musiciens	5585.811	1737.507
		Psychomoteurs	6752.062	1713.544

Tests des effets intersujets				
Facteur	dl	F	p	partial η^2
Audition (PTA OD)	1	4.285	.043*	.064
MOCA	1	.012	.913	.000
Éducation	1	.016	.900	.000
Activité	1	6.526	.013*	.094
Âge	1	10.939	.002**	.148
Âge x Activité	1	2.395	.127	.037
Test des effets intrasujets				
Facteur	dl	F	p	partial η^2
Condition	1	.314	.577	.005
Condition x Audition	1	2.119	.150	.033
Condition x MOCA	1	.115	.735	.002
Condition x Education	1	.107	.745	.002
Condition x Activité	1	.000	.994	.000
Condition x Âge	1	.405	.527	.006
Condition x Activité x Âge	1	.517	.475	.008

Note. dl = degré de liberté, Condition = condition 1 de SNR (-3 dB) ou condition 2 de SNR (0 dB), Audition = PTA OD, MOCA = le score obtenu au MOCA, Éducation = nombre d'années d'éducation, Activité = musiciens ou psychomoteurs, Âge = jeunes ou âgés.

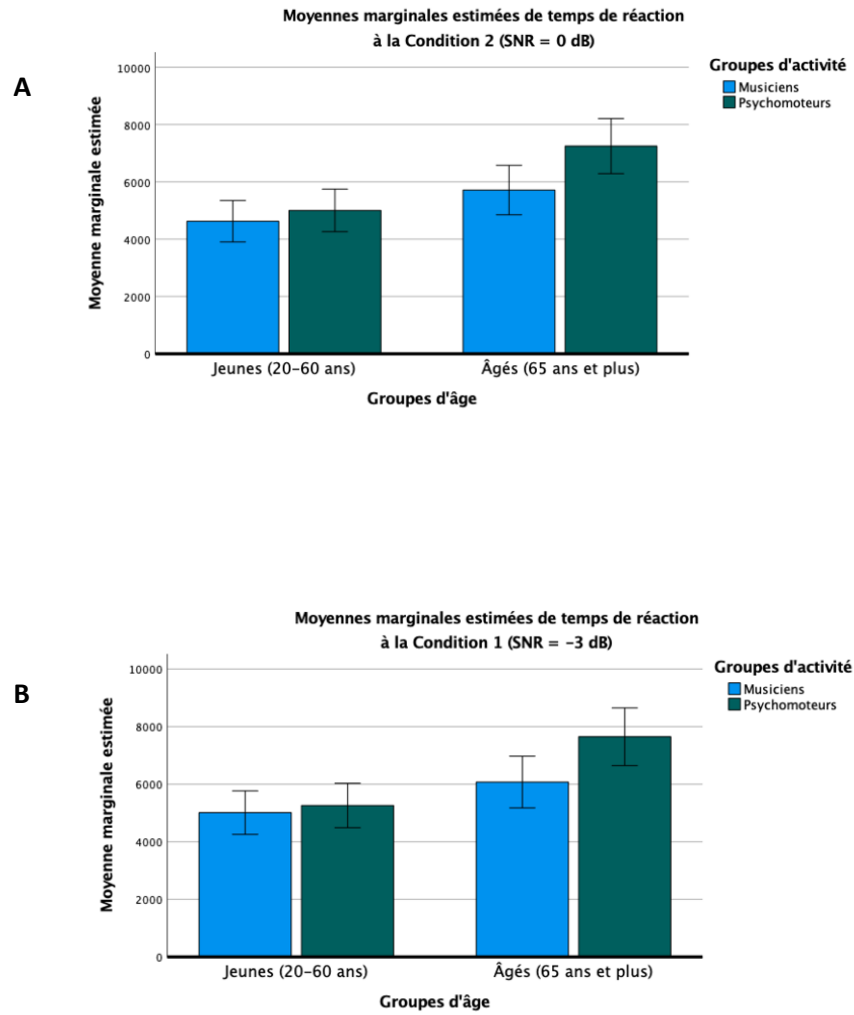


Figure 8. Temps de réaction pour les quatre groupes de participants dans la condition de SNR 2 (A) et dans la condition de SNR 1 (B).

4.4.3 Régressions linéaires mixtes : pourcentage de bonnes réponses et temps de réaction

Au sein du groupe de musiciens, en contrôlant pour l'âge des participants, le nombre d'années de pratique ne prédisaient ni le pourcentage de bonnes réponses ni le temps de réaction pour les conditions 1 et 2 de SNR ($p > .101$). L'âge du début de la pratique n'a été prédicteur ni du pourcentage de bonne réponses dans les deux niveaux de SNR testés ($p = .102$), ni du temps de réaction au niveau 2 de SNR ($p = .102$) (voir tableaux 6 et 7). En revanche, l'âge du début de la pratique musicale a été prédicteur du temps de réponse au niveau 1 de SNR (voir tableaux 6).

Condition de SNR	Temps de réaction	B	IC 95% de B		ES B	β	R2	Δ R2	β 's p
			LI	LS					
1, SNR = - 3 dB	Modèle						.150	.067	
	Constante	5395.694	3658.904	7132.484	851.571				<.001
	Âge	41.344	-3.789	86.478	22.130	.460			.071
	Âge de début	-55.053	-106.120	-3.985	25.039	-.452			.035
	Années de pratique	-40.325	-90.490	9.840	24.597	-.399			.111
2, SNR = 0 dB	Modèle						.127	.043	
	Constante	4734.426	3183.436	6285.415	760.470				<.001
	Âge	39.412	-.893	79.718	19.762	.498			.055
	Âge de début	-37.677	-83.282	7.927	22.360	-.351			.102
	Années de pratique	-37.175	-81.973	7.927	21.965	-.417			.101

Condition de SNR	Pourcentage de bonnes réponses	B	IC 95% de B		ES B	β	R2	Δ R2	β 's p
			LI	LS					
1, SNR = - 3 dB	Modèle						.145	.062	
	Constante	76.768	66.175	87.361	5.194				<.001
	Âge	-.219	-.495	.056	.135	-.402			.114
	Âge de début	-.032	-.259	.353	.153	-.044			.834
	Années de pratique	.047	-.344	.279	.150	.076			.757
2, SNR = 0 dB	Modèle						.172	.091	
	Constante	82.093	71.150	93.036	5.366				<.001
	Âge	-.257	-.542	.027	.139	-.448			.075
	Âge de début	-.066	-.388	.256	.158	-.085			.678
	Années de pratique	.170	-.146	.486	.155	.264			.281

Note. B = coefficient de régression non-standardisé, IC = intervalle de confiance, LI = limite inférieure, LS = limite supérieure, ES = erreur standard, β = coefficient standardisé, R² = coefficient de détermination, Δ R² = coefficient ajusté, β 's p = valeur p associée au coefficient standardisé.

4.5 Discussion

À ce jour, cette étude est la première à avoir comparé l'effet de la pratique musicale et d'activités psychomotrices de niveau amateur sur la PPB. Le premier objectif de cette étude était d'étudier l'effet de ces deux pratiques sur la PPB chez des adultes jeunes et des adultes âgés à différents niveaux de SNR. Le deuxième objectif était d'évaluer la contribution de l'âge du début et le nombre d'années de pratique sur la performance en PPB chez les musiciens. Les résultats ont montré, en accord avec notre hypothèse, que les adultes âgés ont obtenu des performances (pourcentage de bonnes réponses, temps de réaction) significativement moins bonnes que les jeunes adultes, peu importe le type de pratique et peu importe le niveau de SNR. Cette étude a non seulement confirmé notre deuxième hypothèse, en montrant que la pratique musicale était associée à une réduction de l'effet de l'âge chez les adultes de plus de 65 ans, mais a également montré qu'elle était associée à une meilleure performance en PPB chez les jeunes adultes à tous les niveaux de SNR. Finalement, les résultats de cette étude ont montré que l'âge du début de la pratique et le nombre d'années de pratique n'étaient pas prédicteurs de la performance en PPB.

4.5.1 Performance en PPB

Dans la présente étude, la performance en PPB a été étudiée à l'aide du pourcentage de bonnes réponses et du temps de réaction. Les résultats obtenus sur le pourcentage de bonnes réponses seront discutés dans un premier temps, les résultats sur le temps de réaction seront abordés dans un second temps et la performance globale sera discutée dans un troisième temps.

Dans la présente étude, à chaque niveau de SNR, les participants jeunes ont obtenu un plus haut pourcentage de bonnes réponses que les adultes âgés. Ces résultats sont en accord avec les études qui montrent que les adultes âgés obtiennent de moins bonnes

performances que les jeunes adultes en présence d'un BM de parole et les études qui montrent que pour le même niveau de SNR, les adultes âgés obtiennent un pourcentage de bonnes réponses plus bas que les jeunes (voir la section 1.3.1). Toutefois, les musiciens âgés tout comme les musiciens jeunes ont obtenu un plus haut pourcentage de bonnes réponses à tous les niveaux de SNR que les psychomoteurs jeunes et âgés. Dans la littérature, seules trois études existent à ce jour sur l'effet de la pratique musicale sur la perception de syllabes présentées dans du bruit, deux études sur de jeunes adultes présentant des syllabes dans un bruit énergétique (Du & Zatorre, 2017; Varnet et al., 2015) et une étude sur des adultes entre 20 et 87 ans présentant des syllabes dans un BM de parole à quatre locuteurs (Perron et al., 2021). Le plus haut pourcentage de bonnes réponses obtenu par les jeunes musiciens de la présente étude est en accord avec les résultats obtenus par Du and Zatorre (2017) et Varnet et al. (2015), bien que le BM ne soit pas du même type. Toutefois, le bénéfice des musiciens jeunes et âgés obtenus dans la présente étude n'est pas totalement en accord avec les résultats de l'étude de Perron et al., (2021). Dans cette dernière, aucun bénéfice global sur la PPB n'a été trouvé pour le groupe de musiciens (chanteurs de chorale). Toutefois, en contrôlant pour l'âge, une performance supérieure aux non-musiciens a été trouvée chez les chanteurs qui pratiquaient en moyenne au moins une fois par semaine en dehors des rencontres de la chorale et chez les musiciens qui chantaient dans plusieurs langues. Dans l'étude de Perron et al., (2021), les musiciens recrutés étaient exclusivement des chanteurs de chorale qui devaient pratiquer leur activité depuis un minimum de 2 ans et chanter en groupe au moins 60 minutes par semaine. Dans la présente étude, des instrumentistes tout comme des chanteurs ont été recrutés, qui de plus pratiquaient leur activité depuis un minimum de 6 ans avec un minimum de 3 heures de pratique par semaine. Il est donc possible que les effets de la pratique musicale en PPB ne soient visibles que sous certaines conditions de pratique, comme la fréquence de pratique par semaine et/ou le nombre d'année de pratique. Il n'existe à ce jour aucune autre étude sur la performance en PPB avec des syllabes chez des groupes de musiciens et de non-musiciens âgés. Néanmoins, les meilleures performances obtenues dans cette étude par les musiciens âgés comparés au groupe contrôle sont en accord avec la plupart des études publiées sur des adultes d'âge moyen et âgés, avec des phrases ou de mots comme cibles dans un BM de parole (Fostick, 2019; Parbery-Clark et al., 2012; Parbery-

Clark et al., 2011; Zhang et al., 2020), sauf avec celle de Mussoi (2021) qui n'a pas trouvé de meilleurs performances chez les musiciens comparés aux non-musiciens.

