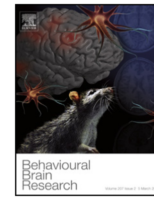
Behavioural Brain Research 291 (2015) 244–252

Contents lists available at ScienceDirect

Behavioural Brain Research

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



Research report

Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program

Jessica Slater[a,b], Erika Skoe[a,1], Dana L. Strait[a,2], Samantha O'Connell[a,3], Elaine Thompson[a,b], Nina Kraus[a,b,c,d,*]

[a] Auditory Neuroscience Laboratory, Northwestern University, Evanston, IL, USA
[b] Department of Communication Sciences, Northwestern University, Evanston, IL, USA
[c] Department of Neurobiology and Physiology, Northwestern University, Evanston, IL, USA
[d] Department of Otolaryngology, Northwestern University, Evanston, IL, USA

HIGHLIGHTS

- Longitudinal evidence of improved speech-in-noise perception with musical training.
- Random-assignment study assesses group instruction in established music program.
- Speech-in-noise perception improved in low-income, bilingual population.

ARTICLE INFO

Article history:
Received 12 January 2015
Received in revised form 9 May 2015
Accepted 13 May 2015
Available online 22 May 2015

Keywords:
Learning
Music
Speech-in-noise perception
Longitudinal
Education
Auditory
Listening

ABSTRACT

Music training may strengthen auditory skills that help children not only in musical performance but in everyday communication. Comparisons of musicians and non-musicians across the lifespan have provided some evidence for a "musician advantage" in understanding speech in noise, although reports have been mixed. Controlled longitudinal studies are essential to disentangle effects of training from pre-existing differences, and to determine how much music training is necessary to confer benefits. We followed a cohort of elementary school children for 2 years, assessing their ability to perceive speech in noise before and after musical training. After the initial assessment, participants were randomly assigned to one of two groups: one group began music training right away and completed 2 years of training, while the second group waited a year and then received 1 year of music training. Outcomes provide the first longitudinal evidence that speech-in-noise perception improves after 2 years of group music training. The children were enrolled in an established and successful community-based music program and followed the standard curriculum, therefore these findings provide an important link between laboratory-based research and real-world assessment of the impact of music training on everyday communication skills.

1. Introduction

Everyday communication rarely occurs in ideal conditions: from busy restaurants to noisy classrooms, the human auditory system is constantly faced with the challenge of picking out a meaningful signal from competing inputs. Understanding speech in noise not only presents an everyday communication challenge, it also provides an informative measure of auditory function under limiting conditions, since comprehension in noise requires the successful integration of cognitive, linguistic and sensory processing in response to novel incoming sounds. There is evidence that speech-in-noise perception can be improved with computer-based auditory training [1–3], and cross-sectional studies have indicated that auditory experts, specifically musicians, outperform age-matched peers in this task when matched for factors such as IQ

* Corresponding author at: Northwestern University, 2240 Campus Drive, Evanston, IL, 60208, USA. Tel.: +1 847 491 3181; fax: +1 847 491 2523.
E-mail address: nkraus@northwestern.edu (N. Kraus).
URLs: http://www.brainvolts.northwestern.edu (J. Slater), http://www.brainvolts.northwestern.edu (E. Skoe), http://www.brainvolts.northwestern.edu (D.L. Strait), http://www.brainvolts.northwestern.edu (S. O'Connell), http://www.brainvolts.northwestern.edu (E. Thompson), http://www.brainvolts.northwestern.edu (N. Kraus).
1 Present address: Department of Speech, Language and Hearing Sciences, Department of Psychology, University of Connecticut, Storrs, CT, USA.
2 Present address: Neural Systems Laboratory, Institute for Systems Research, University of Maryland, College Park, MD, USA.
3 Present address: Center for Cognitive Neuroscience, University of California, Los Angeles, CA, USA.

http://dx.doi.org/10.1016/j.bbr.2015.05.026
0166-4328/

and hearing thresholds [4–10], although recent studies attempting to replicate this advantage have reported mixed outcomes [11–13]. Further understanding of how this skill can be improved with training may provide important insights into the malleability of complex auditory processing, as well as informing educators and clinicians interested in the development of communication skills.

The extent to which the specific auditory skills developed through musical experience may transfer to non-musical domains is a matter of continuing debate [14–18], however numerous components of auditory processing that support the perception of speech in noise have been found to be strengthened in musicians, including syllable discrimination [19–21] and the processing of temporal speech cues [22–24], prosody [25], pitch [26–30] and melodic contour [31]. Further, musicians demonstrate enhanced auditory cognitive function such as working memory [32–34] and attention [6,35,36], as well as enhanced neural representation of speech when presented in acoustically-compromised conditions [6,8,9,37–40]. These findings are consistent with the theoretical framework proposed by Patel suggesting that music training promotes adaptive plasticity in speech-processing networks [41].

In the Parbery-Clark et al. (2009) study, which first reported the musician advantage, better speech perception was observed in both multi-talker babble and continuous speech-shaped noise [4]. The advantage in speech-shaped noise was observed specifically in the most challenging condition, where both speech and masker originated from the same source, and not in conditions where the speech and masker were spatially separated. Other studies aimed to differentiate the potential factors contributing to this observed advantage by using speech stimuli that were modified in various ways (e.g. whispered speech to reduce periodic content [12] or sine-wave vocoded speech to simulate perception through a cochlear implant [13]) and presented under a variety of different masking conditions (e.g. maskers varying in informational content [10,11]). Swaminathan et al. [10] replicated the musician advantage with co-located energetic masking, consistent with Parbery-Clark et al., and further identified that when the masker was intelligible, musicians showed a greater spatial release from masking than non-musicians [10]. Fuller et al. found that musicians outperformed non-musicians in one out of the eight masking conditions that were assessed, specifically the most challenging condition [13]. Zendel et al. found that French-speaking musicians performed better on a words-in-noise task than non-musicians, again in the most difficult listening condition, and this performance advantage was associated with more resilient cortical responses in noise [9]. However, neither Ruggles et al. (2014) nor Boebinger et al. (2015) replicated a musician advantage under any of their masking conditions [11,12], although Ruggles and colleagues observed a significant correlation between years of musical practice and performance on the same clinical measures of speech-in-noise perception that were assessed in the original study (i.e. HINT and QuickSIN).

There are a number of differences in methodology and participant characteristics that may have contributed to divergent outcomes. While all of the recent studies adopted similar musicianship inclusion criteria, there were some differences in average age of onset of training, which was lowest in the Parbery-Clark et al. [4] study (mean = 4.68 years). The Boebinger et al. (2015) study included participants across a much wider age range and did not perform audiometric screening, which raises concern about the possibility of noise exposure and undiagnosed hearing loss within the sample, especially within the musician group (e.g. several participants were reported to play drums).

