Neurobiology of Aging 81 (2019) 102–115

Contents lists available at ScienceDirect

Neurobiology of Aging

journal homepage: www.elsevier.com/locate/neuaging

Musical training improves the ability to understand speech-in-noise in older adults

Benjamin Rich Zendel [a,b,c,d,*], Greg L. West [e], Sylvie Belleville [c,e], Isabelle Peretz [b,e]

[a] Memorial University of Newfoundland, St. John's, Newfoundland and Labrador, Canada
[b] International Laboratory for Brain, Music, and Sound Research, Montréal, Québec, Canada
[c] Centre de Recherche, Institut Universitaire de Gériatrie de Montréal (CRIUGM), Montréal, Québec, Canada
[d] Aging Research Centre-Newfoundland and Labrador, Corner Brook, Newfoundland and Labrador, Canada Aging Research Centre-Newfoundland and Labrador, Corner Brook, Newfoundland and Labrador, Canada
[e] Département de Psychologie, Université de Montréal, Québec, Canada

ARTICLE INFO

Article history:
Received 19 December 2017
Received in revised form 1 May 2019
Accepted 22 May 2019
Available online 29 May 2019

Keywords:
Aging
Musical training
ERPs
Speech perception
Speech-in-noise
Hearing

ABSTRACT

It is well known that hearing abilities decline with age, and one of the most commonly reported hearing difficulties reported in older adults is a reduced ability to understand speech in noisy environments. Older adult musicians have an enhanced ability to understand speech in noise, and this has been associated with enhanced brain responses related to both speech processing and the deployment of attention; however, the causal impact of music lessons in older adults has not yet been demonstrated. To investigate whether a causal relationship exists between short-term musical training and performance on auditory tests in older adults and to determine if musical training can be used to improve hearing in older adult nonmusicians, we conducted a longitudinal training study with random assignment. A sample of older adults was randomly assigned to learn to play piano (Music), to learn to play a visuo-spatially demanding video game (Video), or to serve as a no-contact control (No-contact). After 6 months, the Music group improved their ability to understand a word presented in loud background noise, whereas the other 2 groups did not. This improvement was related to an increase in positive-going electrical brain activity at fronto-left electrodes 200–1000 ms after the presentation of a word in noise. Source analyses suggest that this activity was due to sources located in the left inferior frontal gyrus and other regions involved in the speech-motor system. These findings support the idea that musical training provides a causal benefit to hearing abilities. Importantly, these findings suggest that musical training could be used as a foundation to develop auditory rehabilitation programs for older adults.

1. Introduction

Difficulties with hearing are one of the most commonly reported health issues in older adults. Rates of hearing loss approximately double from the second through seventh decade of life, with the degree of hearing loss defined by a pure-tone average (PTA; 0.5, 1, 2, and 4 kHz) of greater than 25 dB in the better ear (Yamasoba et al., 2013). Nearly 80% of adults over age 80 meet the criteria for hearing loss (Yamasoba et al., 2013). Age-related decline in auditory perception can vary substantially between individuals and often includes difficulties understanding speech in adverse listening situations, such as when there is significant background noise (Pichora-Fuller et al., 1995; Robert Frisina and Frisina, 1997; Schneider et al., 2010). These age-related changes in auditory perception are thought to reflect both bilateral sensorineural hearing loss, due to physical changes in the inner ear (Gates and Mills, 2005; Stenklev and Laukli, 2004), and changes in the central auditory system (Alain et al., 2006; Schneider et al., 2010). In addition to the social isolation that can arise from hearing-related communication difficulties, age-related decline in hearing has also been associated with cognitive decline (Lin et al., 2013; Mick et al., 2014). Given the prevalence and negative outcomes of age-related decline in hearing abilities, finding ways to prevent, mitigate, or delay these changes is of utmost importance, and evidence suggests that musical training may be a useful intervention for preserving or enhancing auditory abilities in older adults.

It is well known that musicians have enhanced auditory processing abilities (Kraus and Chandrasekaran, 2010), and these benefits are paralleled by an enhanced ability to understand speech in noisy environments (Parbery-Clark et al., 2009; Zendel et al.,

* Corresponding author at: Memorial University of Newfoundland, 230 Elizabeth Ave, St. John's, Newfoundland and Labrador A1B 3X9, Canada. Tel.: +1 7098646080; fax: +1 7098644991.
E-mail address: bzendel@mun.ca (B.R. Zendel).

0197-4580/$ – see front matter © 2019 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.neurobiolaging.2019.05.015

2015). Findings from cross-sectional studies suggest that musical training causes neuroplasticity along the auditory pathway from the brainstem to the cortex, and this neuroplasticity leads to enhanced auditory abilities. Recent work suggests that the benefit of musical training may be related to enhanced connections between the auditory and motor systems (Du and Zatorre, 2017). Musical training could therefore facilitate speech processing via the speech-motor system, which has been shown to be an important component of speech perception, particularly when speech is presented in background noise (Du et al., 2014). Electrophysiological studies have shown that musicians exhibit enhanced frequency-following responses and auditory event-related potentials (ERPs; i.e., N1-P2) compared with nonmusicians (e.g., Koelsch et al., 1999; Pantev et al., 1998, 2001; Shahin et al., 2003, 2005; Wong et al., 2007). Longitudinal research in younger adults has provided support for the idea that improved auditory processing observed in musicians is due to neuroplasticity. For instance, several studies with random assignment and control groups have demonstrated that after musical training, participants have enhanced auditory abilities that are related to enhanced neurophysiological measurements (Fujioka et al., 2006; Kraus and White-Schwoch, 2015; Lappe et al., 2008; Tierney et al., 2015). Emerging evidence suggests that these enhanced auditory abilities can persist into old age, with older musicians being able to understand speech in noisy environments better than older nonmusicians (Parbery-Clark et al., 2011; Zendel and Alain, 2012). Further support for this idea comes from longitudinal research that used a nonmusic-based auditory training intervention and found that the ability to understand speech in noise can be improved in older adults with training (Anderson et al., 2013). However, no study has yet examined whether musical training in older adults can improve the ability to understand speech in noise.

Understanding speech in noise is a hierarchical process that occurs in multiple subcortical and cortical structures, and evidence suggests that musicianship and musical training can alter neural functions related to processing speech (Coffey et al., 2017). In older musicians, there is evidence that enhanced endogenous or attention-dependent processing contributes to the auditory benefit (Zendel and Alain, 2013, 2014). Using ERPs recorded during a task that required perceptual segregation of concurrently occurring sounds, older musicians had enhanced late positivities that were dependent on attention being directed toward the task, compared with older nonmusicians and younger nonmusicians (Zendel and Alain, 2013, 2014). Enhancements to subcortical processing of speech-in-noise in older musicians have also been observed (Parbery-Clark et al., 2012); however, these benefits are smaller than the subcortical enhancements in speech-in-noise processing observed in younger musicians, compared with nonmusicians (Parbery-clark, Skoe and Kraus, 2009). This pattern of results suggests that the benefit of musical training shifts from an exogenous processing benefit to an endogenous processing benefit as musicians age (Alain et al., 2014), which could be related to preserved cognitive abilities in older musicians including nonverbal memory and executive processes (Hanna-Pladdy and MacKay, 2011).

Overall, these findings suggest that musical training could be used in older adults as an engaging form of auditory training with the potential to improve the ability to understand speech in noisy environments. To determine if this is a possibility, we conducted a three-arm, single-blind, longitudinal training study with random assignment, where one group of older participants received musical training, one group served as an active control group by playing a 3D video-game, and one served as a no-contact control group. Previous studies where older adults were randomized into a musical training intervention and then compared to control groups have demonstrated that musical training can be used to improve cognitive abilities, mood, and quality of life (Bugos et al., 2007; Seinfeld et al., 2013). Thus, the goal of the present study was to examine if musical training could improve the ability to understand speech in noise by modifying the underlying functional neurophysiology. The ability to understand speech in noise and the associated event-related brain responses were assessed at 3 different timepoints: before training, midway through training, and after training.

There are a number of ways to assess the ability to understand speech in noise by varying both the target and the background noise. For example, different studies have used tones, phonemes, words, or full sentences as the target stimuli, and white noise, filtered white noise, or various forms of single- or multi-talker babble noise as the background noise (e.g., Billings et al., 2009; Kaplan-Neeman et al., 2006; Martin et al., 1997; Parbery-Clark et al., 2012; Pichora-Fuller et al., 1995). To best understand if musical training can improve the ability to understand speech-in-noise in the real world, we aimed to use ecological stimuli when possible. Accordingly, this paradigm included real words as the target stimuli and multi-talker babble as the background noise. This paradigm was similar to one used previously, where younger musicians exhibited an enhanced ability to understand words in background noise compared with younger nonmusicians (Zendel et al., 2015). In the study by Zendel et al. (2015), early event-related brain responses related to stimulus encoding were enhanced, whereas later responses related to semantic processing were reduced, in younger musicians compared with younger nonmusicians. This suggests that musical training enhanced the representation of an incoming speech stimulus, which facilitated later semantic access. This paradigm is therefore well suited to examine the impact of music lessons on older adults, as it is sensitive to the impact of musical training for both behavioral and neurophysiological measures and it is based on natural speech sounds.

2. Methods

2.1. Design

The study was designed as a three-arm single-blind longitudinal training study with random assignment. Participants were randomized into 3 groups using a stratified covariate-adaptive procedure (see below). Participants took part in 3 testing sessions. The pre-training session (Pre) took place before the intervention, the mid-training session (Mid) took place 3 months after the start of the intervention, and the post-training session (Post) took place 6 months after the onset of the intervention. During the Pre, Mid, and Post sessions, participants completed a series of auditory and cognitive assessments. Group differences on the assessments that are not reported here have been published, are under review, or are in preparation for submission elsewhere. These assessments include structural magnetic resonance imaging (MRI), eye-tracking, and short-term memory (Diarra et al., 2019; West et al., 2017); functional MRI (Fleming et al., 2019, unpublished data); IQ and working memory (Zendel et al., unpublished data); and ERP and visual inhibition (West et al., unpublished data). Here, we report the results of a speech-in-noise task that was done while monitoring electrical brain activity (electroencephalography; EEG).

2.2. Participants and randomization procedure

Participants were recruited into the study from the Centre de Recherche, Institut Universitaire de Gériatrie de Montréal participant pool. The study received ethical approval from the Comite conjoint d'evaluation scientifique—Regroupment Neuroimagerie/Quebec (CES-RNQ). All participants were native French speakers, and all testing was conducted in French. Participants were prescreened to ensure that they did not have any present or past major illness that could impact their ability to complete the present study, were not taking any psychiatric medications or medication known to have an impact on cognition, were MRI compatible, were non–video game players, and were nonmusicians. All participants were screened for mild cognitive impairment using the Montreal Cognitive Assessment (MOCA; Nasreddine et al., 2005). All participants scored 23 or higher on the MOCA. This revised cutoff was based on a recent meta-analysis of validation studies on the MOCA, which suggested that the original cutoff of 26 yielded a high number of false-positive diagnoses of mild cognitive impairment (Carson et al., 2018). Average scores for the MOCA are presented in Table 1, and the distribution of scores in the present study (M = 27.4, range 23–30, SD = 2.03 [all participants]) is similar to normal controls (M = 27.4; SD = 2.2; Nasreddine et al., 2005).