Dans notre étude, alors que le pourcentage de réponses correctes reflète la capacité à discriminer deux syllabes, le temps de réaction est une mesure de la vitesse de traitement nécessaire pour effectuer la tâche, c'est à dire le temps de traitement nécessaire pour décider si les syllabes d'une paire sont identiques ou différentes. Nos résultats ont montré que les participants jeunes, peu importe le groupe d'activité, avaient des temps de réaction plus rapide que les deux groupes de participants âgés. La théorie de la vitesse de traitement dans le vieillissement propose qu'une des explications pour la baisse de performance des adultes âgés dans différents tests cognitifs soit le ralentissement général des processus cognitifs (Salthouse, 1996). Salthouse avance que ce ralentissement augmente le temps nécessaire pour réaliser efficacement les différentes étapes d'un processus cognitif complexe. La vitesse de traitement cognitif est associée avec la performance en PPB (pour une revue, voir : Dryden et al., 2017). Un ralentissement cognitif dû à l'âge pourrait alors contribuer à la baisse de performances en PPB des personnes âgées. En effet, la présence d'un signal de parole dégradé en périphérie par une baisse de l'audition ainsi que par les propriétés énergétiques d'un BM, doublé des interférences provoquées par le masquage informationnel, augmente les demandes cognitives. Si la vitesse de traitement est ralentie, il se peut que la mémoire de travail devienne surchargée et que des éléments du signal entrant ne puissent pas être adéquatement traités et par conséquent ne pas être perçus correctement. Toutefois, nos résultats ont montré que le temps de réaction des musiciens était plus rapide que celui des psychomoteurs, peu importe l'âge. Toutefois, les résultats obtenus dans cette étude sur le temps de réaction des adultes jeunes ne sont pas en accord avec ceux trouvés en perception de syllabes par Du and Zatorre (2017) et Varnet et al. (2015). En effet, ces deux études n'ont pas montré de différence entre le temps de réaction des musiciens et celui des non-musiciens, tous des adultes dans la vingtaine. Cependant, ces deux études ont utilisé un bruit blanc comme BM, c'est-à-dire un BM énergétique, tandis que c'est un BM de parole, composé de quatre locuteurs qui a été utilisé dans la présente étude. Plusieurs résultats dans la littérature montrent que, chez de jeunes adultes non-musiciens, le temps de réaction en présence d'un BM de parole est plus long qu'en présence d'un bruit masquant énergétique (Bennett et al., 2012; Carter & Bidelman, 2021). En particulier, l'étude de Carter and Bidelman (2021) a

montré que le temps de réaction pour catégoriser des syllabes en présence d'un bruit à deux locuteurs était significativement plus long qu'en présence d'un bruit énergétique. De plus, l'étude de Bennett et al. (2012), a montré que dans une tâche de discrimination de type « oddball », présentant deux syllabes dans trois BMs différents, le temps de réaction était significativement plus long en présence d'un BM à quatre locuteurs qu'en présence de *speech-shaped noise* continu interrompu (BMs énergétiques). De façon intéressante, les temps de réaction en présence de bruit à 4 locuteurs étaient significativement corrélés au temps de latence de l'apparition de l'onde P3 obtenue en EEG, qui représente un changement volontaire de l'attention. La différence dans le temps de réaction obtenu entre un BM énergétique (plus court) et un BM de parole (plus long) pourrait représenter le temps ajouté par les interférences linguistiques d'un BM de parole. Les interférences linguistiques, comparées aux effets d'un BM énergétique, augmenteraient donc le temps de traitement dans différentes tâches de PPB, comme dans le cas des tâches de catégorisation ou de discrimination. En comparant notre étude avec celle de Du and Zatorre (2017) et de Varnet et al. (2015), il est possible que chez les jeunes adultes, la pratique musicale réduise le temps de réaction des musiciens uniquement en présence d'un bruit de fond dont la charge cognitive est plus élevée, comme le sont les BMs de parole. Cela pourrait impliquer que la pratique musicale augmente la vitesse des traitements nécessaires pour résister aux interférences perceptuelles des BMs de parole. Dans le cas des adultes âgés, il est possible que cet avantage persiste et limite ainsi les effets de l'âge sur la vitesse de traitement et contribue ainsi à préserver la PPB.

Les résultats de cette étude sur la performance en PPB suggèrent que la pratique musicale a un effet supérieur sur la PPB, comparée à une activité psychomotrice non musicale. Ceci suggérerait qu'en plus d'un entraînement cognitif, la présence de la modalité auditive confère sa supériorité à la pratique musicale pour la PPB. Cette hypothèse rejoint les résultats trouvés dans l'étude de Stewart et al. (2020), qui montrent que les joueurs de jeux vidéo ne présentent pas de performances supérieures dans deux tâches de PPB comparés à des participants contrôles. Alternativement, étant donné que le groupe d'activités psychomotrices n'a pas été recruté sur la sélection que d'une seule activité, il est possible que le manque d'homogénéité dans les activités psychomotrices pratiquées par les

psychomoteurs de cette étude ait empêché la détection de bénéfices provenant d'activités psychomotrices qui seraient plus favorables à la PPB que d'autres.

En somme, les meilleures performances obtenues par les musiciens dans cette étude parlent en faveur d'un avantage de la pratique musicale de niveau amateur pour la PPB chez des adultes jeunes et des adultes âgés. Ces résultats suggéreraient que la pratique musicale confère une réserve cognitive qui limiterait le déclin de la capacité à percevoir la parole dans le bruit (Stern, 2009; Stern et al., 2020).

4.5.2 Prédicteurs de la performance en PPB

Dans la présente étude, le nombre d'années de pratique n'a pas été un prédicteur significatif de la performance en PPB (pourcentage de bonnes réponses et temps de réaction). En revanche, l'âge de début de la pratique musicale a prédit significativement le temps de réaction au niveau de SNR 1 mais pas au niveau de SNR 2. Ces résultats sont partiellement en accord avec différentes études ayant évalué la PPB à l'aide de syllabes, de mots ou de phrases. Dans l'étude de Du and Zatorre (2017), une relation entre la performance en perception de syllabes dans du bruit blanc n'a été trouvée ni avec le nombre d'années de pratique, ni avec l'âge de début de la pratique. L'étude de Parbery-Clark et al. (2009) n'a pas montré de corrélation entre la performance au HINT et le nombre d'années de pratique. Dans l'étude de Zhang et al. (2020), la performance dans une tâche de phrases présentées dans deux types de BM n'est corrélée ni avec le nombre d'années de pratique, ni avec le nombre d'heures de pratique par semaine. Dans ces deux études, le niveau de pratique des participants n'était pas mentionné. Dans trois autres études, une corrélation a en revanche été trouvée entre le nombre d'années de pratique et un score composite de la PPB (Anaya et al., 2016), un score au QuickSIN (Yoo & Bidelman, 2019) ou encore un score au HINT (Ruggles et al., 2014). Dans l'étude de Anaya et al. (2016), les musiciens étaient des pianistes professionnels, dans l'étude de Ruggles et al. (2014) la plupart des musiciens étaient également des musiciens experts (étudiants dans un programme de musique à l'université) et dans l'étude de Yoo et al. (2019), le niveau de pratique n'était pas mentionné.

Dans notre étude, comparée à la majorité des études incluses dans la revue systématique du chapitre 4, l'âge de début de la pratique musicale ne figurait pas parmi les critères d'inclusion. Les âges de début de la pratique chez les musiciens de notre étude se situaient entre 5 et 59 ans ($17,61 \pm 13,94$). En revanche, dans les études citées plus haut, les âges de début étaient inférieurs à 10 ans. Ces musiciens avaient donc tous commencé leur pratique musicale pendant l'enfance, période pendant laquelle le cerveau en maturation passe par des périodes « sensibles » de plasticité. Comme vu au chapitre 2, lors de ces périodes, le système nerveux est particulièrement sensible au contact de l'environnement, dont les expériences sensorielles auront des effets à long terme sur la structure et le fonctionnement cérébral (pour une revue, voir : Knudsen, 2004; White et al., 2013). Il se pourrait alors que l'entraînement musical sur un cerveau en maturation apporte des bénéfices plus importants. Différentes études se sont intéressées à l'importance de l'âge du début de la pratique musicale sur la structure cérébrale et le traitement multisensoriel chez des musiciens, en contrôlant pour le nombre d'années de pratique musicale. Watanabe et al. (2007), Bailey and Penhune (2012) et Steele et al. (2013) ont montré que des musiciens ayant commencé leur pratique musicale avant 7 ans obtenaient de meilleures performances dans une tâche testant la synchronisation visuo-auditivo-motrice que des musiciens appariés sur le nombre d'années de pratique, mais ayant commencé après l'âge de 7 ans. De façon intéressante, Bailey and Penhune (2012) ont également évalué certaines fonctions cognitives dans leurs différents groupes et ont montré que les deux groupes de musiciens ne différaient pas sur les tâches évaluant le vocabulaire, la mémoire de travail et la flexibilité mentale. Ces résultats montrent qu'un âge de pratique en dessous de 10 ans confère un avantage pour le traitement multisensoriel, mais pas pour certaines fonctions cognitives.

Les résultats de la présente étude, sur des musiciens ayant commencé leur pratique entre l'enfance et l'âge adulte, suggèrent que des gains de la pratique musicale sur la PPB sont possibles, même en commençant sa pratique à l'âge adulte. Quant au nombre d'années de pratique, il est possible que cette variable isolée ne puisse pas prédire la performance. En effet, il est probable que la performance en PPB soit prédite par une association de différents facteurs, tels que la durée des pratiques et leur fréquence.

4.6 Conclusion

Cette étude a montré que la pratique d'activités musicales de niveau amateur améliore la PPB et limite l'effet de l'âge en présence d'un BM de parole. L'utilisation de syllabes a pu montrer que la pratique musicale agit sur la PPB en amont du traitement lexical et la comparaison de la pratique musicale avec des activités psychomotrices non musicales a montré l'importance de l'entraînement auditif dans l'amélioration de la PPB. Dans cette étude, le nombre d'années de pratique n'ont pas été associées à la performance en PPB des musiciens. Davantage d'études sont par conséquent nécessaires pour confirmer ces résultats chez les musiciens âgés et déterminer quels sont les paramètres de pratique qui sont associés avec une amélioration de la PPB.

Chapitre 5 Discussion

5.1 Rappel des résultats

L'objectif principal de ce mémoire était d'étudier l'effet de la pratique musicale sur la PPB, avec les deux sous-objectifs suivants : 1) effectuer un bilan qualitatif et quantitatif de la littérature et 2) de comparer l'effet de pratiques musicales (chant, instrument de musique) et d'activités psychomotrices non musicales, toutes à niveau amateur, sur la PPB chez des adultes jeunes et des adultes âgés. Les deux études incluses dans ce mémoire ont permis 1) de mettre en lumière des effets bénéfiques de la pratique musicale qui varient en fonction de conditions expérimentales de PPB et 2) de montrer des effets bénéfiques de la pratique musicale à niveau amateur chez des adultes jeunes et chez des adultes âgés comparés à la pratique d'activités psychomotrices non musicales. Les résultats des deux études seront d'abord rappelés, puis la question du transfert d'habileté sera discutée.

