


No Evidence That Music Training Benefits Speech Perception in Hearing-Impaired Listeners: A Systematic Review

Trends in Hearing
Volume 25: 1–16
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216520985678
journals.sagepub.com/home/tia


Colette M. McKay^{1,2,3} 

Abstract

As musicians have been shown to have a range of superior auditory skills to non-musicians (e.g., pitch discrimination ability), it has been hypothesized by many researchers that music training can have a beneficial effect on speech perception in populations with hearing impairment. This hypothesis relies on an assumption that the benefits seen in musicians are due to their training and not due to innate skills that may support successful musicianship. This systematic review examined the evidence from 13 longitudinal training studies that tested the hypothesis that music training has a causal effect on speech perception ability in hearing-impaired listeners. The papers were evaluated for quality of research design and appropriate analysis techniques. Only 4 of the 13 papers used a research design that allowed a causal relation between music training and outcome benefits to be validly tested, and none of those 4 papers with a better quality study design demonstrated a benefit of music training for speech perception. In spite of the lack of valid evidence in support of the hypothesis, 10 of the 13 papers made claims of benefits of music training, showing a propensity for confirmation bias in this area of research. It is recommended that future studies that aim to evaluate the association of speech perception ability and music training use a study design that differentiates the effects of training from those of innate perceptual and cognitive skills in the participants.

Keywords

music training, speech perception, hearing impairment

Received 18 August 2020; Revised 9 December 2020; accepted 12 December 2020

There has been an increase in interest recently in the question of whether music training has causal benefits for a range of speech perception skills in people with and without sensory or other deficits. This review focusses on the question of whether there is evidence that music training provides benefits in speech understanding for people with hearing impairment. Sensorineural hearing loss is associated with difficulty in speech perception, particularly in background noise, a difficulty that is not wholly ameliorated by amplification. This difficulty arises mostly through loss of spectral or temporal information in the periphery (e.g., loss of hair cells and their connections with the auditory nerve—features that are not known to be amenable to plastic changes due to training; Moore, 1996). However, deafness itself can induce plastic changes centrally (Kral et al., 2002; Lee et al., 2003; Sharma et al., 2015). It is therefore of interest to know whether music training can help to overcome the limitations in speech understanding in adults or

children with hearing impairment via induced plasticity in the central language networks. Studies that have tested this hypothesis have made the implicit assumption that central brain plasticity induced by music training will overcome or limit the effects of hearing loss.

The hypothesis that music training has causal benefits for skills outside of the music domain (i.e., far transfer of training—such as improved speech perception) has been based on hypothesized specific and general benefits of

¹Bionics Institute, Melbourne, Australia

²Department of Medical Bionics, The University of Melbourne, Melbourne, Australia

³Department of Audiology and Speech Pathology, The University of Melbourne, Melbourne, Australia

Corresponding Author:

Colette M. McKay, Bionics Institute, 384 Albert St, East Melbourne, VIC 3002, Australia.

Email: cmckay@bionicsinstitute.org



music training that could be induced via brain plasticity. First, specific skills acquired by music training may overlap with the skills needed for success in the other domain. For example, it has been hypothesized that training in musical pitch perception might transfer to an increased ability to perceive voice pitch, which is an auditory cue that can help to distinguish between two simultaneous talkers (Darwin et al., 2003) or to detect emotional prosody in speech (Bulut & Narayanan, 2008). These hypotheses have been evaluated in studies measuring the frequency following response (FFR) to the fundamental frequency (F0) and its harmonics evoked by a speech syllable. A systematic review of these papers (Rosenthal, 2020), however, challenged this hypothesis and concluded that, although the subcortical F0 response tends to be larger in musicians, the response size is not correlated with speech perception ability in noise. The overlap, precision, emotion, repetition, and attention (OPERA) hypothesis of Patel (2014) is an example of proposed transfer of benefits from music to speech domains via plasticity in shared neural networks. However, examples of far transfer of training are extremely rare in the psychology literature, and many scholars express doubt that it is possible (e.g., Melby-Lervag et al., 2016).

Second, it has been proposed that music training might improve general cognitive or academic skills, and these skills can be used to improve performance on any task-related outcome measure, including speech perception. In the educational field, many existing studies about the effect of music training are targeted at school-aged children, with the purpose of testing whether music training can transfer to cognitive ability or academic achievement (literacy or mathematics). However, a recent meta-analysis by Sala and Gobet (2020) found that, when quality of the research design was taken into account, there was no effect of music training on cognitive skills or academic achievement regardless of age or duration of training. Very small positive effects were only seen in studies with poor design (no randomization and using non-active controls), implying that those small positive results are very likely to be false positives. This null result is supported by studies suggesting that innate characteristics of musicians are better predictors of intelligence. For example, a large control study in twins, one of each twin being musically trained, concluded that the association between musicianship and intelligence was not causal (Mosing et al., 2016).

There are many published papers that compare normal-hearing musicians (with a variety of definitions of “musician”) to non-musicians using cross-sectional and correlational research designs and that show that musicians have a range of psychoacoustic skills that are better than those of non-musicians. These skills include those directly related to musicianship: pitch or

melodic contour discrimination (Baskent et al., 2018; Boebinger et al., 2015; Madsen et al., 2017, 2019; Martinez-Montes et al., 2013), temporal beat discrimination (Sares et al., 2018), and ability to attend to stimuli in a complex environment (Tierney et al., 2020; Vanden Bosch der Nederlanden et al., 2020). However, the question of whether these skills are associated with improved speech understanding in musicians in quiet or background noise has not been universally experimentally supported. A number of studies have not found a benefit for speech understanding in musicians in spite of benefits being demonstrated for pitch or intensity discrimination (e.g., Baskent et al., 2018; Boebinger et al., 2015; Madsen et al., 2017, 2019; Ruggles et al., 2014). For a review of studies that investigated speech in noise perception in neurologically normal musicians and non-musicians, the reader is referred to Coffey et al. (2017), who outline the possible theoretical bases of the connection between musicianship and speech in noise perception. However, that review does not discuss the possibility that any advantage for speech perception in noise for musicians may be related not to the music training of musicians, but to their innate skills, except for a note that future studies should “use longitudinal training studies to confirm the causal effects . . .”

Cross-sectional or correlational studies as described earlier, where musicians and non-musicians are compared, cannot distinguish between putative effects of plasticity due to music training and differences that may be due to innate (genetic or developmental) characteristics and skills. For example, superior innate auditory skills or personality characteristics may be a necessary or at least beneficial characteristic for becoming a professional or amateur musician (Swaminathan & Schellenberg, 2018a). A review by Schellenberg (2015) concluded that the association between music training and speech perception may be largely driven by interactions of genes and environment. This conclusion was based on genetic studies (e.g., Hambrick & Tucker-Drob, 2015) showing that musical aptitude (innate musical ability or potential to succeed in musicianship) has more influence on musical achievement than music practice, demonstrating a powerful influence of genes for becoming a successful musician. In addition, genetic linkage studies (e.g., Oikkonen et al., 2015) link musical aptitude with a range of innate auditory and cognitive skills that would also underpin performance of speech understanding. Cognitive skills, such as working memory, non-verbal intelligence, and attentional skills, may similarly contribute to musicianship (Swaminathan & Schellenberg, 2018a, 2018b). Additional group differences, such as socioeconomic status, are likely to be relevant (itself associated with other characteristics such as educational status, parental interaction quality during childhood, and engagement with other social and

intellectual activities). Some authors have argued that a correlation between duration of music training with speech perception outcomes implies a causal effect (e.g., Kraus et al., 2014). However, it is also the case that people with superior auditory and cognitive skills are likely to persist for longer in music training (Swaminathan & Schellenberg, 2018a). To directly test whether music training has a *causal* influence on speech perception (whether in normal-hearing or hearing-impaired populations), an appropriately controlled longitudinal training study in non-musicians, which controls or accounts for potential influences of genes or innate characteristics on the outcomes, is required. For this reason, only longitudinal training studies are reviewed in this systematic review.

