

# Listening to the Brainstem: Musicianship Enhances Intelligibility of Subcortical Representations for Speech

Michael W. Weiss<sup>1</sup>\* and  Gavin M. Bidelman<sup>2,3</sup>\*

<sup>1</sup>Department of Psychology, University of Toronto Mississauga, Mississauga, Ontario L5L 1C6, Canada and <sup>2</sup>Institute for Intelligent Systems and <sup>3</sup>School of Communication Sciences and Disorders, University of Memphis, Memphis, Tennessee 38105

Auditory experiences including musicianship and bilingualism have been shown to enhance subcortical speech encoding operating below conscious awareness. Yet, the behavioral consequence of such enhanced subcortical auditory processing remains undetermined. Exploiting their remarkable fidelity, we examined the intelligibility of auditory playbacks (i.e., “sonifications”) of brainstem potentials recorded in human listeners. We found naive listeners’ behavioral classification of sonifications was faster and more categorical when evaluating brain responses recorded in individuals with extensive musical training versus those recorded in nonmusicians. These results reveal stronger behaviorally relevant speech cues in musicians’ neural representations and demonstrate causal evidence that superior subcortical processing creates a more comprehensible speech signal (i.e., to naive listeners). We infer that neural sonifications of speech-evoked brainstem responses could be used in the early detection of speech–language impairments due to neurodegenerative disorders, or in objectively measuring individual differences in speech reception solely by listening to individuals’ brain activity.

**Key words:** brainstem response; categorical speech perception; experience-dependent plasticity; frequency-following response

## Introduction

There is growing interest to identify human experiences and activities which positively impact brain function and consequently benefit perceptual-cognitive skills (Herholz and Zatorre, 2012). Among those benefiting behavior, musical training (i.e., learning to play an instrument) places high demands on a wide variety of sensory and higher-order processes (Kraus and Chandrasekaran, 2010; Moreno and Bidelman, 2014), and has been shown to produce pervasive neuroplastic changes throughout the brain (Kraus and Chandrasekaran, 2010; Herholz and Zatorre, 2012; Moreno and Bidelman, 2014). Aspects of music and language share common neural substrates (Patel, 2003). Indeed, musical aptitude has been linked to lifelong gains in linguistic skills (Wong et al., 2007; Bidelman and Krishnan, 2010; Bidelman et al., 2011; Skoe and Kraus, 2012; Bidelman and Alain, 2015) including the reception and production of sound structures in language (Slevc and Miyake, 2006). Music-related neuroplasticity has generally been ascribed to changes in neocortical circuitry (Patel, 2003). Yet, measuring frequency-following responses (FFRs), a neuromicrophonic potential generated in the rostral brainstem coding spectrotemporal properties of acoustic signals (Bidelman et al., 2013), recent studies demonstrate that music and language experience tune the subcortical transcription of com-

plex sounds, improving the fidelity with which linguistic features of speech are mapped to brain representations milliseconds after entering the ear (Wong et al., 2007; Bidelman and Krishnan, 2010; Bidelman et al., 2011; Parbery-Clark et al., 2012a). Despite its relevance to neuroscientific, clinical, and educational communities, the direct behavioral consequences of subcortical speech enhancements have not been explored. Identifying a causal relationship between enhanced brainstem function resulting from auditory experiences and improved behavioral outcomes would suggest that subcortical processing could be used as an objective “barometer” of speech and other developmental listening skills (Kraus and Chandrasekaran, 2010; Anderson et al., 2013).

We collected behavioral psychometric responses in a standard categorical perception (CP) task (Pisoni, 1973; Bidelman et al., 2014a). In a novel stimulus approach, rather than classifying speech material, participants listened to and identified digitized audio versions (i.e., “sonifications”) of neural brainstem responses, recorded in our previous study comparing the effects of musicianship on the neurophysiological encoding of speech (Bidelman et al., 2014a). Given that musical training is known to enhance subcortical auditory representations and the brainstem response to speech (Bidelman et al., 2011, 2014a; Parbery-Clark et al., 2012b; Skoe and Kraus, 2012; White-Schwoch et al., 2013; Bidelman and Alain, 2015), we hypothesized that naive listeners would show more robust and faster classification when listening to playbacks of musicians’ relative to nonmusicians’ brainstem potentials.

