

Older Adults Benefit from Music Training Early in Life: Biological Evidence for Long-Term Training-Driven Plasticity

Travis White-Schwoch,^{1,2} Kali Woodruff Carr,^{1,2} Samira Anderson,^{1,2} Dana L. Strait,^{1,3} and Nina Kraus^{1,2,3,4,5}

¹Auditory Neuroscience Laboratory, ²Department of Communication Sciences, ³Institute for Neuroscience, ⁴Department of Neurobiology & Physiology, ⁵Department of Otolaryngology, Northwestern University, Evanston, Illinois 60208

Aging results in pervasive declines in nervous system function. In the auditory system, these declines include neural timing delays in response to fast-changing speech elements; this causes older adults to experience difficulty understanding speech, especially in challenging listening environments. These age-related declines are not inevitable, however: older adults with a lifetime of music training do not exhibit neural timing delays. Yet many people play an instrument for a few years without making a lifelong commitment. Here, we examined neural timing in a group of human older adults who had nominal amounts of music training early in life, but who had not played an instrument for decades. We found that a moderate amount (4–14 years) of music training early in life is associated with faster neural timing in response to speech later in life, long after training stopped (>40 years). We suggest that early music training sets the stage for subsequent interactions with sound. These experiences may interact over time to sustain sharpened neural processing in central auditory nuclei well into older age.

Introduction

Over the past several years, evidence has emerged to suggest that musicians have nervous systems distinct from nonmusicians, potentially due to training-related plasticity (Moreno et al., 2009; Kraus and Chandrasekaran, 2010; Bidelman et al., 2011; Herholz and Zatorre, 2012). One of the most provocative findings in this literature is that lifelong music training may offset age-related declines in cognitive and neural functions (Hanna-Pladdy and MacKay, 2011; Parbery-Clark et al., 2012a; Zendel and Alain, 2012; Alain et al., 2013). To date, the majority of work on music training—including in older adults—has focused on individuals who play an instrument continuously throughout their lives. These are special cases, as many adults have dabbled with an instrument on and off, especially during childhood. An outstanding question is whether limited training early in life leaves a trace in the aging

nervous system, affecting neural function years after training has stopped.

Aging results in pervasive declines throughout the nervous system, including in the auditory system. Even older adults with preserved peripheral function face a loss of central auditory function (Pichora-Fuller et al., 2007; Anderson et al., 2012; Ruggles et al., 2012). These age-driven deficits compound with declines in neocortical function that affect sensory and cognitive circuits (Gazzaley et al., 2005; Recanzone et al., 2011). Together, these declines reduce the ability to form a clear and veridical representation of the sensory world.

In speech, age-related deficits affect the neural encoding of fast-changing sounds, such as consonant–vowel (CV) transitions (Anderson et al., 2012). CV transitions pose perceptual challenges across the lifespan due to fast-changing spectrotemporal content and relatively low amplitudes (compared to vowels; Tallal, 1980). Precise coding of CV transitions supports auditory-cognitive and language-based abilities (Strait et al., 2013; Anderson et al., 2010). Yet age-related declines are not inevitable: older adults with lifelong music training do not exhibit neural timing delays in response to CV transitions (Parbery-Clark et al., 2012a). Moreover, the typical auditory system that has experienced this age-related decline is not immutable: short-term computer training in older adult nonmusicians partially reverses age-related delays in neural timing (Anderson et al., 2013b). This finding is consistent with emerging evidence from humans and animals suggesting a maintained potential for neuroplasticity into older age (Linkenhoker and Knudsen, 2002; Berry et al., 2010; de Villers-Sidani et al., 2010).

In young adults, neural enhancements from limited music training during childhood persist 5–10 years after training has

Received June 17, 2013; revised Aug. 22, 2013; accepted Sept. 18, 2013.

Author contributions: S.A. and N.K. designed research; T.W.-S. and S.A. performed research; T.W.-S. and K.W.C. analyzed data; T.W.-S., K.W.C., S.A., D.L.S., and N.K. wrote the paper.

This work was supported by the National Institutes of Health (R01 DC010016 to N.K., F31 DC011457-01 to D.S.), the Hugh Knowles Center (to N.K.), and the Northwestern Cognitive Science program (to K.W.C.). We thank Hee Jae Choi for her assistance with data analysis and Trent Nicol, Elaine Thompson, and Jennifer Krizman for critical reviews of the manuscript.

Correspondence should be addressed to Nina Kraus, 2240 Campus Dr., Evanston, IL 60208. E-mail: nkraus@northwestern.edu.

S. Anderson's present address: Department of Hearing and Speech Sciences, University of Maryland, College Park, MD 20742.

D. Strait's present address: Neural Systems Laboratory, Institute for Systems Research, University of Maryland, College Park, MD 20742. E-mail: nkraus@northwestern.edu.

DOI:10.1523/JNEUROSCI.2560-13.2013

Copyright © 2013 the authors 0270-6474/13/3317667-08\$15.00/0

stopped (Skoe and Kraus, 2012). The question remains, however, whether enhancements are preserved into older adulthood. Here, we hypothesized that older adults with a moderate amount of instrumental music training during childhood and/or young adulthood retain a neural trace of this early training, manifest as less severe age-related neural timing delays. To test this hypothesis, we measured scalp-recorded auditory brainstem responses to speech in a cohort of older adults who reported varying levels of music training early in life, but with no training after age 25. We predicted that there would be faster neural timing, indicating a more efficient auditory system, in older adults with more years of past music training.

Materials and Methods

Subjects. Forty-four older adults (ages 55–76; 25 female) were recruited from the Chicago area. Audiometric thresholds were measured bilaterally at octave intervals from 0.125 to 8 kHz, including interoctave intervals at 3 and 6 kHz. Subjects had either normal hearing [thresholds from 0.25 to 8 kHz \leq 20 dB hearing level (dB HL) bilaterally] or no more than a mild-to-moderate sensorineural hearing loss; all pure-tone averages (average threshold from 0.5 to 4 kHz) were \leq 45 dB HL bilaterally. No individual threshold was $>$ 40 dB HL at or below 4 kHz or $>$ 60 dB HL at 6 or 8 kHz; no asymmetries were noted ($>$ 15 dB HL difference at two or more frequencies between ears). Stimulus presentation levels for electrophysiology were corrected for hearing loss (see Stimuli, below) and our subject cohort was a mix of older adults with normal hearing and hearing loss, as our group has previously published (Anderson et al., 2013b, 2013c). All subjects had normal wave V click-evoked latencies ($<$ 6.8 ms in response to a 100 μ s click presented at 80 dB SPL at 31.25 Hz), and no history of neurological disorders. All subjects passed the Montreal Cognitive Assessment, a screening for mild cognitive impairment (Nasreddine et al., 2005). One subject was a statistical outlier and so was excluded (see Neural timing, below); all reported statistics, including descriptive statistics for each group, exclude this individual. Informed consent was obtained from all subjects in accordance with Northwestern University’s Institutional Review Board.