5.1.1 Étude 1

L'étude 1 est la première revue systématique et méta-analyse sur l'effet de la pratique musicale sur la perception de la parole dans le bruit. Dans la méta-analyse, l'effet de la pratique musicale a été évalué dans différentes conditions expérimentales de PPB. Un avantage des musiciens a été trouvé en présence de BMs énergétiques, de BMs de parole à 2 et 4 locuteurs, en présence de stimuli séparés spatialement et lorsque les stimuli sont présentés à des niveaux de SNR négatifs et de zéro. En revanche, aucun avantage des musiciens n'a été trouvé lorsqu'un BM de parole est composé d'un seul locuteur. Dans la littérature qui s'intéresse à l'effet de différentes conditions expérimentales sur la PPB dans la population générale, il est montré que la difficulté de la tâche varie, entre autres, en fonction du BM, du SNR et de la présentation spatiale des stimuli. Lorsqu'un BM énergétique ou un BM de parole est colocalisé aux paroles cibles, la perception de la parole diminue significativement. S'agissant des BMs de parole, leur pouvoir masquant dépend de leur contenu énergétique et surtout des interférences linguistiques produites par le/les locuteurs du BM. Par conséquent, plus le contenu linguistique du BM est perceptible, plus le BM de parole est masquant. Plusieurs études ont montré que des BM de parole entre 2 et

3 locuteurs engendraient les moins bonnes performances et qu'augmenter le nombre de locuteurs au-delà de 3 réduisait graduellement le pouvoir masquant en rendant le contenu linguistique moins perceptible (Hoen et al., 2007; Rosen et al., 2013; Tun & Wingfield, 1999). En revanche, un BM composé d'un seul locuteur a le plus faible pouvoir masquant, dû aux variations d'énergie dans le discours du locuteur (Cooke, 2006). Par ailleurs, il a été montré qu'une séparation spatiale diminue le pouvoir masquant des BMs grâce à l'émergence d'indices binauraux de localisation (Glyde et al., 2011). Les résultats obtenus dans la méta-analyse montrent par conséquent que l'avantage des musiciens est plus conséquent en présence des conditions expérimentales d'écoute qui sont les plus difficiles. Ces résultats suggèrent qu'en présence de stimuli colocalisés, la pratique musicale confère un avantage non seulement contre la dégradation de la parole cible en périphérie par un BM énergétique et par la diminution du SNR, mais également contre les interférences perceptuelles provoquées par des BM de parole. Ces avantages pourraient être dûs à une amélioration du traitement auditif central (Deguchi et al., 2012; Donai & Jennings, 2016; Grassi et al., 2017; Hutka et al., 2015; Kishon-Rabin et al., 2001; Kuhn et al., 2013; Kumar et al., 2014; Liang et al., 2016; Sares et al., 2018; Schon et al., 2004; Strait et al., 2010; Zendel & Alain, 2009; B. R. Zendel & C. Alain, 2013) et/ou d'une amélioration dans certaines fonctions cognitives (Grassi et al., 2017; Groussard et al., 2020; Hanna-Pladdy & Gajewski, 2012; Mansens et al., 2018; Zuk et al., 2014). Dû au manque d'étude sur des participants âgés, l'effet de l'âge sur la performance en PPB n'a pas pu être évalué de manière robuste. En effet, la revue de littérature a révélé que la plupart des études portaient sur des adultes jeunes, majoritairement entre 18 et 30 ans. De plus, le manque d'information et d'homogénéité soulevé dans le niveau de pratique (amateur/professionnel) n'a pas permis de déterminer le niveau de pratique à partir duquel ces avantages émergent.

En somme, cette étude a montré l'existence d'un effet de la pratique musicale qui varie en fonction de la difficulté des conditions expérimentales de PPB. Les musiciens présentent un avantage dans les conditions expérimentales les plus difficiles, mais peu ou pas dans les conditions expérimentales plus faciles. Davantage d'études sur les personnes âgées et sur des musiciens amateurs permettront de vérifier si ces avantages peuvent émerger d'une pratique accessible à la population générale et si ces avantages peuvent limiter l'effet du déclin lié à l'âge.

5.1.2 Étude 2

Dans l'étude 2, des participants jeunes (20-59 ans) et des participants âgés (65 ans et plus) pratiquant à niveau amateur une activité psychomotrice ou une activité musicale (chant/instrument de musique) ont été recrutés. En contrôlant pour l'audition, les participants âgés, peu importe le groupe d'activité, ont obtenu de moins bonnes performances (pourcentage de bonne réponse, temps de réaction) que les participants jeunes. Ces résultats sont en accord avec la littérature qui montre que les adultes âgés obtiennent de moins bonnes performances que les jeunes au même niveau de SNR (Buss et al., 2019; Tun & Wingfield, 1999) et que l'effet de l'âge affecte particulièrement la performance dans un BM de parole (Rajan & Cainer, 2008; Taitelbaum-Swead & Fostick, 2016). Toutefois, les musiciens des deux groupes d'âge ont obtenu de meilleures performances que les psychomoteurs, avec plus de bonnes réponses et des temps de réaction plus rapides. Ces résultats suggèrent que la pratique musicale limite l'effet de l'âge sur la PPB. Par ailleurs, l'utilisation de syllabes a permis de montrer que la pratique musicale agissait en amont du traitement lexical. De plus, la comparaison avec un groupe contrôle pratiquant une activité psychomotrice a montré la supériorité de l'entraînement musical comparé à l'entraînement psychomoteur non musical. Par ailleurs, dans l'étude 2, ni l'âge du début ni le nombre d'années de pratique n'étaient associés à la performance en PPB des musiciens. Il est possible que d'autres paramètres prédisent la performance en PPB, comme la fréquence, le type et/ou la durée des entraînements. Ceci rejoindrait l'étude de Perron et al. 2021, qui a montré qu'indépendamment de l'âge, les chanteurs qui pratiquaient plus d'une fois par semaine et qui chantaient dans plusieurs langues obtenaient de meilleures performances en PPB.

En somme, cette étude a montré qu'une pratique musicale de niveau amateur est associée à une meilleure performance en PPB chez des adultes jeunes et à une limitation de l'effet de l'âge chez des adultes âgés. Les paramètres de pratique menant à ces bénéfices seront encore à clarifier pour permettre d'envisager un programme de prévention de la PPB.

Au terme de ces deux études, des questions additionnelles sur l'effet de la pratique musicale sur la PPB se posent. Notamment, par quel biais la pratique musicale améliore-t'-

elle la PPB ? Les bénéfices trouvés dans des tâches expérimentales de PPB se retrouvent-ils dans la vie de tous les jours ? Ces deux questions seront abordées dans la prochaine section dans le cadre d'une réflexion sur le transfert d'apprentissage.

5.2 Transfert d'apprentissage

Comme vu au chapitre sur la plasticité, la notion de transfert d'apprentissage fait référence à la généralisation d'un apprentissage effectué sur domaine donné à un domaine peu ou pas entraîné. La notion de transfert d'apprentissage a émergé au début du 20^e siècle dans le contexte de l'éducation (Woodworth & Thorndike, 1901) et depuis, la notion de transfert d'apprentissage s'est étendue au domaine des sciences cognitives et est présente notamment dans beaucoup d'études qui s'intéressent aux effets de la pratique musicale sur la PPB. Dans ces dernières, il est proposé que l'entraînement de la pratique musicale *transfère* à la PPB. Dans les domaines de l'éducation et des sciences cognitives, les notions de transfert proche et de transfert lointain sont utilisées pour qualifier la distance de transfert entre le domaine entraîné et le domaine non entraîné. Toutefois, la qualité de la distance de transfert ne fait pas l'objet de descriptions établies et est alors qualifiée arbitrairement de proche ou de lointaine en estimant, de manière subjective, la distance entre deux domaines de compétence. Barnett and Ceci (2002) ont alors proposé une taxonomie pour permettre une qualification systématique de la relation de transfert d'apprentissage entre deux domaines. Cette taxonomie a été développée dans le cadre de l'éducation et c'est la seule taxonomie du transfert qui existe à ma connaissance. Cette taxonomie sera présentée dans les prochains paragraphes et pour l'aborder, nous parlerons de domaine entraîné et de domaine receveur pour différencier le domaine entraîné du domaine non entraîné dans lequel le transfert est évalué.

Barnett and Ceci (2002) proposent de qualifier le transfert d'un domaine entraîné à un domaine receveur à l'aide de plusieurs dimensions qui concernent le contenu qui est transféré (*What ?*) et les situations dans lesquelles ce transfère s'effectue (*When ?*). Un continuum entre un transfert proche et un transfert lointain s'applique dans les différentes dimensions du *When ?*.

La description du *What ?* se fait sur trois dimensions :

- 1) Le type de compétence qui est entraîné. La compétence entraînée est-elle une procédure, une représentation ou un principe heuristique ?

2) Le changement de performance qui est attendu sur le domaine receveur. Le changement de performance sur le domaine receveur s'observe-t-il sous forme d'augmentation de la vitesse de réalisation, de la précision ou sous la forme d'un changement d'approche pour la réalisation d'une tâche ?

3) Les demandes mnésiques que requiert le succès du transfert. Le transfert au domaine receveur se fait-il spontanément lors de son évaluation ou, par exemple, est-ce que des instructions précisant qu'un changement de stratégie est attendu sont nécessaires?

En ce qui concerne le *When?*, six dimensions sont proposées : 1) la similarité entre le domaine entraîné et le domaine receveur, 2) le contexte physique de l'entraînement et de l'évaluation du transfert, 3) le contexte temporel entre l'entraînement et l'évaluation, 4) le contexte fonctionnel de l'entraînement et de l'évaluation, 5) leur contexte social et 6) leur modalité. Les propriétés de ces différentes dimensions sont abordées ci-dessous.

1) La similarité entre les domaines fait référence à la présence de caractéristiques communes entre le domaine entraîné et le domaine receveur. Barnett and Ceci (2002) illustre ceci en suggérant que des connaissances sur les souris et des connaissances sur les rats sont deux domaines qui présentent beaucoup de caractéristiques communes, tandis que des connaissances scientifiques et des connaissances artistiques en partageraient moins.

2) S'agissant du contexte physique, les auteurs font référence au type de lieux dans lequel l'apprentissage s'est déroulé et celui dans lequel l'évaluation du transfert sur le domaine receveur s'effectue. S'agit-il, par exemple, d'une classe d'école, d'un laboratoire ou encore de chez soi ?

3) Le contexte temporel fait référence à la fenêtre temporelle entre l'entraînement et l'évaluation du transfert sur le domaine receveur. S'agit-il de minutes, d'heures, de jour, de mois ou encore d'années ?

4) Pour la dimension du contexte fonctionnel, l'entraînement et l'évaluation du transfert sont-ils ancrés dans des activités de la vie de tous les jours ou plutôt dans des contextes contrôlés comme dans un cadre scolaires ou dans un laboratoire ?

5) Avec le contexte social, c'est la présence ou l'absence d'autres individus et d'interactions lors de l'entraînement et de l'évaluation du transfert qui est à relever.

6) Finalement, la dimension de la modalité fait référence, d'une part, à la modalité sensorielle dans laquelle l'entraînement et l'évaluation sont effectués, et d'autre part au moyen utilisé pour l'entraînement et l'évaluation. Par exemple, s'agit-il d'un apprentissage/d'une évaluation informatisée ou par le biais d'exercices sur papier ?

Avec ces différentes dimensions, Barnett and Ceci (2002) proposent une qualification plus riche et nuancée de la relation de transfert et de sa distance, au lieu de la catégorisation binaire entre transfert proche et transfert lointain, basée sur des critères arbitraires et hétérogènes. Bien que proposée pour évaluer le transfert dans le domaine de l'éducation, la taxonomie de Barnett and Ceci (2002) a été appliquée pour interpréter le résultat d'études sur des entraînements cognitifs effectués chez des personnes âgées (Zelinski, 2009). Dans le paragraphe suivant, la relation de transfert entre la pratique musicale et la PPB avancée dans la littérature revue au chapitre 4 sera évaluée en appliquant, lorsque possible, la taxonomie de Barnett et Ceci. Les dimensions du *What ?* seront présentées d'abord, suivies des dimensions du *When ?*.

What ?

Sur le plan du *What ?*, le transfert de la pratique musicale à la PPB peut être qualifié surtout dans les dimensions des compétences et du changement de performance.

Sur le plan des compétences qui transfèrent de la pratique musicale à la PPB, comme vu au chapitre 2, la pratique musicale est une activité qui implique le traitement et l'intégration de plusieurs modalités sensorielles, leur intégration avec le système moteur, et plusieurs fonctions cognitives. Quand la pratique musicale n'est pas instrumentale, mais qu'il s'agit du chant, à ceci s'ajoute un entraînement articulatoire et phonologique. Dans les articles revus au chapitre 3, les deux pôles de compétences musicales « transférables » qui sont les plus mentionnés sont le traitement auditif central et certaines fonctions cognitives. Le traitement auditif central est présenté comme une compétence plus spécifique, tandis que certaines fonctions cognitives sont abordées comme des habiletés plus générales. Sur le plan auditif, il est proposé qu'un encodage sous-cortical plus précis et/ou plus rapide du signal

acoustique et/ou encore qu'un traitement cortical spectrotemporal plus précis soient les compétences qui transfèrent de la pratique musicale vers la PPB. Sur le plan cognitif, il est suggéré dans certains articles que l'entraînement de la mémoire de travail transférerait à la PPB. Dans un article de Du and Zatorre (2017), il est proposé qu'une meilleure intégration sensori-motrice transfère de la pratique musicale à la PPB, tandis que l'intégration multisensorielle n'est proposée dans aucun article en tant que compétence transférée à la PPB. En résumé, il ressort des articles revus dans le chapitre 3 que la pratique musicale pourrait améliorer la PPB grâce à un transfert d'habiletés spécifiques, le traitement auditif central, ou d'habiletés générales, comme certaines fonctions exécutives.