## Materials and Methods

**Participants.** Listeners were 14 young adults (8 female; mean  $\pm$  SD, 28.7  $\pm$  3.2 years) recruited from the University of Memphis student body. All participants reported normal hearing at the time of testing, were native speakers of English without fluency in other languages, were

Received Sept. 3, 2014; revised Nov. 26, 2014; accepted Dec. 3, 2014.

Author contributions: M.W.W. and G.M.B. designed research; M.W.W. and G.M.B. performed research; G.M.B. analyzed data; M.W.W. and G.M.B. wrote the paper.

This work was supported by grants from the GRAMMY Foundation to G.M.B.

\*M.W.W. and G.M.B. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Gavin M. Bidelman, School of Communication Sciences and Disorders, University of Memphis, 807 Jefferson Avenue, Memphis, TN 38105. E-mail: g.bidelman@memphis.edu.

DOI:10.1523/JNEUROSCI.3680-14.2015

Copyright © 2015 the authors 0270-6474/15/351687-05\$15.00/0

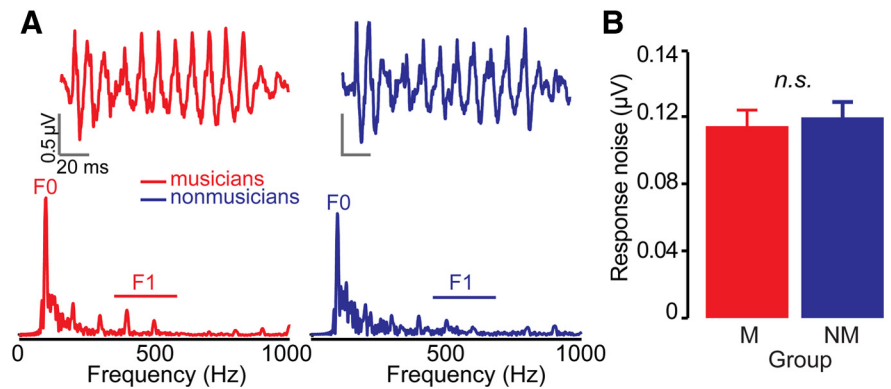
right-handed ( $89.3 \pm 14.4\%$  Edinburgh handedness laterality; Oldfield, 1971), and had obtained similar levels of formal education ( $19.2 \pm 1.9$  years). Musicianship enhances categorical speech perception (Bidelman et al., 2014a); thus, participants were required to have minimal ( $<3$  years) musical training throughout their lifetime ( $1.6 \pm 2.3$  years). Participants provided written informed consent in accordance with a protocol approved by the University of Memphis Institutional Review Board.

**Auditory stimuli derived from human brainstem responses.** Auditory stimuli were created by converting human brainstem FFRs (neural responses) into digital audio files. Brainstem responses were originally recorded in response to a five-step synthetic vowel continuum described in our previous reports (Bidelman et al., 2013, 2014a). These tokens differ minimally acoustically, but are perceived categorically (Pisoni, 1973). Tokens were 100 ms in duration and contained identical voice fundamental (F0), second (F2), and third formant (F3) frequencies (F0: 100, F2: 1090, and F3: 2350 Hz, respectively). The critical stimulus variation was achieved by parameterizing the first formant (F1) over five equal steps between 430 and 730 Hz such that the resultant stimulus set spanned a perceptual phonetic continuum from /u/ to /a/.

Electrophysiological recordings followed typical procedures from our laboratory (Bidelman and Krishnan, 2010; Bidelman et al., 2013, 2014a) and are detailed in our original report (Bidelman et al., 2014a). Briefly, FFRs were elicited in response to the 2000 trials of each of the five speech tokens presented binaurally at 83 dB SPL at a repetition rate of 4/s. EEGs were recorded differentially using a vertical montage with electrodes placed on the high forehead at the hairline ( $\sim$ Fz) referenced to linked mastoids (A1/A2). EEGs were digitized at 20 kHz with an online pass-band of 0.05–3500 Hz (NeuroScan SymAmps2 amplifiers). Traces were segmented ( $-40$  to 210 ms epoch window), baselined to the respective prestimulus period, and subsequently averaged in the time domain to obtain FFRs for each vowel token. Trials exceeding  $\pm 50 \mu\text{V}$  were rejected as artifacts before averaging. Grand averaged evoked responses were then bandpass filtered (80–2500 Hz) to isolate the brainstem FFR from the slower, low-frequency cortical potentials (Bidelman et al., 2013).