Review Methods

A search on PubMed, PsyArticles, and Google Scholar was undertaken on March 31, 2020, and rechecked on November 30, 2020, using the terms *music training* and *speech* contained in the title or abstract. Review articles and references were scanned to find additional papers and to reject those that were not referring to speech perception. The abstracts of 60 papers were then used to identify those papers that used a longitudinal music training program in people with a hearing impairment and with outcome measures related to speech perception. Papers comparing musicians with non-musicians in a cross-sectional design were rejected, as were papers that studied music training in populations other than people with hearing impairment. Reviews and pilot studies were rejected, as were papers that did not contain original research. A total of 13 papers were identified. A flowchart of the selection process is given in Figure 1.

The papers are discussed with respect to the design and analysis principles that are detailed below. A meta-analysis statistical approach was not taken in this review, as the small number of studies, and the large number of potential covariables that need to be accounted for in this population, do not allow the statistical approach to be usefully interpreted. Instead, each paper is separately assessed and general conclusions drawn.

Assessment of Research Design: Did the Authors Use a Valid Research Design (Crossover or Randomized Control Trial)?

The gold standard research design to test the efficacy of a training therapy is a randomized controlled trial (RCT). In an RCT, the randomization of subjects between test and control groups aims to make the two groups equivalent in the pertinent individual characteristics that may confound the interpretation (George et al., 2016). Randomization is particularly important

when some or all of those characteristics are unknown (e.g., in a new drug trial). However, for randomization to be effective, the number of subjects in each group must be large enough to ensure that the two groups end up being equivalent in all potential confounds (i.e., they are both a good representation of the total population). For subjects with hearing loss, there are multiple known potential confounds related to the hearing impairment (hearing loss type and degree, age, age at onset, type of hearing device, etc.) as well as other general features (e.g., IQ, educational level, incidental exposure to music, etc.) and others we may not be able to identify before starting the experiment (such as differences in baseline performance on the outcome measure). This means that the test and control groups may need to be very large or very tightly defined for randomization to be truly statistically effective in this population.

For smaller studies, an alternative research design is a crossover design, in which the same participants undergo *both* test and control training sequentially, with the order of training balanced between two subgroups. This design would seem to overcome the potential problem of individual characteristics of subjects being different for the test and control training. In this case, there is an assumption that the effect of order (first or second training) on session (before, after training) is the same for both test and control training (i.e., no statistically significant *Order* × *Training* interaction). To satisfy this assumption, the test and control training should be carefully selected to avoid this potential interaction, and the analysis of the results should confirm a lack of this interaction before combining the two orders to test the main hypothesis.

Assessment of Control Training: Did the Authors Use a Control Group With an Appropriate Active Training? Did They Pay Attention to Potential Biases of Trainers and Trainees?

Without a control group, any increase in outcome measures between before and after training cannot be attributed to the training. A similar problem arises if the control group is a “no-training” passive control group. In that case, any increase in outcomes of the control group may be due to expectations of the effect of training (placebo effect) or the increased interactions with experimenters, and such a study could not distinguish effects of music training from another type of training. Such studies *fail to validly test the hypothesis*.

The choice of the active-control training is also important: The control and test training must only differ in the feature under test. Wright and Zhang (2009) discuss the distinction between stimulus learning and procedural learning, where stimulus learning refers

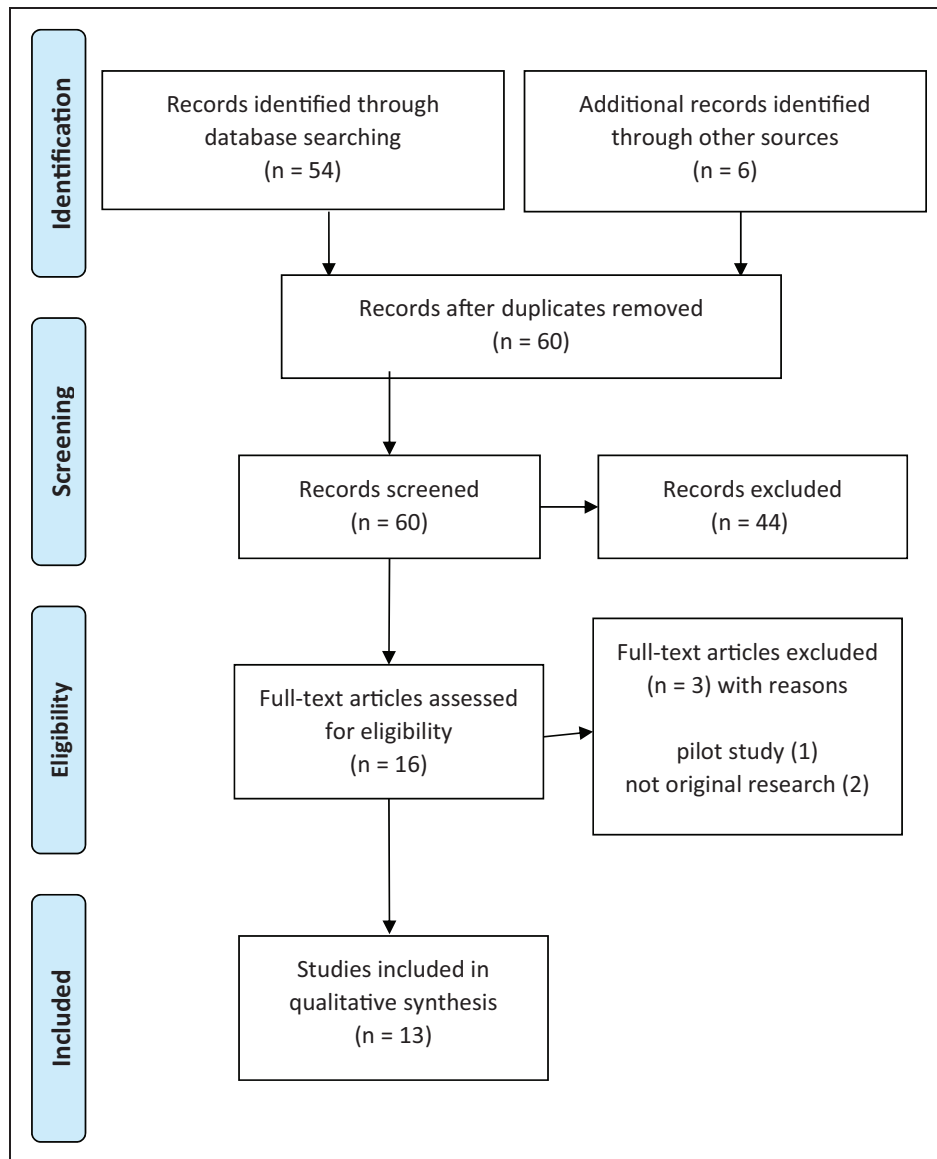


Figure 1. Database Searching and Selection Flowchart of Articles for the Review.

to learning the attributes of a stimulus and procedural learning refers to learning of the factors independent of the trained stimulus. The latter factors include task learning, and environmental factors such as lab environment and interactions with trainers. Without an active control group, the contribution of stimulus learning on its own (usually the aim of the experiment) cannot be assessed. All of the potential factors that could induce confounding differences between effects of test and control training, such as quality and duration of interaction with the trainers, and differences between potential strength of placebo effects need to be carefully controlled by an appropriate choice of control training.

Participant bias is a potential confounding factor that is quite difficult to eliminate or even limit in training

studies, as participants will always know what sort of training they are experiencing. In the case of music training, social and mainstream media often contain stories about benefits of music training, making participant bias and expectation particularly difficult to control. Therefore, to limit potential participant bias, both test and control training should be, as far as possible, equally plausibly associated with improved speech perception and equally enthusiastically proposed to the participant by the research team and/or trainer. In the case of child participants, the same principles apply equally to parents. Similarly, in a crossover design, participants (and/or their parents) should be told that both training types are proposed to improve speech perception and that researchers do not know which one is hypothesized

to be better. The exact script of what participants are told about the study should be predetermined and included in the report. These techniques do not entirely prevent individuals from having their own biases about the benefits of each training method, however.