Concernant l'observation du changement de performance, il varie en fonction de la tâche de PPB utilisée. Dans les articles inclus dans la revue systématique du chapitre 3, selon les études, une meilleure performance est attendue sous forme d'un plus haut pourcentage de bonnes réponses à un niveau de SNR donné, d'un SRT plus bas (donc meilleur) ou encore d'un temps de réaction plus court.

When ?

Concernant le *When ?* Barnett and Ceci (2002) proposent d'appliquer les six dimensions de leur taxonomie aux contextes de l'entraînement et à ceux de l'évaluation du transfert. En ce qui concerne la PPB, les contextes d'évaluation de la PPB et les contextes de PPB de la vie de tous les jours ne sont pas les mêmes, par conséquent, lorsque possible, les dimensions du *When ?* seront appliquées aux contextes expérimentaux ainsi qu'aux contextes « naturels » de la PPB.

Sur le plan des similarités entre le domaine entraîné et le domaine receveur, la PPB et la pratique musicale partagent des étapes de traitements, mais la PPB passe également par des traitements absents de la pratique musicale instrumentale, notamment les étapes du traitement linguistique. Pour aborder la similarité entre le domaine entraîné et le domaine receveur, il faut donc clarifier le ou les niveaux de traitement de la PPB que l'on propose comme domaine(s) receveur(s). Dans la littérature de la revue systématique du chapitre 3, les étapes de la PPB qui sont les plus abordées comme domaine receveur sont le traitement auditif central et le traitement cognitif. En ce qui concerne le traitement auditif central, on

retrouve l'encodage sous-cortical du signal auditif et le traitement cortical spectrotemporal. Le transfert d'habileté entre des traitements auditifs commun à la musique et à la PPB pourrait être considéré comme du transfert proche. En effet, le stimulus auditif n'est pas le même (syllabes, mots, phrases VS note de musique, phrase musicale), mais le traitement et les structures cérébrales impliquées sont similaires (Patel, 2011, 2012, 2014). Ce transfert proche aurait ensuite des conséquences en aval sur des traitements propres à la perception de la parole. Par exemple, un encodage sous-cortical plus robuste et un traitement spectrotemporal plus précis pourraient tous deux contribuer à résister à la dégradation d'un masquage énergétique et faciliter l'accès aux représentations phonologiques internes. En ce qui concerne les fonctions cognitives, l'amélioration de l'attention sélective ou de la mémoire de travail pourraient agir sur des étapes communes entre la pratique musicale et la PPB, mais pourraient également agir directement sur des étapes propres à la perception de la parole, par exemple l'accès aux représentations phonologiques. En effet, selon le modèle *ELU* (Rönnberg et al., 2013) (chapitre 2), la mémoire de travail joue un rôle central dans l'accès aux représentations internes en conditions d'écoute difficile. Un effet direct de l'entraînement des fonctions cognitives par la pratique musicale sur des traitements propres à la perception de la parole pourrait être qualifié de transfert lointain. Par conséquent, selon le niveau de traitement de la PPB proposé comme domaine receveur, le transfert de la pratique musicale à la PPB pourrait être qualifié de proche lorsqu'il a lieu au sein de traitements communs entre la pratique musicale et la PPB, et de lointain lorsque le domaine receveur est un niveau de traitement qui n'est pas partagé entre la pratique musicale et la PPB.

Sur le plan du contexte physique, l'évaluation expérimentale de la PPB se fait dans le contexte contrôlé d'un laboratoire. Le contexte physique de la pratique musicale peut être, par exemple, une salle de classe, une salle de concert ou de encore chez soi. Dans la vie de tous les jours, les contextes physiques de la PPB sont extrêmement variables, par exemple, la parole dégradée par du bruit peut se rencontrer dans une salle de classe, un café, un supermarché, une salle de travail ou encore chez soi. Il y a donc plus de similarités entre les contextes physiques de la PPB dans la vie courante et ceux de la pratique musicale, qu'entre le contexte physique de l'évaluation de la PPB et celui de la pratique musicale et qu'entre celui de l'évaluation de la PPB et les contextes naturels de la PPB. Dans ce contexte, un

transfert de la pratique musicale aux situations de PPB de la vie de tous les jours serait par conséquent plus proche qu'un transfert de la pratique musicale vers l'évaluation expérimentale de la PPB.

Sur le plan du contexte temporel, dans les études transversales revues au chapitre 3, les musiciens ont une pratique musicale régulière, il n'y a donc pas de délai entre la fin d'un apprentissage et l'évaluation du transfert à la PPB. Dans les des études longitudinales, il est mentionné que l'évaluation de la PPB se fait à la fin d'un programme d'apprentissage musical, mais plus de précisions quant au délai entre l'arrêt du programme et l'évaluation de la PPB ne sont pas données. Il n'est donc pas possible d'évaluer si les changements entraînés par la pratique musicale perdurent après l'arrêt de l'apprentissage ou si une pratique en continu est nécessaire pour entretenir les bénéfices apportés.

En ce qui concerne le contexte fonctionnel, la pratique musicale peut s'inscrire dans un contexte plus scolaire lorsqu'il s'agit de cours, mais également dans des contextes de loisirs. La PPB dans la vie de tous les jours se retrouve dans des contextes de la routine comme dans des contextes de loisirs (p. ex. au restaurant, lors de réunions de famille), mais aussi dans des lieux publics comme les épiceries ou dans la rue. Comme mentionné précédemment, l'évaluation de la PPB se fait dans un contexte contrôlé, dissocié des activités de la vie tous les jours. Sur le plan du contexte fonctionnel, le transfert de la pratique musicale le plus proche s'opère sur les contextes naturels de PPB.

S'agissant du contexte social, la pratique musicale peut être individuelle tout comme elle peut se faire en groupe, que ce soit dans le cadre d'un cours ou d'un ensemble vocal ou instrumental. Dans la vie de tous les jours, la PPB se déroule dans la plupart des cas dans un contexte d'interaction sociale. Dans certains cas, il peut toutefois s'agir de contextes non sociaux, comme l'écoute de la radio ou de la télévision. L'évaluation de la PPB, en revanche, se fait toujours de façon individuelle, à l'aide de stimuli enregistrés par un ou deux locuteurs, et de façon décontextualisée.

Finalement, en ce qui concerne la modalité sensorielle, la pratique musicale implique l'audition, la vision et la proprioception. La perception de la parole dans la vie de tous les jours est également multisensorielle. En effet, comme vu au chapitre 1, la perception de la

parole est audiovisuelle ; les mouvements articulatoires, mais également l'expression faciale et le langage corporel sont accessibles à la vision dans la majorité des cas. Il existe toutefois certains contextes dans lesquels la PPB est unimodale, comme dans le cas de conversations téléphoniques ou de l'écoute de la radio en présence de bruit. Les tâches de PPB utilisées dans les études sur l'effet de la pratique musicale sont uniquement unimodales et c'est la modalité auditive qui est présentée exclusivement. Par conséquent, ce n'est qu'une partie des traitements impliqués dans la PPB qui est évaluée. En ce qui concerne le moyen d'évaluation, les tâches de PPB présentent des extraits de parole isolés, séparés par un temps de réponse. Les extraits de parole sont dans la plupart des cas parlés par un même locuteur. Le BM est soit un bruit purement énergétique, soit un BM exclusivement composé de parole. Par ailleurs, les stimuli proviennent soit de la même source sonore, soit de deux ou trois sources statiques séparées spatialement. Or, dans la vie courante, ce sont des conversations entières qu'il faut percevoir en continu et en temps réel, la plupart du temps en présence de plusieurs locuteurs séparés spatialement dont les interventions sont spontanées, et dans un bruit de fond composé de bruits organiques tout comme inorganiques.

L'analyse des différents contextes de la pratique musicale, de l'évaluation de la PPB et de son utilisation dans la vie de tous les jours à travers les différentes dimensions du *When ?* suggère que le transfert entre la pratique musicale et l'évaluation de la PPB soit plus lointain qu'entre la pratique musicale et la PPB dans un contexte naturel. En effet, les tâches de PPB ne sont ancrées dans aucun contexte de la vie de tous les jours. Elles se déroulent dans des environnements statiques, contrôlés, avec des stimuli unimodaux, dont des extraits de paroles isolés, souvent produits par le même locuteur, masqués par des BMs plus simples que les bruits de fond rencontrés dans la plupart des contextes de la vie de tous les jours. Par conséquent, il est possible que les avantages des musiciens trouvés dans des tâches de PPB ne soient pas représentatifs de la capacité à percevoir la parole dans le bruit des musiciens dans un contexte naturel de PPB. Par exemple, sur le plan de la modalité, différentes études ont montré que, comparée à la modalité auditive uniquement, la perception de la parole était améliorée en présence de stimuli audiovisuels (Schubotz et al., 2021; Schwartz et al., 2004). Or, toutes les études ayant trouvé des bénéfices des musiciens en PPB démontrent des

avantages dans des conditions expérimentales de PPB auditives uniquement. En présence de stimuli audiovisuels, les bénéfices apportés par la pratique musicale dans la modalité auditive demeurent-ils ? Il est possible que les bénéfices de la pratique musicale ne demeurent pas en présence de parole audiovisuelle (AV) et que les performances supérieures des musiciens en PPB ne soient qu'un artéfact de la présentation unimodale des tâches. Ces avantages n'apporteraient par conséquent pas d'améliorations dans les contextes de réunions sociales qui se déroulent dans des lieux publics, tels que les restaurants ou les cafés. Ces avantages demeureraient néanmoins pertinents dans les contextes de PPB auditifs uniquement, comme lors de conversations téléphoniques ou de l'écoute de la radio en présence de bruits environnants. Cependant, il est également envisageable que le bénéfice auditif et/ou cognitif de la pratique musicale isolé à l'aide des tâches auditives demeure, même en présence d'un support visuel pour la parole et que les musiciens soient par conséquent avantagés également dans les contextes multisensoriels de PPB.

Alternativement, il est possible que les tâches auditives ne révèlent qu'une partie des bénéfices de la pratique musicale. En effet, en plus d'une amélioration sur le plan auditif, il est envisageable que l'intégration multisensorielle de la parole AV soit également améliorée par la pratique musicale. En effet, des études suggèrent que l'intégration AV est améliorée chez les musiciens. Par exemple, une étude a montré que les musiciens présentaient un meilleur jugement de la synchronicité temporelle de stimuli audiovisuels (Lu et al., 2014) et une autre étude a montré que les musiciens reconnaissaient mieux les émotions présentées avec des stimuli audiovisuels que des non-musiciens (Weijkamp & Sadakata, 2017). Des bénéfices de la pratique musicale sur la parole AV pourraient être particulièrement pertinents pour les personnes âgées, dont les capacités d'intégrations audiovisuelles baissent avec l'âge (Tremblay, Basirat, et al., 2021). Les conditions expérimentales des tâches de PPB utilisées dans les études sur les musiciens et les non-musiciens restreignent par conséquent également l'exploration du transfert de la pratique musicale vers la PPB sur le plan du *What* ?. En effet, les tâches uniquement auditives limitent la possibilité de détecter le transfert d'habiletés multisensorielles.

Pour se rapprocher du réel impact fonctionnel de la pratique musicale sur la PPB et pour pouvoir explorer le transfert d'un plus large éventail d'habiletés de la pratique musicale

à la PPB, des études avec des tâches de PPB plus écologiques sont nécessaires. Il existe par exemple des tâches qui évaluent la perception de la parole AV. Ces tâches consistent généralement en la présentation d'une vidéo montrant un locuteur accompagné du son, et de séquences visuelles et auditives présentées séparément. À ce jour, il n'existe à ma connaissance aucune étude ayant évalué l'effet de la pratique musicale sur la PPB avec une tâche AV. Une tâche AV, bien qu'encore éloignée de la réalité de la PPB, permettrait tout de même d'évaluer le transfert d'habiletés multisensorielles et son impact fonctionnel. Des tâches de perception de la parole en réalité virtuelle ont également été développées. Grâce à cette technologie, qui crée des environnements immersifs, dynamiques, en 3 dimensions et multisensoriels, la perception de la parole pourrait être évaluée de façon plus écologique, tout en gardant un contrôle expérimental (pour une revue, voir : Peeters, 2019). En effet, la réalité virtuelle permettrait de réconcilier plusieurs aspects manquant dans les tâches de PPB auditives. Elles permettraient de présenter un ou plusieurs locuteurs cibles séparés spatialement en simulant un environnement de PPB de la vie de tous les jours autant dans son aspect sonore que visuel.