Musicians FFRs represent the grand average brainstem response obtained from each vowel token from a cohort ( $n = 12$ ) of young, normal-hearing adults with extensive formal musical training ( $13.6 \pm 4.5$  years); nonmusician FFRs reflect average responses obtained from a cohort ( $n = 12$ ) of age-matched, normal hearing individuals with little, if any, musical training ( $0.4 \pm 0.7$  years). Additional demographics (e.g., instrumental backgrounds, education), EEG acquisition parameters, and neurophysiological response description can be found in our previous EEG study (Bidelman et al., 2014a). Under investigation here is the perception of auditory stimuli created from these neural responses. In the current study, stimuli were generated by transforming recorded FFR potentials of each group into CD-quality audio WAV files (44.1 kHz, 16 bit) using the “wavwrite” command in MATLAB (MathWorks). Thus, there were five stimuli (i.e., sonifications) generated from each groups’ brainstem responses (i.e., musicians’ and nonmusicians’ FFRs). Prior work has shown that when replayed as auditory stimuli, FFRs are highly intelligible to human listeners (Galbraith et al., 1995). Before presentation, speech tokens were gated with 5 ms ramps to reduce spectral splatter and audible clicks at the waveforms’ on/offsets. Stimuli were then lowpass filtered (cutoff frequency = 3200 Hz;  $-6$  dB/oct slope) and the rms normalized to equate the overall bandwidth and intensity between tokens.

Evoked potentials are intrinsically noisy and suffer from low signal-to-noise ratio. Differences in the noise level between musician and nonmusician FFR recordings could influence their behavioral classification. To rule out this potential confound, we measured the noise-floor of the original FFR recordings (Skoe et al., 2013; Bidelman et al., 2014a). Noise was computed as the mean amplitude of the time waveforms in the cumulative interstimulus periods ( $-40$  to 0 ms and 150–210 ms win-



**Figure 1.** Brainstem FFRs recorded in musicians and nonmusicians (Bidelman et al., 2014a). **A**, Response time waveforms (top) and frequency spectra (bottom) for the /u/ token of the continuum. **B**, Enhanced speech encoding in musicians is due to increased representation of speech cues [e.g., voice pitch (F0), formant (F1) cues] and not differences in recording noise level.

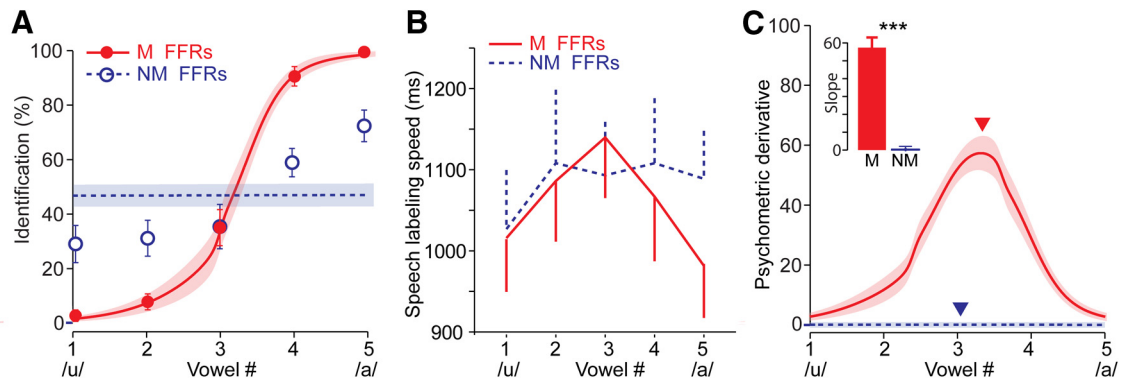
dows). These segments fall immediately before and after the time-locking stimulus, contain no phase-locked neural activity, and thus provided a proxy measure of the noise level of the electrophysiological recordings (Bidelman et al., 2014b). Critically, average noise level (collapsed across the stimulus continuum) did not differ between groups (independent samples *t* test; two-tailed:  $t_{(11)} = -0.35$ ,  $p = 0.74$ ; Fig. 1B) confirming that differences in response fidelity were not due to nonmusicians having noisier recordings than musicians.

