PNAS
www.pnas.org

Supplementary Information for

Inherent auditory skills rather than formal music training shape the neural encoding of speech

Kelsey Mankel[1,2] and Gavin M. Bidelman[1,2,3,↑]
[1]School of Communication Sciences & Disorders, University of Memphis
[2]Institute for Intelligent Systems, University of Memphis
[3] University of Tennessee Health Sciences Center, Department of Anatomy and Neurobiology

↑Email: gmbdlman@memphis.edu

This PDF file includes:

Supplementary text
Figs. S1 to S2
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1811793115

Supplementary Information Text

Contrasting nature vs. nurture effects in behavioral SIN perception

PROMS musicality test

The Profile of Music Perception Skills (PROMS) is a test of receptive auditory skills that shows excellent internal consistency and reliability for both the full (composite Cronbach's α=0.94; McDonald's ω=0.95; test-retest reliability ICC (18)=0.88, $P<0.01$) and brief versions of the test (α=0.84; ω=0.85; ICC=0.82) (1). The PROMS is also highly sensitive in differentiating subpopulations based on their musical skills (e.g., professional musicians vs. amateur musicians vs. nonmusicians). Higher scores are associated with longer durations of musical training, involvement in critical listening activities, music degrees and qualifications, and musicianship status (1). Both the full (1) and abbreviated (2) versions of the test have been cross-validated in large cohort studies and show strong correlations with other well-established musical test batteries including Gordon's Advanced Measures of Music Audiation (AMMA; 3), the Musical Aptitude Profile (MAP; 4), and the Musical Ear Test (MET; 5). Critically, PROMS scores are independent of basic psychoacoustic (temporal gap detection; 1) and cognitive abilities (working memory; 2), thus motivating its use here as a measure of complex, receptive musical skills.

Musician participants

To address whether musical training (nurture) provides an additional boost to degraded speech perception above and beyond inherent auditory skills, we measured QuickSIN scores in an additional sample of n=14 formally trained musicians who were matched in age [$t(26)=-1.91$, $P=0.07$], education [$t(26)=-1.93$, $P=0.064$], and gender (Fisher's exact test: $P=1.0$) to the PROMS listeners (i.e., nonmusicians) in our main experiment. Musician data were collected as part of our ongoing studies on music-related plasticity and speech-language function. Musicians were defined as individuals with least 10 years of continuous, self-reported training (*mean* ± *SD*; 16.0 ± 4.9 yrs) on a musical instrument starting before age 11 (7.14 ± 2.47 yrs). This definition of a musician is identical to prior cross-sectional studies on music-induced neuroplasticity (6-8). As in the experiment proper, all participants showed normal hearing sensitivity (puretone audiometric thresholds ≤ 25 dB HL; 250 to 8000 Hz), had no previous history of brain injury or psychiatric problems, and were native speakers of American English.

Bonferroni corrected independent samples *t*-tests assessed group differences in QuickSIN performance.

Supplemental Results

QuickSIN in musicians vs. nonmusicians

QuickSIN scores were invariant across high- vs. low-PROMS scorers [*t*(26)=1.27, *P*=0.22], indicating that nonmusicians process degraded speech similarly at the behavioral level regardless of their intrinsic auditory skills. However, trained musicians outperformed all nonmusicians regardless of their musicality [*t*(26)=2.75, *P*=0.011], achieving SIN reception thresholds that were ~1.5-2 dB lower (i.e., better) than their nonmusician peers (**Fig. S2**). The fact that actual musicians outperformed even high-PROMS listeners ("musical sleepers") suggests that formal musical experience (nurture) provides an additional boost to degraded speech processing at the behavioral level above and beyond innate differences in auditory system function (i.e., natural propensities).

Relations between FFR and different auditory perceptual sub-domains: GLME regressions

Regression models for the GLMEs between PROMS subtest scores and FFR neural measures were ranked by their Akaike information criterion (AIC; 9) to evaluate the relative predictive value of neural responses for each auditory sub-domain. In terms of relative (ranked) predictive power, FFR neural noise was best predicted by *tuning*, tempo, accent, followed by melody perceptual subtest scores [AICs= -188.03, -184.03, -181.54, and -180.56, respectively]. Similarly, the rank order for predicting FFR F0 amplitude based on behavioral measures was *tuning*, tempo, accent, followed by melody scores [AICs=-156.80, -154.98, -153.10, and -152.98, respectively]. In contrast, latency showed best correspondence with *accent*, followed by melody, tempo, and tuning [AICs=179.22, 183.65, 184.148, 184.149, respectively]. As mentioned in the main text, this dissociation in brain-behavior relationships suggests that spectral measures of the FFR (neural noise, F0 amps.) are more associated with perceptual skills related to fine pitch discrimination (i.e., tuning) (10, 11) whereas neural latencies are more strongly associated with timing perception (i.e., rhythmic accent).