En somme, évaluer de façon plus écologique le transfert de la pratique musicale à la PPB permettrait de d'évaluer plus précisément son bénéfice fonctionnel et par conséquent, permettrait également d'évaluer la pertinence pour la qualité de vie des aînés de la pratique musicale comme outil de prévention contre le déclin de la PPB lié à l'âge.

Conclusion

Dans ce mémoire, des bénéfices de la pratique musicale sur la PPB ont été montrés de manière empirique et grâce à une revue de littérature. Les résultats présentés ont mis en lumière les différents contextes expérimentaux de PPB dans lesquels les musiciens présentent un avantage. De plus, les résultats de ce mémoire ont montré qu'une pratique de niveau amateur était associée à des bénéfices en PPB ainsi qu'avec une limitation de l'effet de l'âge sur de la PPB. En somme, les résultats de ce mémoire affinent les connaissances sur les effets comportementaux de la plasticité induite par la pratique musicale. Ces nouvelles connaissances contribueront à guider l'élaboration de futures études sur l'effet de la pratique musicale sur la PPB.

Bibliographie

- Akeroyd, M. A. (2008, 2008/01/01). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology*, 47(sup2), S53-S71. <https://doi.org/10.1080/14992020802301142>
- Alain, C., & Bernstein, L. J. (2015). Auditory Scene Analysis. *Music Perception*, 33(1), 70-82. <https://doi.org/10.1525/mp.2015.33.1.70>
- Amer, T., Kalender, B., Hasher, L., Trehub, S. E., & Wong, Y. (2013). Do older professional musicians have cognitive advantages? *PLoS One*, 8(8), e71630. <https://doi.org/10.1371/journal.pone.0071630>
- Anaya, E. M., Pisoni, D. B., & Kronenberger, W. G. (2016). Long-term musical experience and auditory and visual perceptual abilities under adverse conditions. *The Journal of the Acoustical Society of America*, 140(3), 2074-2081. <https://doi.org/10.1121/1.4962628>
- Arbogast, T. L., Mason, C. R., & Kidd, G., Jr. (2005, Apr). The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*, 117(4 Pt 1), 2169-2180. <https://doi.org/10.1121/1.1861598>
- Arnold, P., & Hill, F. (2001, May). Bisensory augmentation : A speechreading advantage when speech is clearly audible and intact. *Br J Psychol*, 92 Part 2, 339-355.
- Baddeley, A. (1992). Working Memory. *Science*, 255(5044), 556-559. <https://doi.org/doi:10.1126/science.1736359>
- Bailey, J., & Penhune, V. B. (2012, Apr). A sensitive period for musical training: contributions of age of onset and cognitive abilities. *Ann N Y Acad Sci*, 1252, 163-170. <https://doi.org/10.1111/j.1749-6632.2011.06434.x>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612-637. <https://doi.org/10.1037//0033-2909.128.4.612>
- Barulli, D., & Stern, Y. (2013, Oct). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends Cogn Sci*, 17(10), 502-509. <https://doi.org/10.1016/j.tics.2013.08.012>
- Baumann, P. S., Hagmann, P., Marquet, P., & Conus, P. (2010). De l'eau et des neurones... IRM de diffusion, cartographie cérébrale et brèves perspectives psychiatriques. *L'information psychiatrique*, 86(6). <https://doi.org/10.3917/inpsy.8606.0493>
- Bédard, P., Audet, A. M., Drouin, P., Roy, J. P., Rivard, J., & Tremblay, P. (2017, Oct). SyllabO+: A new tool to study sublexical phenomena in spoken Quebec French. *Behav Res Methods*, 49(5), 1852-1863. <https://doi.org/10.3758/s13428-016-0829-7>

- Benichov, J., Cox, L. C., Tun, P. A., & Wingfield, A. (2012, Mar-Apr). Word recognition within a linguistic context: effects of age, hearing acuity, verbal ability, and cognitive function. *Ear Hear*, 33(2), 250-256. <https://doi.org/10.1097/AUD.0b013e31822f680f>
- Bennett, K. O., Billings, C. J., Molis, M. R., & Leek, M. R. (2012, Mar-Apr). Neural encoding and perception of speech signals in informational masking. *Ear Hear*, 33(2), 231-238. <https://doi.org/10.1097/AUD.0b013e31823173fd>
- Bidelman, G. M., & Alain, C. (2015, Jan 21). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J Neurosci*, 35(3), 1240-1249. <https://doi.org/10.1523/jneurosci.3292-14.2015>
- Bidelman, G. M., Weiss, M. W., Moreno, S., & Alain, C. (2014, Aug). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci*, 40(4), 2662-2673. <https://doi.org/10.1111/ejn.12627>
- Billings, C. J., & Madsen, B. M. (2018, Nov). A perspective on brain-behavior relationships and effects of age and hearing using speech-in-noise stimuli. *Hear Res*, 369, 90-102. <https://doi.org/10.1016/j.heares.2018.03.024>
- Bilodeau-Mercure, M., Lortie, C. L., Sato, M., Guitton, M. J., & Tremblay, P. (2015, Mar). The neurobiology of speech perception decline in aging. *Brain Struct Funct*, 220(2), 979-997. <https://doi.org/10.1007/s00429-013-0695-3>
- Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015, Jan). Musicians and non-musicians are equally adept at perceiving masked speech. *J Acoust Soc Am*, 137(1), 378-387. <https://doi.org/10.1121/1.4904537>
- Boersma, P., & Weenink, D. (2018). *Praat: doing phonetics by computer [Computer program]*.
- Bowe, K. L., Chrobak, Q. M., & Karst, A. T. (2021, Apr 9). Active maintenance of musical and linguistic information as a function of musical experience. *J Gen Psychol*, 1-18. <https://doi.org/10.1080/00221309.2021.1908945>
- Boyke, J., Driemeyer, J., Gaser, C., Büchel, C., & May, A. (2008, Jul 9). Training-induced brain structure changes in the elderly. *J Neurosci*, 28(28), 7031-7035. <https://doi.org/10.1523/jneurosci.0742-08.2008>
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. The MIT Press.
- Brouwer, S., Van Engen, K. J., Calandruccio, L., & Bradlow, A. R. (2012, Feb). Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content. *J Acoust Soc Am*, 131(2), 1449-1464. <https://doi.org/10.1121/1.3675943>
- Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, 109(3), 1101-1109. <https://doi.org/10.1121/1.1345696>

- Brunner, D., Abramovitch, A., & Etherton, J. (2017). A yoga program for cognitive enhancement. *PLoS One*, *12*(8), e0182366. <https://doi.org/10.1371/journal.pone.0182366>
- Bugos, J. A., Perlstein, W. M., McCrae, C. S., Brophy, T. S., & Bedenbaugh, P. H. (2007, Jul). Individualized piano instruction enhances executive functioning and working memory in older adults. *Aging Ment Health*, *11*(4), 464-471. <https://doi.org/10.1080/13607860601086504>
- Buss, E., Hodge, S. E., Calandruccio, L., Leibold, L. J., & Grose, J. H. (2019, Sep/Oct). Masked Sentence Recognition in Children, Young Adults, and Older Adults: Age-Dependent Effects of Semantic Context and Masker Type. *Ear Hear*, *40*(5), 1117-1126. <https://doi.org/10.1097/AUD.0000000000000692>
- Cabeza, R., Albert, M., Belleville, S., Craik, F. I. M., Duarte, A., Grady, C. L., Lindenberger, U., Nyberg, L., Park, D. C., Reuter-Lorenz, P. A., Rugg, M. D., Steffener, J., & Rajah, M. N. (2018, Nov). Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat Rev Neurosci*, *19*(11), 701-710. <https://doi.org/10.1038/s41583-018-0068-2>
- Carter, J. A., & Bidelman, G. M. (2021, May 15). Auditory cortex is susceptible to lexical influence as revealed by informational vs. energetic masking of speech categorization. *Brain Res*, *1759*, 147385. <https://doi.org/10.1016/j.brainres.2021.147385>
- Cervera, T., & Rosell, V. (2015, Dec). The Effects of Linguistic Context on Word Recognition in Noise by Elderly Listeners Using Spanish Sentence Lists (SSL). *J Psycholinguist Res*, *44*(6), 819-829. <https://doi.org/10.1007/s10936-014-9321-7>
- Cisneros-Franco, J. M., Voss, P., Thomas, M. E., & de Villers-Sidani, E. (2020). Critical periods of brain development. *Handb Clin Neurol*, *173*, 75-88. <https://doi.org/10.1016/B978-0-444-64150-2.00009-5>
- Colcombe, S. J., Erickson, K. I., Scalf, P. E., Kim, J. S., Prakash, R., McAuley, E., Elavsky, S., Marquez, D. X., Hu, L., & Kramer, A. F. (2006, Nov). Aerobic exercise training increases brain volume in aging humans. *J Gerontol A Biol Sci Med Sci*, *61*(11), 1166-1170. <https://doi.org/10.1093/gerona/61.11.1166>
- Cooke, M. (2006). A glimpsing model of speech perception in noise. *The Journal of the Acoustical Society of America*, *119*(3), 1562-1573. <https://doi.org/10.1121/1.2166600>
- Cools, R. (2010). Listening Span Test. In I. P. Stolerman (Ed.), *Encyclopedia of Psychopharmacology* (pp. 713-713). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-68706-1_1626
- Courchesne, E., Chisum, H. J., Townsend, J., Cowles, A., Covington, J., Egaas, B., Harwood, M., Hinds, S., & Press, G. A. (2000, Sep). Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology*, *216*(3), 672-682. <https://doi.org/10.1148/radiology.216.3.r00au37672>
- D'Souza, A. A., Moradzadeh, L., & Wiseheart, M. (2018). Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cogn Res Princ Implic*, *3*(1), 11. <https://doi.org/10.1186/s41235-018-0095-6>

- Dahlin, E., Nyberg, L., Backman, L., & Neely, A. S. (2008, Dec). Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol Aging*, 23(4), 720-730. <https://doi.org/10.1037/a0014296>
- De Cheveigné, A. (2002). L'analyse de scènes auditives computationnelle. In Hermès & J. Mariani (Eds.), *La parole, des modèles cognitifs aux machines communicantes - Analyse, synthèse et codage de la parole*, None (pp. 175-196). <https://hal.archives-ouvertes.fr/hal-01105769>
- Deguchi, C., Boureux, M., Sarlo, M., Besson, M., Grassi, M., Schon, D., & Colombo, L. (2012, May 21). Sentence pitch change detection in the native and unfamiliar language in musicians and non-musicians: behavioral, electrophysiological and psychoacoustic study. *Brain Res*, 1455, 75-89. <https://doi.org/10.1016/j.brainres.2012.03.034>
- Deroche, M. L. D., Limb, C. J., Chatterjee, M., & Gracco, V. L. (2017). Similar abilities of musicians and non-musicians to segregate voices by fundamental frequency. *The Journal of the Acoustical Society of America*, 142(4), 1739-1755. <https://doi.org/10.1121/1.5005496>
- Desjardins, J. L., & Doherty, K. A. (2013, May-Jun). Age-related changes in listening effort for various types of masker noises. *Ear Hear*, 34(3), 261-272. <https://doi.org/10.1097/AUD.0b013e31826d0ba4>
- Ding, Y., Gray, K., Forrence, A., Wang, X., & Huang, J. (2018). A behavioral study on tonal working memory in musicians and non-musicians. *PLoS One*, 13(8), e0201765. <https://doi.org/10.1371/journal.pone.0201765>
- Donai, J. J., & Jennings, M. B. (2016). Gaps-in-noise detection and gender identification from noise-vocoded vowel segments: Comparing performance of active musicians to non-musicians. *The Journal of the Acoustical Society of America*, 139(5), EL128-EL134. <https://doi.org/10.1121/1.4947070>
- Douglas, H., Georgiou, A., & Westbrook, J. (2017, Aug). Social participation as an indicator of successful aging: an overview of concepts and their associations with health. *Aust Health Rev*, 41(4), 455-462. <https://doi.org/10.1071/AH16038>
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004, Jan 22). Neuroplasticity: changes in grey matter induced by training. *Nature*, 427(6972), 311-312. <https://doi.org/10.1038/427311a>
- Draganski, B., & May, A. (2008, Sep 1). Training-induced structural changes in the adult human brain. *Behav Brain Res*, 192(1), 137-142. <https://doi.org/10.1016/j.bbr.2008.02.015>
- Dryden, A., Allen, H. A., Henshaw, H., & Heinrich, A. (2017, Jan-Dec). The Association Between Cognitive Performance and Speech-in-Noise Perception for Adult Listeners: A Systematic Literature Review and Meta-Analysis. *Trends Hear*, 21, 2331216517744675. <https://doi.org/10.1177/2331216517744675>
- Du, Y., & Zatorre, R. J. (2017, Dec 19). Musical training sharpens and bonds ears and tongue to hear speech better. *Proc Natl Acad Sci U S A*, 114(51), 13579-13584. <https://doi.org/10.1073/pnas.1712223114>