The mixed outcomes highlight that the influence of musical expertise on speech perception in noise may vary according to the characteristics of the speech and masker, as well as the presentation conditions. Ruggles et al. (2014) is the only study that used the same speech-in-noise perception measures as the original Parbery-Clark et al. [4] study (i.e. HINT and QuickSIN), however Ruggles and colleagues administered the HINT (sentences in speech-shaped noise) with headphones (diotic), whereas Parbery-Clark et al. presented the test in sound field (binaural). Although the target and masker were co-located, musicians may still have gained a greater benefit in this more difficult condition by using binaural cues to improve the perception of acoustic details.

There have been divergent interpretations regarding the importance of cognitive vs. sensory factors: Interestingly, Ruggles et al. reported no significant correlation between speech-in-noise perception and IQ, and Parbery-Clark et al.'s outcomes were observed despite the lack of group differences in IQ, however Boebinger et al. reported that speech-in-noise performance is significantly predicted by non-verbal IQ. Fuller et al. [13] suggest that musician benefits are "mainly due to better processing of low-level acoustic cues" and not cognitive factors, while Boebinger et al. emphasize the importance of considering general cognitive abilities, given the relationship with IQ in their data. These mixed experimental outcomes also highlight some of the inherent limitations of cross-sectional comparisons and further demonstrate the importance of longitudinal assessments to determine the impact of musical training on speech-in-noise perception within individual subjects.

Research assessing the impact of musical skill on more general aspects of auditory and cognitive processing has also yielded mixed results, for example a recent study did not find any difference between musicians and non-musicians in multi-modal sequencing or auditory scene analysis, and the authors emphasize the importance of task context as a factor that may influence the transfer of musical skills to non-musical domains [18]. The complex processing demands of both speech and music may point to similarities that are important for transfer between these domains. As with spoken language, musical communication relies not only on the ability to detect and process specific acoustic cues, but on the ability to integrate these components into *meaningful* sounds through the engagement of cognitive, sensory and emotional brain circuitry. This integrated neural activation across multiple brain areas may help to explain why musical expertise has been associated with neural and perceptual advantages not only for music but for other forms of communication, such as speech [see 41 for review].

In this study we performed a longitudinal investigation of the effect of group music instruction on performance on a standard clinical measure of speech-in-noise perception. In contrast to previous studies in which music training regimens have been initiated by researchers specifically for the purposes of an experimental study, the present study uses a rigorous scientific approach to assess the impact of an established and successful music program, and therefore takes an important step in bridging the gap between laboratory and real-world application. All study participants received the standard curriculum of musical training provided by Harmony Project, a non-profit organization offering free music education to children in the gang reduction zones of Los Angeles. Harmony Project has provided music education to underserved children in the Los Angeles area for more than 10 years, garnering significant national acclaim as well as impressive musical and academic outcomes. We hypothesized that music training improves the ability to process novel soundscapes and extract meaningful information from competing auditory streams, and that this transfers to non-musical communication contexts. Specifically, we hypothesized that music training engages and strengthens neural circuitry that is important for speech perception, consistent with Patel's theoretical framework [41]. This is supported by cross-sectional evidence of enhanced neural encoding of speech in musically-trained children [21,42,43] as well as a small number of longitudinal studies showing effects of music training on speech processing

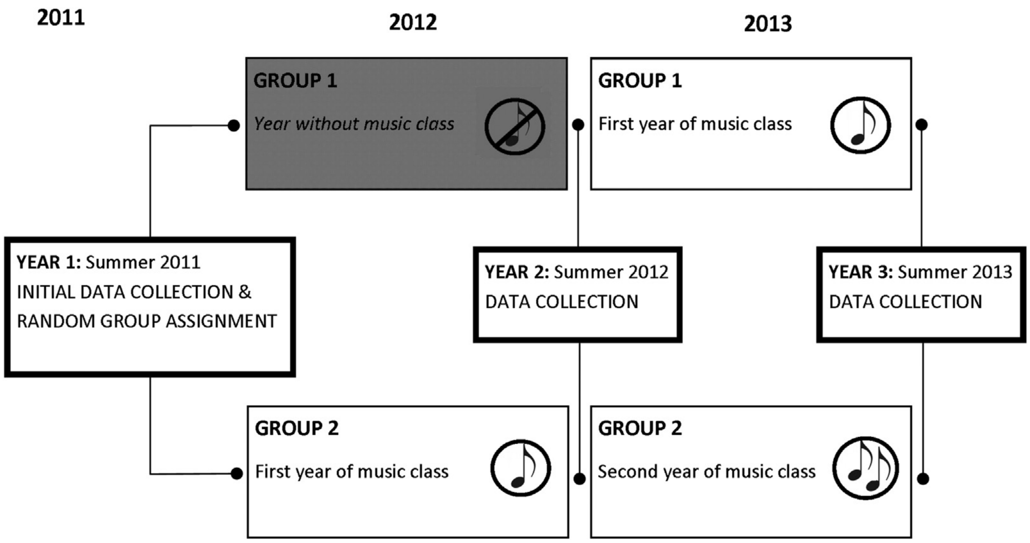


Fig. 1. Schematic of the experimental design.

[44–49]. However, the time course of musical training's impact on speech-in-noise perception in children has not previously been assessed. In line with three recent longitudinal studies demonstrating enhanced neural processing of speech after 2 academic years of music training, but not after a single year [47–49], we predicted significant improvements in speech-in-noise perception with 2 years of musical training.

2. Materials and methods

2.1. Experimental design

Data collection spanned 3 years, with one of the two experimental groups serving as a control group for the first full year of the study (an overview of the experimental design is provided in Fig. 1). Researchers from Northwestern University traveled to Los Angeles, California, for 3 weeks in June–July of 2011. Following this initial assessment, children were randomly assigned by the research team to one of two groups. Group 1 served as the control group during the first year of the study and did not begin musical training until August 2012, while Group 2 began training immediately in August 2011. The research team returned the following two summers to collect another round of data. By the third testing session in the summer of 2013, Group 1 had completed 1 year of musical training whereas Group 2 had completed 2 years of training. This experimental design allowed between-subject and within-subject assessments of the effects of 1 year of musical training, providing a baseline of maturational changes in the control group during the first year of the study, and the assessment of 2 years of musical training in Group 2.