Supplemental Discussion

The generators of auditory FFR and ERPs are often described as brainstem vs. cortical origin, respectively (6, 7, 12-15). Under this view, our data imply that pre-existing differences in neural processing are more apparent at subcortical levels. While the low F0 pitch (100 Hz) of our stimulus evokes FFRs from brainstem and cortical sources (12, 16), this choice was intentional to replicate the vast majority of FFR studies in musicians, which have similarly used low-pitched speech tokens (6-8, 16, 17). As far-field potentials including the FFR reflect distributed activity from a wide array of subcortical and cortical sources (12, 16), our data cannot definitively circumscribe where pre-existing enhancements in auditory processing are located within the auditory system. Yet, as discussed in the main text, regardless of where our scalp-recorded responses originate, neural differences were still observed among people without formal music training, and some of these individuals' brains produce phase-locked neural responses that better capture the acoustic information of speech.

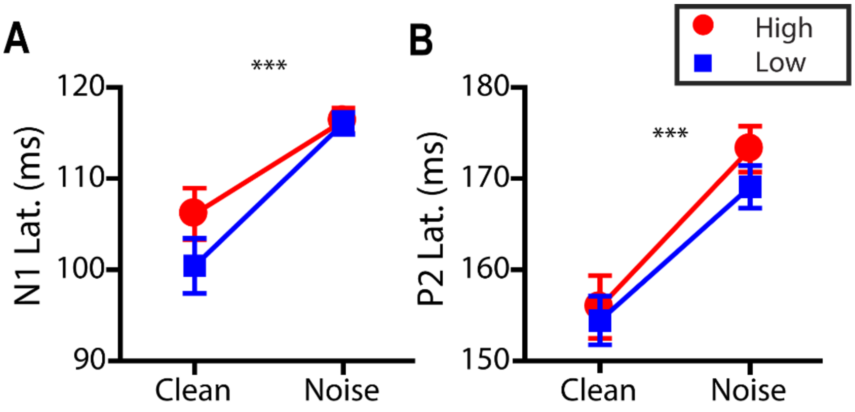

Fig. S1. Individual (**A**) N1 and (**B**) P2 component latencies indicate noise-related changes in neural activity but no group differences between high and low PROMS listeners. Error bars = ±1 s.e.m. *** $P < 0.001$

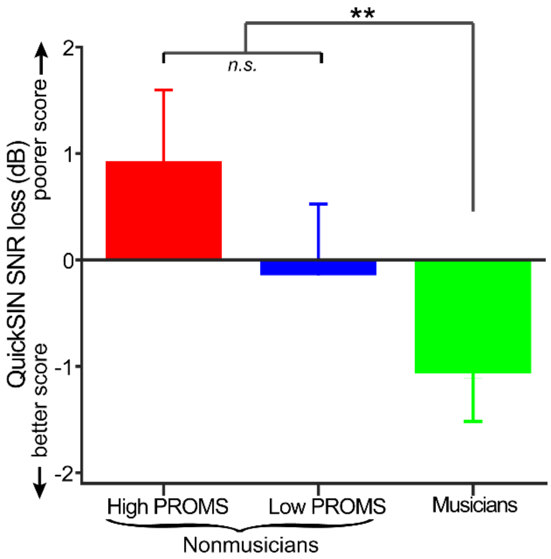

Fig. S2. Behavioral QuickSIN scores for high and low PROMS scoring nonmusicians vs. actual trained musicians. SIN perception is invariant among PROMS scores. Yet, trained musicians outperform all nonmusicians regardless of their inherent musical listening skills. Thus, formal musical experience (nurture) provides an additional boost to degraded speech processing at the behavioral level above and beyond natural listening skills. Error bars = ±1 s.e.m. ** $P \leq 0.01$

References

1. Law LNC & Zentner M (2012) Assessing musical abilities objectively: construction and validation of the Profile of Music Perception Skills. *PLoS One* 7(12):1-15.
2. Kunert R, Willems RM, & Hagoort P (2016) An Independent Psychometric Evaluation of the PROMS Measure of Music Perception Skills. *PLoS One* 11(7):e0159103.
3. Gordon EE (1989) *Advanced measures of music audiation* (GIA, Chicago, IL).
4. Gordon EE (1965) *Musical aptitude profile manual* (Houghton Mifflin, Boston, MA).
5. Wallentin M, *et al.* (2010) The Musical Ear Test, a new reliable test for measuring musical competence. *Learn. Individ. Differ.* 20(3):188-196.
6. Wong PC, *et al.* (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10(4):420-422.
7. Musacchia G, *et al.* (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* 104(40):15894-15898.
8. 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(2):1240 –1249.
9. Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6):716–723.
10. Bidelman GM & Krishnan A (2010) Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* 1355:112-125.
11. Bidelman GM, Krishnan A, & Gandour JT (2011) Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *Eur. J. Neurosci.* 33(3):530-538.
12. Bidelman GM (2018) Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage* 175:56–69.
13. Sohmer H, Pratt H, & Kinarti R (1977) Sources of frequency-following responses (FFR) in man. *Electroencephalogr. Clin. Neurophysiol.* 42:656-664.
14. Bidelman GM (2015) Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hear. Res.* 323:68-80.
15. Zhao TC & Kuhl PK (2018) Linguistic effect on speech perception observed at the brainstem. *Proc. Natl. Acad. Sci. USA* 115(35):8716-8721.
16. Coffey EB, *et al.* (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat. Commun.* 7:11070.
17. Bidelman GM, *et al.* (2014) Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur. J. Neurosci.* 40:2662-2673.