**CP listening task and experimental procedures.** We conducted a rapid speech classification task (Pisoni, 1973; Bidelman et al., 2013, 2014a) using the audio versions of FFRs recorded in musicians and nonmusicians (Bidelman et al., 2014a). Listeners were tested individually in a sound-attenuating booth (Industrial Acoustics) using high-fidelity headphones (Sennheiser HD 280). Stimuli were presented on a PC at a comfortable volume ( $\sim 70$  dB SPL). On each trial, listeners heard one of the five tokens of a given set (musician vs nonmusician neural responses) drawn randomly from the continuum and were asked to label it via a binary button press as quickly and as accurately as possible (/u/ or /a/). Subjects were encouraged to emphasize accuracy over speed. Listeners heard musician and nonmusician FFRs in different blocks, ordered randomly and counterbalanced across participants. Fifty repetitions were obtained for each vowel (250 total trials) in a given block. A brief silence (jittered between 800 and 1200 ms; rectangular distribution) separated consecutive trials. Stimulus presentation and response collection was achieved using a custom GUI coded in MATLAB.

Customarily, a pairwise (e.g., 1 vs 2; 2 vs 3, etc.) discrimination task is a complement to identification functions in CP paradigms. A discrimination task is undesirable in the current study given the nonuniform background (neural) noise between the FFR stimulus tokens. Differences in stimulus noise would have inadvertently provided listeners a cue for making discrimination judgments independent of F1 (the cue of interest). This confound would have rendered discrimination judgments largely uninterpretable so we explicitly avoided a discrimination task in this study.

**Behavioral data analysis.** We calculated identification scores at each step in the continuum by computing the percentage of trials for which listeners selected each vowel class (i.e., /u/ vs /a/). Individual identification scores were also fit with a two-parameter sigmoidal function. We used standard logistic regression:  $p = 1/[1 + e^{-\beta_1(x - \beta_0)}]$ , where  $p$  is the proportion of trials identified as a given vowel,  $x$ , the step number along the stimulus continuum, and  $\beta_0$  and  $\beta_1$  the location and slope of the logistic fit estimated using nonlinear least-squares regression (Bidelman et al., 2013, 2014a; Fig. 2A).

We quantified the “steepness” of the categorical speech boundary when listening to musician and nonmusician continua by computing the first derivative of fitted psychometric functions (Fig. 2C). The peak maxima of these curves were used to quantify the degree to which musician and nonmusician brainstem responses produced categorical speech perception; higher slope values are indicative of more dichotomous pho-



**Figure 2.** Enhanced subcortical encoding supports categorical speech percepts. **A**, Grand-average psychometric functions. Identification scores were fit with sigmoidal logistic functions for each listener's perception of musician and nonmusician brainstem potentials. **B**, Speech labeling speeds for vowel classification. When listening to acoustic playbacks of musicians' neural FFRs, listeners hear a clear perceptual shift in the phonetic category (/u/ vs /a/) and are much slower at labeling stimuli near the categorical boundary (vw3) compared with within-category tokens (e.g., vws1–2 or 4–5), the hallmarks of categorical perception (Pisoni and Tash, 1974; Bidelman et al., 2013). In contrast, nonmusicians' FFRs are perceived at near chance levels and identified with the same RT. As such, they do not allow listeners to accurately classify speech. **C**, Slopes of psychometric functions derived from the derivative of listeners' identification functions in **A**. Inset, Maximal slope of the categorical boundary (max slope =  $\blacktriangledown$ ). Musicians' neural speech representations offer steeper, more dichotomous speech percepts that delineate a clearer categorical boundary than nonmusicians';  $***p < 0.001$ ; error bars/shading =  $\pm$ SEM.

netic boundaries and categorical brain organization (Bidelman et al., 2014a). Higher psychometric slopes when listening to musician responses would suggest that their subcortical representations carry more behaviorally relevant information of the speech signal (Bidelman et al., 2014a). In addition to percentage identification, we also analyzed stimulus labeling speeds (i.e., reaction times), computed from each listener's mean response latency across trials for a given vowel condition within each FFR type (musician vs nonmusician responses). Prior work has shown that when perceived categorically, listeners take longer to identify stimuli near the CP boundary compared with tokens near each end of the continuum (Pisoni and Tash, 1974; Bidelman et al., 2013, 2014a).