2.2. Participants

Participants were recruited from the waitlist of Harmony Project and from local elementary schools with active Harmony Project programs, with participation in the research study ensuring a place in the Harmony Project program either immediately following the first assessment or after 1 year. The mission of Harmony Project is to provide free music education to children from low-income communities, and their programs are established in schools where at least 90% of the children qualify for free or reduced lunch. Because U.S. government guidelines qualify children for reduced lunch if the family income is 185% or less of Federal poverty guidelines, and for free lunch if the family income is less than 130% of poverty level, it can be assumed that the participants in this study were predominantly of low income. All the participants had normal hearing (pure-tone air-conduction thresholds ≤20 dB normal hearing level for octaves from 250 to 8000 Hz), no formal diagnosis of learning, audiological or neurological impairments based on parental report, and no prior musical training.

Of the original 80 participants tested in 2011, 64 returned in both 2012 and 2013 for further testing. Eighteen of these 64 participants were excluded from the present analyses for a variety of reasons: failed audiometric screening at the second or third testing session (two participants), formal diagnosis of learning or developmental disability confirmed after the initial study enrollment (two participants), or failure to fulfill the musical training assignment (10 Group 1 participants failed to join the program after their initial year without training; two Group 1 participants and one Group 2 participant left the program before the end of the study; one Group 1 participant received music training outside Harmony Project during the first year of the study). This left a total of 46 children (Group 1, $n = 19$ and Group 2, $n = 27$) for whom data were collected across all 3 years and who met the training requirements for their assigned groups. To create a comparison group of equal size, we assigned random numbers to each participant in group 2 and then sorted in ascending order and selected the first 19 children from Group 2 to match the smaller Group 1 ($n = 19$). All remaining analyses were carried out with these two groups of 19 children.

There were no significant differences between the two groups at the initial assessment (prior to training) with regard to age, sex, verbal IQ and non-verbal IQ (assessed using the vocabulary and matrix reasoning subtests of the Wechsler Abbreviated Scale of Intelligence; [50]), and maternal education (an index of SES: see Stevens et al. [51] for discussion regarding the predictive value of maternal education for inferring a child's socioeconomic status), speech-in-noise perception (assessed by Hearing in Noise Test (HINT)) and age of acquisition of English ($p > 0.2$ for all comparisons; see Table 1 for a summary of participant characteristics by group based on the first year of assessments). The participants were predominantly Spanish–English bilinguals based on both parent and child report, reflecting the demographics of the neighborhood in which services are provided: Group 1 included 5 simultaneous bilinguals (both languages learned from birth), 12 sequential bilinguals (second language acquired later) and 2 English monolinguals; Group 2 included 6 simultaneous bilinguals, 12 sequential bilinguals and 1 English monolingual.

Table 1
Group matching characteristics, based on year 1 data collection.

	Group 1	Group 2	Statistic
	(n = 19)	(n = 19)	
Age (months)	Mean 97.32 (SD = 7.20)	100.05 (8.98)	$F_{(1,36)} = 1.075$, $p = 0.307$
Sex	12 females	9 females	$\chi^2 = 0.958$, $p = 0.515$
Maternal education (years)	11.26 (4.34)	11.32 (3.77)	$F_{(1,36)} = 0.002$, $p = 0.968$
Verbal IQ	45.16 (11.54)	48.58 (10.34)	$F_{(1,36)} = 0.926$, $p = 0.342$
Non-verbal IQ	52.42 (10.76)	52.74 (10.15)	$F_{(1,36)} = 0.009$, $p = 0.926$
Speech-in-noise perception (SNR/dB)	−0.226 (1.67)	0.289 (1.77)	$F_{(1,36)} = 0.855$, $p = 0.361$
Age of acquisition of English (years)	2.16 (2.09)	1.68 (1.95)	$F_{(1,36)} = 0.523$, $p = 0.474$

2.3. *Experimental procedures*

Testing was carried out in English, in quiet rooms within the Harmony Project's Los Angeles offices. At the beginning of each testing session, informed written consent was obtained from legal guardians in their language of preference (either English or Spanish) on behalf of the children participating in the study and informed written assent was obtained in English from the child participants. These data were collected as part of a larger study assessing the transfer effects of musical training to auditory processing and language development. All forms and experimental procedures were approved by the Northwestern University Institutional Review Board. Participants were monetarily compensated for their testing time.

2.4. *Speech-in-noise perception*

The Hearing in Noise Test (HINT) is a clinical measure of speech-in-noise perception [52]. Participants were instructed to repeat back short, semantically and syntactically simple English sentences from the Bamford–Kowal–Bench corpus [53] presented via Sennheiser HD 25-1 headphones. Participants were presented with 10 sentences, in accordance with the manufacturer's specifications for administering the test to children [54,55]. These sentences contain vocabulary appropriate for a first grade reading level: sample sentences include "Sugar is very sweet" and "Children like strawberries." The speech stimuli were spoken by a man and presented in speech-shaped background noise which matched the spectra of the test sentences. Participants were asked to repeat the sentence they heard; sentences were marked as correct only if *all* words were repeated accurately. The experimenter waited for the child's response before continuing to the next sentence and the amount of time taken did not factor into whether the sentence was scored as correct. The intensity level of the noise was fixed at 65 dB SPL and the intensity level of the target sentences was adaptively adjusted by the HINT software until a threshold signal-to-noise ratio (SNR) was obtained. The standard HINT adaptive procedure includes the following steps: the first sentence is presented until the subject responds correctly, and then the presentation level of each following sentence is adjusted based on the participant's response, such that the level is lowered after a correct response and raised after an incorrect response. The standard HINT protocol was used in which a larger step size (4 dB) is used for the first four sentences to obtain a rough estimate of the subject's threshold. This estimate is calculated by averaging the presentation levels of the first four sentences and the level at which the fifth sentence would have been presented based on the response to the fourth sentence. Then the fifth sentence is presented at this estimated threshold and a smaller step size (2 dB) is used for the remaining sentences to obtain a more precise measure of the threshold. The final threshold is calculated as the mean of the presentation levels starting with sentence 5 and including the level at which the 11th sentence would have been presented. A lower threshold indicates a greater ability to perceive speech in adverse listening conditions. Participants were presented with one list of 10 sentences for each testing session, randomly selected from a total of 10 possible lists. Each participant was given five practice sentences in quiet at a fixed level (65 dB SPL) and five practice sentences in noise before the assessment.