Pure-tone thresholds (PTTs) were measured by air conduction using ER-2 insert earphones before training to evaluate hearing status. All participants had a PTA for speech frequencies (500, 1000, 2000, and 4000 Hz) below 35 dB HL (i.e., normal hearing or mild hearing loss) in their better ear (Humes, 2019). Although this pattern of PTTs is suggestive of sensorineural hearing loss, without bone-conduction thresholds and an otoscopic examination, we cannot rule out the possibility of mild conductive hearing loss in some participants. To ensure participants had similar thresholds in both ears, PTA from the left ear was compared with the PTA from the right ear in all participants. There was no significant difference between the PTA for the right and left ears [$t(32) = 0.84$, $p = 0.41$]. PTA asymmetries (i.e., the difference between PTA in the left and right ears) ranged from 0 to 15 dB (M = 3.9 dB, SD = 3.8). To ensure the groups were matched in terms of PTTs, a mixed design analysis of variance (ANOVA) was calculated that included Group (Music, Video, No-contact [see below for details]), Frequency (250, 500, 1000, 2000, 4000, and 8000 Hz), and Ear (Left, Right) as factors. As expected, the effect of Frequency was significant, $F(5,150) = 69.6$, $p < 0.001$, $\eta^2 = 0.70$, with PTTs being larger as Frequency increased. Most critical, the effect of Group and its interactions with Frequency and Ear were not significant (p's > 0.33), indicating that hearing thresholds were similar between the groups. No participants reported any difficulties meeting the motor demands of the laboratory tasks or the training protocols. To be considered a nonmusician, participants must not currently play a musical instrument and not have had more than 3 years of formal musical training in their life. Music lessons that were part of the normal education curriculum were not included in the exclusion criteria. To be considered a non–video game player, participants must have had little to no experience with commercial video games (e.g., games played on a computer or game console) during their lifetime. Casual games such as computerized card or puzzle games were not considered to be video games.

All participants were randomized into one of 3 groups. Randomization was done by an independent research assistant, using a predefined randomization table before contacting participants to ensure that participants were blind to the existence of the other 2 groups. Randomization was stratified using a covariate-adaptive randomization procedure. Each factor was stratified into 2 categories. For the factor of age, there was "younger" (55–64 years) and "older" (65–75 years) groups, for the factor of education, there was low (<16 years) and high (>16 years) groups, and for the factor of gender, there was female and male groups. Because participants were recruited from a database, age, education level, and gender of each participant were known before they were contacted and it was thus possible to stratify randomization on the basis of these 3 factors. This stratification led to 6 possible stratification groups. Based on the stratification variables, each participant was assigned to 1 of the 8 stratification groups based on the demographics available from the participant database (e.g., female, younger, high education; male, older, low education; etc…). Each of these 6 groups was assigned a random but balanced order to determine experimental group assignment. For example, the first person contacted from the participant database who was in the "female, younger, higher education" stratification group was invited to participate in the experimental music training group (Music; see below for details). If this participant accepted, she would become the first participant in the Music group. If she refused, she would be excluded from further participation in the study, and the next person contacted in the "female, younger, higher education" stratification group was invited to participate in the Music group. This was repeated until a person in this stratification group volunteered to participate in the Music group. Next, people in the same stratification group were invited to participate in the no-contact control group (No-contact; see below for details) until one person volunteered to participate. Finally, a person in this stratification group was invited to participate in the video game training group (Video; see below for details) until one person volunteered to participate. Once one participant was recruited for each of the 3 groups (i.e., Music, Video, and No-contact), the procedure was then repeated, except the order of recruitment for each experimental group was randomized for each cycle. That is, each participant was recruited into 1 of the 3 groups (Music, Video, and Control), but the order in which each participant was recruited was random. The orders were also randomized across all the stratification groups. Participants who chose not to participate were not included in the randomization matrix.

Forty-five participants in total were recruited into the study. Using the stratified randomization procedure, 15 participants were assigned to the Music group, 15 participants were assigned to the Video group, and 15 participants were assigned to the No-contact group. During the study, 2 participants withdrew from the Music group, 11 withdrew from the Video group, and 2 withdrew from the No-contact group. To compensate for the higher attrition rate within the Video group, an additional 4 participants were assigned to the Video group who were matched for the age, gender, and education level of the other 2 groups; however, the stratified randomization procedure was not used. This resulted in a total of 8 participants completing the training within the Video group. The demographics of the participants within each group are presented in Table 1. EEG data were unusable from 1 participant in the Music group because of the hairstyle; however, this participant still completed the behavioral tasks. For more details about the withdrawal rate from the Video group, see the study by West et al. (2017).

Table 1
Participant demographics

Group	Age (SD)	Gender	Years of education (SD)	MOCA (SD)	PTA[a] (SD)
Music (n = 13)	67.5 (4.2)	10 females; 3 males	14.5 (2.2)	28.3 (1.9)	16.5 (6.8)
Video (n = 8)	66.9 (3.9)	4 females; 4 males	17.5 (2.3)	26.9 (1.3)	16.8 (7.1)
No-contact (n = 13)	69.3 (5.7)	10 females; 3 males	15.2 (3.1)	26.6 (2.1)	14.4 (7.7)

Key: MOCA, Montreal Cognitive Assessment; PTA, pure-tone average; SD, standard deviation.
[a] PTA (500, 1000, 2000, 4000 Hz; left and right ear).

2.3. Training procedure

2.3.1. Piano training group (Music)

Piano training was carried out at each participant's home using the Synthesia software and an 88-key M-Audio MIDI piano that was brought to the participant's home. Synthesia is a piece of software that uses a nonstandard form of musical notation that can be understood within a few minutes. This was critical as learning to read traditional music notation can take much longer. Notes in Synthesia are presented as colored bars that fall from the top of the computer screen. At the bottom of the screen, there was an image of a piano keyboard, and when a colored bar hit a certain note, the participant played that note. The length of the bar indicated how long to hold the note for, and a metronome click helped the participant follow the rhythm of the task. Accordingly, the musical training involved auditory, visual, and motor systems.

First, a research assistant (RA) installed and calibrated the piano to work on each participant's home computer. Next, each participant completed an introductory lesson that included introductory information about music, detailed instructions on how to use Synthesia, and directions on how to record their progress. Introductory music information included lessons about note names, how to position their hands on the piano, and how to synchronize performance with the information on the screen and the metronome. The RA worked with each participant until they were comfortable navigating the software interface, were able to complete the first lesson, and were able to load new songs or lessons. This typically took around 30–60 minutes. A set of introductory lessons and beginner piano music was installed on each participant's computer. The sixty-five introductory lessons were created by Artur Gajewski for use with Synthesia (2011). These lessons begin by playing a single note in time with a metronome, on each hand. They progress through alternating notes, playing notes simultaneously with both hands, scales (e.g., major and minor in different keys), chords, and the lessons end with a few etudes. Participants were told to start with the lessons, and once they were comfortable with the lessons, to try out some of the introductory songs. Participants were encouraged to move at their own pace, but to try to master a given lesson or song before moving on. Sometimes participants would work on a lesson and song simultaneously. The goal was to keep participants as engaged as possible in the piano lessons. Although the selection of songs was at the discretion of the participant, all participants completed the introductory lessons. At the end of the training period, success was assessed by the highest lesson number each participant could complete with an accuracy score above 80%. Accuracy was defined as the number of notes the participant hit correctly during the lesson. For the participants in the Music group, the highest lesson with 80% accuracy ranged from lesson 10 to 60 (M = 30.01, SD = 16.21).

2.3.2. Video game training group (Video)

Video game training was done at each participant's home using a Nintendo Wii console system equipped with a Wii Classic Controller. All participants in this group trained on the Nintendo game "Super Mario 64." Two participants completed all tasks within Super Mario 64 before the completion of the 6-month training period. In these cases, they continued to train on a very similar game called Super Mario Galaxy, until the end of the training period. Super Mario 64 and Super Mario Galaxy are three-dimensional platform games where the player is tasked with exploring a virtual environment to search for stars (i.e., tokens). When enough stars are collected through completing in-game goals, the player can then progress further into the game and will encounter new environments to explore.

After each participant completed the pre-training assessments, an RA installed the Nintendo Wii on the participant's home television. The RA then gave an initial orientation to teach each participant how to turn on the Nintendo Wii and access the Super Mario 64 game. This was followed by a custom in-game orientation, which taught the participant to move the character around the virtual environment. At this point, some participants encountered certain challenges associated with maneuvering the character. Some had issues with understanding the game's mechanics. Furthermore, Super Mario 64 has a very long learning curve, as it was not originally designed to be played by someone with little to no video game experience. For this reason, the RA returned to each participant's home for up to 3 additional supervised two-hour training sessions to help the participant learn how to properly maneuver the character and progress through the game. All participants were able to properly maneuver the character and progress through the game at the conclusion of these training sessions. After this, participants were given a custom-made instruction booklet that outlined how and where to collect all the stars for the first 4 levels. This allowed participants to learn the game's mechanics in further detail and practice the basic motor coordination that was required. After this point, participants had to search for and obtain the stars within each remaining level without any assistance from the research team. All participants were able to find stars on their own and progress through the game. Participants collected between 57 and 90 stars (M = 69.1, SD = 13.5) during the six-month training period.

2.3.3. No-contact control group (No-contact)

The No-contact group had no contact with the research team during the 6-month period other than to complete the pre-training, mid-term, and post-training sessions.

2.3.4. Training procedure summary

Music and video game training lasted 6 months. In all cases, participants kept a record of their daily training progress and were asked to complete a minimum of 30 minutes of training at least 5 days a week, although some completed more than this amount. To equalize potential placebo effects across groups, all participants were told that they were expected to improve in performance. Participants in the Video group were told that there was evidence that video game training enhances cognitive abilities and that video game training in older adults was expected to improve those abilities. Participants in the Music group were told that there was evidence that musicians have enhanced cognitive abilities and that we expected musical training to improve those abilities. Finally, the No-contact group was told that we were investigating test-retest effects and that they were expected to improve on the experimental tasks. All participants were debriefed about the other groups at the end of the final testing session.

2.4. Stimuli

Stimuli were 150 French words spoken by a male, from a list of phonetically balanced, equally understandable monosyllabic words (see Picard, 1984). Words were presented binaurally at about 75 decibel sound pressure level (dB SPL), through insert earphones (Etymotic ER-2), as measured by a sound level meter (Quest Technologies) that measured the amplitude of the stimuli presented from one insert earphone. Words were presented in 3 different noise conditions. In the first condition, words were presented in isolation with no background noise ("No-Noise"). In the other 2 conditions, multi-talker babble noise was presented with the words at about 60 dB SPL for the quiet noise condition ("Quiet"; 15 dB signal-to-noise ratio [SNR]) and about 75 dB SPL for the loud noise condition ("Loud"; 0 dB SNR). The same words were also presented in each listening condition (active, passive; see below), and in each session (Pre, Mid, and Post). To minimize learning effects within a session, the Loud, active listening condition (i.e., the most difficult) was always presented first. Although there may have been learning effects because of the repetition of words across the sessions, this effect would be equal across the groups. The multi-talker babble was created by individually recording 4 native speakers of French (2 females and 2 males) reading a rehearsed monologue in a sound-attenuated room for 10 minutes. The recordings were made at a sampling rate of 44.1 KHz at 16 bits, using an Audio-Technica 4040 condenser microphone. The individual recordings of each monologue were normalized to −8 dBFS and then combined into a single monaural sound, which was then normalized to −2 dBFS using Adobe Audition (version 10). The 10-minute multi-talker babble noise was looped repeatedly during listening conditions where the multi-talker babble was present.