Biases arising from both the people supplying the training and the researchers testing the outcome measurements also need to be considered and controlled. The interactions between trainer and trainee should be carefully controlled to make sure that participants in each group are receiving equal quality and type of interaction and encouragement to complete the training. The researchers who are measuring the outcome measures (e.g., speech perception) should be blind as to which training group the participant belongs. This last requirement is the easiest for researchers to implement in their research design.

Assessment of Randomization: Did the Authors Randomize Participants to the Test and Control Training? Did the Authors Report and Statistically Handle Dropouts Appropriately?

Without very careful randomization, it is not possible to control effects of innate abilities on the outcomes. For example, studies that compare people who *choose* to undertake music training with those who do not cannot separate the effects of training from inherent characteristics. In addition, the analysis should take into consideration the participants who do not complete the training and/or drop out of the study. For example, if only a subgroup of participants in the music training group complete the training, they are likely to be the subgroup that had innate auditory and cognitive skills to enable enjoyment of, and success in, music training. Therefore, they may have, in a sense, self-selected based on their innate ability and interest. Therefore, the final groups for analysis *are no longer randomly selected and can be affected by all the biases that non-random selection makes likely*. In general, it has been shown that compliant participants in clinical trials (whether in test or active control groups) have better outcomes than non-compliant participants (Sommer & Zeger, 1991). The gold standard analysis for medical clinical trials is “Intention to Treat” analysis, in which *all originally assigned randomized participants* must be included in the outcome analysis, regardless of whether there was failure of compliance, deviation of protocol, or drop out of participants from the study. This analysis tests the pragmatic effectiveness of the treatment as potentially applied in clinical practice. The challenge for Intention to Treat analysis occurs when there is missing outcome data. In this case, the statistical best practice is to do the analysis multiple times with different estimates of the missing

data to calculate a confidence interval that takes into account both the uncertainty of the missing data and the variance in the real data (George et al., 2016).

In a crossover study design, the challenge of dropouts and non compliance is somewhat different from that in an RCT. Leaving out subjects based on any missing data will have the same effect for both training methods, as the hypothesis is tested by a within-subject comparison. However, if many participants drop out, the results of the analysis may only apply to people who have characteristics likely to make them compliant. In this case, rate of dropouts and/or non compliance should also be measured and reported for each training method. Given equal compliance levels and dropout levels, the analysis can be performed with the remaining subjects (unlike in an RCT) to test the hypothesis.

Assessment of the Statistics: Did the Authors Use an Appropriate Statistical Test to Test Their Hypothesis and Did They Take Into Consideration Multiple Comparisons?

An excellent review of common scientific and statistical errors in clinical trial analysis has been published by George et al. (2016), and several pertinent topics relevant to this review are summarized here. First, when comparing test and control training outcomes, the appropriate test statistic is always one that compares *across groups*, not statistics that compare before and after training within groups. For example, a significant *Group (control, test) × Session (before, after) interaction* shows *the effect of training differs in the two groups*. Many papers fall into the statistical fallacy of assuming that the hypothesis is supported if the effect of session is significant in the test group and not the control group. Such a difference in significance level does not imply that the training effect is significantly larger in the test group: Only if the interaction term is significant will the hypothesis be supported. For example, the statistical significance of session in one group may be driven by a smaller variance between subjects in that group rather than by a larger effect size. Equivalently to using the interaction factor, the change in speech perception could be calculated for each person and the two groups then compared (e.g., using *t* test) to see if there are group differences in the change measure. A third method (baseline-adjusted analysis of covariance) uses the final performance level as the outcome measure and baseline values and training method as independent variables. All these methods test the hypothesis *between groups*, not within groups. In the crossover design, there are three factors: order (first or second), training method, and session (before, after training). Again, the statistic that tests the hypothesis is the *Training × Session*

interaction term (with the two orders combined). One should also check that there is no significant interaction between effects of order and training method before combining the orders.

Two other common errors are mentioned by George et al. (2016): failure to account for multiple comparisons and effects of regression to the mean. Many studies test hypotheses about multiple outcome measures, therefore increasing the chance of type I error. “Regression to the mean” describes a phenomenon whereby, in repeated measures in the same subject, those people with the highest scores tend to have lower scores when retested and those with the lowest scores tend to have improved scores on retest. This can be due to both random variation in test results (experimental error) but also due to real effects. This phenomenon can significantly bias the interpretation of statistical results when the baseline measures of the test and control groups differ and can be measured using a *Baseline × Training* effect interaction.

Results

Thirteen papers were identified that met the inclusion criteria. Eleven of the 13 addressed the population of users of cochlear implants (CIs; seven in adults, four in children) with the remaining two addressing children with either hearing aids or CIs and adults with mild to normal hearing impairment. The 13 papers are listed in Table 1, in which the columns detail information about the research design (type of design, control group and training, randomization, and outcome measures tested).

Of the 13 papers identified, four (Cheng et al., 2018; Firestone et al., 2020; Hutter et al., 2015; Smith et al., 2017) used no control group. Although all these papers claimed improvements in speech perception in CI users after training, with no control group, these improvements cannot be validly attributed to training or any aspect specifically about music. Cheng et al. (2018) studied 22 pediatric Mandarin-speaking CI users who were trained using the melodic contour identification (MCI) test for 30–60 hours over 8 weeks. Outcome measures were MCI performance, lexical tone identification, and sentence understanding in quiet, all of which improved over the five test sessions 2 weeks apart. Hutter et al. (2015) studied 12 newly implanted adult CI users, who undertook ten 45-minute music therapy sessions over an average of 134 days. The music therapy consisted of five modules that included both music (pitch, rhythm, timbre) and speech perception training. Outcome measures were three questionnaires (with sub measures) of sound quality, self-concept, and therapy satisfaction, along with three music assessments (pitch discrimination, timbre identification, and melody recognition). Assessment of musical timbre and melody identification, but not pitch discrimination, improved after music

therapy, as did estimated sound quality and self-concept. However, in a new CI user, most aspects of hearing improve rapidly in the first 3 months after implantation (Blamey et al., 2013; Blamey et al., 2001; Lazard et al., 2012), so any effect of music therapy cannot be deduced from this study. Smith et al. (2017) studied 21 experienced adult CI users who undertook music training using self-administered melody training software. Speech outcome measures were sentences in quiet and in noise. Participants were assessed before and after training and at 6 months post-training and were split for analysis between low and high levels of baseline music experience. Results showed improvements in both speech perception measures only for the low-musical-experience group. However, no between-group statistics were presented, and the music-experienced group had better baseline speech perception scores, suggesting that regression to the mean may have affected the difference between groups. No statistics were presented for both groups combined. Firestone et al. (2020) studied 11 experienced adult CI users, who were instructed to listen to music of their choice for 40 minutes a day, 5 days a week, for 4 or 8 weeks. No training tasks were involved. Outcome measures were obtained from three speech perception tests (words, sentences in quiet, and sentences in noise), a hearing questionnaire, a frequency change detection test, audiometric thresholds, and four cortical acoustic change response parameters for three sizes of frequency change. Significant differences between pre- and post-training were observed in all behavioral measures but not in the electroencephalogram (EEG) measures. It should be noted that the audiograms showed lower thresholds in the post-test session compared with the pre-test session, making the improvements in speech perception potentially caused by better audibility (e.g., due to higher volume or sensitivity setting in the CI or test environment changes).

A further paper compared two different music trainings without any non-music control group. Lo et al. (2015) studied 16 experienced adult CI users and randomized them between two types of melodic contour training, one of which manipulated the difficulty using pitch intervals, whereas the other manipulated duration cues. They hypothesized that both types would improve speech perception due to improved prosodic cues and F0 tracking but that the duration group would have additional specific benefit for identification of stop consonants due to better perception of voice onset time and formant transitions. Unfortunately, as there was no non-music control group, the research design only allows inferences about the *differences* between the two types of training (the authors’ second hypothesis), not the overall effect of training (the authors’ first hypothesis). Results showed improvements after training in

Table 1. List of Papers Reviewed With Information About Research Design and Outcome Measures.