2.5. *Music training*

The music training followed the Harmony Project's standard curriculum. All students attend an introductory musicianship class before progressing to instrumental classes. The musicianship classes meet for 1 h, twice a week, and the learning objectives include fundamental pitch and rhythm skills, vocal performance, basic improvisation and composition, and awareness of musical styles and notation (details are provided in Table 2) as well as basic recorder playing. Students attend the musicianship class for up to 1 year; they may progress to instrumental instruction in less than 1 year if the teacher considers them ready based on their classroom performance and/or instrument availability (the instruments are provided at no cost to the students and availability is dependent on donation). Instrument and ensemble opportunities vary based on teacher and program availability but typically include at least 4 h per week of group instrumental instruction. For the students involved in this research study, once they began instrumental classes they continued their training on that same instrument for the duration of the study. A summary of the instruments played by study participants is provided in Table 3.

2.6. *Data analysis*

A repeated-measures analysis of covariance was conducted to assess the impact of musical training on speech-in-noise (SIN) perception using musical training group as the between-subject factor and year as the within-subject factor, controlling for age of training onset, sex and age of acquisition of English. *Post hoc* paired t-tests assessed the extent of improvement in speech-in-noise perception in each training group. We also compared SIN performance of Group 1 and Group 2 using one-way ANOVAs. Pearson correlations were conducted to explore relations between speech-in-noise perception and the number of hours of instrumental training. All statistical analyses were performed using SPSS 19 (SPSS, Inc., Chicago, IL, USA) and reflect 2-tailed p-values. Bonferroni corrections were used to adjust for multiple comparisons in *post hoc* analyses.

3. Results

During the first year of the study, Group 1 did not participate in musical training whereas Group 2 completed their first year of training. By the third and final assessment, Group 1 had completed 1 year of musical training while Group 2 had completed 2 full years of training (see Fig. 1 for overview of the experimental design). There was a significant effect of training group on speech-in-noise

Table 2
Musicianship class – learning objectives.

Rhythm
Identify, read and perform basic note/rest values, measures, bar lines, meters
Identify and perform simple rhythmic patterns with a steady beat
Pitch
Name lines/spaces on the staff
Identify and perform simple melodic patterns, follow pitch direction through movement
Sing and identify major and minor scales
Performance
Sing and perform independently and in groups, on pitch and in rhythm, blending timbres
Follow a conductor for dynamics, tempo, and cues; exhibit appropriate rehearsal etiquette
Echo short rhythms and melodic patterns (call and response)
Improvisation and composition
Improvise "answers" in the same style to given rhythmic and melodic phrases
Improvise simple rhythmic variations and melodic embellishments on familiar melodies
Create and arrange short songs and instrumental pieces within specified guidelines
Write and perform simple compositions
Musical awareness
Explain personal preferences for music and styles using appropriate terminology
Identify instruments and their sounds, including instruments from various cultures
Listen to music, analyze and describe structure/emotion
Musical terms
Melody, rhythm, and harmony
Beat, measure, bar line, repeat sign, double bar, Grand Staff, treble clef, bass clef, C clef
Tempo, dynamics, time signature
Verse, chorus, tutti, solo, duet
Scale, chord, sharp, flat, key signature
Orchestra instrumentation
Conductor
Strings: Violin, Viola, Cello, Double Bass, Guitar, Harp
Winds: Recorder, Flute, Clarinet, Oboe, Bassoon, Saxophone
Brass: French horn, Trumpet, Trombone, Tuba
Percussion: Timpani, Cymbals, Snare Drum, Drum Set, etc.

perception over the three assessments (years 1, 2 and 3: repeated measures ANCOVA, controlling for age of training onset, sex and age of acquisition of English: group × year interaction: $F_{(2,34)} = 4.066$, $p = 0.022$, partial $\eta^2 = 0.110$). There was no significant effect of age of training onset or age of English acquisition (p's > 0.3).

Post hoc analyses revealed that Group 2 showed significant improvement after 2 years of musical training (year 1 pre-training assessment to year 3 post-training assessment: paired $t_{(18)} = 3.958$, $p = 0.001$), with a mean SNR change of −2.1 dB (95% CI [0.98,3.22]), and significantly outperformed Group 1 at the third assessment (year 3: independent $t_{(36)} = 3.419$, $p = 0.002$, $d = 1.140$, 95% CI [0.525, 2.054]). See Fig. 2. Significance thresholds were adjusted using Bonferroni corrections to allow for multiple comparisons.

The total hours of instrumental training received related to HINT performance at the end of the study ($r = -0.448$, $p = 0.005$, $d = 1.002$, 95% CI [0.15,0.67]), with more hours of training linked to better speech-in-noise perception (see Fig. 3). There was no such relationship with pre-training HINT performance ($r = 0.149$, $p = 0.372$, $d = 0.3013$).

3.1. Significance of improvement for everyday listening

To assess the potential impact of training on everyday listening, we considered the number of children in each group who improved their SNR thresholds by at least 1 dB, which can equate to an improvement in speech recognition performance by as much as 10–15% [56]. In their baseline year without training, only 3 out of 19 children (16%) in Group 1 improved their HINT performance by 1 dB or more. After 1 year of training, 7 out of 19 (37%) children in Group 1 improved by 1 dB or more. In Group 2, 10 out of 19 participants (53%) improved by 1 dB or more after 1 year of training, and after 2 years of training, 12 participants (63% of the total) improved their thresholds by 1 dB or more, with nine of those participants improving by 2 dB or more.

4. Discussion

We provide the first random-assignment, longitudinal evidence for improved hearing in noise with music training, validating the relationship between music training and speech-in-noise perceptual advantages and indicating that this "musician advantage" observed in previous cross-sectional studies [4,9,10,57,58] is not simply a reflection of pre-existing differences between those who pursue music and those who do not. Further, we reveal this improvement in the context of an established and successful music program providing free, group music instruction to underserved children. For the first time, we demonstrate an effect of musical experience on hearing speech in noise in bilinguals, a population for whom the task presents a particular challenge [59–62], and reveal that the extent of training benefit is not influenced by the age of acquisition of English. Taken together, this study bridges an important research gap by extending previous assessments of computer-based training or experimenter-initiated music programs on speech perception [2,45,63–65] to an established program that has demonstrated its viability, sustainability and effectiveness in the community.

Playing music involves not only the accurate perception and production of sound, but *communication* through sound. Whether tuning in to the sound of one's instrument in an ensemble, or tracking a talker's voice in a noisy background, the meaningful signal must be extracted from a complex soundscape. Both activities

Table 3
Music program characteristics and instrumental training summary.