Training groups. Participants were divided into three groups based on self-report of formal instrumental music training: “None” (0 years, $N = 15$), “Little” (1–3 years, $N = 13$), and “Moderate” (4–14 years, $N = 13$). Music training was private and/or group instrumental instruction. Although the groups were separated based upon total years of music training, there were converging factors that motivated dividing the music-trained groups into 1–3 years versus 4–14 years. In the United States, 1–3 years of training corresponds approximately to training during middle school or junior high school. Many of the subjects in the Little group completed their music training during this time, whereas many subjects in the Moderate group continued training into high school or the beginning of college. In general, subjects in the Little group also rated themselves as less proficient on their instruments than those with \geq 4 years of training. Therefore, we believe that this cutoff is meaningful and reflects the length, intensity, and success of subjects’ music training. No subject reported any instrumental practice, performance, or instruction after age 25.

The groups did not differ on sex distribution, age, hearing, intelligence quotient (IQ), educational attainment, current levels of exercise, or age of training onset. Since groups were matched on these criteria, we did not model them as covariates in comparisons of neural function. IQ was measured with the Wechsler Abbreviated Scale of Intelligence vocabulary (verbal) and matrix reasoning (nonverbal) subtests (Zhu and Garcia, 1999). Educational attainment was inferred from a four-item Likert scale (highest academic grade completed: 1, middle school; 2, high school/equivalent; 3, college; 4, graduate or professional degree). To guard against group differences driven by general health or lifestyle factors (Anderson et al., 2013c), current levels of exercise were evaluated by having subjects provide the frequency (number of times per week) with which they engage in physical activities (cycling, walking, gardening, etc.) and then averaged, in an effort to control for health factors that may support better auditory function (Anderson et al., 2013c). See Tables 1 and 2 for group characteristics.

Table 1. Groups are matched on a wide array of demographic criteria, including: sex, age, intelligence, hearing, educational attainment, and levels of current exercise (RM-ANOVAs, all $p > 0.15$)^a

Criterion	None (0 years; $N = 15$)	Little (1–3 years; $N = 14$)	Moderate (4–14 years; $N = 13$)
Number of females	6	9	9
Age (years)	63.56 (3.76)	65.46 (4.54)	65.15 (5.46)
IQ (standard score)	120.40 (8.09)	118.69 (8.44)	122.38 (5.37)
IQ-verbal (<i>t</i> -score)	63.13 (7.18)	63.42 (6.13)	62.41 (5.07)
IQ-nonverbal (<i>t</i> -score)	58.13 (10.25)	58.91 (6.77)	62.75 (4.69)
PTA (dB HL; average hearing 0.5–4 kHz)	20.61 (11.51)	17.19 (11.39)	24.95 (8.90)
Click V latency (ms)	6.25 (0.32)	5.98 (0.49)	5.97 (0.33)
Education	3.47 (0.64)	3.23 (0.56)	3.77 (0.44)
Exercise levels	2.62 (1.39)	2.81 (0.98)	2.90 (1.19)
Years of music training		1.71 (0.61)	8.14 (3.68)
Years since music training		52.01 (4.10)	47.46 (6.92)

^aMeans are reported with SDs. PTA, Pure-tone average hearing threshold (dB HL) measured bilaterally from 0.5 to 4 kHz. Refer to Materials and Methods for information on how education and exercise were rated.

Table 2. Music practice histories, based on self-report, for individuals from the two relevant groups^a

Group; instruments	Years of Training	Years since Training
Little (1–3 years; $N = 14$)		
Piano	1	49
Piano	1	55
Piano	1	53
Piano	1	^b
Piano	2	49
Piano	2	58
Piano	2	53
Trumpet	2	52
Accordion, organ	3	49
Clarinet	3	51
Piano	3	53
Piano, French horn	3	46
Piano, guitar	3	51
Violin, piano	3	61
Moderate (4–14 years; $N = 13$)		
Clarinet	4	43
Viola, clarinet	4	50
Piano, trumpet	5	59
Trumpet	5	51
Guitar	5	50
Clarinet, guitar, piano	6	54
Piano	7	42
Piano	8	51
Piano	9	58
Piano	11	47
Piano, guitar	12	45
Piano, flute	14	40
Trombone	14	37

^aFor each subject, the training instrument(s), total years of training, and years since training stopped are reported.

^bThis individual only reported that training stopped at a “very young” age.

Electrophysiology. To compare groups’ neural representation of speech, we used scalp electrodes to measure auditory brainstem responses to a synthesized speech sound [da] (Fig. 1). This subcortical electrophysiologic response is a variant of the auditory brainstem response that is elicited in response to complex sounds instead of simple clicks or tones (Skoe and Kraus, 2010) and is generated by synchronous firing of midbrain nuclei, predominantly inferior colliculus (IC; for review, see Chandrasekaran and Kraus, 2010). Despite its subcortical origin, this response is malleable with experience and similar techniques have revealed changes due to developmental (Johnson et al., 2008; Anderson et al., 2012) and experience-dependent plasticity (Krishnan et al., 2009; Anderson et al., 2013b), including plasticity resulting from music

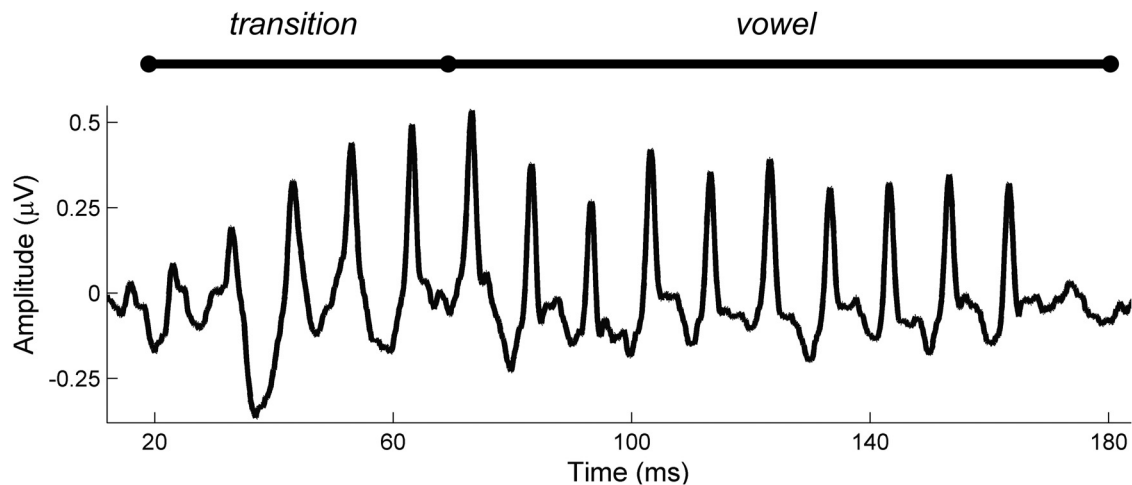


Figure 1. Grand average response from all subjects to the speech sound [da] presented in quiet. The CV transition (/d/) and vowel (/a/) regions are marked.

training (Kraus and Chandrasekaran, 2010; Bidelman et al., 2011; Strait et al., 2013).