	Design	Control group and training	Quality of randomization	Outcome measures
Smith et al. (2017) Adult CI users, N = 21 Self-administered melody training 3.5 hours/week for 4 weeks	No control	NA	NA	Three music assessments: music enjoyment, pitch and timbre discrimination, and complex melody and patterns Sentences in noise and quiet
Hutter et al. (2015) Newly implanted CI adults, N = 12 Music therapy (10 45-minute sessions)	No control	NA	NA	3 questionnaires are sound quality, self-concept, and satisfaction Music (pitch timbre, melody)
Cheng et al. (2018) CI children, N = 22 MCI training (30–60 hours over 8 weeks)	No control	NA	NA	MCI Lexical tones Sentences in quiet
Firestone et al. (2020) CI adults, N = 11 Active music listening 40 mins/day for 4 or 8 weeks	No control	NA	NA	Speech perception (words, sentences in quiet and noise) Hearing questionnaire EEG (acoustic change response)
Lo et al. (2015) CI adults, N = 8 in each group 2 music trainings 1–2 hours/week for 6 weeks	RCT	MCI interval training MCI duration training (No non-music control)	Subjects were randomized between the two music training groups only	MCI Speech perception in noise Consonant discrimination in quiet and in noise Prosody (question/statement)
Yücel et al. (2009) 18 newly implanted child CI users, N = 9 each group 2 years home training on computer/keyboard	CT	No-training control	No randomization (Self-selected test group)	Closed-set word identification and two open-set sentence identification tests Music questionnaires
Lo et al. (2020) 9 hearing-impaired children with hearing aids or CIs 12 weeks of weekly 40-minute group music therapy, plus online exercises	CT	No-training control A subgroup waited 12 weeks before starting training	Pseudorandom allocation: Parents could opt for a different group for convenience	Sentences in noise Pitch Timbre Modulation detection Emotional prosody Question/statement prosody
Petersen et al. (2012) 18 newly implanted CI adults, N = 9 in each group 1 hour/week for 6 months face-to-face training (singing, playing, listening) plus home practice using computer apps	CT	No-training control	No randomization: groups matched on hearing factors.	Music tests (5 subtests) Sentences in noise Emotional prosody

(continued)

Table 1. Continued.

	Design	Control group and training	Quality of randomization	Outcome measures
Dubinsky et al. (2019) Older adults with normal to mild hearing impairment, N = 45 (test), 30 (control) Group 2-hour/week singing course plus 1-hour/week online musical and vocal training exercises for 10 weeks	CT	No-training control	No randomization: self-selected music group with 9 withdrawals	Sentences in noise Frequency difference limens Frequency following response (EEG)
Chari et al. (2020) Adult CI users, N = 7 in two music training groups 1 month of auditory-only MCI 1 month of auditory-motor MCI N = 4 in no-training control	RCT	No-training control	Randomized: but some withdrew and very small groups	Consonant perception, sentences in noise Speech prosody Pitch perception
Fuller et al. (2018) Adult CI users, N = 6, 7, 6 Two types of music training: MCI training Music therapy Six 2-hour weekly sessions	RCT	Group therapy consisting of writing, cooking, and woodworking	Randomized: but very small groups and did not say how randomized	Speech understanding in quiet and noise Vocal emotion identification MCI Quality of life
Good et al. (2017) Child CI users 6 months piano lessons N = 9 in active control and test groups	RCT	Art classes	Partial randomization: also 7/25 withdrew	Music ability Emotional speech prosody
Bedoin et al. (2018) Child CI users, N = 10 Music (rhythmic) primes vs. non-music primed in syntax training Eight 20-minute sessions over 4 weeks	Crossover	Nonmusical auditory primes	NA	Syntactic judgment Grammatical processing

Note. (R)CT = (randomized) controlled trial; MCI = melodic contour identification; CI = cochlear implant.

consonant perception in quiet (but not in noise) and in prosody perception, but not in sentence understanding in noise, and no significant Test \times Group interaction for any measure. Multiple comparisons were not taken into account (only the prosody perception result would have survived Bonferroni correction). Neither of the authors' hypotheses were supported by the data: There was no difference (significant interaction term) between types of music training, and the lack of a non-music control did not allow the overall effects of music training to be deduced. The authors invalidly concluded, however, that both trainings had a benefit for consonant discrimination in quiet and that noise reduced the benefit of the training.

Four papers used a passive "no-training" control group (Dubinsky et al., 2019; Lo et al., 2020; Petersen et al., 2012; Yucel et al., 2009). Again, although any improvements seen in such studies might be due to the training, any difference before and after training or between groups may be due not to the intervention but to expectations of participants or researchers (placebo-type effects), or due to the additional beneficial interactions between trainers and participants that would happen with any training scheme. In addition, none of these four studies randomized participants to control and test groups, making it possible or likely that the two groups differed on important innate characteristics.

Dubinsky et al. (2019) studied older adults with normal hearing or mild hearing impairment. Forty-five participants in a group singing course were recruited for the test group (nine withdrew from the study), and the passive control group was made up of age- and audiometrically-matched other adults. Because the test group participants were self-selected, the possibility of innate differences between groups was high, especially considering that nine withdrew from the study. The 10-week singing course was supplemented by online musical and vocal training exercises. Outcome measures were sentence perception in noise, frequency difference limens, and two EEG measures (FFR amplitude and phase coherence). Group \times Session interactions were significant for speech perception in noise and frequency difference limens and non-significant for FFR amplitude but not phase coherence. The trend for significance seen in the EEG amplitude data seemed to be driven by the unequal baseline EEG amplitudes (i.e., regression to the mean). Although the test group gained more improvement in speech perception in noise than the control group, the use of a passive control and self-selection in the test group makes the interpretation of this result problematic.

Lo et al. (2020) recruited 14 children with moderate to profound hearing loss who used a variety of hearing aids and CIs. Five were assigned to start 12 weeks of music training immediately (but two withdrew before

completing the study), and nine were assigned to start 12 weeks later. Three of the 14 did not have music training, making the final composition of the groups unclear. Although group allocation was "pseudorandom," parents could opt for a different group for convenience. Changes in outcome measures over the first 12 weeks in the wait-list group were used as a passive (no training) control data; however, the main hypothesis testing did not include Group (trained vs. untrained) \times Session interactions. Outcome measures were sentence perception in noise, spectrotemporal modulation detection, emotional prosody, question/statement prosody, and music perception (pitch and timbre subtests). Music training consisted of weekly 40-minute group music therapy sessions (activities such as drumming, singing, dancing) and online exercises three times a week. Within the trained group, sentences in noise, modulation detection, and timbre perception improved compared with baseline, but not emotional prosody or pitch perception, and only sentences in noise and modulation detection remained significantly above baseline 12 weeks after training finished. There are several problems related to statistical analyses in this article, foremost being that no across-group analysis of trained versus untrained children was presented. In addition, the analysis of many outcome measures in a very small group without taking into account multiple comparisons makes the chance of type I error very high. Finally, without an active control group, any actual effect in the trained group may not be due to music training per se.

Petersen et al. (2012) studied 18 newly implanted adult CI users, divided into a test group who undertook 6 months of music training, and a no-training group. The groups were matched on hearing factors and not randomized. It is unclear whether there was any self-selection for the music training arm. The music training consisted of 1-hour/week face-to-face training plus home practice using computer applications. Outcome measures were a music test battery with five subtests, speech perception in noise, and emotional prosody. Group \times Session interactions were calculated for each outcome measure, with three of the five music tests showing greater improvement in the test group compared with control group (one of two music instrument identification tests, rhythm detection and MCI). Gains in pitch ranking, speech perception, and emotional prosody identification were all not different between the test and control groups. Multiple comparisons were not taken into account.

Yucel et al. (2009) studied 18 newly implanted children (mean age of implantation around 4 years) who were assessed preimplantation and over 2 years following implantation. The test group was enrolled in a program that included music training carried out at home with a computer and electronic keyboard, consisting of

pitch and rhythm tasks and color-coded playing of tunes (mean time approximately 2–3 hours per month). The control group was selected from a different research program that did not include music training. No information about how the control group was selected was given. Outcome measures at each test point were speech sound detection, closed-set word identification, and two types of open-set sentence perception tests, along with a parent questionnaire about music perception after 12 and 24 months. In addition, parent questionnaires were administered to assess use of sound in everyday situations. Separate statistics compared speech perception of music and control groups at each of 6 time points between preimplant and 24 months postimplant (a total of 18 tests without control of multiple comparisons), with only one instance of a p value $< .05$ at 3 months postimplant. The preimplant scores (median of zero in all speech tests) were taken as baseline, which is inappropriate because the effect of implantation itself will vary greatly among children. In spite of the flaws in research design, no influence of music training on development of speech perception could be found in this study.