Harmony project program	Typical weekly instrumental instruction	Number of students	
		Group 1 ($n = 19$)	Group 2 ($n = 19$)
Alexandria elementary school	One-hour instrumental classes twice a week plus a 2 h string ensemble rehearsal each week (4 h/wk)	10 (10 musicianship only)	3 (3 bass)
Beyond the bell	Twice-weekly 2 h ensemble rehearsals. These include pull-out sectional rehearsals, which are similar to large instrumental classes at other sites (4 h/wk)	NA	8 (2 flute, 5 trumpet, 1 clarinet)
EXPO Center (YOLA)	One-hour instrumental music classes each week and a 3 h ensemble rehearsal each week (4 h/wk)	9 (9 viola)	3 (2 cello, 1 trumpet)
Hollywood	One-hour instrumental classes twice a week plus a 3-h ensemble rehearsal (concert band) each week (5 h/wk)	NA	5 (5 trumpet)

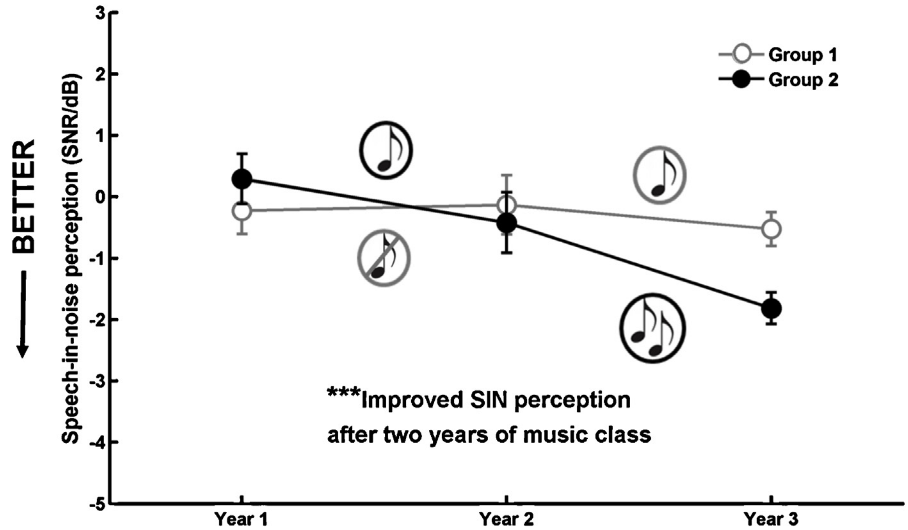


Fig. 2. Musical training strengthens the ability to hear speech in challenging listening environments: children with 2 years of musical training (Group 2, black) improved their speech-in-noise SNR thresholds by an average of 2.1 dB (asterisks indicate statistical significance: $^{***}p \leq 0.001$, $^{**}p \leq 0.01$, $^{*}p \leq 0.05$). Children with 1 year of musical training (Group 1, gray) did not show significant improvement.

involve the use of fine-grained acoustic cues to segregate distinct input streams, as well as the ability to retain information in working memory, extrapolate from patterns and regularities within the signal, and make use of prior experience and context to disambiguate degraded input. Many of the same aspects of auditory processing that are important to speech-in-noise perception have been shown to be strengthened in musicians compared with non-musicians [see 66,67,68,69 for review], with previous research indicating that musicians' superior auditory skills arise from more precise neural encoding of sound [67–71] coupled with strengthened cognitive function [72,73]. Musicians demonstrate selective enhancements based on their specific experience, such as strengthened neural encoding of the timbre of their own instrument [74–76], but also show advantages in processing speech [4,27,41,42,58,67,77] and non-verbal communication sounds [78]. These outcomes suggest that musical experience not only strengthens the specific components of sound that are meaningful within musical practice (e.g. the sound of a musician's own instrument) but also promotes more precise and effective processing of meaningful sound in other communication contexts [41,67,68,79]. This may be due to the fact that musical practice provides experience not only with the specific ingredients of musical sound, but also with the process of integrating those ingredients together during communication. This ability to extract meaning from a complex auditory scene may be an important factor in the transfer of skills to non-musical domains.

Understanding speech in a noisy background is difficult for anyone, but in a low-income, bilingual population there are additional factors that compound these challenges. On language-based tests of perception, bilinguals are more adversely affected by noise than monolinguals, despite normal hearing and intelligence [59–62]. Given evidence that bilingual experience can enhance the neural representation of sound [80–82], it seems unlikely that greater difficulties with speech perception in noise in bilinguals stem from sensory deficits. Rather, these difficulties may result from the effects of reduced exposure, lower competency and smaller vocabulary in the target language [83,84], as well as interference from activation of similar words in the non-target language in the case of ambiguous speech [62]. The training effects reported here were not influenced by age of English language acquisition, and therefore represent the first evidence that musical experience confers

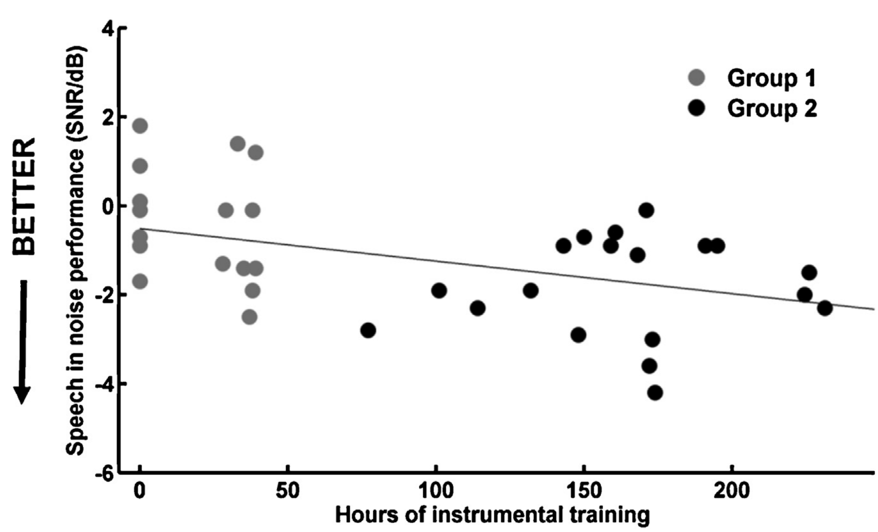


Fig. 3. Hours of instrumental training tracks with speech-in-noise perception abilities after training, across all participants.

advantages for speech perception in noise in bilinguals, as well as monolinguals. For children from low-SES households, difficulties with speech perception are exacerbated further, with low-SES children typically demonstrating smaller vocabularies than children from high-SES backgrounds [85,86] due to a combination of impoverished home literacy environment [87] and reduced access to print materials [88]. Our outcomes suggest that the additional auditory enrichment provided through musical engagement may help to counterbalance some of the obstacles to everyday listening that are encountered in a low-income, bilingual population.