2.5. Procedure

After being fitted with an EEG cap, participants were seated in a double-walled, electrically shielded sound-attenuating booth where all testing took place. The words were presented in a random order, in each of the 3 levels of multi-talker babble noise. In the "No-Noise" condition, words were presented without multi-talker babble noise. In the "Quiet" condition, words were presented with multi-talker babble noise that was 15 dB below the level of word (i.e., 15 dB SNR), whereas in the "Loud" condition, words were presented with multi-talker babble noise that was at the same level as the word (i.e., 0 dB SNR). In addition, all 3 noise level conditions were presented in 2 listening conditions: active and passive. In the passive condition, participants were told to ignore the words and watched a self-selected silent subtitled movie. Words were presented with a stimulus onset asynchrony that was randomized between 2500 and 3500 ms. The use of muted subtitled movies has been shown to effectively capture attention without interfering with auditory processing of speech stimuli (Pettigrew et al., 2004). In the active condition, participants were told to repeat the word aloud and to look at a blank computer screen. To avoid muscle artifacts in the ERPs, participants were told to delay their response until they saw a small LED light flash 2000 ms after the presentation of the word. Word correctness was judged immediately by a native French speaker. The RA performing the judgment had the text of the word presented in front of them on a screen and did not hear background noise during the word repetition. The RA was told to score the word as correct if it was understandable as the word written on the screen. If there was any confusion, the RA asked the participant to repeat the word until the RA was sure the repetition was a match or not. We chose to use word repetitions because it required an accurate lexical match of the incoming word to correctly repeat it back, and pilot testing confirmed that the delayed oral response did not contaminate the ERPs with muscle artifacts. An alternative would have been to use a forced choice procedure; however, this would likely create a biased estimate of word understanding because the presentation of choices limits what a participant can report and may bias their performance if they were able to hear part of the word. Overall, the experiment took about 70 minutes to complete during each session, with each passive block taking about 8 minutes and each active block taking about 15 minutes.

2.6. Recording and averaging of electrical brain activity

Neuroelectric brain activity was digitized continuously from 70 active electrodes at a sampling rate of 1024 Hz using a Biosemi ActiveTwo system (Biosemi, Inc., Netherlands). Sixty-four electrodes were placed on the scalp according to the standard 10–20 system. In addition, 6 electrodes were placed bilaterally at mastoid, inferior ocular, and lateral ocular sites (i.e., M1, M2, IO1, IO2, LO1, LO2). All averages were computed using Brain Electrical Source Analysis (BESA) software, version 6.1. The analysis epoch included 200 milliseconds of prestimulus activity and 1000 milliseconds of poststimulus activity. Trials containing excessive noise (>120 μV) at electrodes not adjacent to the eyes (i.e., IO1, IO2, LO1, LO2, FP1, FP2, FPz, FP9, and FP10) were rejected before averaging. Continuous EEG was then averaged separately for each condition, into 18 different ERPs. That is, the active and passive conditions included 3 ERPs each for the No-Noise, Quiet, and Loud conditions, and these ERPs were collected for the pre-, mid-, and post-training sessions. Prototypical eye blinks and eye movements were recorded before the start of the experiment, in each of the 3 sessions. A principal component analysis (PCA) of these averaged recordings provided a set of components that best explained the eye movements. These components were then decomposed into a linear combination along with topographical components that reflect brain activity. This linear combination allowed the scalp projections of the artifact components to be subtracted from the experimental ERPs to minimize ocular contamination such as blinks, as well as vertical and lateral eye movements for each individual average with minimal effects on brain activity (Berg and Scherg, 1994). After this correction, trials with greater than 120 μV of activity were considered artifacts and excluded from further analysis. In addition, during the active listening condition, trials where the participant did not correctly repeat the word were excluded from the analysis.

To determine if any of the experimental manipulations had an impact on the number of trials accepted after artifacts were corrected, a mixed design ANOVA was calculated that included session, noise level, and listening condition as within-subject factors, and group as a between-subject factor. As expected, there were significant main effects of both noise level, $F(2, 60) = 55.79$, $p < 0.001$, and listening condition, $F(1, 30) = 85.0$, $p < 0.001$. In addition, there was a significant interaction between noise level and listening condition $F(2, 60) = 36.2$, $p < 0.001$. No other effects or interactions were significant. The ERP analysis included an average of 121, 118, and 90 trials during the No-Noise, Quiet, and Loud noise levels, respectively, during active listening, and 135, 134, and 131 trials during the No-Noise, Quiet, and Loud noise levels during passive listening. This means that there were fewer observations during the active listening condition, and fewer still during the loud condition during active listening. The advantage to this approach was that the brain activity during active listening trials only included successful understanding of the speech. This permits us to connect brain activity to correct understanding of speech during active listening. The disadvantage of this approach is that the individual averaged ERPs from the Loud active listening trials are likely more variable than trials in the other conditions. By using a pre-post design, and comparing differences within the same condition, however, the effect of this variability is minimized. Finally, ERPs were band-pass filtered to attenuate frequencies below 0.1 Hz, and above 30 Hz, and referenced to the linked mastoid.

3. Results

The critical effects in the data analyses were based on Group by Session interactions, followed-up by a significant main effect of Session in the Music group, and nonsignificant effects of Session in the other 2 groups. To fully explore the data, other effects are reported as well. The 2 training groups (Music & Video) trained for a similar amount of time during the six-month training period [Music: 86.4 hours, SD = 34.4; Video: 72.3 hours, SD = 11.3; $t(18) = 1.12$, $p = 0.28$]. Alpha for statistical tests was set at 0.05; however, when p was between 0.05 and 0.1 and the effect size (partial eta-squared; η^2) was above 0.06 (medium effect size), the effect was considered to be marginally significant. In general, data were normally distributed. In situations where normality was violated, p-values were adjusted using the Greenhouse-Geisser correction, and the adjusted p-values are reported, whereas the unadjusted, original degrees of freedom are reported.

3.1. Behavioral data

Data were analyzed using a mixed design ANOVA that included Session (Pre, Mid, Post) and Noise Level (No-Noise, Quiet, Loud) as within-subject factors and Group (Music, Video, No-contact) as a between-subject factor. There was no behavioral data collected during passive trials, so listening condition was not included as a factor here. Accuracy, defined as the percentage of words repeated correctly, was affected by noise level, $F (2, 62) = 263.63$, $p < 0.001$, $\eta^2 = 0.9$, with accuracy being higher in the No-Noise condition than Quiet ($p < 0.001$) and accuracy being higher in Quiet than Loud ($p < 0.001$; see Fig. 1). The three-way Noise level by Session by Group interaction was marginally significant, $F (8, 124) = 1.91$, $p = 0.064$, $\eta^2 = 0.11$. Follow-up simple two-way ANOVAs revealed a significant Noise level by Session interaction for the Music group, $F (4, 48) = 3.25$, $p = 0.019$, $\eta^2 = 0.21$, but not for the Video or No-contact groups ($p = 0.90$ and 0.16, respectively), demonstrating the main effect of Session was only significant for the Music group. Further follow-up tests in the Music group revealed that accuracy improved from Pre to Post Sessions in the Loud condition, $F (2, 24) = 3.99$, $p = 0.032$, $\eta^2 = 0.25$, but not in the Quiet or No-Noise conditions ($p = 0.45$ & 0.88, respectively). To ensure groups were balanced at baseline (i.e., during the Pre-Training Session), a series of one-way between-subjects' ANOVAs were carried out for each noise level and none were significant ($p = 0.45$, 0.84, and 0.56 for Loud, Quiet, and No-Noise, respectively).

3.2. ERP results

To determine an appropriate electrode montage, a spatial independent component analysis (ICA) was calculated on the Pre-training ERPs (6 ERPs: No-Noise, Quiet, Loud during both active and passive listening) averaged across groups to identify the spatial distribution of the independent components that make up the averaged data. The first independent component (IC1) accounted for 30.7% of the spatial variance in the data. This component had a spatial distribution that was maximal over fronto-left electrodes and inverted over occipital electrodes with a peak at electrode F5/F7 (Fig. 2A). An examination of the averaged ERPs at these electrodes revealed one major negative peak during the pre-training ERPs, which was not part of the N1-P2 response. The negative-going activity started around 200 ms, when the direction of the ERP started to become increasingly negative, peaked around 600 ms, stayed negative until the end of the analysis epoch, and was only evident during active listening. The second independent component (IC2) accounted for 12.8% of the variance in the data. This component had a spatial distribution that was maximal over the frontocentral electrode (FCz) and was inverted over mastoid electrodes. This spatial distribution is likely related to the N1-P2 auditory evoked response (Fig. 2B). Other independent components accounted for less than 10% of the spatial variability. Accordingly, the ERP analysis focused on 2 electrode montages, one based on IC1 (i.e., FP1, AF7, AF3, F7, F5, F3, FT7, FC5, FC3) and one based on IC2 (i.e., F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2). These 2 electrode montages are highlighted in yellow in Fig. 2.

One important question is what cortical sources contributed to each of these 2 spatial ICs. IC2 was likely related to the N1-P2 response and thus likely has sources along the superior temporal plane. IC1 was not as obvious, so the topography was submitted to a local autoregressive average (LAURA) analysis (de Peralta Menendez et al., 2001). The LAURA analysis of IC1 revealed 4 local maxima. These maxima were located in the (1) left inferior frontal gyrus (IFG; Brodmann area [BA] 45); (2) right subcallosal/para-hippocampal gyrus (PhG; BA 34); (3) right cerebellum (CB); and (4) posterior portion of the left middle temporal gyrus (MTG; BA 22). These sources, negative-going polarity, and scalp topography are consistent with either an N400 response related to lexical access (Silva-Pereyra et al., 2003) or a processing negativity related to attention (Näätänen, 1982; Näätänen and Picton, 1987). Although the N400 tends to peak around 400 ms (Kutas and Federmeier, 2011), the processing negativity can start as early as 50 ms after stimulus onset (Näätänen, 1982).