Stimuli. A 170 ms six-formant speech syllable [da] was synthesized using a Klatt-based synthesizer at a 20 kHz sampling rate, with an initial 5 ms stop burst and a steady fundamental frequency ($F_0 = 100$ Hz). During the first 50 ms (transition between the stop burst /d/ and the vowel /a/), the first, second, and third formants change (F_1 , 400 → 720 Hz; F_2 , 1700 → 1240 Hz; F_3 , 2580 → 2500 Hz) but stabilize for the subsequent 120 ms steady-state vowel. The higher formants are stable in frequency throughout the entire 170 ms (F_4 , 3300 Hz; F_5 , 3750 Hz; F_6 , 4900 Hz). The syllable was presented alone (“quiet” condition) and masked by a two-talker babble track (“noise” condition, adapted from Van Engen and Bradlow, 2007). In cases of hearing loss ($N = 31$ total; None group, $N = 13$; Little group, $N = 7$; Moderate group, $N = 11$; see Subjects for hearing classifications) the [da] stimulus was selectively frequency-amplified with the NAL-R (National Acoustic Laboratories—Revised) algorithm (Byrne and Dillon, 1986), a procedure our group has shown improves the morphology and replicability of the response while retaining important timing properties (Anderson et al., 2013a).

Recording parameters. Auditory brainstem responses were recorded at a 20 kHz sampling rate using four Ag-AgCl electrodes in a vertical montage (Cz active, Fpz ground, and earlobe references). Stimuli were presented binaurally in alternating polarities (stimulus waveform was inverted 180°) with an 83 ms interstimulus interval through electromagnetically shielded insert earphones (ER-3A, Etymotic Research) at 80 dB SPL. During the recording session, participants were seated in a sound-attenuated booth and watched a muted, captioned movie of their choice to facilitate an alert but restful state. Sixty-three hundred sweeps were collected.

Data reduction. Responses were bandpass filtered offline from 70 to 2000 Hz (12 dB/octave, zero-phase shift). The responses were epoched using a −40 to 213 ms time window referenced to stimulus onset (0 ms). Sweeps with amplitudes greater than ± 35 μ V were considered artifact and rejected, resulting in 6000 response trials for each participant. Responses to the two polarities were added to limit the influence of cochlear microphonic and stimulus artifact. Amplitudes of responses were baseline-corrected to the prestimulus period.

Neural timing. To gauge the timing of subjects’ neural responses to speech, response peaks corresponding to the CV transition of the stimulus (peaks occurring ~23, 33, 43, 53, 63 ms) and the steady-state sustained vowel period (peaks occurring ~73, 83, 93, 103, 113, 123, 133, 143, 153, 163 ms) were identified by an automated program. Peak selections were confirmed by a rater blind to the subject group or condition and checked by a secondary peak picker. “Relative” peak timings (Figs. 2, 3) were computed by subtracting expected latency (i.e., 43 ms) from group average peak latency (for example, were 43.69 ms the mean latency in a given group for a peak occurring ~43 ms, that group would be plotted as having a “relatively latency” of 0.69 ms). One subject with 8 years of music training was a statistical outlier (four of five latencies for

noise transition peaks were >2.5 SDs beyond the mean) and was excluded from analysis.

Statistical analyses. To compare neural timing between groups and conditions, peak latencies were submitted to a repeated-measures ANOVA (RM-ANOVA) with condition (quiet/noise) as a within-subjects factor and group (None/Little/Moderate) as a between-subjects factor. Main effects of group are reported for each condition. Normality was confirmed using the Shapiro–Wilk test and Levene’s test ensured homogeneity of variance. The RM-ANOVA was run with $\alpha = 0.05$, and all follow-up tests were strictly Bonferroni-corrected for multiple comparisons; all p values refer to two-tailed tests. In the Little and Moderate groups, nonparametric Spearman’s correlations were used to relate years of training to neural timing, since the distribution of years of training was slightly skewed. Analyses were performed in SPSS 20.0 (SPSS).

Results

Summary of results

A greater amount of music training early in life was associated with the most efficient auditory function decades after training stopped. Older adults in the Moderate training group had the fastest neural timing in response to the [da] presented in quiet and noise. The Moderate group was also the most resilient to noise-induced timing delays. Group differences were only seen in the region of the response corresponding to the CV transition of the syllable (the region between 20 and 60 ms in the response); in the steady-state vowel portion of the response, the groups were equivalent.

Neural resistance to noise degradation on response timing

In a typical auditory system, noise slows neural responses to speech (Hall, 2006; Anderson et al., 2010), although musicians experience less drastic timing delays than nonmusicians (Parbery-Clark et al., 2009; Strait et al., 2012; Kraus and Nicol, 2013). To compare the degrading effect of noise on neural timing across all three groups, latencies in quiet and noise for the CV transition region of the responses (phase-locked peaks occurring at ~23, 33, 43, 53, and 63 ms) were compared. There was a significant main effect of group, revealing fastest timing in the Moderate group in both quiet and noise ($F_{(1,71)} = 2.85$, $p = 0.005$). There was also a significant group \times condition interaction, demonstrating that the Moderate group was least affected by latency delays due to noise ($F_{(1,71)} = 2.17$, $p = 0.030$). There were no effects of group in response to the steady-state vowel (all $p > 0.1$) meaning that past music training was only associated with improved timing in the CV transition.