In the present study we found that within group 2, some children showed significant improvement after only 1 year of training, whereas others showed improvement only after 2 years. It is notable that in the Chobert et al. [48] study (2014), enhanced preattentive processing of acoustic and phonological cues in syllables was observed after 2 academic years of music training but not after 1 year, and this was also the case with a recent study from our laboratory showing enhanced neural processing of speech in adolescents who completed 2 years of music training, compared with active controls participating in fitness-based training [49]. These findings support Patel's suggestion that the plasticity of speech-processing networks with musical training requires repetition [41] and indicate that the transfer of training benefits from music to language not only takes time, but that the time required will vary across individuals. However, this slower rate of plasticity may also result in longer-lasting effects, supported by evidence that childhood music-making confers benefits for neural function that endure into adulthood, years after training has ceased [89], including the processing of speech [90].

We would expect further improvement in Group 1 after a second year of training, following a similar trajectory to Group 2. It is possible that the slightly older age of onset of training would result in a smaller effect in Group 1, since age of onset of music training can influence the extent of its impact on auditory skills [see 91 for review], however there was no significant effect of age of onset of training in our dataset. Further research is necessary to disambiguate the relative effect of frequency vs. duration of training, as well as the impact of age of onset of training, and how these factors contribute to training retention. Further research should also investigate which specific aspects of musical training most directly impact speech processing in noise, such as rhythm vs. melody, style of musical playing, and type of instrument. It would be particularly interesting to determine whether group singing could elicit similar benefits to the instrumental classes at this age [49]: this would have important implications for the development of programs in low-income settings since instruments are costly and therefore not always available.

Our study design allowed for both within- and across-groups comparison, with Group 1 serving as the control group to Group 2 in year one, and then as their own controls in the second year of the study. However, one limitation of the present study is the absence of an *active* control group, since the provision of an alternative form of educational enrichment within this population was not logistically or financially feasible. It is possible that the improvements in speech-in-noise perception observed in our study could result from participation in additional enrichment activities but not specifically from engagement with music, for example, as a result of additional linguistic interaction with teachers, mentors and fellow students in the course of musical training. A recent longitudinal study assessing the impact of music training on academic development failed to identify significant improvements beyond musical achievement and second language development, and encouraged caution in the interpretation of the non-musical cognitive benefits of music training [92]. However, previous cross-sectional studies indicate there may be specific links between music training and speech processing, for example demonstrating superior speech-in-noise abilities in musically-trained children compared with non-musician children matched for extra-curricular activities [58], and random-assignment longitudinal studies comparing music training with painting classes demonstrated enhanced speech processing only in the musically-trained group [48,65]. A recent study in adolescents provided the first longitudinal evidence that music training improves neural timing in response to speech in noise whereas fitness-based training does not [49]. Taken together, these outcomes suggest that music training influences speech processing to a greater degree than other forms of enrichment, likely due to overlap in the neural networks involved in processing speech and music [41]. We would therefore not have expected to see the same benefits for speech-in-noise perception in an active control group that did not receive music training.

Further, music offers a framework for life-long enrichment since participation in musical activities is something that can be enjoyed throughout a lifetime, and even short periods of musical training can have lasting effects on the brain's response to sound [46,89,90,93]. Musical engagement activates emotion and reward centers of the brain [94–97] and serves an important function in social bonding [98]. The same aspects of playing music that make it emotionally and socially engaging may also make it a powerful vehicle for neuroplasticity [99–101]. Music may therefore offer particular and lasting benefits for communication skills that other forms of auditory training, such as computer-based training, do not.

5. Conclusions

In summary, we provide longitudinal evidence that 2 years of music instruction is associated with modest but clinically meaningful gains in the ability to understand speech in noise. Importantly, we assessed the impact of music training within an established music program, since our study participants followed the same curriculum and engaged in the same activities as every other student in the program. These outcomes therefore provide an important bridge between laboratory-based research and real-world educational settings. Outcomes indicate that 2 years of group music training are sufficient to confer significant benefits for speech perception in noise. Further, these benefits are observed in a low-income, predominantly bilingual population, and were not influenced by the age at which English was acquired. In addition to the intrinsic benefits of musical engagement, music training may therefore provide important support for the development of fundamental auditory skills such as the ability to perceive speech in noise.

Conflict of interest

The authors declare no competing financial interests.

Acknowledgments

The authors wish to thank the children who participated in this study and their families. We would like to thank Dr. Margaret Martin, Myka Miller, Monk Turner, Sara Flores and all other Harmony Project staff for their support. We also acknowledge Jason Thompson, Shivani Bhatia and Emily Spitzer for their contributions to data collection, Yea Lee for onsite testing support, and Trent Nicol who provided comments on an earlier version of this manuscript. This work was supported by the Knowles Hearing Center, the National Association of Music Merchants (NAMM) Foundation and the GRAMMY Foundation.