Finally, given that the Video group was smaller than the Music and No-contact groups, a comparison between the Video and No-contact groups was done to determine if there were any training-related differences between these 2 control groups. No significant Group by Session interactions were observed for these 2 groups (analyses were the same as the ones reported below, except the Group factor included only Video and No-contact). Accordingly, to

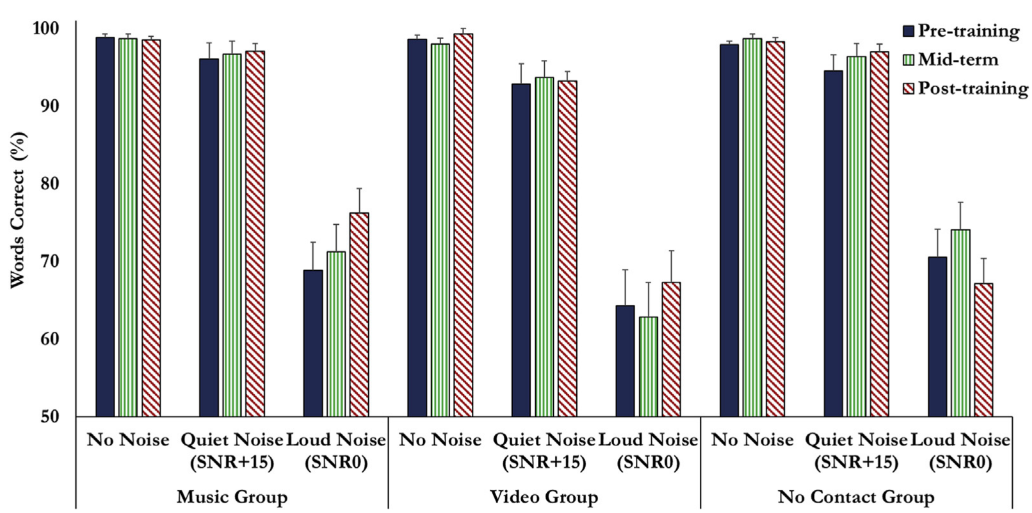

Fig. 1. Accuracy during the active listening task. Participants in all 3 groups were able to repeat nearly all the words accurately during the No-Noise and Quiet conditions. Accuracy improved during the Loud condition in the Music group from Pre- to Post-training ($p = 0.032$, $\eta^2 = 0.25$). No improvement was observed in either of the other groups. Error bars represent one standard error.

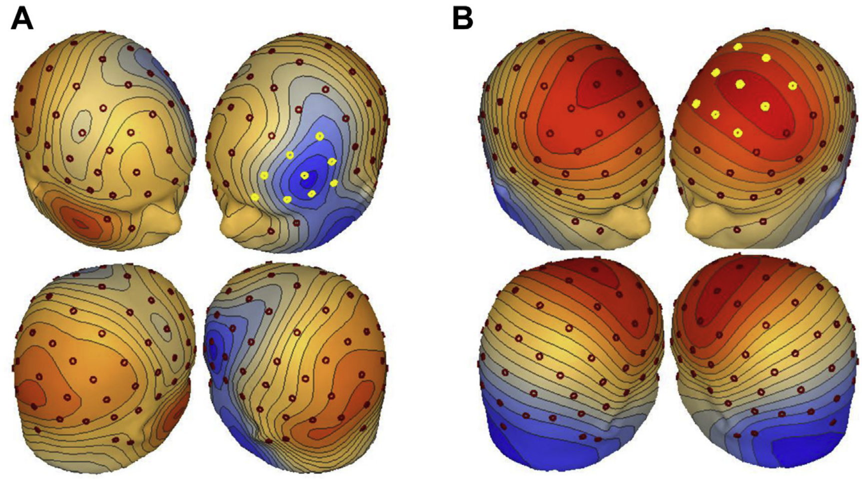

Fig. 2. (A) Scalp topography for independent component 1 (IC1). This component accounted for 30.7% of the variance in the ERP data. The electrodes chosen for the ERP analysis are highlighted in yellow on the top right plot. (B) Scalp topography for independent component 2 (IC2). This component accounted for 12.8% of the variance in the ERP data. The electrodes chosen for the ERP analysis are highlighted in yellow on the top right plot. Abbreviation: ERP, event-related potential. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

improve statistical power, we combined the Video and No-contact groups into a single control group (Control) for all subsequent analysis of ERP data. Thus, all EEG data were analyzed using Session (Pre, Mid, Post), Listening Condition (Active, Passive), Noise Level (No-Noise, Quiet, Loud), and Electrode (different montages were used for different components) as within-subject factors and Group (Music, Control) as a between-subject factor. Interactions and main effects of each electrode are not reported as multiple electrodes were used to gain a stable and reliable estimate of the relevant component, as calculated by the ICA.

3.2.1. ERP: N1

The first analysis focused on the N1 from the electrode montage based on IC2. Peak N1 amplitude and latency were extracted from a montage of frontocentral electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2). For this analysis, we chose to examine the peak amplitude because the N1 was clear across all participants in all listening conditions. A long extraction window (90–275 ms) was used because increasing the level of background noise has been shown to delay ERP latencies (e.g., Billings et al., 2009; Kaplan-Neeman et al., 2006; Martin et al., 1997; Zendel et al., 2015). Peaks were extracted automatically and then visually verified. The N1 can be seen in Fig. 3, and its topography is presented in Fig. 4.

3.2.2. N1 amplitude

Overall N1 amplitude was affected by noise level during both active and passive Listening, $F(2, 62) = 74.12.6$, $p < 0.001$, $\eta^2 = 0.71$; polynomial decompositions revealed that N1 amplitude decreased as noise level increased, $F(1, 31) = 84.35$, $p < 0.001$, $\eta^2 = 0.73$. The N1 was larger during active listening than passive listening, but this was only marginally significant, $F(1, 31) = 3.20$, $p = 0.084$, $\eta^2 = 0.09$. Both these main effects were qualified by a significant Noise Level by Listening Condition interaction, $F(2, 62) = 4.86$, $p = 0.011$, $\eta^2 = 0.14$. This interaction was due to a greater decline in N1 amplitude as noise level increased during passive listening compared with active listening. Polynomial decompositions across noise level revealed a larger decline in N1 amplitude during passive listening, $F(1, 32) = 94.4$, $p < 0.001$, $\eta^2 = 0.75$, than active listening, $F(1, 32) = 69.1$, $p < 0.001$, $\eta^2 = 0.68$.

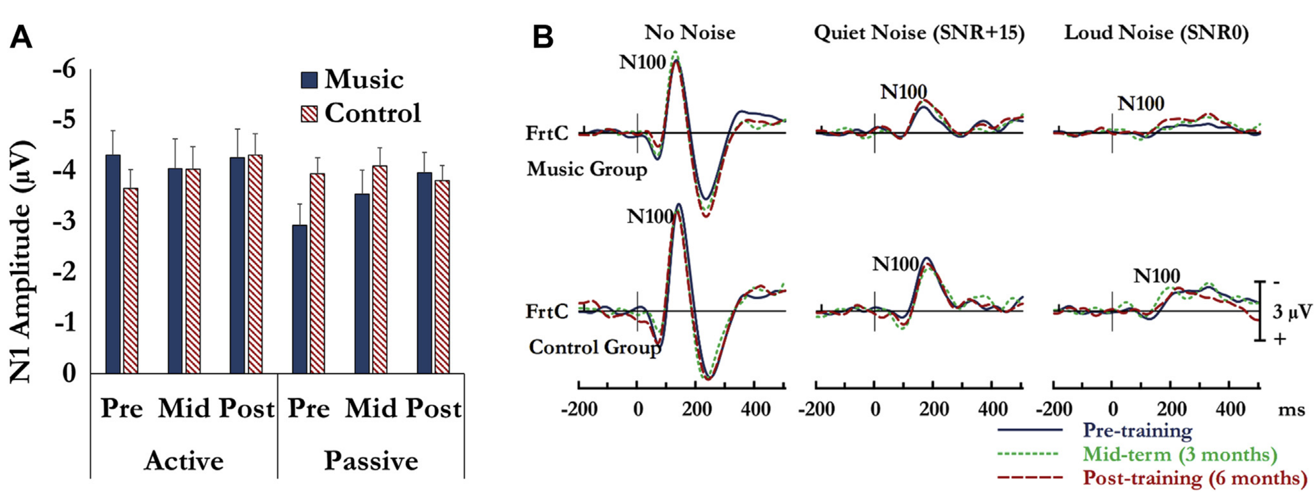

Fig. 3. (A) N1 amplitude averaged across frontocentral electrodes. N1 amplitude increased in the Music group from Pre to Mid to Post during passive listening ($p = 0.04$, $\eta^2 = 0.33$). (B) ERPs recorded over frontocentral electrodes during passive listening. The plots present the average of the 9 electrodes included in the montage. Abbreviations: ERP, event-related potential; Pre, pre-training session; Mid, mid-training session; Post, post-training session.

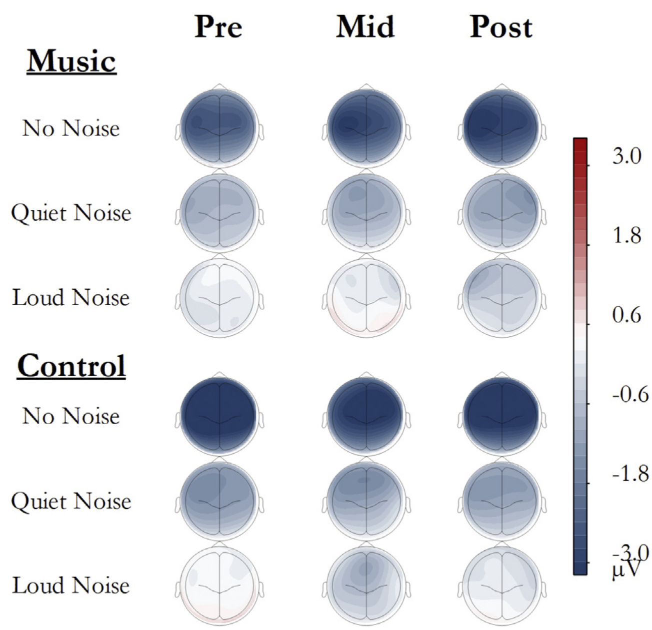

Fig. 4. Topographic headplots for the N1 evoked during passive listening. Plots present the mean topography for the ±10 ms window around the peak of the N1 for each noise level (No-Noise: 134–154 ms; Quiet: 188–208 ms; Loud: 206–226 ms). Plots are presented from the top view to highlight the montage used in the data analysis.

3.2.3. N100

3.2.3.1. Amplitude: effect of training. Most importantly, the Group by Session by Listening Condition interaction was significant, F (2, 62) = 6.17, $p = 0.004$, $\eta^2 = 0.17$. Follow-up tests revealed that the Session by Listening Condition interaction was significant in the Music group, F (2, 22) = 4.75, $p = 0.019$, $\eta^2 = 0.30$, whereas the Session by Listening Condition interaction and main effect of Session were not significant in the control group ($p = 0.50$, 0.37, respectively). In the Music group, polynomial decompositions were calculated to model the linear trend of N1 amplitude across testing sessions. These revealed that N1 amplitude increased linearly from the Pre to Mid to Post sessions during passive listening, F (1, 11) = 5.43, $p = 0.04$, $\eta^2 = 0.33$, but not active listening ($p = 0.70$).

3.2.4. N1 latency

Overall N100 latency was affected by noise level, F (2, 62) = 243.74, $p < 0.001$, $\eta^2 = 0.89$; polynomial decompositions revealed that N1 latency increased linearly as noise level increased, F (1, 31) = 415.52.73, $p < 0.001$, $\eta^2 = 0.93$ (144 ms, 198 ms, and 216 ms for No-Noise, Quiet, and Loud, respectively).

3.2.5. N1 latency: effect of training

No main effect of session ($p = 0.38$) or interactions with session were significant for N1 latency (for all $p > 0.14$) across all 3 groups.