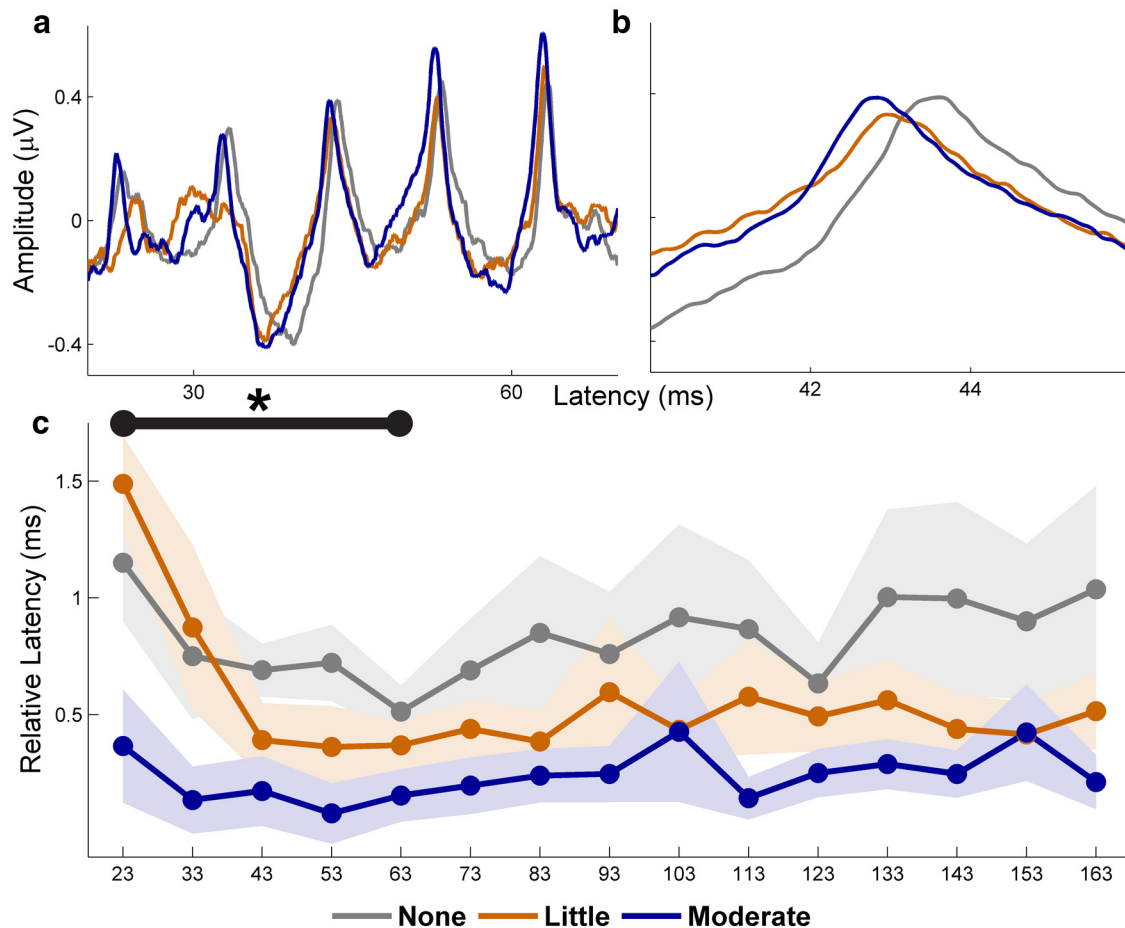


Figure 2. Neural response timing to the [da] presented in quiet. The Moderate training group (blue) has the fastest neural timing. *a*, Group average responses in the CV transition region. *b*, The transition peak at 43 ms is highlighted. *c*, Latency relative to the expected peak timing. Lower relative latency values are faster. Shaded regions represent ± 1 SEM. * $p < 0.05$.

Absolute neural timing in response to speech in quiet

There was a significant effect of group membership on peak timing in response to speech in quiet, with the Moderate training group having the fastest timing, on average, followed by the Little and None groups, respectively ($F_{(1,71)} = 2.13$, $p = 0.033$). See Table 3 for mean peak latencies in quiet and Figure 2 for responses in the transition and an illustration of mean neural timing by group.

Follow-up tests revealed that this difference was driven by the Little versus Moderate comparison ($F_{(1,20)} = 3.28$, $p = 0.025$), with no significant group differences in the None versus Moderate ($F_{(1,22)} = 1.74$, $p = 0.168$) or None versus Little comparisons ($F_{(1,22)} = 1.49$, $p = 0.237$). However, Bonferroni-corrected assessments at an individual peak level revealed differences between the None and Moderate groups for peaks 43 ($t_{(26)} = 2.79$, $p = 0.036$) and 53 ($t_{(26)} = 3.02$, $p = 0.018$) with trending differences for peaks 23 ($t_{(26)} = 2.24$, $p = 0.067$) and 63 ($t_{(26)} = 2.26$, $p = 0.075$). This suggests that there was a None versus Moderate difference in quiet but that our sample size was underpowered to bring this out at a multivariate level. This is consistent with previous work that suggests training effects in quiet are relatively weak compared to those in noise (Russo et al., 2010; Anderson et al., 2013b).

There were no effects of group on timing in response to the steady-state vowel (overall: $F_{(1,62)} = 0.82$, $p = 0.685$; all individual peaks $p > 0.1$), which is consistent with previous findings establishing a selective effect of music training for encoding spectrotemporally dynamic regions of speech (Parbery-Clark et al., 2009, 2012a; Strait et al., 2012, 2013).

Absolute neural timing in response to speech in noise

There was a significant effect of group membership on peak timing in response to speech in noise, with the Moderate training group having the fastest timing, on average, followed by the Little and None groups, respectively ($F_{(1,71)} = 3.35$, $p = 0.001$). See Table 4 for mean peak latencies in noise and Figure 3 for responses in the transition and an illustration of mean neural timing by group.

Follow-up tests revealed that this difference was driven by the None versus Moderate and Little versus Moderate group comparisons (None vs Moderate: $F_{(1,22)} = 4.94$, $p = 0.004$; Little vs Moderate: $F_{(1,20)} = 3.57$, $p = 0.018$), with no significant group difference in the None versus Little comparison ($F_{(1,22)} = 2.07$, $p = 0.107$).

As in quiet, there were no effects of group on timing in response to the steady-state vowel (overall: $F_{(1,62)} = 0.97$, $p = 0.504$; all individual peaks $p > 0.1$). Again, this is consistent with previous findings suggesting a selective enhancement conferred by extensive music training.

Linear relationship between years of training and neural timing

In addition to group-wise comparisons, we correlated years of training and neural timing within subjects with some degree of music training to determine whether the relationship between extent of training and neural timing reflected more years of training. For every transition peak in quiet and noise there was a