**Statistical analysis.** We analyzed behavioral measures (percentage identification and labeling speed) using mixed model ANOVAs with subjects as a random factor and two fixed, within-subject factors: group (2 levels: musician vs nonmusician FFRs) and stimulus (5 levels: vowels 1–5). *Post hoc* pairwise comparisons were corrected for type I error inflation using Bonferroni adjustments. Noise level of the original FFR recordings was evaluated by computing the noise floor of each individual subject's original FFR response (Bidelman et al., 2014a) per stimulus and group. For each group, we then collapsed noise measurements across the five stimuli and conducted an independent samples *t* test to confirm that inherent neural noise levels between musicians' and nonmusicians' FFRs did not differ. We evaluated statistical significance between psychometric slopes when listening to musician and nonmusician response continua using a two-tailed, paired samples *t* test.

## Results

Exploiting the fact that scalp-recorded FFRs (a neural response) are highly intelligible when replayed to human listeners (Galbraith et al., 1995), we asked observers to rapidly judge audio playbacks (i.e., "sonifications") of brainstem responses originally recorded in cohorts of individuals with and without extensive auditory experience (i.e., musicians with decades of training vs nonmusicians) in response to a speech vowel continuum (/u/ to /a/; Bidelman et al., 2014a). We have previously shown superior brainstem encoding for these categorical speech sounds in trained musicians (Bidelman et al., 2014a). Under investigation here is whether or not brainstem responses to these stimuli, elicited from listeners with varying auditory expertise, contain sufficient acoustic information for lay observers to properly identify the original speech input.

When judging musicians' neuroelectric response sonifications, listeners showed clear signs of categorical speech perception in both identification and speed not observed for nonmusicians' brainstem

response sonifications (Fig. 2). We found a group  $\times$  stimulus interaction on identification scores ( $F_{(4,52)} = 85.32$ ,  $p < 0.001$ ; Fig. 2A). Whereas each of the five syllables was identified at near chance levels (i.e., relatively invariantly) when listening to nonmusician FFRs, listeners' heard a clear phonetic shift (/u/ vs /a/) when listening to musicians' brainstem responses. We further observed differences in the pattern of speech labeling speeds when judging musicians' and nonmusicians' speech-FFRs (group  $\times$  stimulus:  $F_{(4,52)} = 2.71$ ,  $p = 0.039$ ). Planned comparisons revealed that for musicians' FFRs, identification was faster at ends of the continuum and slowed near the CP boundary [(vw1, vw2, vw4, vw5) vs vw3 contrast:  $t_{(52)} = 3.83$ ,  $p = 0.0007$ ], indicative of categorical perception (Pisoni and Tash, 1974; Bidelman et al., 2013, 2014a). However, speech labeling speeds were invariant when judging nonmusicians' brainstem responses ( $t_{(52)} = 0.38$ ,  $p = 1.0$ ; Fig. 2B).

We analyzed the slope of listeners' psychometric functions for the two continua to quantify the degree to which neural FFRs were perceived categorically and hence, maintained the original phonetic characteristics of the evoking stimuli (Bidelman et al., 2014a). We computed slopes via the first derivative of psychometric identification functions. Behavioral identification curves were considerably steeper when performing the CP task listening to musicians' FFRs ( $t_{(13)} = 16.43$ ,  $p < 0.0001$ ; Fig. 2C). Collectively, these findings suggest that listeners were poorer at categorizing original signal identity when listening to nonmusicians' speech-evoked brainstem responses whereas the neural code in musicians fully supported robust phonetic (i.e., categorical) decisions.