3.2.6. ERP: fronto-left activity

The ICA revealed a component at fronto-left electrodes (FP1, AF7, AF3, F7, F5, F3, FT7, FC5, FC3). Visual inspection of the ERPs at these electrodes revealed a large negative complex that extended from ~400 to 1000 ms, with negative-going activity starting at 200 ms. To capture the late negative complex, the mean amplitude (MA) of this component was analyzed during 3 epochs: 200–400 ms, 400–700 ms, and 700–1000 ms. The ERPs at frontal-left electrodes can be seen in Fig. 5, and the topographies for each analysis epoch can be seen in Figs. 6–8.

3.2.7. Mean amplitude 200–400 ms: effects of noise and condition

Electrical brain activity during the 200–400 ms epoch was significantly impacted by noise level, F (2, 62) = 84.69, $p < 0.001$, $\eta^2 = 0.73$. Polynomial decompositions on noise level revealed a quadratic trend, F (1, 31) = 92.72, $p < 0.001$, $\eta^2 = 0.75$. Amplitude during this epoch was smaller in the Quiet condition than the No-Noise condition (1.86 vs. −1.03 μV) but was similar between the Quiet and Loud conditions (−1.03 vs. −0.73 μV). No significant main effects or interaction with listening condition were observed (except a training effect—see below).

3.2.8. Mean amplitude 200–400 ms: effect of training

Most importantly, the Group by Session by Listening Condition interaction was significant, F (2, 62) = 3.63, $p = 0.032$, $\eta^2 = 0.11$. Follow-up simple two-way interactions revealed a significant Session by Listening Condition interaction in the Music group, F (2, 22) = 3.89, $p = 0.036$, $\eta^2 = 0.26$, but not in the control group ($p = 0.62$). Follow-up tests in the Music group revealed a significant effect of session during active listening, F (2, 22) = 4.10, $p = 0.031$, $\eta^2 = 0.27$. The amplitude increased (i.e., became more positive) from Pre to Mid to Post, as revealed by a significant linear trend, F (1, 11) = 5.28, $p = 0.04$. The effect of session was not significant in the Music group during passive listening ($p = 0.95$).

3.2.9. Mean amplitude 400–700 ms: effects of noise and condition

Electrical brain activity during this epoch was impacted by Noise Level, F (2, 62) = 23.10, $p < 0.001$, $\eta^2 = 0.43$. Polynomial decompositions on Noise Level revealed a quadratic trend, F (1, 31) = 69.04, $p < 0.001$, $\eta^2 = 0.69$. Amplitude during this epoch was smallest in the Quiet condition (−1.9 μV) compared to both No-noise (−0.98 μV) and Loud (−0.44 μV) conditions. This effect was qualified by a significant interaction with Listening Condition, F (2, 62) = 6.27, $p = 0.003$, $\eta^2 = 0.17$. Follow-up tests in the Quiet condition revealed that the amplitude was smaller (i.e., more negative) during Active Listening compared to Passive Listening ($p = 0.013$). During the No-noise and Loud conditions, there was no significant impact of Listening Condition ($p = 0.2$ & 0.32).

3.2.10. Mean amplitude 400–700 ms: effect of training

Most importantly, the Group by Session by Listening Condition interaction was significant, F (2, 62) = 3.70, $p = 0.030$, $\eta^2 = 0.11$. Follow-up simple two-way interactions revealed a significant Listening Condition by Session interaction in the Music group, F (2, 22) = 4.26, $p = 0.027$, $\eta^2 = 0.28$, but not the control group ($p = 0.46$). Further follow-ups in the Music group revealed a marginally significant effect of session during active listening, F (2, 22) = 3.41, $p = 0.051$, $\eta^2 = 0.24$, but not during passive listening, ($p = 0.4$). During active listening, the amplitude increased (i.e., became more positive) from Pre to Mid to Post, as revealed by a marginally significant linear trend, F (1, 11) = 4.12, $p = 0.066$, $\eta^2 = 0.28$.

3.2.11. Mean amplitude 700–1000 ms: effects of noise and condition

Electrical brain activity during this epoch was impacted by noise level, F (2, 62) = 6.32, $p = 0.003$, $\eta^2 = 0.17$. Polynomial decompositions on noise level revealed a quadratic trend, F (1, 31) = 18.30, $p < 0.001$, $\eta^2 = 0.37$. Amplitude during this epoch was smallest in the Quiet condition (−1.6 μV) compared with both No-Noise (−0.77 μV) and Loud (−0.73 μV) conditions. No other main effects or higher-order interactions between listening condition and noise level were significant.

3.2.12. Mean amplitude 700–1000 ms: effect of training

At the 700–1000 ms epoch, the Session by Listening Condition interaction was significant, F (2, 62) = 3.15, $p = 0.050$, $\eta^2 = 0.09$, but

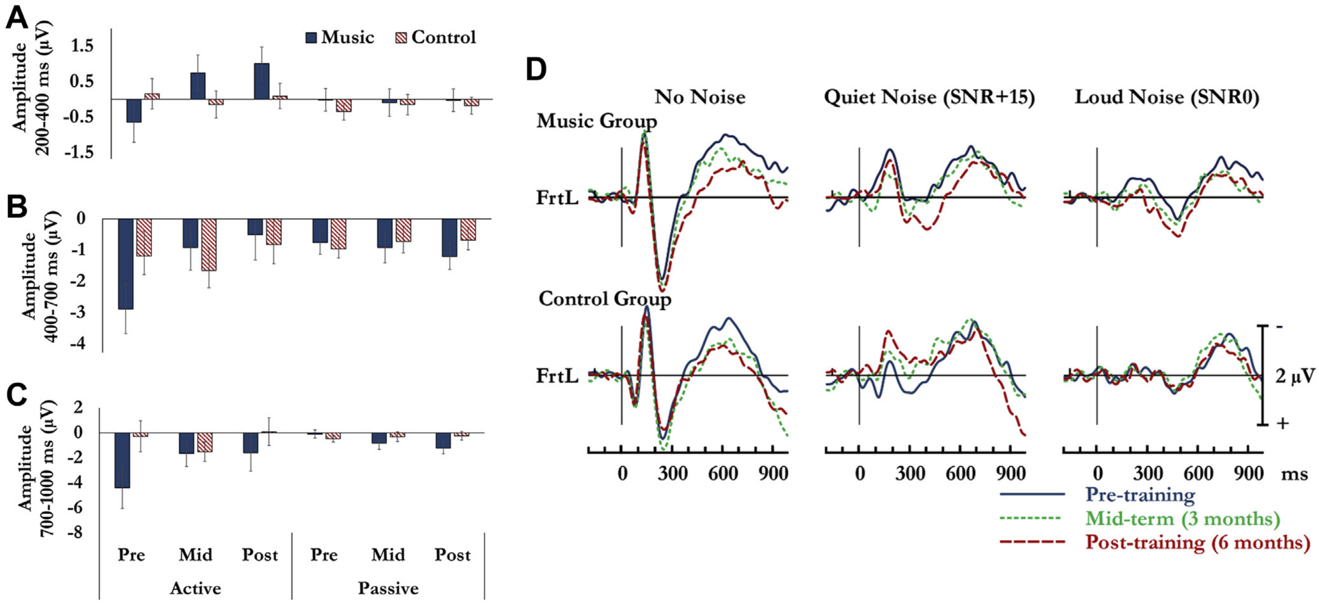

Fig. 5. ERP amplitude at fronto-left electrodes. (A) Mean amplitude from 200 to 400 ms averaged across fronto-left electrodes. There was an increase in positive-going electrical brain activity in the Music group during active listening ($p = 0.031$, $\eta^2 = 0.27$). (B) Mean amplitude from 400 to 700 ms averaged across fronto-left electrodes. There was an increase in positive-going electrical brain activity in the Music group during active listening ($p = 0.051$, $\eta^2 = 0.24$). (C) Mean amplitude from 700 to 1000 ms averaged across fronto-left electrodes. There was an increase in positive-going electrical brain activity in the Music group during active listening ($p = 0.051$, $\eta^2 = 0.24$). (D) ERPs recorded over fronto-left electrodes during active listening. The plots present the average of the 9 electrodes included in the montage. Abbreviation: ERP, event-related potential.

this was qualified by a significant three-way interaction between group, listening condition, and session, $F\,(2, 62) = 5.08$, $p = 0.009$, $\eta^2 = 0.14$. Follow-up simple two-way interactions revealed a significant Listening Condition by Session interaction in the Music group, $F\,(2, 22) = 6.14$, $p = 0.008$, $\eta^2 = 0.36$, but not in the control group ($p = 0.25$). Further follow-ups in the Music group revealed a marginally significant effect of session during active listening, $F\,(2, 22) = 3.43$, $p = 0.051$, $\eta^2 = 0.24$. During active listening, the amplitude increased (i.e., became more positive) from Pre to Mid to Post, as revealed by a marginally significant linear trend, $F\,(1, 11) = 4.42$, $p = 0.059$, $\eta^2 = 0.29$. During passive listening, there was also a significant effect of session in the Music group, $F\,(2, 22) = 7.49$, $p = 0.003$, $\eta^2 = 0.41$. More specifically, during passive listening, the amplitude decreased (i.e., became more negative) from Pre to Mid to Post, as revealed by a significant linear trend, $F\,(1, 11) = 15.04$, $p = 0.003$, $\eta^2 = 0.58$.

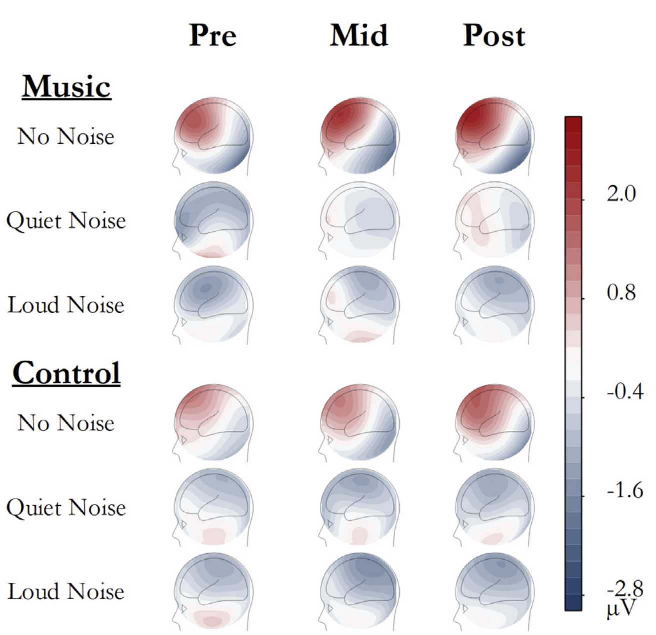

Fig. 6. Topographic headplots for the mean amplitude between 200 and 400 ms during active listening. Plots are presented from the left view to highlight the montage used in the data analysis. Abbreviations: Pre, pre-training session; Mid, mid-training session; Post, post-training session.