- Dubno, J. R., Dirks, D. D., & Morgan, D. E. (1984, Jul). Effects of age and mild hearing loss on speech recognition in noise. *J Acoust Soc Am*, 76(1), 87-96. <https://doi.org/10.1121/1.391011>
- Egan, J. P. (1971, 1971/01/01). Auditory Masking and Signal Detection Theory. *Audiology*, 10(1), 41-47. <https://doi.org/10.3109/00206097109072539>
- El-Assal, H. A., El-Gharib, A. M., Kolkaila, E. A., & Elmahallawy, T. H. (2020, 2020/04/02). Encoding of speech in noise in adults using hearing aids: effect of noise reduction algorithm. *Hearing, Balance and Communication*, 18(2), 98-104. <https://doi.org/10.1080/21695717.2019.1667687>
- Ericsson, K. A. (2020, Jan). Towards a science of the acquisition of expert performance in sports: Clarifying the differences between deliberate practice and other types of practice. *J Sports Sci*, 38(2), 159-176. <https://doi.org/10.1080/02640414.2019.1688618>
- Ericsson, K. A., & Harwell, K. W. (2019). Deliberate Practice and Proposed Limits on the Effects of Practice on the Acquisition of Expert Performance: Why the Original Definition Matters and Recommendations for Future Research. *Front Psychol*, 10, 2396. <https://doi.org/10.3389/fpsyg.2019.02396>
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363-406. <https://doi.org/10.1037/0033-295X.100.3.363>
- Fennell, A. M., Bugos, J. A., Payne, B. R., & Schotter, E. R. (2020, Dec 2). Music is similar to language in terms of working memory interference. *Psychon Bull Rev*, 1-14. <https://doi.org/10.3758/s13423-020-01833-5>
- Ferreri, L., Mas-Herrero, E., Zatorre, R. J., Ripolles, P., Gomez-Andres, A., Alicart, H., Olive, G., Marco-Pallares, J., Antonijoan, R. M., Valle, M., Riba, J., & Rodriguez-Fornells, A. (2019, Feb 26). Dopamine modulates the reward experiences elicited by music. *Proc Natl Acad Sci U S A*, 116(9), 3793-3798. <https://doi.org/10.1073/pnas.1811878116>
- Fostick, L. (2019, Dec). Card playing enhances speech perception among aging adults: comparison with aging musicians. *Eur J Ageing*, 16(4), 481-489. <https://doi.org/10.1007/s10433-019-00512-2>
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *Journal of the Acoustical Society of America*, 109(5,Pt1), 2112-2122. <https://doi.org/10.1121/1.1354984>
- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *The Journal of the Acoustical Society of America*, 115(5), 2246-2256. <https://doi.org/10.1121/1.1689343>
- Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999, Dec). The role of perceived spatial separation in the unmasking of speech. *J Acoust Soc Am*, 106(6), 3578-3588. <https://doi.org/10.1121/1.428211>

- Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., Den Ouden, D. B., & Rorden, C. (2016, Dec 27). Revealing the dual streams of speech processing. *Proc Natl Acad Sci U S A*, *113*(52), 15108-15113. <https://doi.org/10.1073/pnas.1614038114>
- Frisina, D. R., & Frisina, R. D. (1997, Apr). Speech recognition in noise and presbycusis: relations to possible neural mechanisms. *Hear Res*, *106*(1-2), 95-104. [https://doi.org/10.1016/s0378-5955\(97\)00006-3](https://doi.org/10.1016/s0378-5955(97)00006-3)
- Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010, Jul 1). Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage*, *51*(3), 943-951. <https://doi.org/10.1016/j.neuroimage.2010.03.004>
- Glyde, H., Hickson, L., Cameron, S., & Dillon, H. (2011, Sep). Problems hearing in noise in older adults: a review of spatial processing disorder. *Trends Amplif*, *15*(3), 116-126. <https://doi.org/10.1177/1084713811424885>
- Gordon-Salant, S., & Cole, S. S. (2016, Sep-Oct). Effects of Age and Working Memory Capacity on Speech Recognition Performance in Noise Among Listeners With Normal Hearing. *Ear Hear*, *37*(5), 593-602. <https://doi.org/10.1097/AUD.0000000000000316>
- Gordon-Salant, S., & Fitzgibbons, P. J. (1997, Apr). Selected cognitive factors and speech recognition performance among young and elderly listeners. *J Speech Lang Hear Res*, *40*(2), 423-431. <https://doi.org/10.1044/jslhr.4002.423>
- Grassi, M., Meneghetti, C., Toffalini, E., & Borella, E. (2017). Auditory and cognitive performance in elderly musicians and nonmusicians. *PLoS One*, *12*(11), e0187881. <https://doi.org/10.1371/journal.pone.0187881>
- Groussard, M., Coppalle, R., Hinault, T., & Platel, H. (2020). Do Musicians Have Better Mnemonic and Executive Performance Than Actors? Influence of Regular Musical or Theater Practice in Adults and in the Elderly. *Front Hum Neurosci*, *14*, 557642. <https://doi.org/10.3389/fnhum.2020.557642>
- Guinchard, A.-C., Estoppey, P., & Maire, R. (2017). Presbyacousie : baisse de l'ouïe chez la personne âgée. *Forum Médical Suisse*, *17*, 230-235.
- Hanna-Pladdy, B., & Gajewski, B. (2012). Recent and past musical activity predicts cognitive aging variability: direct comparison with general lifestyle activities. *Front Hum Neurosci*, *6*, 198. <https://doi.org/10.3389/fnhum.2012.00198>
- Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013, Nov). Normal cognitive aging. *Clin Geriatr Med*, *29*(4), 737-752. <https://doi.org/10.1016/j.cger.2013.07.002>
- Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004, Feb). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J Acoust Soc Am*, *115*(2), 833-843. <https://doi.org/10.1121/1.1639908>
- Heald, S., & Nusbaum, H. (2014, 2014-March-17). Speech perception as an active cognitive process [Hypothesis and Theory]. *Frontiers in Systems Neuroscience*, *8*(35). <https://doi.org/10.3389/fnsys.2014.00035>

- Hedden, T., & Gabrieli, J. D. (2004, Feb). Insights into the ageing mind: a view from cognitive neuroscience. *Nat Rev Neurosci*, 5(2), 87-96. <https://doi.org/10.1038/nrn1323>
- Heine, C., & Browning, C. J. (2002, Oct 15). Communication and psychosocial consequences of sensory loss in older adults: overview and rehabilitation directions. *Disabil Rehabil*, 24(15), 763-773. <https://doi.org/10.1080/09638280210129162>
- Helfer, K. S., & Freyman, R. L. (2014). Stimulus and listener factors affecting age-related changes in competing speech perception. *The Journal of the Acoustical Society of America*, 136(2), 748-759. <https://doi.org/10.1121/1.4887463>
- Hennessy, S., Wood, A., Wilcox, R., & Habibi, A. (2021, Apr 6). Neurophysiological improvements in speech-in-noise task after short-term choir training in older adults. *Aging (Albany NY)*, 13(7), 9468-9495. <https://doi.org/10.18632/aging.202931>
- Hickok, G., & Poeppel, D. (2007, 2007/05/01). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393-402. <https://doi.org/10.1038/nrn2113>
- Hickok, G., & Poeppel, D. (2016). Neural Basis of Speech Perception. In *Neurobiology of Language* (pp. 299-310). <https://doi.org/10.1016/b978-0-12-407794-2.00025-0>
- Hoen, M., Meunier, F., Grataloup, C.-L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., & Collet, L. (2007). Phonetic and lexical interferences in informational masking during speech-in-speech comprehension. *Speech Communication*, 49(12), 905-916. <https://doi.org/10.1016/j.specom.2007.05.008>
- Holman, C., & de Villers-Sidani, E. (2014). Indestructible plastic: the neuroscience of the new aging brain. *Front Hum Neurosci*, 8, 219. <https://doi.org/10.3389/fnhum.2014.00219>
- Humes, L. E., Wilson, D. L., Barlow, N. N., & Garner, C. (2002, Aug). Changes in hearing-aid benefit following 1 or 2 years of hearing-aid use by older adults. *J Speech Lang Hear Res*, 45(4), 772-782. [https://doi.org/10.1044/1092-4388\(2002/062\)](https://doi.org/10.1044/1092-4388(2002/062))
- Hutchinson, K. M. (1989, Mar). Influence of sentence context on speech perception in young and older adults. *J Gerontol*, 44(2), P36-44. <https://doi.org/10.1093/geronj/44.2.p36>
- Hutka, S., Bidelman, G. M., & Moreno, S. (2015, May). Pitch expertise is not created equal: Cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia*, 71, 52-63. <https://doi.org/10.1016/j.neuropsychologia.2015.03.019>
- Jain, S., & Nataraja, N. P. (2019, Aug 28). The Effect of Fatigue on Working Memory and Auditory Perceptual Abilities in Trained Musicians. *Am J Audiol*, 28(2S), 483-494. https://doi.org/10.1044/2019_AJA-IND50-18-0102
- Jonides, J. (2004, 2004/01/01). How does practice makes perfect? *Nature Neuroscience*, 7(1), 10-11. <https://doi.org/10.1038/nrn0104-10>
- Juslin, P. N., & Laukka, P. (2004). Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *Journal of New Music Research*, 33(3), 217-238. <https://doi.org/10.1080/0929821042000317813>