Two further papers (Chari et al., 2020; Fuller et al., 2018) used a design that compared two types of music training with a non-music control group. Chari et al. (2020) tested the hypothesis that auditory-motor music training is better than auditory-alone music training for adult CI users. Subjects were randomly assigned to the three groups. However, it should be noted that (a) with such small numbers (4–7 in each group), randomization is unlikely to make the groups equivalent on confounding characteristics in CI users and (b) two subjects were excluded for “failure to complete the training,” but it was not stated which group(s) these subjects originally belonged to. Home-based training consisted of 30 minutes a day for 5 days per week for 4 weeks (10 hours total). Both music trainings involved training MCI using commercial software. Outcome measures included speech perception in quiet and noise, speech prosody perception, and two musical tests (pitch perception and melodic contour perception). To test the hypothesis, the “change measure” was calculated for each outcome measure, and a one-way analysis of variance was used to test differences in changes across the three groups. No differences were found between groups for any outcome measure except for MCI using the piano tones (a test directly associated with the training applied), where the auditory-motor training group had greater benefit than the auditory-alone group. However, this last result appears doubtful because the auditory-motor group happened to have lower baseline scores than the auditory-alone group on this measure, making the result susceptible to regression to the mean. This study had an appropriate design and analyses for testing

whether auditory-motor was better than auditory-alone music training for the outcome measures, except for the rather small N (15 across three groups), which made the randomization ineffective (especially when participants dropped out), and no specific actions to limit participant or experimenter bias. The study design did not allow the valid testing of the hypothesis that music training per se benefited speech perception, as there was no *active non-music* training control group included. Nevertheless, despite the potential influence of confounds associated with the use of the “no training” control group, the changes in speech perception were not statistically different between the passive control group and both the music groups.

Fuller et al. (2018) evaluated the effects of two different types of music training (pitch/timbre training and music therapy) in adult CI users in an RCT. An active non-musical control training was also included which consisted of writing, cooking, and woodwork classes. Training consisted of weekly 2-hour sessions for 6 weeks. The outcome measures tested were speech perception in quiet and noise and vocal emotion identification, as well as MCI and quality of life. Although the group allocation was randomized, the authors confirmed that the final groups could not be equivalent on average for all relevant features as there were too many variables and low numbers of participants (6 or 7 per group). They reported that there was a non-significant Group \times Session interaction for both the speech perception outcome measures. For vocal emotional identification, there was also a non-significant Group \times Session interaction, indicating no significant difference on training outcome between the three groups. However, they then reported that the music therapy group had a significant within-group training effect, different from the other groups, and invalidly claimed this as evidence of intramodal training (as the music therapy included emotion identification training). The conclusion of this article that “. . . computerized music training or group music therapy may be useful additions to rehabilitation programs for CI users . . .” was therefore not substantiated by the data. This article had a better design than the ones reported earlier, in that they used an active control group, randomized participants, and assessed the Training \times Group interaction terms. However, the very low sample size would make randomization ineffective in CI users, and the choice of active control was not ideal to help limit biases of participants or experimenters. Nevertheless, the results did not support the benefit of music training for speech perception.

Good et al. (2017) tested the hypothesis that music (piano and singing) training was superior to visual art training for development of perception of emotional speech prosody in 18 children with CIs. Students were mostly assigned to two different test locations for the

two trainings, based on geographical preference, with the remainder being randomized. However, 7 out of 25 students recruited dropped out of the program before completion. The authors did not report which group they were originally in. The test group had 6 months (half hour per week, total 12 hours plus weekly practice) of music training, and the control group had the same hours of visual art training. Most children in both groups also participated in school-based musical activities. Outcome measures were musical abilities (Montreal Battery for Evaluation of Musical Abilities: Peretz et al., 2013) and perception of emotional prosody. For musical abilities, the interaction of Session \times Group was significant, showing that the music training group improved musical outcome measures more than the art training group. For emotional prosody tests, there was a non-significant interaction of Group \times Session, showing that music training did not improve emotional prosody perception more than did art training. Unfortunately, in spite of the non-significant interaction term, the authors then proceeded to commit the statistical fallacy of interpreting differences in significance of training effects in individual groups as evidence of differences in benefit of the training methods and incorrectly inferred that “music training improved emotional speech prosody.”

Bedoin et al. (2018) tested the hypothesis that musical (rhythmic) primes (played by percussion instruments) used in morphosyntactic training exercises would improve syntax processing in 10 children with CIs more than the same training using primes of (non-rhythmic) environmental sounds. Primes are stimuli preceding the training stimuli that are intended to draw the attention of the trainee to the relevant features in the training stimuli. Musical primes were rhythmic structures taken from four, 30-second musical sequences. Non-musical primes were 30-second sequences of environmental sounds (street, cafeteria, playground, market). In the grammatical judgment training sessions, 10 sentences were presented (preceded by the primes), and participants were asked to detect and correct morphological errors. In training of morphosyntactic comprehension, participants had to follow the instructions in five sentences. The musical primes were not matched in meter with the sentences used. The authors used a crossover design, and the outcome measures were grammatical processing, syntax processing, non-word repetition, attention, and memory. The two groups for different training order were selected based on “best balance” of age and performance on the two morphosyntactic outcome measures. Each child had eight 20-minute sessions of each training method, two per week. The Training \times Session interaction was significant in favor of the rhythmic primes over environmental primes for grammatical judgment but not the syntactic processing (in contrast to the hypothesis, which was that the syntactic processing would be more

benefited by rhythmic primes than was grammar judgment). For the non-word repetition and all the cognitive tests, the interaction term was not reported (likely because it was not significant—certainly so in the case of non-word repetition, where individual data were presented), and hence the reported within-training analyses of outcome measures cannot be interpreted in terms of which training was better. However, the authors claimed that “. . . musical primes enhanced the processing of training syntactic material, thus enhancing the training effects on grammatical processing as well as phonological processing and sequencing of speech signals.” All of these claims are unsubstantiated by the data: There was no test of a causal relation between syntactic training material processing and grammatical processing, and the claimed benefits to phonological processing and sequencing of speech were unsupported by cross-group analyses. This article did not assess the effect of music training per se but rather the use of rhythmic primes to improve syntax training. There was no mention of blinding of participants or researchers, and the interaction of training order and training method was not checked before orders were combined for analysis.

Overall, it is notable that none of the 13 papers, including those with better research designs (use of active controls and randomization), mentioned attempts to limit bias of participants, trainers, or testers. In addition, the non-music active control training choices (visual art, woodworking, cooking, and writing) were likely to introduce bias via higher expectations of participants, trainers, and testers for music training. In spite of these design limitations, none of the 13 papers produced statistically valid evidence to support the specific hypothesis that music training improves speech perception in hearing-impaired listeners. Of the papers that included valid analyses of their data, none found that music training improved speech perception: There was no evidence that music training was better than visual art training (Good et al., 2017) or writing/cooking/woodwork training (Fuller et al., 2018), no evidence that auditory-motor music training was better than auditory-alone music training (Chari et al., 2020), and no evidence that rhythmic primes improved syntactic processing in speech or non-word repetition more than did environmental primes (Bedoin et al., 2018).