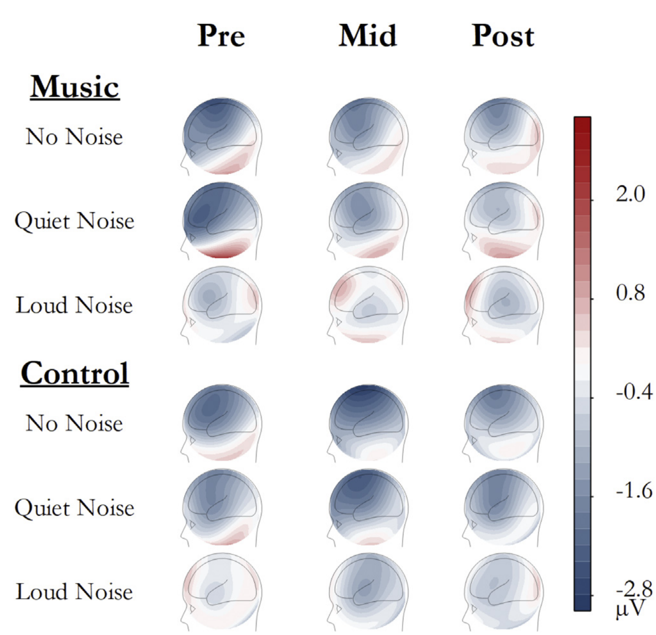

Fig. 7. Topographic headplots for the mean amplitude between 400 and 700 ms during active listening. Plots are presented from the left view to highlight the montage used in the data analysis. Abbreviations: Pre, pre-training session; Mid, mid-training session; Post, post-training session.

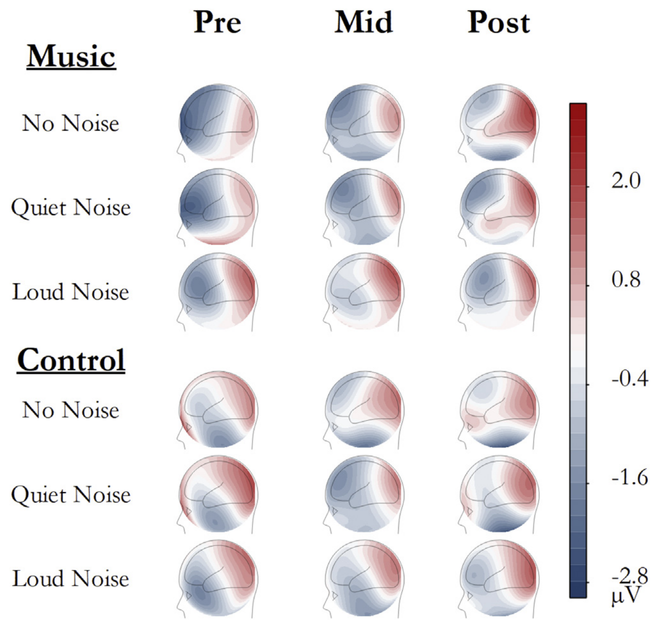

Fig. 8. Topographic headplots for the mean amplitude between 700 and 1000 ms during active listening. Plots are presented from the left view to highlight the montage used in the data analysis. Abbreviations: Pre, pre-training session; Mid, mid-training session; Post, post-training session.

3.3. Brain-behavior Pre-Post difference correlations

The next step in the analysis was to determine if the training-related changes in neurophysiology were related to changes in behavioral performance. Behavioral responses during No-Noise and Quiet conditions were both near ceiling and did not improve in any group. Accordingly, this analysis focused on behavior and neurophysiology recorded during the Loud condition. For all 3 electrophysiological measurements (i.e., MA epochs of 200–400 ms, 400–700 ms, and 700–1000 ms), data were averaged across the analysis montage separately for the pre- and post-training conditions. Next the pre-post difference was calculated for the 3 MA epochs and for the behavioral data. As a first step, we calculated a regression with the pre-post difference in behavior performance entered as a dependent variable and the pre-post difference in the MA during the 200–400 ms, 400–700 ms, and 700–1000 ms epochs as predictor variables. The overall regression model was significant, $F(3, 29) = 3.90$, $p = 0.019$, $R^2 = 0.29$, while none of the individual factors independently predicted a significant portion of the variance in the pre-post difference in behavior, MA 200–400: $\beta = -0.028$, $p = 0.134$; MA 400–700: $\beta = 0.027$, $p = 0.054$; MA 700–1000: $\beta = 0.008$, $p = 0.33$. This pattern of results—a significant model without any significant predictors—suggests multicollinearity in our predictor variables. Accordingly, a PCA was conducted on the pre-post difference in MA during the 3 epochs. The PCA revealed a single component that accounted for 77.7% of the variance (pre-post change in brain activity at fronto-left electrodes from 200 to 1000 ms). This component was then calculated for each individual participant. A correlation between this component and the change in behavior revealed that participants who had the most change in brain activity also had the greatest improvement in the ability to understand speech when background noise was loud, $r(33) = 0.40$, $p = 0.02$ (Fig. 9). Importantly, this correlation remained significant after removing outliers, $r(29) = 0.38$, $p = 0.04$, was significant in the Music group, $r(12) = 0.67$, $p = 0.018$, and nearly significant in the control group, $r(21) = 0.43$, $p = 0.055$.

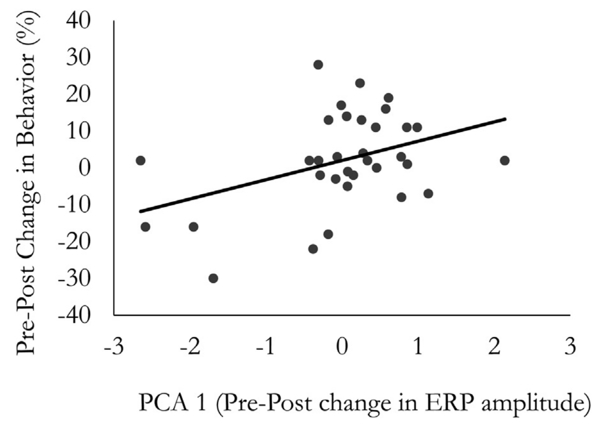

Fig. 9. Post minus Pre difference in accuracy as a function of the first PCA component extracted from the Post minus Pre difference in the mean ERP amplitude between 200 and 1000 ms. This component accounts for 77.7% of the variance in the Post minus Pre ERP data. Abbreviations: ERP, event-related potential; Pre, pre-training session; Mid, mid-training session; Post, post-training session.

4. Discussion

Six months of self-directed music lessons improved the ability to understand speech in background noise in older adults. This improvement was related to an increased positivity over frontal-left electrodes from 200 to 1000 ms after the onset of a word. No increased positivity was observed during the passive listening condition, suggesting that the benefit was due to an attention-dependent cognitive mechanism. There was, however, a musical training–related increase in N1 amplitude during passive listening, which suggests that musical training may have improved early obligatory processing of speech. Overall, these results suggest that musical training can be used to improve the ability to understand speech-in-noise in older adults by improving how older adults process speech and deploy their attentional resources to speech stimuli.

The most critical finding from this study is that music lessons improved the ability to understand speech in noise for older adults. This finding is critical for at least 2 reasons. First, it demonstrates that musical training may have a causal role in improving hearing abilities in older adults. This provides support for previous cross-sectional work demonstrating that older musicians have an advantage in understanding speech in noise compared with older nonmusicians (Parbery-Clark et al., 2011; Zendel et al., 2012). Second, it highlights that the brain remains plastic in older adults and is susceptible to modification based on one's life experiences. In the current context, this is critical as hearing difficulties in older adults are nearly universal (Gates and Mills, 2005; Quaranta et al., 2015), and one of the most commonly reported hearing difficulties is understanding speech in noisy environments. Accordingly, the results of this study demonstrate that music-based auditory rehabilitation is likely to be successful in older adults in improving their hearing abilities.

4.1. Musical training and the N1

Musical training enhanced the N1 during passive listening. The enhanced N1 suggests that musical training may improve the cortical encoding of incoming acoustic information, as it is well established that the N1 is associated with encoding the physical properties of a stimulus (Näätänen and Picton, 1987). One challenge with interpreting this result is that earlier neuroelectric responses (i.e., mid-latency and brainstem) were not measured in this study. The enhanced N1 during passive listening could be indicative of neuroplasticity along the superior temporal plane (Näätänen and

Picton, 1987; Scherg et al., 1989). Alternatively, the enhanced N1 could be due to more robust encoding of the incoming word at an earlier stage of processing, for example, in the brainstem. Support for the latter hypothesis comes from previous work that demonstrated enhanced speech encoding at the level of the brainstem, in musicians, young adults given musical training, and older adults who underwent auditory rehabilitation (Anderson et al., 2013; Parbery-Clark et al., 2012; Slater et al., 2015). Another interesting finding in the present study was that the enhancement in N1 was only observed during passive listening. The lack of enhancement during active listening may be due to the engagement of attentional processes that masked the ability to observe these differences during this condition. For example, attention-dependent cognitive processes are known to be reflected by an increased frontal processing negativity that overlaps the N1 epoch (Hansen and Hillyard, 1980; Näätänen, 1982). Overall, enhanced encoding or processing of basic acoustic features would have implications for later stages of processing related to orienting attention and extracting meaning from the auditory stimulus.

4.2. Musical training and attention

Participants in the Music group exhibited a post-training increase in positive-going electrical brain activity during active listening, at fronto-left electrodes, from 200 to 1000 ms. No such change was observed in the control group, suggesting that the increase in amplitude was related to musical training. Interestingly, the training-related change was not moderated by the level of background noise, suggesting that the benefit of musical training was related to speech processing and not the ability to filter out background noise per se. Moreover, this training-related change was only observed during active listening, suggesting that the auditory benefit of musical training is dependent on attention, and not solely related to enhanced automatic encoding of incoming speech information as indicated by the N1.

4.2.1. Musical training and attention—neural sources

The scalp topography associated with the fronto-left electrode montage used in the analysis was consistent with sources that had peak activations at 4 locations: left IFG, right PhG, right CB, and left MTG. The purpose of the source analysis was to identify putative brain regions that contributed to the scalp-recorded brain activity to identify which brain regions may have been affected by musical training. The next sections will consider the involvement of these brain regions in language- and speech-related processes and how neuroplasticity within each region might contribute to improved abilities to understand speech in noise. Although functional MRI data have been able to reveal specific neuroanatomical locations related to speech processing, (Vigneau et al., 2006, 2011), EEG data are not as precise spatially. Accordingly, the discussion will focus on gross anatomical regions.

The left IFG, left MTG, right PhG, and right CB have all been associated with processing incoming speech and the production of speech (Ackermann et al., 2007; Hwang et al., 2006; Silva-Pereyra et al., 2003; Vigneau et al., 2006, 2011). This pattern is not surprising, as the task in the present study involved listening to a word and repeating it aloud. The speech production portion of the task would have involved motor planning and eventual execution of the motor commands. Participants were told to delay their motor response until a light flashed 2000 ms after the onset of the word. The current data analysis only focused on the first 1000 ms. Thus, the actual production of speech was unlikely to contribute to the current data, but motor planning might be involved. The speech perception portion of the task required segregating the phonology of the word from the background noise (in the Quiet and Loud conditions), processing the phonology of the word, and then accessing the mental lexicon to identify the target word, so that it could be repeated aloud. The right PhG has been shown to be specifically active during speech-processing tasks that include background noise, but not when there is no background noise (Hwang et al., 2006), suggesting that this region may have been involved in separating speech from background noise. The left IFG and MTG are involved in the perception of speech and are part of a network that matches phonology to meaning (Nishitani et al., 2005; Vigneau et al., 2006; Watkins and Paus, 2004). Moreover, the left MTG and PhG are underlying sources of the N400 response; an ERP component associated with matching phonology to meaning (Friederici, 2002; Silva-Pereyra et al., 2003; Tse et al., 2007; Kutas and Federmeier, 2011). It is therefore possible that musical training enhanced the ability to match the incoming speech stimulus to the mental lexicon.