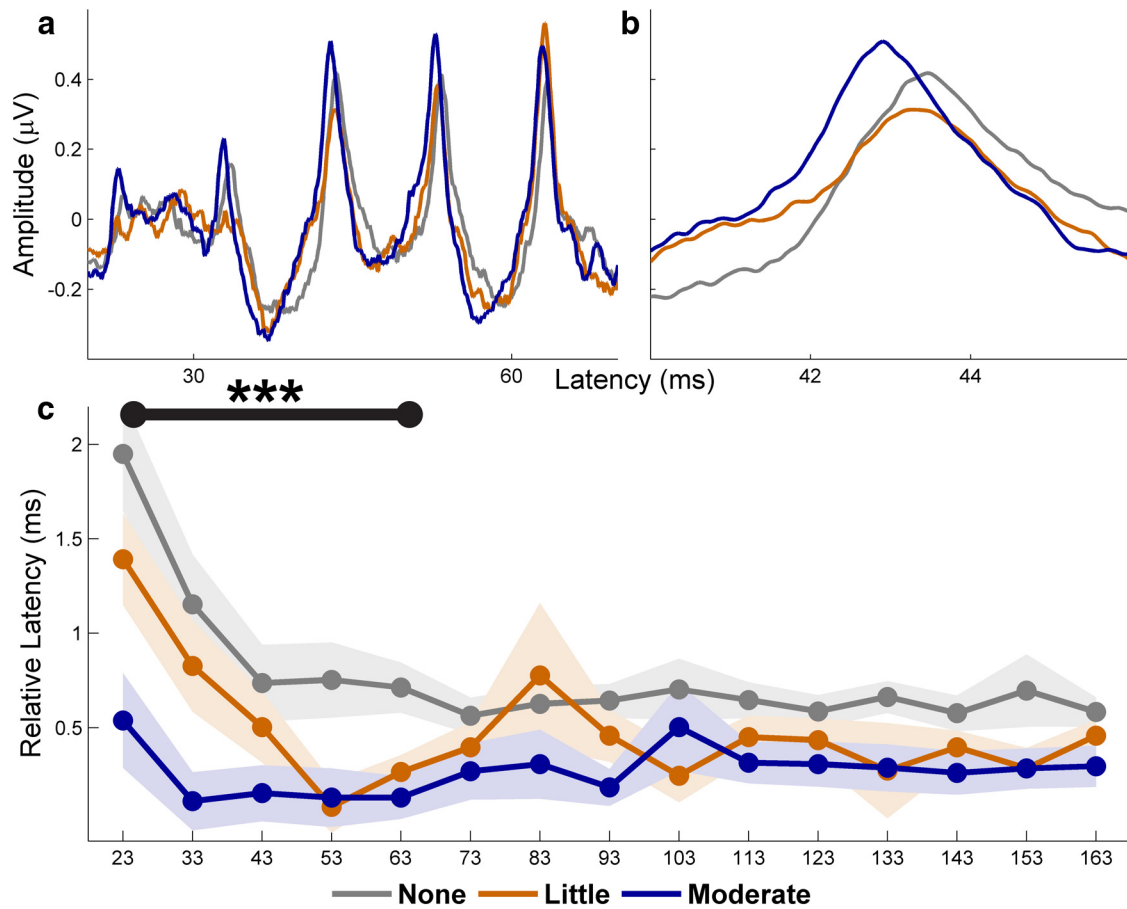


Figure 3. Neural response timing to the [da] presented in noise. The Moderate training group (blue) has the fastest neural timing. *a*, Group average responses in the CV transition region. *b*, The transition peak at 43 ms is highlighted. *c*, Latency relative to the expected peak timing. Lower relative latency values are faster. Shaded regions represent ± 1 SEM. *** $p = 0.001$.

Table 3. Peak latencies in response to the [da] presented in quiet^a

Peak	None	Little	Moderate
23	24.15 (0.97)	24.49 (0.74)	23.37 (0.87)
33	33.75 (1.04)	33.87 (1.28)	33.13 (0.52)
43	43.69 (0.44)	43.39 (0.58)	43.17 (0.54)
53	53.72 (0.63)	53.36 (0.63)	53.08 (0.47)
63	63.51 (0.43)	63.37 (0.37)	63.15 (0.40)
73	73.69 (0.85)	73.44 (0.46)	73.20 (0.44)
83	83.85 (1.27)	83.38 (0.48)	83.23 (0.41)
93	93.76 (1.03)	93.60 (1.15)	93.24 (0.43)
103	103.92 (1.54)	103.43 (0.46)	103.43 (1.08)
113	113.87 (1.14)	113.58 (0.89)	113.14 (0.33)
123	123.63 (0.68)	123.49 (0.54)	123.25 (0.37)
133	134.00 (1.45)	133.56 (0.63)	133.29 (0.39)
143	143.10 (1.61)	143.44 (0.55)	143.25 (0.37)
153	153.90 (1.29)	153.42 (0.50)	153.42 (0.74)
163	164.04 (1.72)	163.52 (0.58)	163.21 (0.42)

^aFirst five rows (peaks 23, 33, 43, 53, 63) refer to the CV transition region of the response where timing differs as a function of group. The group with the most training (Moderate) exhibits the fastest timing in response to the CV transition. Means are reported with SDs.

Table 4. Peak latencies in response to the [da] presented in noise^a

Peak	None	Little	Moderate
23	24.95 (1.17)	24.40 (0.84)	23.54 (0.91)
33	34.15 (1.03)	33.88 (0.87)	33.11 (0.55)
43	43.74 (0.79)	43.60 (0.74)	43.15 (0.54)
53	53.75 (0.77)	52.26 (0.83)	53.13 (0.56)
63	63.71 (0.52)	63.56 (1.15)	63.13 (0.41)
73	73.56 (0.37)	73.40 (0.49)	73.27 (0.54)
83	83.63 (0.36)	83.78 (1.38)	83.13 (0.70)
93	93.64 (0.35)	93.46 (0.50)	93.18 (0.35)
103	103.70 (0.63)	103.25 (0.51)	103.50 (0.83)
113	113.65 (0.27)	113.45 (0.42)	113.32 (0.39)
123	123.59 (0.34)	123.43 (0.42)	123.31 (0.43)
133	133.66 (0.33)	133.27 (0.91)	133.29 (0.45)
143	143.58 (0.37)	143.40 (0.32)	143.26 (0.42)
153	153.70 (0.74)	153.28 (0.40)	153.28 (0.39)
163	163.58 (0.30)	163.46 (0.31)	163.30 (0.40)

^aFirst five rows (peaks 23, 33, 43, 53, 63) refer to the CV transition region of the response where timing differs as a function of group. The group with the most training (Moderate) exhibits the fastest timing in response to the CV transition. Means are reported with SDs.

negative association between years of training and neural timing, indicating that subjects with more years of music training had faster neural responses to speech. However, in quiet, the correlation only reached statistical significance for peak 23 ($\rho_{(26)} = -0.600, p = 0.001$) and in noise for peaks 23 ($\rho_{(26)} = -0.438, p = 0.025$) and 33 ($\rho_{(26)} = -0.494, p = 0.010$). A scatterplot with years of these peak latencies in noise is presented in Figure 4.

Discussion

We compared neural responses to speech in three groups of older adults who reported varying degrees of music training early in their lives. The group with the most music training displayed the fastest neural timing in the response region corresponding to the information-bearing and spectrotemporally dynamic region in a speech syllable, the CV transition, most notably in noise. While this neural enhancement has been observed in older adults with

lifelong music training (Parbery-Clark et al., 2012a), and to a lesser extent in those who have undergone intensive short-term computer training (Anderson et al., 2013b), here we observe this benefit in adults with past music experience ~ 40 years after training stopped.