References

[1] Song JH, Skoe E, Banai K, Kraus N. Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex 2011;22:1180–90.
[2] Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N. Reversal of age-related neural timing delays with training. Proc Natl Acad Sci USA 2013;110:4357–62.
[3] de Boer J, Thornton ARD. Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. J Neurosci 2008;28:4929–37.
[4] Parbery-Clark A, Skoe E, Lam C, Kraus N. Musician enhancement for speech-in-noise. Ear Hear 2009;30:653–61.
[5] Zendel BR, Alain C. Musicians experience less age-related decline in central auditory processing. Psychol Aging 2012;27:410.
[6] Strait DL, Kraus N. Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. Front Psychol 2011;2:113.
[7] Parbery-Clark A, Strait DL, Anderson S, Hittner E, Kraus N. Musical experience and the aging auditory system: implications for cognitive abilities and hearing speech in noise. PLoS ONE 2011;6:e18082.
[8] Strait DL, Parbery-Clark A, Hittner E, Kraus N. Musical training during early childhood enhances the neural encoding of speech in noise. Brain Lang 2012;123:191–201.
[9] Zendel BR, Tremblay CD, Belleville S, Peretz I. The impact of musicianship on the cortical mechanisms related to separating speech from background noise. J Cogn Neurosci 2015;27:1044–59.
[10] Swaminathan J, Mason CR, Streeter TM, Kidd Jr G, Patel AD. Spatial release from masking in musicians and non-musicians. J Acoust Soc Am 2014;135:2281–2.
[11] Boebinger D, Evans S, Rosen S, Lima CF, Manly T, Scott SK. Musicians and non-musicians are equally adept at perceiving masked speech. J Acoust Soc Am 2015;137:378–87.
[12] Ruggles DR, Freyman RL, Oxenham AJ. Influence of musical training on understanding voiced and whispered speech in noise. PLoS One 2014;9:e86980.
[13] Fuller CD, Galvin III JJ, Maat B, Free RH, Başkent D. The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? Front Neurosci 2014;8:179.
[14] Patel AD, Peretz I. Is music autonomous from language? In: A neuropsychological appraisal; 1997.
[15] Peretz I. Music, language, and modularity in action. In: Rebuschat P, Rohmeier M, Hawkins JA, Cross I, editors. Language and music as cognitive systems. 2012. p. 254–68.
[16] Limb CJ. Structural and functional neural correlates of music perception. Anat Rec A: Discov Mol Cell Evol Biol 2006;288A:435–46.
[17] Patel AD. Music, language, and the brain. USA: Oxford University Press; 2010.
[18] Carey D, Rosen S, Krishnan S, Pearce MT, Shepherd A, Aydelott J, et al. Generality and specificity in the effects of musical expertise on perception and cognition. Cognition 2015;137:81–105.
[19] Zuk J, Ozernov-Palchik O, Kim H, Lakshminarayanan K, Gabrieli JD, Tallal P, et al. Enhanced syllable discrimination thresholds in musicians. PloS One 2013;8:e80546.
[20] Parbery-Clark A, Tierney A, Strait DL, Kraus N. Musicians have fine-tuned neural distinction of speech syllables. Neuroscience 2012;219:111–9.
[21] Strait DL, O'Connell S, Parbery-Clark A, Kraus N. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. Cereb Cortex 2013;24(9):2512–21.
[22] Kuhnis J, Elmer S, Meyer M, Jancke L. The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study. Neuropsychologia 2013;51:1608–18.
[23] Jeon JY, Fricke FR. Duration of perceived and performed sounds. Psychol Music 1997;25:70–83.
[24] Rammsayer T, Altenmüller E. Temporal information processing in musicians and nonmusicians; 2006.
[25] Magne C, Schon D, Besson M. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. J Cogn Neurosci 2006;18:199–211.
[26] Wong PC, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 2007;10:420–2.
[27] Bidelman GM, Gandour JT, Krishnan A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci 2011;23:425–34.
[28] Tervaniemi M, Just V, Koelsch S, Widmann A, Schroger E. Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. Exp Brain Res 2005;161:1–10.
[29] Spiegel MF, Watson CS. Performance on frequency-discrimination tasks by musicians and nonmusicians. J Acoust Soc Am 1984;76:1690–5.
[30] Micheyl C, Delhommeau K, Perrot X, Oxenham AJ. Influence of musical and psychoacoustical training on pitch discrimination. Hear Res 2006;219:36–47.
[31] Fujioka T, Trainor LJ, Ross B, Kakigi R, Pantev C. Musical training enhances automatic encoding of melodic contour and interval structure. J Cogn Neurosci 2004;16:1010–21.
[32] Pallesen KJ, Brattico E, Bailey CJ, Korvenoja A, Koivisto J, Gjedde A, et al. Cognitive control in auditory working memory is enhanced in musicians. PloS One 2010;5:e11120.
[33] Schulze K, Zysset S, Mueller K, Friederici AD, Koelsch S. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Hum Brain Mapp 2011;32:771–83.
[34] George EM, Coch D. Music training and working memory: an ERP study. Neuropsychologia 2011;49:1083–94.
[35] Baumann S, Meyer M, Jancke L. Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. J Cogn Neurosci 2008;20:2238–49.
[36] Strait DL, Kraus N, Parbery-Clark A, Ashley R. Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance. Hear Res 2010;261:22–9.
[37] Chandrasekaran B, Kraus N. Music, noise-exclusion, and learning. Music Percep Interdis J 2010;27:297–306.
[38] Strait DL, Parbery-Clark A, O'Connell S, Kraus N. Biological impact of preschool music classes on processing speech in noise. Dev Cogn Neurosci 2013;6:51–60.
[39] Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. J Neuro 2009;29:14100–7.
[40] Bidelman GM, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Res 2010;1355:112–25.
[41] Patel AD. Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. Front Psychol 2011;2:142.
[42] Besson M, Chobert J, Marie C. Transfer of training between music and speech: common processing, attention, and memory. Front Psychol 2011;2:94.
[43] Schön D, Magne C, Besson M. The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology 2004;41:341–9.
[44] Kraus N, Slater J, Thompson EC, Hornickel J, Strait DL, Nicol T, et al. Auditory learning through active engagement with sound: biological impact of community music lessons in at-risk children. Audit Cogn Neurosci 2014;8:351.
[45] Moreno S, Besson M. Musical training and language-related brain electrical activity in children. Psychophysiology 2006;43:287–91.
[46] Moreno S, Marques C, Santos A, Santos M, Castro S, Besson M. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. Cereb Cortex 2009;19:712–23.
[47] Kraus N, Slater J, Thompson EC, Hornickel J, Strait DL, Nicol T, et al. Music enrichment programs improve the neural encoding of speech in at-risk children. J Neurosci 2014;34:11913–8.
[48] Chobert J, Francois C, Velay JL, Besson M. Twelve months of active musical training in 8- to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. Cereb Cortex 2014;24(4):956–67.
[49] Tierney A, Krizman J, Skoe E, Johnston K, Kraus N. High school music classes enhance the neural processing of speech. Front Psychol 2013;4:855.
[50] Woerner C, Overstreet K. Wechsler abbreviated scale of intelligence (WASI). San Antonio, TX: The Psychological Corporation; 1999.
[51] Stevens C, Lauinger B, Neville H. Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. Dev Science 2009;12:634–46.
[52] Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 1994;95:1085.
[53] Bench J, Kowal Å, Bamford J. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br J Audiol 1979;13:108–12.
[54] Soli SD, Wong LL. Assessment of speech intelligibility in noise with the Hearing in Noise Test. Int J Audiol 2008;47:356–61.
[55] Vermiglio AJ. The American English hearing in noise test. Int J Audiol 2008;47:386–7.
[56] Middelweerd M, Festen J, Plomp R. Difficulties with Speech Intelligibility in noise in spite of a normal pure-tone audiogram: original papers. Int J Audiol 1990;29:1–7.
[57] Parbery-Clark A, Anderson S, Hittner E, Kraus N. Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. Front Aging Neurosci 2012;4.
[58] Strait DL, Parbery-Clark A, Hittner E, Kraus N. Musical training during early childhood enhances the neural encoding of speech in noise. Brain Langu 2012;123:191–201.
[59] Blasingame M, Bradlow AR. Perception of speech-in-noise for second language learners and heritage speakers in both first language and second language. J Acoust Soc Am 2012;132:1935.
[60] Mayo LH, Florentine M, Buus S. Age of second-language acquisition and perception of speech in noise. J Speech Lang Hear Res 1997;40:686.
[61] Nelson P, Kohnert K, Sabur S, Shaw D. Classroom noise and children learning through a second language: Double Jeopardy? Lang Speech Hear Serv School 2005;36:219.
[62] Rogers CL, Lister JJ, Febo DM, Besing JM, Abrams HB. Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. Appl Psycholinguistics 2006;27:465.
[63] Anderson S, White-Schwoch T, Choi HJ, Kraus N. Training changes processing of speech cues in older adults with hearing loss. Front syst neurosci 2013;7.
[64] Song J, Skoe E, Wong P, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 2008;20:1892–902.
[65] François C, Chobert J, Besson M, Schön D. Music training for the development of speech segmentation. Cereb Cortex 2012;23(9):2038–43.