One possible explanation for how musical training enhances speech processing is that it appears to strengthen connections between the auditory and motor systems. Support for this idea comes from the right CB and left IFG, which both contributed to enhanced electrical brain activity at fronto-left electrodes in the present study. The right CB has tight functional connectivity with left frontal regions involved in speech perception and production and has been associated with the temporal organization of a prearticulatory verbal code (Ackermann et al., 2007). This prearticulatory code can be thought of as the “inner speech” that is used during subvocal rehearsal in verbal working memory tasks (Ackermann et al., 2007). It has been demonstrated that the motor system can simulate the articulatory motor commands required to produce a speech utterance to aid in understanding (Galantucci et al., 2006), and there is articulatory specificity in motor regions during speech perception (Ausilio et al., 2009; Pulvermüller et al., 2006). Musical training may therefore increase the connection between the auditory and motor systems to help understand the incoming stimulus. Accordingly, this pattern of results suggests that musical training may help people understand speech by increasing activity in regions that feed into the motor system, which then simulates the articulatory gestures that would mimic incoming speech.

4.2.2. Musical training and attention—temporal dynamics

Another way to consider the musical training–related increase in positive-going electrical brain activity at fronto-left electrodes is in terms of the temporal dynamics. Participants in the Music group exhibited an increase in late positive activity over fronto-left electrodes. P3 responses occur during the post-200 ms epoch and are typically related to allocation of cognitive resources for conscious information processing (Dinteren et al., 2014; Kok, 2001). Auditory P3 responses shift from posterior to anterior sites in older adults, reflecting an age-related change in how attention is allocated (Amenedo and Díaz, 1998; Anderer et al., 1996; Dinteren et al., 2014). This anterior increase in positive activity is often observed in older adults when processing sensory information (Davis et al., 2008; Grady et al., 1994). The frontal shift is likely a compensatory mechanism that offsets age-related decline in the functional and structural properties of posterior brain regions. In the present study, a musical training–related enhanced positivity was observed over frontal-left electrodes, during the active listening condition. Given the timing overlap with P3, this pattern of results suggests that musical training improved the ability to orient attention toward salient features in the incoming auditory stimulus, by enhancing brain responses from regions responsible for processing features of the incoming acoustic stimulus. When considered with the source-based interpretation, this could represent an enhanced phonological classification of the word, followed by enhanced orientation of attention toward acoustic features that are related to articulatory gestures.

4.3. Musical training and the speech-motor system

The process of learning to play the piano requires developing reciprocal connections between auditory and motor regions to ensure finger movements line up with metronome clicks, other instruments, or other key presses (i.e., harmony). Given the pattern of neural sources that contributed to the enhanced activity at fronto-left electrodes, it is possible that the mechanism related to improved understanding of speech in noise after musical training is related to plasticity in the connections between the auditory and motor systems. There is a long history of work supporting the idea that speech perception relies, in part, on the speech-motor system (Liberman and Mattingly, 1985; Pulvermüller and Fadiga, 2010) and left frontal regions have been associated with using an articulatory model to distinguish phonemic sounds (Zatorre et al., 1992). The results of the present study suggest that musical training strengthened auditory-motor connections, which resulted in an improved ability to understand speech in noise. Further support for the idea that musical training enhances auditory-motor connections comes from research in younger musicians. For instance, Du and Zatorre (2017) found that musicianship was associated with enhanced connectivity between right primary auditory regions and left IFG during a syllable-in-noise task. Further analyses revealed that there was greater phoneme specificity associated with the left IFG in musicians compared with nonmusicians (Du and Zatorre, 2017). Using a similar task in older adults, Du et al. (2016) found that older adults engaged the left IFG more than younger adults, and the magnitude of this increase was related to improved ability to understand the target speech syllable. Therefore, it is possible that musical training sharpens auditory-motor connections necessary for music performance, and the neurophysiological changes associated with this training transfers to speech tasks that rely on the same network.

Overall, musical training may have strengthened the ability to generate an articulatory model of the incoming word, based on a more robust neural representation of the incoming word (enhanced N1). The enhanced articulatory model of the incoming speech stimulus would then facilitate semantic processing of the incoming word when participants were attending to the incoming speech (enhanced late positive activity). The facilitated semantic processing of the incoming word could be due to either enhanced ability to deploy attention to acoustic features of the incoming speech or by generating a more robust articulatory model of the incoming speech.

4.4. Limitations

The purpose of including the Video group was to control for learning a new skill and to highlight the specific benefits of musical training; that is, it was musical training that caused the benefits and not learning a new skill or engaging in a new activity. Given the high withdrawal rate from the Video group, it is possible that this group did not serve as an adequate control group and the benefit of musical training was in fact related to learning a new skill or engaging in a new activity, and not specifically to musical training. The similarity between the Video and No-contact groups on their ERP and behavioral data suggest that this is not the case; however, given the small, nonrandomized sample in the Video group, this claim is tentative. Future studies performed with more participants and with an active control group that is more comparable to musical training would strengthen the findings reported here. Another important consideration is why the withdrawal rate was higher in the Video group. The high rate of withdrawal may be related to a cohort effect. In general, the current cohort of older adults may have a more positive association with playing music and a more negative association with playing video games. For example, in 1982, the United States Surgeon General identified video games as one of the main factors that contributes to interfamily violence and that video games may detract from a child's social development (Kestenbaum and Weinstein, 1985). The participants in the Video group that remained in the study are likely more representative of future cohorts of older adults who will see video game playing as a normal activity. Analyzing a group (albeit a small one) of older adults with little exposure to video games will become progressively more difficult as time passes. Therefore, although this group is small, the findings may be of critical importance. A more detailed explanation about the withdrawal rate can be found in West et al. (2017). West et al. (2017) presented data from the same study but examined the benefits of video game training on visuospatial processing and used the Music group as the active control.

A second limitation was in the examination of hearing status. In this study, bone-conduction thresholds were not measured and no otoscopic examination was carried out; therefore, participants with elevated PTTs could have either sensorineural or conductive hearing issues. The main findings of the study relate to central processing of acoustic information, and all participants had typical patterns of PTTs given their age. Accordingly, it is unlikely that mild conductive or mild sensorineural hearing loss moderates the benefit of musical training on the central auditory system. However, future work could examine if the effects of musical training differ between people with sensorineural and conductive hearing loss.

4.5. Summary

Musical training improved the ability to understand speech-in-noise in older adults. There was a general increase in positive-going electrical brain activity over fronto-left electrodes in the group that received musical training during the active listening task. This activity was related to underlying cortical sources that are associated with the speech-motor system. One possible explanation for this pattern of results is that musical training strengthened auditory-motor connections, which resulted in positive transfer to speech understanding. Overall these findings are most interesting from a rehabilitation perspective, as they suggest that music-based forms of auditory rehabilitation can be successful in older adults.

Improved hearing abilities could have a cascading effect on other cognitive abilities, as people with hearing difficulties in middle age have higher rates of cognitive decline as they get older (Lin et al., 2013). However, the link between hearing loss and cognitive decline has not yet been demonstrated causally. Musical training is also relatively inexpensive and enjoyable, suggesting that it could be easily incorporated into many people's lives. It remains possible that the benefits of musical training could be strengthened by identifying individual differences in one's susceptibility to the benefits of musical training and by examining how different types of musical training or music-based rehabilitation contribute to enhanced auditory abilities.

Disclosure statement

The authors have no actual or potential conflicts of interest.

Acknowledgements

The authors thank Olivier Dussault, Charles-David Tremblay, Samira Mellah, and Mihaela Felezeu for assistance with data collection. Support for this research came from the Canada Research Chairs program, the GRAMMY Foundation, Fondation Caroline Durant, Fonds de Recherche du Québec—Santé, and Natural Sciences and Engineering Research Council of Canada

Collaborative Research and Training Experience Program in Auditory Cognitive Neuroscience (NSERC–CREATE–ACN).

References

Ackermann, H., Mathiak, K., Riecker, A., 2007. The contribution of the cerebellum to speech production and speech perception: Clinical and functional imaging data. Cerebellum 6, 202–213.

Alain, C., Snyder, J.S., Dyson, B.J., 2006. Aging and the Perceptual Organization of Sounds: A Change of Scene? In: Conn, P.M. (Ed.), Handbook of Models for Human Aging. Elsevier Academic Press, San Diego, CA, pp. 759–769. https://doi.org/10.1016/B978-012369391-4/50065-5.

Alain, C., Zendel, B.R., Hutka, S., Bidelman, G.M., 2014. Turning down the noise: the benefit of musical training on the aging auditory brain. Hear. Res. 308, 162–173.

Amenedo, E., Díaz, F., 1998. Aging-related changes in processing of non-target and target stimuli during an auditory oddball task. Biol. Psychol. 48, 235–267.

Anderer, P., Semlitsch, H.V., Saletu, B., 1996. Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes. Electroencephalogr. Clin. Neurophysiol. 99, 458–472.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., Kraus, N., 2013. Reversal of age-related neural timing delays with training. Proc. Natl. Acad. Sci. U. S. A. 110, 4357–4362.

Ausilio, A.D., Pulvermuller, F., Salmas, P., Bufalari, I., Begliomini, C., Fadiga, L., 2009. The motor somatotopy of speech perception. Curr. Biol. 19, 381–385.

Berg, P., Scherg, M., 1994. A multiple source approach to the correction of eye artifacts. Electroencephalogr. Clin. Neurophysiol. 90, 229–241.

Billings, C.J., Tremblay, K.L., Stecker, G.C., Tolin, W.M., 2009. Human evoked cortical activity to signal-to-noise ratio and absolute signal level. Hear. Res. 254, 15–24.

Bugos, J.A., Perlstein, W.M., McCrae, C.S., Brophy, T.S., Bedenbaugh, P.H., 2007. Individualized piano instruction enhances executive functioning and working memory in older adults. Aging Ment. Health 11, 464–471.

Carson, N., Leach, L., Murphy, K.J., 2018. A re-examination of Montreal Cognitive Assessment (MoCA) cut off scores. Int. J. Geriatr. Psychiatry 33, 379–388.

Coffey, E.B.J., Mogilever, N.B., Zatorre, R.J., 2017. Speech-in-noise perception in musicians : a review. Hear. Res. 352, 49–69.

Davis, S.W., Dennis, N.A., Daselaar, S.M., Fleck, M.S., Cabeza, R., 2008. Que PASA ? The posterior–anterior shift in aging. Cereb. Cortex 18, 1201–1209.

de Peralta Menendez, R.G., Andino, S.G., Lantz, G., Michel, C.M., Landis, T., 2001. Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. Brain Topogr. 14, 131–137.