There are a number of potential mechanisms driving these group differences decades after training. It may be that early music instruction instills a fixed change in the central auditory system that is retained throughout life. Early acoustic experience can have lasting consequences for neural function. For example, rearing rat pups in noise decreases auditory cortical synchrony into adulthood, even after noise exposure has stopped (Zhou and Merzenich, 2008), whereas enrichment early in life promotes auditory processing in adulthood (Engineer et al., 2004; Threlkeld et al., 2009; Sarro and Sanes, 2011). Similar effects are observed in owl auditory midbrain, specifically in central nucleus of IC (Linkenhoker et al., 2005). These principles may apply to the group differences observed here: a moderate amount of training early in life changed subcortical auditory function such that the system responded with faster timing many years later. There are, however, other mechanisms to consider, especially in light of recent evidence that sensory and neocortical systems retain the potential for substantial plasticity into older adulthood (Smith et al., 2009; Berry et al., 2010; Anderson et al., 2013b), including in response to passive stimulus exposure (Kral and Eggermont, 2007).

In addition to instilling fixed changes early in life, music instruction may “set the stage” for future interactions with sound, driving the timing differences we observed between groups (i.e., fastest timing in the group with the greatest amount of past training). Past experiences mapping sounds to meaning—as through music training—may prime the auditory system to interact more dynamically with sound. This could account for the selective relationship between extent of past music training and response timing to the CV transition in speech but not the steady-state vowel. By priming the auditory system to encode sound according to informational saliency and acoustic complexity (Strait et al., 2009; Kraus and Chandrasekaran, 2010), increased neural resources could be devoted to the most relevant acoustic features of auditory scenes. In fact, this may account for the inconsistency in group effects for neural timing in response to speech presented in quiet versus speech presented in noise: whereas the group effect in noise was driven by the Moderate group having faster timing than the Little and None groups, in quiet this was driven by poorer timing in the Little group. Previous studies on musicians have identified stronger training effects in the neural encoding of speech in noise, which may explain these results. This “setting of the stage” may apply especially to degraded acoustic environments that place increased demands on sensory circuitry for accurate target signal transcription.

The nature of an individual’s interactions with sound can alter auditory processing; enriched auditory experiences may improve subcortical auditory function. Assistive listening devices that en-

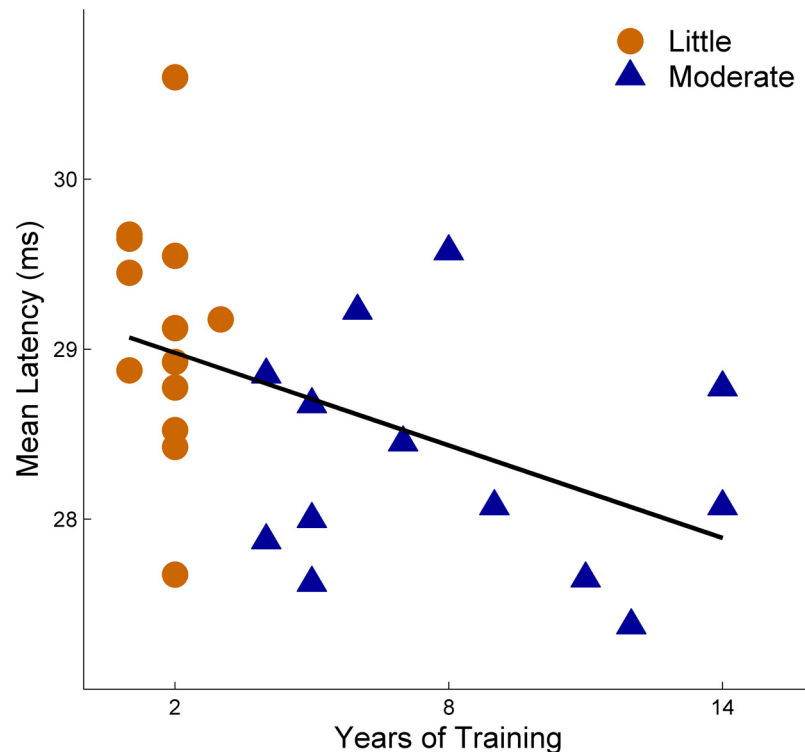


Figure 4. A correlation is observed between years of music training and neural timing. For illustrative purposes, latencies for peaks 23 and 33 in response to the [da] presented in noise were averaged. Faster timing is associated with more years of training. $\rho_{(26)} = -0.539, p = 0.005$.

hance classroom signal-to-noise ratios, for example, improve the stability of children’s subcortical speech-sound processing even after the children stop using the devices (Hornickel et al., 2012). Such devices may assist students in directing attention to important auditory streams even when audibility is returned to normal signal-to-noise levels. Music training may function similarly; after all, a key element of playing music is dynamically directing attention to certain auditory objects or features. Using similar reasoning, Bavelier, Green, and colleagues have pursued the use of action video games to train attentional systems (Green and Bavelier, 2003), arguing that video games enhance attention and thus provide the means to modify neural function (Green and Bavelier, 2012).

Evidence in older adults supports the hypothesis that early or limited training experiences influence subsequent cognitive function and/or learning. For example, older adults with past music training rely more on cognitive mechanisms (such as memory and attention) to understand speech in noise than do older adults with no past music training (Anderson et al., 2013c). In the current study, the members of the Moderate training group may engage with sound using subtly different cognitive mechanisms that modulate auditory processing (cf. Gaab and Schlaug, 2003; Wong et al., 2009). Future work should explore differences in the relative mechanisms used to achieve similar performance on perceptual tasks as a function of previous training experiences. These mechanisms may dictate the nature of interactions with sound after training has stopped to reinforce modulations to subcortical processing in the long term (as reviewed in Kraus and Chandrasekaran, 2010; Strait and Kraus, 2013). Work in the human perceptual learning literature has explored what happens to learning when active training has stopped. Taking breaks from learning (Molloy et al., 2012) or

interspersing active and passive exposure (Wright et al., 2010) may facilitate “latent learning” and improve end performance.

If initial training experiences set the stage for subsequent learning, past music training may promote auditory learning ability or increase the benefits accrued from short-term experiences. In fact, adults with previous music experience perform better on statistical learning tasks (Shook et al., 2013; Skoe et al., 2013), which model naturalistic environmental language learning. Similarly, initial exposure to different statistical “languages” benefits subsequent novel language learning (Graf Estes et al., 2007). Together, this work supports the proposal that limited, initial training experiences interact with later learning to guide the final course of nervous system function. There is also physiologic evidence for early training’s influence on later learning in adulthood. Prior experiences prime the capacity for and efficiency of future plasticity in neocortex (Zelcer et al., 2006; Abraham, 2008; Hofer et al., 2009). Acoustic experiences early in life may also be complemented or counteracted by later training experiences (Threlkeld et al., 2009; Zhou and Merzenich, 2009). Music training may then prime the auditory system to benefit from subsequent auditory experiences, complementing enhancements that are primary effects of the training itself, thereby recapitulating past experiences to guide future interactions with sound (cf. Salimpoor et al., 2013; Skoe and Kraus, 2013).