Discussion

The review has found no evidence to support the hypothesis that music training has a significant causal effect on speech understanding or speech processing in hearing-impaired populations. In fact, the papers with a higher quality research design showed no significant benefit when valid statistics were used. Although this null result may be contributed to by insufficient

statistical power in the studies reviewed, an alternative interpretation is that music training does not transfer to speech perception benefits in hearing-impaired people to any clinically relevant degree. Therefore, either music training is ineffective for improvement of speech understanding in general, or the limitations imposed by hearing loss or listening with a CI cannot be overcome using music training, or both of these proposals are true. The first proposal is supported by the studies reviewed in the introduction that show that innate characteristics (such as IQ, musical aptitude, and personality) predict musicianship and independently predict better speech perception (Schellenberg, 2015; Swaminathan & Schellenberg, 2018a, 2018b). The distinction between plastic effects of training and innate characteristics of the person being trained is particularly important when music training is being proposed as a therapy in clinical populations and where the hypothesis is solely about exploiting plasticity induced by the training. In that case it is extremely important that the experiment (which is essentially a clinical trial) is specifically designed to limit any genetic or innate differences in participants in the comparison groups, for example, by using careful randomization or a crossover design. A comparison of children who choose to take music lessons, or who engage more in music lessons, with those that do not, is just as much a test of genetic differences as it is a test of influence of music training.

All but two of the papers reviewed addressed questions in the profoundly deaf population who use CIs. All of those studies except Bedoin et al. (2018) based their hypotheses at least partially on the proposal that improving (musical) pitch perception would translate to speech perception benefits via better perception of voice pitch (F0). CIs are unable to convey complex pitch (such as musical pitch or voice pitch) to the degree or salience that people with normal hearing can hear pitch (Fielden et al., 2015; McDermott & McKay, 1997). The design of CIs means that fine details about harmonic structure in speech and music (such as precise frequency and spacing) are not represented, and resolved harmonics are what the normal-hearing system processes to extract salient pitch from sounds. A very weak pitch can be heard related to periodicity in the electrical signal (McKay et al., 1994); however, many studies have shown that this periodicity pitch cannot be reliably heard in real-life situations, as the modulations in different electrical channels can interact with each other (McDermott & McKay, 1997; McKay & Carlyon, 1999; McKay & McDermott, 1996). Many signal processing strategies have attempted to improve the transmission of complex pitch for CI users, but the results have been equivocal or unsuccessful (Wouters et al., 2015). The most successful way to transmit complex pitch to a CI user may be via concurrent use of any low-frequency residual acoustic

hearing that the person may have (Chen et al., 2014; Straatman et al., 2010; Visram, Azadpour, et al., 2012; Visram, Kluk, et al., 2012). Because the difficulty of representing pitch in CIs is related to inbuilt limitations of how a CI works, it must be doubtful whether pitch training can have a beneficial effect on pitch perception in real-life listening, especially one that can be transferred to speech understanding, or the perception of emotional prosody. In addition, most implant systems do not transmit modulations above 300 Hz, making F0 of high female voices and children's voices, and the F0 of many musical notes, unrepresented in the modulations of the signal (Wouters et al., 2015). To add to the difficulty of the concept that music training can transfer to better speech understanding in CI users, many studies have shown that, even though musicians may have better pitch discrimination than non-musicians, better pitch discrimination does not necessarily translate to better speech perception in noise, or better use of voice pitch to distinguish between two simultaneous talkers.

Six of the papers hypothesized that music training would improve the perception of speech prosody (emotion detection and/or differentiation between questions and statements) by CI users based on the fact that F0 provides cues for prosody in people with normal hearing. However, in everyday conversations, there are concurrent alternative or correlated cues to prosody such as changes in intensity, duration, or timbre (Coutinho & Dibben, 2013) that are more reliably perceived than voice pitch by a CI user. Therefore, direct training in prosody perception is likely to be a more fruitful way than pitch training of improving this feature of speech perception by CI users.

The Bedoin et al. (2018) paper investigated whether the rhythm features of music could translate to better processing of speech for CI users. This idea follows a series of papers in which it has been shown that preceding a spoken utterance with a rhythmic prime that matches the meter (stress pattern and number of units) of the utterance leads to improved, or faster, processing of the speech, due to the expectations set up by the prime, compared with a prime that does not match the utterance (Cason, Astesano, et al., 2015; Cason & Schon, 2012). That is, knowing ahead of time the meter structure of the sentence you are going to hear helps to process the sentence better. This effect is not surprising, as the matched prime is providing useful cues to the following sentence or multisyllable utterance and focusses attention on the stressed syllables in the utterance. Similarly, matched primes improved the speech production of a following heard sentence in hearing-impaired children (Cason, Hidalgo, et al., 2015). However, in the Bedoin et al. paper, the rhythmic primes used in the training sessions were not related to the training material that immediately followed the prime, so it is unclear how

the rhythmic prime could improve the within-session effectiveness of the training more than the non-rhythmic one (or with no prime). In addition, no paper has reported that the effect of using primes has any longer-term effect than that on the immediately following utterance.

This review has highlighted the difficulty in undertaking high-quality research in the hearing-impaired population to investigate potential benefits of music training. First, there is the challenge of participant and experimenter bias, given the popular community view of music benefit for “brain training,” as promulgated in the media. Experimenters are not immune to this bias, as evidenced by the unsubstantiated claims of benefit made in the majority of the papers reviewed. Confirmation bias, where researchers only report data or analyses that support their original belief and ignore data or analyses that do not fit, seems ubiquitous. This same point was also strongly made by the authors of the large meta-analysis of music training for cognitive benefit in children (Sala & Gobet, 2020) who state, “We conclude that researchers’ optimism about the benefits of music training is empirically unjustified and stems from misinterpretation of the empirical data and, possibly, confirmation bias.” Along similar lines, a study of whether papers about music training inferred causality from correlation found that 72/109 of the papers reviewed invalidly made this inference, a ratio that increased up to 81% for papers written by neuroscientists (Schellenberg, 2020). If research studies are undertaken with the mindset of “demonstrating a known fact” instead of testing a hypothesis, it is no surprise that researchers are tempted to find whatever they can in the data to present to the reader as supporting the “fact.”

The second challenge faced by experimenters is to find a way to ensure that test and control groups are truly equivalent on all relevant factors. Randomization will only be effective if large numbers of participants are used, and few people withdraw from the program. For small groups of profoundly deaf children or adults, this task is virtually impossible. Trying to match groups poses the same problem and opens the opportunity for bias of group assignment. A crossover research design would seem a better choice in this population, provided effects of test order can be carefully controlled.

A different experimental approach using correlational studies has been proposed by Swaminathan and Schellenberg (2018b), in which effects of training and musical aptitude (and other potential confounding factors) are partialled out. When examining the association between music training and non-music abilities such as speech understanding, or cognitive abilities, the association is tested using partial correlations with multiple known factors that account for overlapping variance. In particular, they use a music aptitude test (which is

correlated with cognitive abilities and predicts success in music training) as one such factor. For example, when music training was kept constant, an association between music aptitude and phoneme perception in adults was found (Swaminathan & Schellenberg, 2017); however, when musical aptitude was kept constant, no association between music training and speech perception or intelligence remained. The authors concluded that *pre-existing* differences in cognitive ability and musical aptitude predict whether a person will undergo music training, and hence those differences contribute to the association between music training and non-music abilities. Similarly, a study in children by the same authors using the same partial correlation methods (Swaminathan & Schellenberg, 2019) concluded that musical aptitude predicts language ability independently from IQ and other confounds and that the association between musicianship and language arises from pre-existing factors rather than formal training.

The finding in this review that there is no evidence that music training benefits speech perception in hearing-impaired people does not detract from the evident social, cultural, and enjoyment benefits that music education can confer. Music is an important feature of human cultural and social life, with the large majority of people enjoying listening to music even if they do not perform music or undergo music education. The cultural relevance of music is particularly relevant for the hearing-impaired population, who may be unfairly discouraged from participating in musical activities. It is important that equity of access to music education is sought and engagement in music making is encouraged for all young people with hearing impairment. There is no necessity to promote music education on grounds of a supposed benefit for academic or speech and language benefit, as it has values of its own that make it worth including in education for everyone, including those with hearing impairment.

Conclusions and Recommendations

It is clear that the benefits of music training for speech perception in the hearing-impaired population have not been convincingly demonstrated and probably do not exist to any practically relevant degree. Future researchers who wish to scientifically study whether music training can lead to benefits in non-music domains of cognition such as language perception or language development need to plan their studies to robustly and validly test their hypotheses and to actively distinguish between plasticity effects of training and innate characteristics.


Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Colette M. McKay  <https://orcid.org/0000-0002-1659-9789>

References

- Baskent, D., Fuller, C. D., Galvin, J. J., 3rd, Schepel, L., Gaudrain, E., & Free, R. H. (2018). Musician effect on perception of spectro-temporally degraded speech, vocal emotion, and music in young adolescents. *Journal of the Acoustical Society of America*, *143*(5), EL311. <https://doi.org/10.1121/1.5034489>
- Bedoin, N., Besombes, A. M., Escande, E., Dumont, A., Lalitte, P., & Tillmann, B. (2018). Boosting syntax training with temporally regular musical primes in children with cochlear implants. *Annals of Physical and Rehabilitation Medicine*, *61*(6), 365–371. <https://doi.org/10.1016/j.rehab.2017.03.004>
- Blamey, P., Artieres, F., Baskent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B., Gallego, S., Govaerts, P. J., Green, K., Huber, A. M., Kleine-Punte, A., Maat, B., Marx, M., Mawman, D., Mosnier, I., O'Connor, A. F., Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology and Neuro-Otology*, *18*(1), 36–47. <https://doi.org/10.1159/000343189>
- Blamey, P.J., Sarant, J.Z., Paatsch, L.E., Barry, J.G., Bow, C. P., Wales, R.J., Wright, M., Psarros, C., Rattigan, K. and Tooher, R., (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264–285. [https://doi.org/10.1044/1092-4388\(2001\)022](https://doi.org/10.1044/1092-4388(2001)022)
- Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *Journal of the Acoustical Society of America*, *137*(1), 378–387. <https://doi.org/10.1121/1.4904537>
- Bulut, M., & Narayanan, S. (2008). On the robustness of overall F0-only modifications to the perception of emotions in speech. *Journal of the Acoustical Society of America*, *123*(6), 4547–4558. <https://doi.org/10.1121/1.2909562>
- Cason, N., Astesano, C., & Schon, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica*, *155*, 43–50. <https://doi.org/10.1016/j.actpsy.2014.12.002>
- Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schon, D. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, *29*(1), 102–107. <https://doi.org/10.1037/neu0000115>
- Cason, N., & Schon, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, *50*(11), 2652–2658. <https://doi.org/10.1016/j.neuropsychologia.2012.07.018>
- Chari, D. A., Barrett, K. C., Patel, A. D., Colgrove, T. R., Jiradejvong, P., Jacobs, L. Y., & Limb, C. J. (2020). Impact of auditory-motor musical training on melodic pattern recognition in cochlear implant users. *Otology & Neurotology*, *41*(4), e422–e431. <https://doi.org/10.1097/MAO.0000000000002525>
- Chen, J. K., Chuang, A. Y., McMahan, C., Tung, T. H., & Li, L. P. (2014). Contribution of nonimplanted ear to pitch perception for prelingually deafened cochlear implant recipients. *Otology & Neurotology*, *35*(8), 1409–1414. <https://doi.org/10.1097/MAO.0000000000000407>
- Cheng, X., Liu, Y., Shu, Y., Tao, D. D., Wang, B., Yuan, Y., Galvin III, J. J., Fu, Q. J., & Chen, B. (2018). Music training can improve music and speech perception in pediatric mandarin-speaking cochlear implant users. *Trends in Hearing*, *22*, 2331216518759214. <https://doi.org/10.1177/2331216518759214>
- Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research*, *352*, 49–69. <https://doi.org/10.1016/j.heares.2017.02.006>
- Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition and Emotion*, *27*(4), 658–684. <https://doi.org/10.1080/02699931.2012.732559>
- Darwin, C. J., Brungart, D. S., & Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *Journal of the Acoustical Society of America*, *114*(5), 2913–2922. <https://doi.org/10.1121/1.1616924>
- Dubinsky, E., Wood, E. A., Nespole, G., & Russo, F. A. (2019). Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults with age-related hearing loss. *Frontiers in Neuroscience*, *13*, 1153. <https://doi.org/10.3389/fnins.2019.01153>
- Fielden, C. A., Kluk, K., Boyle, P. J., & McKay, C. M. (2015). The perception of complex pitch in cochlear implants: A comparison of monopolar and tripolar stimulation. *Journal of the Acoustical Society of America*, *138*(4), 2524–2536. <https://doi.org/10.1121/1.4931910>
- Firestone, G. M., McGuire, K., Liang, C., Zhang, N., Blankenship, C. M., Xiang, J., & Zhang, F. (2020). A preliminary study of the effects of attentive music listening on cochlear implant users' speech perception, quality of life, and behavioral and objective measures of frequency change detection. *Frontiers in Human Neuroscience*, *14*, 110. <https://doi.org/10.3389/fnhum.2020.00110>
- Fuller, C. D., Galvin, J. J., 3rd, Maat, B., Baskent, D., & Free, R. H. (2018). Comparison of two music training approaches on music and speech perception in cochlear implant users. *Trends in Hearing*, *22*, 2331216518765379. <https://doi.org/10.1177/2331216518765379>