## Discussion

Our results offer the first empirical evidence that differences in human subcortical processing (Kraus and Chandrasekaran, 2010; Bidelman et al., 2011; Parbery-Clark et al., 2012a) causally affect speech perception by providing a more comprehensible speech signal to naive listeners. We found that neural representations for speech were more intelligible to listeners when judging human brainstem responses recorded from trained musicians than those of nonmusicians (Bidelman et al., 2014a). These findings provide compelling evidence that experience-dependent plasticity of the brainstem allows the nervous system to carry more behaviorally relevant information of communicative signals (Bidelman et al., 2014a) and

support faster and clearer identification of categorical speech percepts. In contrast, nonmusicians' poorer auditory brainstem responses led to more "fuzzy" neural sonifications of speech, which were much harder for external listeners to classify. Our findings are highly relevant to recent studies documenting changes in subcortical speech encoding with short- (Chandrasekaran et al., 2012; Anderson et al., 2013) and long-term (Bidelman and Krishnan, 2010; Kraus and Chandrasekaran, 2010; Bidelman et al., 2011; Parbery-Clark et al., 2012a; Skoe and Kraus, 2012; Bidelman and Alain, 2015) auditory training by confirming that preattentive brainstem response enhancements support perceptually clearer speech signals.

Although our data show that enhanced brainstem speech representations confer more salient perceptual cues for external listeners' comprehension, whether or not superior brainstem processing is responsible for a listener's own improved perception cannot be directly assessed in the current data. However, recent brainstem studies do indicate that improved subcortical representations correlate well with individual differences in speech identification in both younger (Bidelman et al., 2014a) and older adults (Bidelman et al., 2014b; Bidelman and Alain, 2015). Moreover, training-related improvements in auditory perceptual learning (Carcagno and Plack, 2011) and speech recognition tasks (Chandrasekaran et al., 2012; Song et al., 2012) are directly related to the degree of neural changes that occur in brainstem processing during learning. Together, these studies indicate that perceptual benefits with speech are determined, at least in part, by the quality of nascent sensory representations operating below cerebral cortex and at preattentive stages of auditory processing.

The current study used auditory playbacks of brainstem potentials to infer the degree to which neural responses carried categorical information of speech. Invoking a *reductio ad absurdum* argument, one can ask at what point recursive brainstem responses would cease to be perceived categorically. That is, would sonifications of brainstem responses to the original FFR sonifications continue to be identifiable by listeners? Although a provocative thought experiment, we find it unlikely that perception would survive another iteration of sonifications. Spectrotemporal characteristics of the brainstem FFR appear as a quasi-low-pass filtered and noisier version of the acoustic input. Given the relatively low SNR of scalp-recorded potentials, neural responses to "FFR stimuli" would contain substantially higher degrees of noise; noise levels would only be exacerbated in additional iterations. Additive noise causes prominent disruptions (i.e., timing delays and reduced magnitudes) in the speech-evoked FFR (Bidelman and Krishnan, 2010; Song et al., 2012). Thus, it is likely that recursive sonifications would fail to elicit strong enough phase-locking to promote the iterative encoding of acoustic speech features necessary for robust identification.

A natural question emerges as to what aspects of the original speech signal are better maintained in musicians' brainstem response sonifications to support more robust translation to perception. That is, are there objective measures that could be used to quantify musicians' and nonmusicians' sonifications *vis a vis* the original stimuli? To elucidate this question, we compared acoustic characteristics of the original speech tokens to each group's brainstem sonifications. Collapsed across the continuum, the spectra (i.e., FFT) of musicians' brainstem responses showed higher correlations with stimulus spectra than those of nonmusicians [Pearson's correlations ( $r$ ):  $r_M = 0.55, p < 0.0001$ ;  $r_{NM} = 0.46, p < 0.0001$ ; Fisher  $r$ -to- $z$  transform test on difference in correlation between groups:  $z = 8.15, p < 0.0001$ ]. This finding confirms a more faithful depiction of speech acoustics in musi-

cians' subcortical brain representations. Stronger representation for the spectral features of speech stimuli may have supported the improved behavioral categorization of neural sonifications observed here. We infer that musical expertise may act to "warp" or restrict the neural encoding of speech sounds near categorical speech boundaries (Bidelman and Alain, 2015), and consequently, offer stronger, more dichotomous phonetic cues in sonification playbacks. This notion is supported by our recent neuroimaging studies which demonstrate not only more dichotomous perceptual classification but also more salient categorical neural encoding of speech in musically trained individuals (Bidelman et al., 2014a; Bidelman and Alain, 2015).