Up to now, we have proposed that group differences were driven by an initial enhancement from music training that was reinforced throughout life due to music leading to more effortful and meaningful interactions with sound in other contexts. Of course there are other factors that may drive our group differences. For example, we cannot rule out innate differences in central auditory physiology; individuals with better subcortical function may be drawn to play a musical instrument for a longer period of time. We think this unlikely, however, since there are many other factors (family history, personal interest, environment, etc.) that prompt an individual to study music. We also strived to carefully match our groups demographically, especially on factors as IQ and education (Table 1). Finally, we find it particularly striking that the group difference we found (neural timing in response to CV transitions in speech presented in noise) has also been seen in older adults who were randomly assigned to complete short-term training (Kraus and Nicol, 2013; Anderson et al., 2013b); this may be a metric especially suited to demonstrate learning-related enhancements, which would further suggest that the neural timing differences were due to training, and were not innate. We do note, however, that the neural enhancements in the Moderate group, which we attribute to music training, are not as pervasive as those seen in lifelong older adult musicians (Parbery-Clark et al., 2012b; Zendel and Alain, 2012).

It is worth noting that there was a large degree of heterogeneity in the Moderate training group. In this group there was a range of 10 years of training (minimum 4, maximum 14) and a wide array of instruments and means of instruction (classroom, private teacher, etc.) that subjects reported. Unfortunately, it is difficult to quantify the precise amount of training (i.e., number of hours per week) that these subjects pursued since we have to rely on recollections about training many decades ago. That said, it is striking that despite the diversity of this group there was an effect of past training on neural timing many years after the training stopped, with the extent of past training correlating with the extent of neural response enhancement (Fig. 4).

Our findings have important consequences for education and social policy. Today, music education is at high risk for being cut from schools in the United States. School districts with limited

financial resources prioritize science, math, and reading since music is often considered a nonessential component of the curriculum (Rabkin and Hedberg, 2011). Here, we show a neural enhancement linked to a moderate amount of music training many decades after training has stopped. Importantly, this enhancement was for encoding CV transitions in speech, which are especially vulnerable to the effects of age, yet important for everyday communication. These findings support current efforts to reintegrate arts education into schools (President’s Committee on the Arts and the Humanities, 2011; Yajima and Nadarajan, 2013) by suggesting that music training in adolescence and young adulthood may carry meaningful biological benefits into older adulthood.

References

- Abraham WC (2008) Metaplasticity: tuning synapses and networks for plasticity. *Nat Rev Neurosci* 9:387. [CrossRef Medline](#)
- Alain C, Zendel BR, Hutka S, Bidelman GM (2013) Turning down the noise: the benefit of musical training on the aging auditory brain. *Hear Res*. In press.
- Anderson S, Skoe E, Chandrasekaran B, Kraus N (2010) Neural timing is linked to speech perception in noise. *J Neurosci* 30:4922–4926. [CrossRef Medline](#)
- Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N (2012) Aging affects neural precision of speech encoding. *J Neurosci* 32:14156–14164. [CrossRef Medline](#)
- Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N (2013a) Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. *J Speech Lang Hear Res* 56:31–43. [CrossRef Medline](#)
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013b) Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci* 110:4357–4362. [CrossRef Medline](#)
- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013c) A dynamic auditory-cognitive system supports speech-in-noise perception in older adults. *Hear Res* 300:18–32. [CrossRef Medline](#)
- Berry AS, Zanto TP, Clapp WC, Hardy JL, Delahunt PB, Mahncke HW, Gazzaley A (2010) The influence of perceptual training on working memory in older adults. *PLoS One* 5:e11537. [CrossRef Medline](#)
- Bidelman GM, Gandour JT, Krishnan A (2011) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J Cogn Neurosci* 23:425–434. [CrossRef Medline](#)
- Byrne D, Dillon H (1986) The National Acoustic Laboratories’ (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear* 7:257–265. [CrossRef Medline](#)
- Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology* 47:236–246. [CrossRef Medline](#)
- de Villers-Sidani E, Alzghoul L, Zhou X, Simpson KL, Lin RC, Merzenich MM (2010) Recovery of functional and structural age-related changes in the rat primary auditory cortex with operant training. *Proc Natl Acad Sci U S A* 107:13900–13905. [CrossRef Medline](#)
- Engineer ND, Percaccio CR, Pandya PK, Moucha R, Rathbun DL, Kilgard MP (2004) Environmental enrichment improves response strength, threshold, selectivity, and latency of auditory cortex neurons. *J Neurophysiol* 92:73–82. [CrossRef Medline](#)
- Gaab N, Schlaug G (2003) The effect of musicianship on pitch memory in performance matched groups. *Neuroreport* 14:2291–2295. [CrossRef Medline](#)
- Gazzaley A, Cooney JW, Rissman J, D’Esposito M (2005) Top-down suppression deficit underlies working memory impairment in normal aging. *Nat Neurosci* 8:1298–1300. [CrossRef Medline](#)
- Graf Estes K, Evans JL, Alibali MW, Saffran JR (2007) Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol Sci* 18:254–260. [CrossRef Medline](#)
- Green CS, Bavelier D (2012) Learning, attentional control, and action video games. *Curr Biol* 22:R197–R206. [CrossRef Medline](#)
- Green CS, Bavelier D (2003) Action video game modifies visual selective attention. *Nature* 423:534–537. [CrossRef Medline](#)
- Hall JW III (2006) *New handbook for auditory evoked responses*. Harlow, UK: Pearson.