- Kelly, M. E., Duff, H., Kelly, S., McHugh Power, J. E., Brennan, S., Lawlor, B. A., & Loughrey, D. G. (2017, Dec 19). The impact of social activities, social networks, social support and social relationships on the cognitive functioning of healthy older adults: a systematic review. *Syst Rev*, 6(1), 259. <https://doi.org/10.1186/s13643-017-0632-2>
- Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., & Durlach, N. I. (2008). Informational Masking. In W. A. Yost, A. N. Popper, & R. R. Fay (Eds.), *Auditory Perception of Sound Sources* (pp. 143-189). Springer US. https://doi.org/10.1007/978-0-387-71305-2_6
- Kim, S., Frisina, R. D., Mapes, F. M., Hickman, E. D., & Frisina, D. R. (2006). Effect of age on binaural speech intelligibility in normal hearing adults. *Speech Communication*, 48(6), 591-597. <https://doi.org/10.1016/j.specom.2005.09.004>
- Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: are professional musicians better than non-musicians? *J Basic Clin Physiol Pharmacol*, 12(2 Suppl), 125-143. <https://doi.org/10.1515/jbcpp.2001.12.2.125>
- Knudsen, E. I. (2004). Sensitive Periods in the Development of the Brain and Behavior. *Journal of Cognitive Neuroscience*, 16(8), 1412-1425. <https://doi.org/10.1162/0898929042304796>
- Kolb, B. (2018). Brain Plasticity and Experience. In *The Neurobiology of Brain and Behavioral Development* (pp. 341-389). <https://doi.org/10.1016/b978-0-12-804036-2.00013-3>
- Kolb, B., Mychasiuk, R., Muhammad, A., & Gibb, R. (2013). Brain plasticity in the developing brain. *Prog Brain Res*, 207, 35-64. <https://doi.org/10.1016/B978-0-444-63327-9.00005-9>
- Krampe, R. T., & Ericsson, K. A. (1996, Dec). Maintaining excellence: deliberate practice and elite performance in young and older pianists. *J Exp Psychol Gen*, 125(4), 331-359. <https://doi.org/10.1037//0096-3445.125.4.331>
- Krueger, K. R., Wilson, R. S., Kamenetsky, J. M., Barnes, L. L., Bienias, J. L., & Bennett, D. A. (2009, Jan-Mar). Social engagement and cognitive function in old age. *Exp Aging Res*, 35(1), 45-60. <https://doi.org/10.1080/03610730802545028>
- Kuhnis, J., Elmer, S., Meyer, M., & Jancke, L. (2013, Jul). The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study. *Neuropsychologia*, 51(8), 1608-1618. <https://doi.org/10.1016/j.neuropsychologia.2013.04.007>
- Kumar, P. V., Rana, B., & Krishna, R. (2014). Temporal processing in musicians and non-musicians [journal article]. *Journal of Hearing Science*, 4(3), 35-42. <https://www.journalofhearing.com/TEMPORAL-PROCESSING-IN-MUSICIANS-AND-NON-MUSICIANS,120632,0,2.html>
- Larouche, E., Tremblay, M. P., Potvin, O., Laforest, S., Bergeron, D., Laforce, R., Monetta, L., Boucher, L., Tremblay, P., Belleville, S., Lorrain, D., Gagnon, J. F., Gosselin, N., Castellano, C. A., Cunnane, S. C., Macoir, J., & Hudon, C. (2016, Nov 22). Normative Data for the Montreal Cognitive Assessment in Middle-Aged and Elderly Quebec-French People. *Arch Clin Neuropsychol*, 31(7), 819-826. <https://doi.org/10.1093/arclin/acw076>

- Le Bihan, D. (2003, Jun). Looking into the functional architecture of the brain with diffusion MRI. *Nat Rev Neurosci*, 4(6), 469-480. <https://doi.org/10.1038/nrn1119>
- Liang, C., Earl, B., Thompson, I., Whitaker, K., Cahn, S., Xiang, J., Fu, Q. J., & Zhang, F. (2016). Musicians Are Better than Non-musicians in Frequency Change Detection: Behavioral and Electrophysiological Evidence. *Front Neurosci*, 10, 464. <https://doi.org/10.3389/fnins.2016.00464>
- Lu, Y., Paraskevopoulos, E., Herholz, S. C., Kuchenbuch, A., & Pantev, C. (2014). Temporal processing of audiovisual stimuli is enhanced in musicians: evidence from magnetoencephalography (MEG). *PLoS One*, 9(3), e90686. <https://doi.org/10.1371/journal.pone.0090686>
- Madsen, S. M. K., Marschall, M., Dau, T., & Oxenham, A. J. (2019, Jul 18). Speech perception is similar for musicians and non-musicians across a wide range of conditions. *Sci Rep*, 9(1), 10404. <https://doi.org/10.1038/s41598-019-46728-1>
- Maguire, E. A., Woollett, K., & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus*, 16(12), 1091-1101. <https://doi.org/10.1002/hipo.20233>
- Mansens, D., Deeg, D. J. H., & Comijs, H. C. (2018, Aug). The association between singing and/or playing a musical instrument and cognitive functions in older adults. *Aging Ment Health*, 22(8), 964-971. <https://doi.org/10.1080/13607863.2017.1328481>
- Markham, J. A., & Greenough, W. T. (2004, Nov). Experience-driven brain plasticity: beyond the synapse. *Neuron Glia Biol*, 1(4), 351-363. <https://doi.org/10.1017/s1740925x05000219>
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7-8), 953-978. <https://doi.org/10.1080/01690965.2012.705006>
- May, A. (2011, Oct). Experience-dependent structural plasticity in the adult human brain. *Trends Cogn Sci*, 15(10), 475-482. <https://doi.org/10.1016/j.tics.2011.08.002>
- McGuire, G. L. (2010). Perception A Brief Primer on Experimental Designs for Speech Perception Research.
- McGurk, H., & Macdonald, J. (1976, 1976/12/01). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748. <https://doi.org/10.1038/264746a0>
- Meha-Bettison, K., Sharma, M., Ibrahim, R. K., & Mandikal Vasuki, P. R. (2018, Jan). Enhanced speech perception in noise and cortical auditory evoked potentials in professional musicians. *Int J Audiol*, 57(1), 40-52. <https://doi.org/10.1080/14992027.2017.1380850>
- Merten, N., Fischer, M. E., Dillard, L. K., Klein, B. E. K., Tweed, T. S., & Cruickshanks, K. J. (2021, Jul 16). Benefit of Musical Training for Speech Perception and Cognition Later in Life. *J Speech Lang Hear Res*, 64(7), 2885-2896. https://doi.org/10.1044/2021_JSLHR-20-00588

- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983, 1983/01/01). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, 6, 414-417. [https://doi.org/https://doi.org/10.1016/0166-2236\(83\)90190-X](https://doi.org/https://doi.org/10.1016/0166-2236(83)90190-X)
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007, Oct 2). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc Natl Acad Sci U S A*, 104(40), 15894-15898. <https://doi.org/10.1073/pnas.0701498104>
- Mussoi, B. S. (2021, Sep 29). The Impact of Music Training and Working Memory on Speech Recognition in Older Age. *J Speech Lang Hear Res*, 1-11. https://doi.org/10.1044/2021_JSLHR-20-00426
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005, Apr). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc*, 53(4), 695-699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- Nguyen, L., Murphy, K., & Andrews, G. (2019, Aug). Cognitive and neural plasticity in old age: A systematic review of evidence from executive functions cognitive training. *Ageing Res Rev*, 53, 100912. <https://doi.org/10.1016/j.arr.2019.100912>
- Nilsson, J., Berggren, R., Garzon, B., Lebedev, A. V., & Lovden, M. (2021). Second Language Learning in Older Adults: Effects on Brain Structure and Predictors of Learning Success. *Front Aging Neurosci*, 13, 666851. <https://doi.org/10.3389/fnagi.2021.666851>
- Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: gauging the generality of cognitive intervention effects. *Restor Neurol Neurosci*, 27(5), 435-453. <https://doi.org/10.3233/rnn-2009-0496>
- Oldfield, R. C. (1971, Mar). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Palmer, A. D., Carder, P. C., White, D. L., Saunders, G., Woo, H., Graville, D. J., & Newsom, J. T. (2019, Jan 30). The Impact of Communication Impairments on the Social Relationships of Older Adults: Pathways to Psychological Well-Being. *J Speech Lang Hear Res*, 62(1), 1-21. https://doi.org/10.1044/2018_JSLHR-S-17-0495
- Palmer, A. D., Newsom, J. T., & Rook, K. S. (2016, Jul-Aug). How does difficulty communicating affect the social relationships of older adults? An exploration using data from a national survey. *J Commun Disord*, 62, 131-146. <https://doi.org/10.1016/j.jcomdis.2016.06.002>
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998, Apr 23). Increased auditory cortical representation in musicians. *Nature*, 392(6678), 811-814. <https://doi.org/10.1038/33918>
- Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front Aging Neurosci*, 4, 30. <https://doi.org/10.3389/fnagi.2012.00030>
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician Enhancement for Speech-In-Noise. *Ear and Hearing*, 30(6), 653-661. <https://doi.org/10.1097/AUD.0b013e3181b412e9>

- Parbery-Clark, A., Strait, D. L., Anderson, S., Hittner, E., & Kraus, N. (2011, May 11). Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. *PLoS One*, 6(5), e18082. <https://doi.org/10.1371/journal.pone.0018082>
- Park, D. C., & Bischof, G. N. (2013, Mar). The aging mind: neuroplasticity in response to cognitive training. *Dialogues Clin Neurosci*, 15(1), 109-119. <https://doi.org/10.31887/DCNS.2013.15.1/dpark>
- Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Front Psychol*, 2, 142. <https://doi.org/10.3389/fpsyg.2011.00142>
- Patel, A. D. (2012, Apr). The OPERA hypothesis: assumptions and clarifications. *Ann NY Acad Sci*, 1252, 124-128. <https://doi.org/10.1111/j.1749-6632.2011.06426.x>
- Patel, A. D. (2014, Feb). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hear Res*, 308, 98-108. <https://doi.org/10.1016/j.heares.2013.08.011>
- Peeters, D. (2019, Jun). Virtual reality: A game-changing method for the language sciences. *Psychon Bull Rev*, 26(3), 894-900. <https://doi.org/10.3758/s13423-019-01571-3>
- Peng, Z. E., & Wang, L. M. (2019, Apr 15). Listening Effort by Native and Nonnative Listeners Due to Noise, Reverberation, and Talker Foreign Accent During English Speech Perception. *J Speech Lang Hear Res*, 62(4), 1068-1081. https://doi.org/10.1044/2018_JSLHR-H-17-0423
- Perrin, F., & Grimaud, N. (2005). *Fonds sonores*(Laboratoire Unités Mixtes de Recherche, Centre National de la Recherche Scientifique 5020, Lyon, France).
- Plomb, R., & Mimpen, A. M. (1980). Effect of the orientation of the speaker's head and azimuth of a noise source on the speech reception threshold for sentences.
- Pronk, M., Deeg, D. J. H., Festen, J. M., Twisk, J. W., Smits, C., Comijs, H. C., & Kramer, S. E. (2013). Decline in Older Persons' Ability to Recognize Speech in Noise: The Influence of Demographic, Health-Related, Environmental, and Cognitive Factors. *Ear and Hearing*, 34(6), 722-732. <https://doi.org/10.1097/AUD.0b013e3182994eee>
- Rajan, R., & Cainer, K. E. (2008, Jun 23). Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers. *Neuroscience*, 154(2), 784-795. <https://doi.org/10.1016/j.neuroscience.2008.03.067>
- Rauschecker, J. P., & Scott, S. K. (2009, 2009/06/01). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718-724. <https://doi.org/10.1038/nn.2331>
- Rehfeld, K., Luders, A., Hokelmann, A., Lessmann, V., Kaufmann, J., Brigadski, T., Muller, P., & Muller, N. G. (2018). Dance training is superior to repetitive physical exercise in inducing brain plasticity in the elderly. *PLoS One*, 13(7), e0196636. <https://doi.org/10.1371/journal.pone.0196636>