[66] Anderson S, Kraus N. Sensory-cognitive interaction in the neural encoding of speech in noise: a review. J Am Acad Audiol 2010;21:575–85.
[67] Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nat Rev Neurosci 2010;11:599–605.
[68] Strait DL, Kraus N. Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. Hear Res 2014;308: 109–21.
[69] Kraus N, Strait DL. Emergence of biological markers of musicianship with school-based music instruction. Ann NY Acad Sci 2015;1337: 163–9.
[70] Koelsch S, Schröger E, Tervaniemi M. Superior pre-attentive auditory processing in musicians. Neuroreport 1999;10:1309–13.
[71] Tervaniemi M, Just V, Koelsch S, Widmann A, Schröger E. Pitch discrimination accuracy in musicians vs. nonmusicians: an event-related potential and behavioral study. Exp Brain Res 2005;161:1–10.
[72] Forgeard M, Winner E, Norton A, Schlaug G. Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. PLoS ONE 2008;3:e3566.
[73] Kraus N, Strait DL, Parbery-Clark A. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. Ann N Y Acad Sci 2012;1252:100–7.
[74] Pantev C, Roberts LE, Schulz M, Engelien A, Ross B. Timbre-specific enhancement of auditory cortical representations in musicians. Neuroreport 2001;12:169–74.
[75] Strait DL, Chan K, Ashley R, Kraus N. Specialization among the specialized: auditory brainstem function is tuned in to timbre. Cortex 2012;48: 360–2.
[76] Margulis EH, Mlsna LM, Uppunda AK, Parrish TB, Wong PC. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. Hum Brain Mapp 2009;30:267–75.
[77] Patel AD, Iversen JR. The linguistic benefits of musical abilities. Trends Cogn Sci 2007;11:369–72.
[78] Strait DL, Kraus N, Skoe E, Ashley R. Musical experience and neural efficiency–effects of training on subcortical processing of vocal expressions of emotion. Eur J Neurosci 2009;29:661–8.
[79] Patel AD. Can nonlinguistic musical training change the way the brain processes speech? The expanded opera hypothesis. Hear Res 2013;308:98–108.
[80] Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci USA 2012;109:7877–81.
[81] Krizman J, Skoe E, Marian V, Kraus N. Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. Brain Lang 2014;128(1):34–40.
[82] Krizman J, Slater J, Skoe E, Marian V, Kraus N. Neural processing of speech in children is influenced by extent of bilingual experience. Neurosci Lett 2015;585:48–53.
[83] Ezzatian P, Avivi M, Schneider BA. Do nonnative listeners benefit as much as native listeners from spatial cues that release speech from masking. Speech Commun 2010;52:919–29.
[84] Auer P, Wei L. Handbook of multilingualism and multilingual communication. Walter de Gruyter; 2007.
[85] Risley TR, Hart B. Meaningful differences in the everyday experience of young American children. Baltimore, MD: Paul H Brookes; 1995.
[86] Skoe E, Krizman J, Kraus N. The impoverished brain: disparities in maternal education affect the neural response to sound. J Neurosci 2013;33: 17221–31.
[87] Payne AC, Whitehurst GJ, Angell AL. The role of home literacy environment in the development of language ability in preschool children from low-income families. Early Child Res Q 1994;9:427–40.
[88] Neuman SB, Celano D. Access to print in low-income and middle-income communities: an ecological study of four neighborhoods. Read Res Q 2001;36: 8–26.
[89] Skoe E, Kraus N. A little goes a long way: how the adult brain is shaped by musical training in childhood. J Neurosci 2012;32:11507–10.
[90] White-Schwoch T, Carr KW, Anderson S, Strait DL, Kraus N. Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. J Neurosci 2013;33:17667–74.
[91] Penhune VB. Sensitive periods in human development: evidence from musical training. Cortex 2011;47:1126–37.
[92] Yang H, Ma W, Gong D, Hu J, Yao D. A Longitudinal study on children's music training experience and academic development. Sci Rep 2014;4.
[93] Moreno S, Besson M. Influence of musical training on pitch processing: event-related brain potential studies of adults and children. Ann N Y Acad Sci 2005;1060:93–7.
[94] Salimpoor VN, van den Bosch I, Kovacevic N, McIntosh AR, Dagher A, Zatorre RJ. Interactions between the nucleus accumbens and auditory cortices predict music reward value. Science 2013;340:216–9.
[95] Salimpoor VN, Zatorre RJ. Neural interactions that give rise to musical pleasure. Psychol Aesthetics Creat Arts 2013;7:62.
[96] Salimpoor VN, Benovoy M, Longo G, Cooperstock JR, Zatorre RJ. The rewarding aspects of music listening are related to degree of emotional arousal. PloS One 2009;4:e7487.
[97] Menon V, Levitin DJ. The rewards of music listening: response and physiological connectivity of the mesolimbic system. NeuroImage 2005;28:175–84.
[98] Chanda ML, Levitin DJ. The neurochemistry of music. Trends Cogn Sci 2013;17:179–93.
[99] Wan CY, Schlaug G. Music making as a tool for promoting brain plasticity across the life span. Neuroscientist 2010;16:566–77.
[100] Herholz Sibylle C, Zatorre Robert J. Musical training as a framework for brain plasticity: behavior, function, and structure. Neuron 2012;76:486–502.
[101] Munte TF, Altenmuller E, Jancke L. The musician's brain as a model of neuroplasticity. Nat Rev Neurosci 2002;3:473–8.