Our results further highlight a common (subcortical) mechanism to account for the widespread linguistic benefits associated with intense musical training (Slevc and Miyake, 2006; Kraus and Chandrasekaran, 2010; Herholz and Zatorre, 2012; Moreno and Bidelman, 2014) and underscore the important role of early, preattentive auditory sensory processing in determining language-learning success (Chandrasekaran et al., 2012; Anderson et al., 2013). More broadly, our novel stimulus approach resonates with emerging assistive technologies, the translation of data into sound, that detect medical symptoms or diseases based on aural "sonification" of nonauditory data (e.g., identifying abnormal tissue by listening to acoustic analogs of MRI images; Barrass and Kramer, 1999). Auditory EEG displays, as used here, could be applied to earlier detection of speech-language impairments due to neurodegenerative disorders (Barrass and Kramer, 1999) or to objectively measure individual differences in the ability to understand speech, solely by listening to individuals' brain activity.

## References

- Anderson S, White-Schwoch T, Parbery-Clark A, Kraus N (2013) Reversal of age-related neural timing delays with training. *Proc Natl Acad Sci U S A* 110:4357–4362. [CrossRef Medline](#)
- Barrass S, Kramer G (1999) Using sonification. *Multimedia Systems* 7:23–31. [CrossRef](#)
- Bidelman GM, Alain C (2015) Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J Neurosci* 35:1240–1249. [CrossRef](#)
- Bidelman GM, Krishnan A (2010) Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res* 1355:112–125. [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](#)
- Bidelman GM, Moreno S, Alain C (2013) Tracing the emergence of categorical speech perception in the human auditory system. *Neuroimage* 79: 201–212. [CrossRef Medline](#)
- Bidelman GM, Weiss MW, Moreno S, Alain C (2014a) Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur J Neurosci* 40:2662–2673. [CrossRef Medline](#)
- Bidelman GM, Villafuerte JW, Moreno S, Alain C (2014b) Age-related changes in the subcortical-cortical encoding and categorical perception of speech. *Neurobiol Aging* 35:2526–2540. [CrossRef Medline](#)
- Carcagno S, Plack CJ (2011) Subcortical plasticity following perceptual learning in a pitch discrimination task. *J Assoc Res Otolaryngol* 12:89–100. [CrossRef](#)
- Chandrasekaran B, Kraus N, Wong PC (2012) Human inferior colliculus activity relates to individual differences in spoken language learning. *J Neurophysiol* 107:1325–1336. [CrossRef Medline](#)
- Galbraith GC, Arbagey PW, Branski R, Comerci N, Rector PM (1995) Intelligent speech encoded in the human brain stem frequency-following response. *Neuroreport* 6:2363–2367. [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](#)

- Kraus N, Chandrasekaran B (2010) Music training for the development of auditory skills. *Nat Rev Neurosci* 11:599–605. [CrossRef Medline](#)
- Moreno S, Bidelman GM (2014) Understanding neural plasticity and cognitive benefit through the unique lens of musical training. *Hear Res* 308:84–97. [CrossRef Medline](#)
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113. [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 e14. [CrossRef Medline](#)
- Parbery-Clark A, Tierney A, Strait DL, Kraus N (2012b) Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* 219:111–119. [CrossRef Medline](#)
- Patel AD (2003) Language, music, syntax and the brain. *Nat Neurosci* 6:674–681. [CrossRef Medline](#)
- Pisoni DB (1973) Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Percept Psychophys* 13:253–260. [CrossRef Medline](#)
- Pisoni DB, Tash J (1974) Reaction times to comparisons within and across phonetic categories. *Percept Psychophys* 15:285–290. [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, Krizman J, Anderson S, Kraus N (2013) Stability and plasticity of auditory brainstem function across the lifespan. *Cereb Cortex*. Advance online publication. Retrieved Dec. 22, 2013. [CrossRef Medline](#)
- Slevc LR, Miyake A (2006) Individual differences in second-language proficiency: does musical ability matter? *Psychol Sci* 17:675–681. [CrossRef Medline](#)
- Song JH, Skoe E, Banai K, Kraus N (2012) Training to improve hearing speech in noise: biological mechanisms. *Cereb Cortex* 22:1180–1190. [CrossRef Medline](#)
- White-Schwoch T, Woodruff Carr K, Anderson S, Strait DL, Kraus N (2013) Older adults benefit from music training early in life: biological evidence for long-term training-driven plasticity. *J Neurosci* 33:17667–17674. [CrossRef Medline](#)
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422. [CrossRef Medline](#)