- Roberts, B., Summers, R. J., & Bailey, P. J. (2011, May 22). The intelligibility of noise-vocoded speech: spectral information available from across-channel comparison of amplitude envelopes. *Proc Biol Sci*, 278(1711), 1595-1600. <https://doi.org/10.1098/rspb.2010.1554>
- Rönnerberg, J. (2003, 2003/01/01). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *International Journal of Audiology*, 42(sup1), 68-76. <https://doi.org/10.3109/14992020309074626>
- Rönnerberg, J., Holmer, E., & Rudner, M. (2019, May). Cognitive hearing science and ease of language understanding. *Int J Audiol*, 58(5), 247-261. <https://doi.org/10.1080/14992027.2018.1551631>
- Rönnerberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., Dahlstrom, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci*, 7, 31. <https://doi.org/10.3389/fnsys.2013.00031>
- Rönnerberg, J., Rudner, M., Foo, C., & Lunner, T. (2008, 2008/01/01). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, 47(sup2), S99-S105. <https://doi.org/10.1080/14992020802301167>
- Rosen, S., Souza, P., Ekelund, C., & Majeed, A. A. (2013, Apr). Listening to speech in a background of other talkers: effects of talker number and noise vocoding. *J Acoust Soc Am*, 133(4), 2431-2443. <https://doi.org/10.1121/1.4794379>
- Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One*, 9(1), e86980. <https://doi.org/10.1371/journal.pone.0086980>
- Saklofske, D. H., & Schoenberg, M. R. (2011). Wechsler Adult Intelligence Scale (All Versions). In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 2675-2680). Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1073
- Salthouse, T. A. (1996, Jul). The processing-speed theory of adult age differences in cognition. *Psychol Rev*, 103(3), 403-428. <https://doi.org/10.1037/0033-295x.103.3.403>
- Santini, Z. I., Jose, P. E., York Cornwell, E., Koyanagi, A., Nielsen, L., Hinrichsen, C., Meilstrup, C., Madsen, K. R., & Koushede, V. (2020). Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *The Lancet Public Health*, 5(1), e62-e70. [https://doi.org/10.1016/s2468-2667\(19\)30230-0](https://doi.org/10.1016/s2468-2667(19)30230-0)
- Sares, A. G., Foster, N. E. V., Allen, K., & Hyde, K. L. (2018, Mar 15). Pitch and Time Processing in Speech and Tones: The Effects of Musical Training and Attention. *J Speech Lang Hear Res*, 61(3), 496-509. https://doi.org/10.1044/2017_JSLHR-S-17-0207
- Schlaug, G. (2001, Jun). The brain of musicians. A model for functional and structural adaptation. *Ann N Y Acad Sci*, 930, 281-299.
- Schlaug, G. (2015). Musicians and music making as a model for the study of brain plasticity. *Prog Brain Res*, 217, 37-55. <https://doi.org/10.1016/bs.pbr.2014.11.020>

- Schlaug, G., Jancke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. *Science*, 267(5198), 699-701. <https://doi.org/10.1126/science.7839149>
- Schon, D., Magne, C., & Besson, M. (2004, May). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology*, 41(3), 341-349. <https://doi.org/10.1111/1469-8986.00172.x>
- Schubotz, L., Holler, J., Drijvers, L., & Ozyurek, A. (2021, Jul). Aging and working memory modulate the ability to benefit from visible speech and iconic gestures during speech-in-noise comprehension. *Psychol Res*, 85(5), 1997-2011. <https://doi.org/10.1007/s00426-020-01363-8>
- Schwartz, J. L., Berthommier, F., & Savariaux, C. (2004, Sep). Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition*, 93(2), B69-78. <https://doi.org/10.1016/j.cognition.2004.01.006>
- Shinn-Cunningham, B. (2013). *Understanding informational masking from a neural perspective*
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012, Jul). Is working memory training effective? *Psychol Bull*, 138(4), 628-654. <https://doi.org/10.1037/a0027473>
- Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017, Jan). The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang*, 164, 77-105. <https://doi.org/10.1016/j.bandl.2016.10.004>
- Slater, J., & Kraus, N. (2016, Feb). The role of rhythm in perceiving speech in noise: a comparison of percussionists, vocalists and non-musicians. *Cogn Process*, 17(1), 79-87. <https://doi.org/10.1007/s10339-015-0740-7>
- Souza, P. E., & Turner, C. W. (1994, Jun). Masking of speech in young and elderly listeners with hearing loss. *J Speech Hear Res*, 37(3), 655-661. <https://doi.org/10.1044/jshr.3703.655>
- Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013, Jan 16). Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. *J Neurosci*, 33(3), 1282-1290. <https://doi.org/10.1523/JNEUROSCI.3578-12.2013>
- Stern, Y. (2002, Mar). What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc*, 8(3), 448-460.
- Stern, Y. (2006, Apr-Jun). Cognitive reserve and Alzheimer disease. *Alzheimer Dis Assoc Disord*, 20(2), 112-117. <https://doi.org/10.1097/01.wad.0000213815.20177.19>
- Stern, Y. (2009, Aug). Cognitive reserve. *Neuropsychologia*, 47(10), 2015-2028. <https://doi.org/10.1016/j.neuropsychologia.2009.03.004>
- Stern, Y., Arenaza-Urquijo, E. M., Bartres-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., Vuoksimaa, E., the Reserve, R., Protective Factors, P. I. A. E. D., & Conceptual Frameworks, W. (2020, Sep). Whitepaper:

- Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimers Dement*, 16(9), 1305-1311. <https://doi.org/10.1016/j.jalz.2018.07.219>
- Stewart, H. J., Martinez, J. L., Perdeu, A., Green, C. S., & Moore, D. R. (2020, Sep 1). Auditory cognition and perception of action video game players. *Sci Rep*, 10(1), 14410. <https://doi.org/10.1038/s41598-020-71235-z>
- Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010, Mar). Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. *Hear Res*, 261(1-2), 22-29. <https://doi.org/10.1016/j.heares.2009.12.021>
- Sussman, E. S. (2017, Oct 17). Auditory Scene Analysis: An Attention Perspective. *J Speech Lang Hear Res*, 60(10), 2989-3000. https://doi.org/10.1044/2017_JSLHR-H-17-0041
- Taitelbaum-Swead, R., & Fostick, L. (2016). The Effect of Age and Type of Noise on Speech Perception under Conditions of Changing Context and Noise Levels. *Folia Phoniatr Logop*, 68(1), 16-21. <https://doi.org/10.1159/000444749>
- Tardif, C. L., Gauthier, C. J., Steele, C. J., Bazin, P. L., Schafer, A., Schaefer, A., Turner, R., & Villringer, A. (2016, May 1). Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. *Neuroimage*, 131, 55-72. <https://doi.org/10.1016/j.neuroimage.2015.08.047>
- Taylor, H. O., Taylor, R. J., Nguyen, A. W., & Chatters, L. (2018, Feb). Social Isolation, Depression, and Psychological Distress Among Older Adults. *J Aging Health*, 30(2), 229-246. <https://doi.org/10.1177/0898264316673511>
- Teixeira-Santos, A. C., Moreira, C. S., Magalhaes, R., Magalhaes, C., Pereira, D. R., Leite, J., Carvalho, S., & Sampaio, A. (2019, Aug). Reviewing working memory training gains in healthy older adults: A meta-analytic review of transfer for cognitive outcomes. *Neurosci Biobehav Rev*, 103, 163-177. <https://doi.org/10.1016/j.neubiorev.2019.05.009>
- Tremblay, P., Basirat, A., Pinto, S., & Sato, M. (2021, Aug 20). Visual prediction cues can facilitate behavioural and neural speech processing in young and older adults. *Neuropsychologia*, 159, 107949. <https://doi.org/10.1016/j.neuropsychologia.2021.107949>
- Tremblay, P., Brisson, V., & Deschamps, I. (2021, Feb 15). Brain aging and speech perception: Effects of background noise and talker variability. *Neuroimage*, 227, 117675. <https://doi.org/10.1016/j.neuroimage.2020.117675>
- Tun, P. A. (1998, Sep). Fast noisy speech: age differences in processing rapid speech with background noise. *Psychol Aging*, 13(3), 424-434. <https://doi.org/10.1037//0882-7974.13.3.424>
- Tun, P. A., & Wingfield, A. (1999, Sep). One voice too many: adult age differences in language processing with different types of distracting sounds. *J Gerontol B Psychol Sci Soc Sci*, 54(5), P317-327. <https://doi.org/10.1093/geronb/54b.5.p317>
- Varnet, L., Wang, T., Peter, C., Meunier, F., & Hoen, M. (2015, Sep 24). How musical expertise shapes speech perception: evidence from auditory classification images. *Sci Rep*, 5, 14489. <https://doi.org/10.1038/srep14489>

- Vermeire, K., Knoop, A., De Sloovere, M., Bosch, P., & van den Noort, M. (2019, Sep 20). Relationship Between Working Memory and Speech-in-Noise Recognition in Young and Older Adult Listeners With Age-Appropriate Hearing. *J Speech Lang Hear Res*, 62(9), 3545-3553. https://doi.org/10.1044/2019_jslhr-h-18-0307
- Wambach, D., Lamar, M., Swenson, R., Penney, D. L., Kaplan, E., & Libon, D. J. (2011). Digit Span. In J. S. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 844-849). Springer New York. https://doi.org/10.1007/978-0-387-79948-3_1288
- Wan, C. Y., & Schlaug, G. (2010). Music Making as a Tool for Promoting Brain Plasticity across the Life Span. *The Neuroscientist*, 16(5), 566-577. <https://doi.org/10.1177/1073858410377805>
- Watanabe, D., Savion-Lemieux, T., & Penhune, V. B. (2007, Jan). The effect of early musical training on adult motor performance: evidence for a sensitive period in motor learning. *Exp Brain Res*, 176(2), 332-340. <https://doi.org/10.1007/s00221-006-0619-z>
- Weijkamp, J., & Sadakata, M. (2017). Attention to affective audio-visual information: Comparison between musicians and non-musicians. *Psychology of Music*, 45(2), 204-215. <https://doi.org/10.1177/03057356166654216>
- Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., Grydeland, H., Tamnes, C. K., Ostby, Y., & Fjell, A. M. (2010, Sep). Life-span changes of the human brain white matter: diffusion tensor imaging (DTI) and volumetry. *Cereb Cortex*, 20(9), 2055-2068. <https://doi.org/10.1093/cercor/bhp280>
- White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013, Nov 20). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Front Syst Neurosci*, 7, 90. <https://doi.org/10.3389/fnsys.2013.00090>
- Woodworth, R. S., & Thorndike, E. L. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, 8(3), 247-261.
- Working Group on Speech, U., & Aging. (1988, 1988/03/01). Speech understanding and aging. *The Journal of the Acoustical Society of America*, 83(3), 859-895. <https://doi.org/10.1121/1.395965>
- Worschech, F., Marie, D., Junemann, K., Sinke, C., Kruger, T. H. C., Grossbach, M., Scholz, D. S., Abdili, L., Kliegel, M., James, C. E., & Altenmuller, E. (2021). Improved Speech in Noise Perception in the Elderly After 6 Months of Musical Instruction. *Front Neurosci*, 15, 696240. <https://doi.org/10.3389/fnins.2021.696240>
- Yoo, J., & Bidelman, G. M. (2019, Jun). Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hear Res*, 377, 189-195. <https://doi.org/10.1016/j.heares.2019.03.021>
- Yost, W. A. (2017, Mar). Spatial release from masking based on binaural processing for up to six maskers. *J Acoust Soc Am*, 141(3), 2093. <https://doi.org/10.1121/1.4978614>

- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007, 2007/07/01). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547-558. <https://doi.org/10.1038/nrn2152>
- Zatorre, R. J., & Salimpoor, V. N. (2013, Jun 18). From perception to pleasure: music and its neural substrates. *Proc Natl Acad Sci U S A*, 110 Suppl 2, 10430-10437. <https://doi.org/10.1073/pnas.1301228110>
- Zelinski, E. M. (2009). Far transfer in cognitive training of older adults. *Restor Neurol Neurosci*, 27(5), 455-471. <https://doi.org/10.3233/RNN-2009-0495>
- Zendel, B. R., & Alain, C. (2009). Concurrent Sound Segregation Is Enhanced in Musicians. *Journal of Cognitive Neuroscience*, 21(8), 1488-1498. <https://doi.org/10.1162/jocn.2009.21140>
- Zendel, B. R., & Alain, C. (2013). The Influence of Lifelong Musicianship on Neurophysiological Measures of Concurrent Sound Segregation. *Journal of Cognitive Neuroscience*, 25(4), 503-516. https://doi.org/10.1162/jocn_a_00329
- Zendel, B. R., & Alain, C. (2013, Apr). The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. *J Cogn Neurosci*, 25(4), 503-516. https://doi.org/10.1162/jocn_a_00329
- Zhang, L., Fu, X., Luo, D., Xing, L., & Du, Y. (2020, Aug 11). Musical Experience Offsets Age-Related Decline in Understanding Speech-in-Noise: Type of Training Does Not Matter, Working Memory Is the Key. *Ear Hear*, 42(2), 258-270. <https://doi.org/10.1097/AUD.0000000000000921>
- Zuk, J., Benjamin, C., Kenyon, A., & Gaab, N. (2014). Behavioral and neural correlates of executive functioning in musicians and non-musicians. *PLoS One*, 9(6), e99868. <https://doi.org/10.1371/journal.pone.0099868>