- George, B. J., Beasley, T. M., Brown, A. W., Dawson, J., Dimova, R., Divers, J., Goldsby, T. U., Heo, M., Kaiser, K. A., Keith, S. W., & Kim, M. Y. (2016). Common scientific and statistical errors in obesity research. *Obesity, 24*(4), 781–790. <https://doi.org/10.1002/oby.21449>
- Good, A., Gordon, K. A., Papsin, B. C., Nespoli, G., Hopyan, T., Peretz, I., & Russo, F. A. (2017). Benefits of music training for perception of emotional speech prosody in deaf-children with cochlear implants. *Ear and Hearing, 38*(4), 455–464. <https://doi.org/10.1097/AUD.0000000000000402>
- Hambrick, D. Z., & Tucker-Drob, E. M. (2015). The genetics of music accomplishment: Evidence for gene-environment correlation and interaction. *Psychonomic Bulletin & Review, 22*(1), 112–120. <https://doi.org/10.3758/s13423-014-0671-9>
- Hutter, E., Argstatter, H., Grapp, M., & Plinkert, P. K. (2015). Music therapy as specific and complementary training for adults after cochlear implantation: A pilot study. *Cochlear Implants International, 16*(Suppl 3), S13–S21. <https://doi.org/10.1179/1467010015Z.00000000000261>
- Kral, A., Hartmann, R., Tillein, J., Heid, S., & Klinke, R. (2002). Hearing after congenital deafness: Central auditory plasticity and sensory deprivation. *Cerebral Cortex, 12*(8), 797–807. <https://doi.org/10.1093/cercor/12.8.797>
- Kraus, N., Hornickel, J., Strait, D. L., Slater, J., & Thompson, E. (2014). Engagement in community music classes sparks neuroplasticity and language development in children from disadvantaged backgrounds. *Frontiers in Psychology, 5*, 1403. <https://doi.org/10.3389/fpsyg.2014.01403>
- Lazard, D. S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P. H., Skarzynski, H., Schauwers, K., O’Leary, S., Mawman, D., Maat, B., Kleine-Punte, A., Huber, A. M., Green, K., Govaerts, P. J., Fraysse, B., Dowell, R., Dillier, N., ... Blamey, P. J. (2012). Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: A new conceptual model over time. *PLoS One, 7*(11), e48739. <https://doi.org/10.1371/journal.pone.0048739>
- Lee, J. S., Lee, D. S., Oh, S. H., Kim, C. S., Kim, J. W., Hwang, C. H., Koo, J., Kang, E., Chung, J. K., & Lee, M. C. (2003). PET evidence of neuroplasticity in adult auditory cortex of postlingual deafness. *Journal of Nuclear Medicine, 44*(9), 1435–1439.
- Lo, C. Y., Looi, V., Thompson, W. F., & McMahan, C. M. (2020). Music training for children with sensorineural hearing loss improves speech-in-noise perception. *Journal of Speech, Language, and Hearing Research, 63*(6), 1990–2015. https://doi.org/10.1044/2020_JSLHR-19-00391
- Lo, C. Y., McMahan, C. M., Looi, V., & Thompson, W. F. (2015). Melodic contour training and its effect on speech in noise, consonant discrimination, and prosody perception for cochlear implant recipients. *Behavioural Neurology, 2015*, 352869. <https://doi.org/10.1155/2015/352869>
- Madsen, S. M. K., Marschall, M., Dau, T., & Oxenham, A. J. (2019). Speech perception is similar for musicians and non-musicians across a wide range of conditions. *Scientific Reports, 9*(1), 10404. <https://doi.org/10.1038/s41598-019-46728-1>
- Madsen, S. M. K., Whiteford, K. L., & Oxenham, A. J. (2017). Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Scientific Reports, 7*(1), 12624. <https://doi.org/10.1038/s41598-017-12937-9>
- Martinez-Montes, E., Hernandez-Perez, H., Chobert, J., Morgado-Rodriguez, L., Suarez-Murias, C., Valdes-Sosa, P. A., & Besson, M. (2013). Musical expertise and foreign speech perception. *Frontiers in Systems Neuroscience, 7*, 84. <https://doi.org/10.3389/fnsys.2013.00084>
- McDermott, H. J., & McKay, C. M. (1997). Musical pitch perception with electrical stimulation of the cochlea. *Journal of the Acoustical Society of America, 101*(3), 1622–1631. <https://doi.org/10.1121/1.418177>
- McKay, C. M., & Carlyon, R. P. (1999). Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains. *Journal of the Acoustical Society of America, 105*(1), 347–357. <https://doi.org/10.1121/1.424553>
- McKay, C. M., & McDermott, H. J. (1996). The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *Journal of the Acoustical Society of America, 100*(2 Pt 1), 1081–1092. <https://doi.org/10.1121/1.416294>
- McKay, C. M., McDermott, H. J., & Clark, G. M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *Journal of the Acoustical Society of America, 96*(5 Pt 1), 2664–2673. <https://doi.org/10.1121/1.411377>
- Melby-Lervag, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of “Far Transfer”: Evidence from a meta-analytic review. *Perspectives on Psychological Science, 11*(4), 512–534. <https://doi.org/10.1177/17456916166635612>
- Moore, B. C. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear and Hearing, 17*(2), 133–161. <https://doi.org/10.1097/00003446-199604000-00007>
- Mosing, M. A., Madison, G., Pedersen, N. L., & Ullen, F. (2016). Investigating cognitive transfer within the framework of music practice: Genetic pleiotropy rather than causality. *Developmental Science, 19*(3), 504–512. <https://doi.org/10.1111/desc.12306>
- Oikkonen, J., Huang, Y., Onkamo, P., Ukkola-Vuoti, L., Rajjas, P., Karma, K., Vieland, V. J., & Järvelä, I. (2015). A genome-wide linkage and association study of musical aptitude identifies loci containing genes related to inner ear development and neurocognitive functions. *Molecular Psychiatry, 20*(2), 275–282. <https://doi.org/10.1038/mp.2014.8>
- Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research, 308*, 98–108. <https://doi.org/10.1016/j.heares.2013.08.011>
- Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Beland, R. (2013). A novel tool for evaluating children’s musical abilities across age and culture. *Frontiers in Systems Neuroscience, 7*, 30. <https://doi.org/10.3389/fnsys.2013.00030>
- Petersen, B., Mortensen, M. V., Hansen, M., & Vuust, P. (2012). Singing in the key of life: A study on effects of

- musical ear training after cochlear implantation. *Psychomusicology*, 22(2), 134–151. <https://doi.org/10.1037/a0031140>
- Rosenthal, M. A. (2020). A systematic review of the voice-tagging hypothesis of speech-in-noise perception. *Neuropsychologia*, 136, 107256. <https://doi.org/10.1016/j.neuropsychologia.2019.107256>
- 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>
- Sala, G., & Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, 48, 1429–1441. <https://doi.org/10.3758/s13421-020-01060-2>
- Sares, A. G., Foster, N. E. V., Allen, K., & Hyde, K. L. (2018). Pitch and time processing in speech and tones: The effects of musical training and attention. *Journal of Speech, Language, and Hearing Research*, 61(3), 496–509. https://doi.org/10.1044/2017_JSLHR-S-17-0207
- Schellenberg, E. G. (2015). Music training and speech perception: A gene-environment interaction. *Annals of the New York Academy of Sciences*, 1337, 170–177. <https://doi.org/10.1111/nyas.12627>
- Schellenberg, E. G. (2020). Correlation = Causation? Music training, psychology, and neuroscience. *Psychology of Aesthetics, Creativity, and the Arts*, 14(4), 475–480. <https://doi.org/10.1037/aca0000263>
- Sharma, A., Campbell, J., & Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: Evidence from the P1 and N1 event related potentials in cochlear implanted children. *International Journal of Psychophysiology*, 95(2), 135–144. <https://doi.org/10.1016/j.ijpsycho.2014.04.007>
- Smith, L., Bartel, L., Joglekar, S., & Chen, J. (2017). Musical rehabilitation in adult cochlear implant recipients with a self-administered software. *Otology & Neurotology*, 38(8), e262–e267. <https://doi.org/10.1097/MAO.0000000000001447>
- Sommer, A., & Zeger, S. L. (1991). On estimating efficacy from clinical trials. *Statistics in Medicine*, 10(1), 45–52. <https://doi.org/10.1002/sim.4780100110>
- Straatman, L. V., Rietveld, A. C., Beijen, J., Mylanus, E. A., & Mens, L. H. (2010). Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. *Journal of the Acoustical Society of America*, 128(4), 1884–1895. <https://doi.org/10.1121/1.3474236>
- Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin & Review*, 24(6), 1929–1934. <https://doi.org/10.3758/s13423-017-1244-5>
- Swaminathan, S., & Schellenberg, E. G. (2018a). Musical competence is predicted by music training, cognitive abilities, and personality. *Scientific Reports*, 8(1), 9223. <https://doi.org/10.1038/s41598-018-27571-2>
- Swaminathan, S., & Schellenberg, E. G. (Eds.). (2018b). *Music training and cognitive abilities: Associations, causes, and consequences*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198804123.013.26>
- Swaminathan, S., & Schellenberg, E. G. (2019). Musical ability, music training, and language ability in childhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. <https://doi.org/10.1037/xlm0000798>
- Tierney, A., Rosen, S., & Dick, F. (2020). Speech-in-speech perception, nonverbal selective attention, and musical training. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(5), 968–979. <https://doi.org/10.1037/xlm0000767>
- Vanden Bosch der Nederlanden, C. M., Zaragoza, C., Rubio-Garcia, A., Clarkson, E., & Snyder, J. S. (2020). Change detection in complex auditory scenes is predicted by auditory memory, pitch perception, and years of musical training. *Psychological Research*, 84(3), 585–601. <https://doi.org/10.1007/s00426-018-1072-x>
- Visram, A. S., Azadpour, M., Kluk, K., & McKay, C. M. (2012). Beneficial acoustic speech cues for cochlear implant users with residual acoustic hearing. *Journal of the Acoustical Society of America*, 131(5), 4042–4050. <https://doi.org/10.1121/1.3699191>
- Visram, A. S., Kluk, K., & McKay, C. M. (2012). Voice gender differences and separation of simultaneous talkers in cochlear implant users with residual hearing. *Journal of the Acoustical Society of America*, 132(2), EL135–EL141. <https://doi.org/10.1121/1.4737137>
- Wouters, J., McDermott, H. J., & Francart, T. (2015). Sound coding in cochlear implants. *IEEE Signal Processing Magazine*, 32(2), 67–80. <https://doi.org/10.1109/Msp.2014.2371671>
- Wright, B. A., & Zhang, Y. (2009). A review of the generalization of auditory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1515), 301–311. <https://doi.org/10.1098/rstb.2008.0262>
- Yucel, E., Sennaroglu, G., & Belgin, E. (2009). The family oriented musical training for children with cochlear implants: Speech and musical perception results of two year follow-up. *International Journal of Pediatric Otorhinolaryngology*, 73(7), 1043–1052. <https://doi.org/10.1016/j.ijporl.2009.04.009>