Diarra, M., Zendel, B.R., Benady-Chorney, J., Blanchette, C., Lepore, F., Peretz, I., Belleville, S., West, G.L., 2019. Playing Super Mario increases oculomotor inhibition and frontal eye field grey matter in older adults. Exp. Brain Res. 237, 723–733.

van Dinteren, R., Arns, M., Jongsma, M.L.A., Kessels, R.P.C., 2014. P300 development across the Lifespan : a Systematic review and meta-analysis. PLoS One 9, e87347.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2014. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc. Natl. Acad. Sci. U. S. A. 111, 7126–7131.

Du, Y., Buchsbaum, B.R., Grady, C.L., Alain, C., 2016. Increased activity in frontal motor cortex compensates impaired speech perception in older adults. Nat. Commun. 7, 12241.

Du, Y., Zatorre, R.J., 2017. Musical training sharpens and bonds ears and tongue to hear speech better. Proc. Natl. Acad. Sci. U. S. A. 114, 13579–13584.

Fleming, D., Belleville, S., Peretz, I., West, G., Zendel, B.R., 2019. The effects of short-term musical training on the neural processing of speech-in-noise in older adults. Brain Cogn.

Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., Trainor, L.J., 2006. One year of musical training affects development of auditory cortical-evoked fields in young children. Brain 129, 2593–2608.

Gajewski, A., 2011. Piano Lesson Package for Synthesia. Availabl at: www.arturgajewski.com. (Accessed 19 January 2012).

Galantucci, B., Fowler, C., Turvey, M.T., 2006. The motor theory of speech perception reviewed. Psychon. Bull. Rev. 13, 361–377.

Gates, G.A., Mills, J.H., 2005. Presbycusis. Lancet 366, 1111–1120.

Grady, C.L., Maisog, J.M., Horwitz, B., Ungerleider, L.G., Mentis, M.J., Salerno, J., Haxby, J.V., 1994. Age-related processing changes in cortical blood flow activation of faces and location during visual processing of Faces and location. J. Neurosci. 14, 1450–1462.

Hanna-Pladdy, B., MacKay, A., 2011. The relation between instrumental musical activity and cognitive aging. Neuropsychology 25, 378.

Hansen, J.C., Hillyard, S. a., 1980. Endogenous brain potentials associated with selective auditory attention. Electroencephalogr. Clin. Neurophysiol. 49, 277–290.

Humes, L.E., 2019. The World Health Organization's hearing-impairment grading system: an evaluation for unaided communication in age-related hearing loss. Int. J. Audiol. 58, 12–20.

Hwang, J.H., Wu, C.W., Chen, J.H., Liu, T.C., 2006. The effects of masking on the activation of auditory-associated cortex during speech listening in white noise. Acta Otolaryngol. 126, 916–920.

Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., Muchnik, C., 2006. Identification of syllables in noise: electrophysiological and behavioral correlates. J. Acoust. Soc. Am. 120, 926.

Kestenbaum, G.I., Weinstein, L., 1985. Personality, psychopathology, and developmental issues in male adolescent video game use. J. Am. Acad. Child Psychiatry 24, 329–333.

Koelsch, S., Schröger, E., Tervaniemi, M., 1999. Superior pre-attentive auditory processing in musicians. Neuroreport 10, 1309–1313.

Kok, A., 2001. On the utility of P3 amplitude as a measure of processing capacity. Psychophysiology 38, 557–577.

Kraus, N., Chandrasekaran, B., 2010. Music training for the development of auditory skills. Nat. Rev. Neurosci. 11, 599–605.

Kraus, N., White-Schwoch, T., 2015. Unraveling the biology of auditory learning: a cognitive–sensorimotor–Reward Framework. Trends Cogn. Sci. 19, 642–654.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Ann. Rev. Psychol. 62, 621–647.

Lappe, C., Herholz, S.C., Trainor, L.J., Pantev, C., 2008. Cortical plasticity induced by short-term unimodal and multimodal musical training. J. Neurosci. 28, 9632–9639.

Liberman, A., Mattingly, I., 1985. The motor theory of speech perception revised. Cognition 21, 1–36.

Lin, F.R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T.B., Purchase-Helzner, E., Simonsick, E.M., 2013. Hearing loss and cognitive decline in older adults. JAMA Intern. Med. 173, 293–299.

Martin, B.A., Sigal, A., Kurtzberg, D., Stapells, D.R., 1997. The effects of decreased audibility produced by high-pass noise masking on cortical event-related potentials to speech sounds/ba/and/da. J. Acoust. Soc. Am. 101, 1585–1599.

Mick, P., Kawachi, I., Lin, F.R., 2014. The association between hearing loss and social isolation in older adults. Otolaryngol. Head Neck Surg. 150, 378–384.

Näätänen, R., 1982. Processing negativity: an evoked-potential reflection. Psychol. Bull. 92, 605–640.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375–425.

Nasreddine, Z.S., Phillips, N.A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Chertkow, H., 2005. The Montreal cognitive assessment, MoCA : a brief screening. J. Am. Geriatr. Soc. 53, 695–699.

Nishitani, N., Amunts, K., Hari, R., 2005. Broca's region: from action to language. Physiology 20, 60–69.

Pantev, C., Oostenveld, R., Engelien, a., Ross, B., Roberts, L.E., Hoke, M., 1998. Increased auditory cortical representation in musicians. Nature 392, 811–814.

Pantev, C., Roberts, L.E., Schulz, M., Engelien, a., Ross, B., 2001. Timbre-specific enhancement of auditory cortical representations in musicians. Neuroreport 12, 169–174.

Parbery-Clark, A., Anderson, S., Hittner, E., Kraus, N., 2012. Musical experience offsets age-related delays in neural timing. Neurobiol. Aging 33, 1483.e1-4.

Parbery-clark, A., Skoe, E., Kraus, N., 2009. Musical experience limits the degradative effects of background noise on the neural processing of sound. J. Neurosci. 29, 14100–14107.

Parbery-Clark, A., Skoe, E., Lam, C., Kraus, N., 2009. Musician enhancement for speech-in-noise. Ear Hear. 30, 653–661.

Parbery-Clark, A., Strait, D.L., Anderson, S., Hittner, E., Kraus, N., 2011. Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. PLoS One 6, e18082.

Pettigrew, C.M., Murdoch, B.E., Ponton, C.W., Kei, J., Chenery, H.J., Alku, P., 2004. Subtitled videos and mismatch negativity (MMN) investigations of spoken word processing. J. Am. Acad. Audiol. 15, 469–485.

Pichora-Fuller, M.K., Schneider, B.A., Daneman, M., 1995. How young and old adults listen to and remember speech in noise. J. Acoust. Soc. Am. 97, 593–608.

Pulvermüller, F., Huss, M., Kherif, F., Martin, F., Hauk, O., Shtyrov, Y., 2006. Motor cortex maps articulatory features of speech sounds. Proc. Natl. Acad. Sci. U. S. A. 103, 7865–7870.

Pulvermüller, F., Fadiga, L., 2010. Active perception: sensorimotor circuits as a cortical basis for language. Nat. Rev. Neurosci. 11, 351–360.

Quaranta, N., Coppola, F., Casulli, M., Barulli, M.R., Panza, F., Tortelli, R., Solfrizzi, V., Sabbà, C., Logroscino, G., 2015. Epidemiology of age related hearing loss: a review. Hearing Balance Commun. 13, 77–81.

Robert Frisina, D., Frisina, R.D., 1997. Speech recognition in noise and presbycusis: Relations to possible neural mechanisms. Hear. Res. 106, 95–104.

Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the late human auditory evoked potentials. J. Cogn. Neurosci. 1, 336–355.

Schneider, B.A., Pichora-Fuller, M.K., Daneman, M., 2010. Effects of Senescent changes in audition and cognition on spoken language comprehension. In: Gordon-Salant, S., Frisina, R.D., Popper, A.N., Fay, R.R. (Eds.), Aging Auditory System. Springer US, New York, pp. 167–209.

Seinfeld, S., Figueroa, H., Ortiz-Gil, J., Sanchez-Vives, M.V., 2013. Effects of music learning and piano practice on cognitive function, mood and quality of life in older adults. Front. Psychol. 4, 810.

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E., 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J. Neurosci. 23, 5545–5552.

Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B., 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport 16, 1781–1785.

Silva-Pereyra, J., Rivera-Gaxiola, M., Aubert, E., Bosch, J., Galán, L., Salazar, A., 2003. N400 during lexical decision tasks: a current source localization study. Clin. Neurophysiol. 114, 2469–2486.

Slater, J., Skoe, E., Strait, D.L., O'Connell, S., Thompson, E., Kraus, N., 2015. Music training improves speech-in-noise perception: longitudinal evidence from a community-based music program. Behav. Brain Res. 291, 244–252.
Stenklev, N.C., Laukli, E., 2004. Presbyacusis—hearing thresholds and the ISO 7029. Int. J. Audiol. 43, 295–306.
Tierney, A.T., Krizman, J., Kraus, N., 2015. Music training alters the course of adolescent auditory development. Proc. Natl. Acad. Sci. U. S. A. 112, 10062–10067.
Tse, C.-Y., Lee, C.-L., Sullivan, J., Garnsey, S.M., Dell, G.S., Fabiani, M., Gratton, G., 2007. Imaging cortical dynamics of language processing with the event-related optical signal. Proc. Natl. Acad. Sci. U. S. A. 104, 17157–17162.
Vigneau, M., Beaucousin, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Tzourio-Mazoyer, N., 2006. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. Neuroimage 30, 1414–1432.
Vigneau, M., Beaucousin, V., Hervé, P.Y., Jobard, G., Petit, L., Crivello, F., Tzourio-Mazoyer, N., 2011. What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. Neuroimage 54, 577–593.
Watkins, K., Paus, T., 2004. Modulation of motor Excitability during speech perception: the role of Broca's area. J. Cogn. Neurosci. 16, 978–987.
West, G.L., Zendel, B.R., Konishi, K., Benady-Chorney, J., Bohbot, V.D., Peretz, I., Belleville, S., 2017. Playing Super Mario 64 increases hippocampal grey matter in older adults. PLoS One 12, e0187779.
Wong, P.C.M., Skoe, E., Russo, N.M., Dees, T., Kraus, N., 2007. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat. Neurosci. 10, 420–422.
Yamasoba, T., Lin, F.R., Someya, S., Kashio, A., Sakamoto, T., Kondo, K., 2013. Current concepts in age-related hearing loss: epidemiology and mechanistic pathways. Hear. Res. 303, 30–38.
Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch Discrimination in speech processing. Science 256, 846–849.
Zendel, B.R., Alain, C., 2012. Musicians experience less age-related decline in central auditory processing. Psychol. Aging 27, 410–417.
Zendel, B.R., Alain, C., 2013. The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. J. Cogn. Neurosci. 25, 503–516.
Zendel, B.R., Alain, C., 2014. Enhanced attention-dependent activity in the auditory cortex of older musicians. Neurobiol. Aging 35, 55–63.
Zendel, B.R., Tremblay, C., Belleville, S., Peretz, I., 2015. The impact of musicianship on the cortical mechanisms related to seperating speech from background noise. J. Cogn. Neurosci. 27, 1044–1059.