- Hanna-Pladdy B, MacKay A (2011) The relation between instrumental musical activity and cognitive aging. *Neuropsychology* 25:378–386. [CrossRef Medline](#)
- Herholz SC, Zatorre RJ (2012) Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron* 76:486–502. [CrossRef Medline](#)
- Hofer SB, Mrcic-Flogel TD, Bonhoeffer T, Hübener M (2009) Experience leaves a lasting structural trace in cortical circuits. *Nature* 457:313–317. [CrossRef Medline](#)
- Hornickel J, Zecker SG, Bradlow AR, Kraus N (2012) Assistive listening devices drive neuroplasticity in children with dyslexia. *Proc Natl Acad Sci U S A* 109:16731–16736. [CrossRef Medline](#)
- Johnson KL, Nicol T, Zecker SG, Kraus N (2008) Developmental plasticity in the human auditory brainstem. *J Neurosci* 28:4000–4007. [CrossRef Medline](#)
- Kral A, Eggermont JJ (2007) What's to lose and what's to learn: development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Res Rev* 56:259–269. [CrossRef Medline](#)
- Kraus N, Chandrasekaran B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605. [CrossRef Medline](#)
- Kraus N, Nicol T (2013) The cognitive auditory system. In: *Springer handbook of auditory research* (Fay R, Popper A, eds), in press. New York: Springer.
- Krishnan A, Swaminathan J, Gandour JT (2009) Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *J Cogn Neurosci* 21:1092–1105. [CrossRef Medline](#)
- Linkenhoker BA, Knudsen EI (2002) Incremental training increases the plasticity of the auditory space map in adult barn owls. *Nature* 419:293–296. [CrossRef Medline](#)
- Linkenhoker BA, von der Ohe CG, Knudsen EI (2005) Anatomical traces of juvenile learning in the auditory system of adult barn owls. *Nat Neurosci* 8:93–98. [Medline](#)
- Molloy K, Moore DR, Sohoglu E, Amitay S (2012) Less is more: latent learning is maximized by shorter training sessions in auditory perceptual learning. *PLoS ONE* 7:e36929. [CrossRef Medline](#)
- Moreno S, Marques C, Santos A, Santos M, Castro SL, Besson M (2009) Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. *Cereb Cortex* 19:712–723. [CrossRef Medline](#)
- Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H (2005) The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 53:695–699. [CrossRef Medline](#)
- Parbery-Clark A, Skoe E, Kraus N (2009) Musical experience limits the degradative effects of background noise on the neural processing of sound. *J Neurosci* 29:14100–14107. [CrossRef Medline](#)
- Parbery-Clark A, Anderson S, Hittner E, Kraus N (2012a) Musical experience offsets age-related delays in neural timing. *Neurobiol Aging* 33:1483.e1–1483.e4. [CrossRef Medline](#)
- Parbery-Clark A, Anderson S, Hittner E, Kraus N (2012b) Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Front Aging Neurosci* 4:30. [CrossRef Medline](#)
- Pichora-Fuller MK, Schneider BA, Macdonald E, Pass HE, Brown S (2007) Temporal jitter disrupts speech intelligibility: a simulation of auditory aging. *Hear Res* 223:114–121. [CrossRef Medline](#)
- President's Committee on the Arts and the Humanities (2011) Re-investing in arts education: winning America's future through creative schools. Washington, D.C.
- Rabkin N, Hedberg E (2011) Arts education in America: What the declines mean for arts participation. Washington, D.C.: National Endowment for the Arts.
- Recanzone GH, Engle JR, Juarez-Salinas DL (2011) Spatial and temporal processing of single auditory cortical neurons and populations of neurons in the macaque monkey. *Hear Res* 271:115–122. [CrossRef Medline](#)
- Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2012) Why middle-aged listeners have trouble hearing in everyday settings. *Curr Biol* 22:1417–1422. [CrossRef Medline](#)
- Russo NM, Hornickel J, Nicol T, Zecker S, Kraus N (2010) Biological changes in auditory function following training in children with autism spectrum disorders. *Behav Brain Funct* 6:60. [CrossRef Medline](#)
- Salimpoor VN, van den Bosch I, Kovacevic N, McIntosh AR, Dagher A, Zatorre RJ (2013) Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* 340:216–219. [CrossRef Medline](#)
- Sarro EC, Sanes DH (2011) The cost and benefit of juvenile training on adult perceptual skill. *J Neurosci* 31:5383–5391. [CrossRef Medline](#)
- Shook A, Marian V, Bartolotti J, Schroeder SR (2013) Musical experience influences statistical learning of a novel language. *Am J Psychol* 126:95–104. [Medline](#)
- Skoe E, Kraus N (2010) Auditory brain stem response to complex sounds: a tutorial. *Ear Hear* 31:302–324. [CrossRef Medline](#)
- Skoe E, Kraus N (2012) A little goes a long way: how the adult brain is shaped by musical training in childhood. *J Neurosci* 32:11507–11510. [CrossRef Medline](#)
- Skoe E, Kraus N (2013) Musical training heightens auditory brainstem function during sensitive periods in development. *Front Psychol* 4:622. [CrossRef Medline](#)
- Skoe E, Krizman J, Spitzer E, Kraus N (2013) The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience* 243:104–114. [CrossRef Medline](#)
- Smith GE, Housen P, Yaffe K, Ruff R, Kennison RF, Mahncke HW, Zelinski EM (2009) A cognitive training program based on principles of brain plasticity: results from the Improvement in Memory with Plasticity-Based Adaptive Cognitive Training (IMPACT) study. *J Am Geriatr Soc* 57:594–603. [CrossRef Medline](#)
- Strait DL, Kraus N (2013) Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hear Res*. In press.
- Strait DL, Kraus N, Skoe E, Ashley R (2009) Musical experience and neural efficiency—effects of training on subcortical processing of vocal expressions of emotion. *Eur J Neurosci* 29:661–668. [CrossRef Medline](#)
- Strait DL, Parbery-Clark A, Hittner E, Kraus N (2012) Musical training during early childhood enhances the neural encoding of speech in noise. *Brain Lang* 123:191–201. [CrossRef Medline](#)
- Strait DL, O'Connell S, Parbery-Clark A, Kraus N (2013) Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages three to thirty. *Cereb Cortex*. Advance online publication. Retrieved Sept. 24, 2013. [Medline](#)
- Tallal P (1980) Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang* 9:182–198. [CrossRef Medline](#)
- Threlkeld SW, Hill CA, Rosen GD, Fitch RH (2009) Early acoustic discrimination experience ameliorates auditory processing deficits in male rats with cortical developmental disruption. *Int J Dev Neurosci* 27:321–328. [CrossRef Medline](#)
- Van Engen KJ, Bradlow AR (2007) Sentence recognition in native-and foreign-language multi-talker background noise. *J Acoust Soc Am* 121:519–526. [CrossRef Medline](#)
- Wong PC, Jin JX, Gunasekera GM, Abel R, Lee ER, Dhar S (2009) Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* 47:693–703. [CrossRef Medline](#)
- Wright BA, Sabin AT, Zhang Y, Marrone N, Fitzgerald MB (2010) Enhancing perceptual learning by combining practice with periods of additional sensory stimulation. *J Neurosci* 30:12868–12877. [CrossRef Medline](#)
- Yajima R, Nadarajan G (2013) Benefits beyond beauty: integration of art and design into STEM education and research. Symposium: Feb. 18, 2013. Boston: American Association for the Advancement of Science.
- Zelcer I, Cohen H, Richter-Levin G, Lebiosn T, Grossberger T, Barkai E (2006) A cellular correlate of learning-induced metaplasticity in the hippocampus. *Cereb Cortex* 16:460–468. [Medline](#)
- Zendel BR, Alain C (2012) Musicians experience less age-related decline in central auditory processing. *Psychol Aging* 27:410–417. [CrossRef Medline](#)
- Zhou X, Merzenich MM (2008) Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. *Proc Natl Acad Sci U S A* 105:4423–4428. [CrossRef Medline](#)
- Zhou X, Merzenich MM (2009) Developmentally degraded cortical temporal processing restored by training. *Nat Neurosci* 12:26–28. [CrossRef Medline](#)
- Zhu J, Kaufman E (1999) The Weschler Abbreviated Scale of Intelligence (WASI). New York: Psychological.