July 2014

Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians

Gavin Bidelman
University of Memphis

Jonathan Schug
University of Memphis

Skyler Jennings
University of Utah

Shaum Bhagat
University of Memphis

Recommended Citation

Gavin Bidelman, Jonathan Schug, Skyler Jennings, and Shaum Bhagat. "Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians" *Journal of the Acoustical Society of America* (2014): 33-39. https://doi.org/10.1121/1.4885484

Psychophysical auditory filter estimates reveal sharper cochlear tuning in musicians

Gavin M. Bidelman[a]
Institute for Intelligent Systems, University of Memphis, Memphis, Tennessee 38105
g.bidelman@memphis.edu

Jonathan M. Schug
School of Communication Sciences & Disorders, University of Memphis, Memphis, Tennessee 38105
jschug@memphis.edu

Skyler G. Jennings
Department of Communication Sciences & Disorders, University of Utah, Salt Lake City, Utah 84112
skyler.jennings@hsc.utah.edu

Shaum P. Bhagat
School of Communication Sciences & Disorders, University of Memphis, Memphis, Tennessee 38105
sbhagat@memphis.edu

Abstract: Musicianship confers enhancements to hearing at nearly all levels of the auditory system from periphery to percept. Musicians' superior psychophysical abilities are particularly evident in spectral discrimination and noise-degraded listening tasks, achieving higher perceptual sensitivity than their nonmusician peers. Greater spectral acuity implies that musicianship may increase auditory filter selectivity. This hypothesis was directly tested by measuring both forward- and simultaneous-masked psychophysical tuning curves. Sharper filter tuning (i.e., higher Q_{10}) was observed in musicians compared to nonmusicians. Findings suggest musicians' pervasive listening benefits may be facilitated, in part, by superior spectral processing/decomposition as early as the auditory periphery.

PACS numbers: 43.75.St, 43.66.Dc, 43.66.Mk, 43.66.Hg [DD]
Date Received: April 26, 2014 **Date Accepted:** June 16, 2014

1. Introduction

Musicianship is linked to enhanced spectrotemporal acuity[1,2] including the identification and discrimination of speech.[2,3] These pervasive benefits extend to real-world perception and auditory scene analysis, as musicians are better able to extract speech cues within noisy listening environments than their nonmusician peers.[2,4] The current study investigated whether behavioral estimates of frequency selectivity are consistent with musicians' well-documented advantages in spectral processing.

It is widely believed that auditory spectral acuity is limited by peripheral filtering at the level of the cochlea.[5–7] Basilar membrane processing is typically conceived as a bank of overlapping bandpass filters that performs a spectral decomposition of the sound input. Cochlear filter bandwidth (BW) contributes to the frequency resolution of the system, and thus, the perceptual acuity for detecting changes in the spectral

[a]Author to whom correspondence should be addressed. Also at: School of Communication Sciences & Disorders, University of Memphis, Memphis, TN 38105.

soundscape. Behaviorally, auditory filters can be estimated via the measurement of psychophysical tuning curves (PTCs).[5,7–11] In this approach, the detection of a low-level probe tone is used to infer auditory filter shape at a given location along the basilar membrane by examining the effectiveness of various narrowband signals to mask the target probe. PTCs can be measured using either simultaneous or forward masking paradigms; both approaches provide a means to measure auditory filter shapes behaviorally and estimate cochlear tuning noninvasively in humans.[5,10,11]

Greater spectral sensitivity in musically trained ears suggests that musicianship may increase auditory filter selectivity. Increased cochlear tuning (i.e., auditory filter resolution) in musicians would help explain their enhanced pitch[1,12–14] and timbre[2,3] discrimination abilities and spectral acuity observed in countless behavioral studies. The possibility of musicians having sharper auditory filters has been suggested[15] but not validated empirically (cf. Ref. 16). In the current study, we directly tested this hypothesis by comparing auditory filters in musician and nonmusician listeners. We measured both forward- and simultaneous-masked PTCs to obtain complementary measures of cochlear tuning using two widely accepted approaches. Results reveal increased filter tuning (i.e., higher Q_{10}) in musicians suggesting that their pervasive auditory benefits may be facilitated by physiological mechanisms as early as the cochlea.

2. Methods

2.1 Participants

Nineteen young adults (age range: 18–35 yrs) participated in the experiment: Ten musicians (Ms) (7 female) and 9 nonmusicians (NMs) (6 female). All participants were native speakers of English, right-handed, had normal hearing (i.e., audiometric thresholds $\leq$25 dB hearing level; 500–4000 Hz), and reported no previous history of neuropsychiatric illnesses. Ms were amateur instrumentalists who had received $\geq$5 yrs of continuous private instruction on their principal instrument (mean $\pm$ standard deviation; 10.4 ± 4.0 yrs), beginning prior to age 12 (9.6 ± 2.3 yrs), and were currently active in music practice or ensemble engagement. This definition of "musician" is consistent with previous reports.[2,12,14] NMs had $\leq$3 years of self-directed music training (1.6 ± 1.2 yrs) with no instruction within the past 5 years. The two groups were otherwise closely matched in age (Ms: 24.6 ± 3.5 yrs, NMs: 24.8 ± 1.4 yrs) and formal education (Ms: 18.9 ± 2.8 yrs, NMs: 18.8 ± 1.2 yrs). Participants were paid and gave written informed consent in compliance with a protocol approved by the Institutional Review Board of The University of Memphis.

2.2 PTCs

2.2.1 Procedure

PTCs were measured for each listener at two characteristic frequencies (CFs), 1 and 4 kHz, using both simultaneous and forward masking approaches.[10,11] These relatively high CFs were selected to circumvent difficulties in measuring auditory filters with short duration tones at lower frequencies where signal BW can exceed that of the filter.[11] Listeners sat comfortably in a sound attenuating chamber in front of a computer monitor. Auditory stimuli were generated through custom built graphical user interfaces (GUIs) coded in MATLAB (The MathWorks, Natick, MA). The output signal was routed through a LynxTWO soundcard (Lynx Studio Technology, Inc., Costa Mesa, CA) and delivered monaurally to the right ear through insert earphones (ER-3 A; Etymotic Research, Elk Grove Village, IL). The contralateral ear was occluded with a foam plug. Stimulus intensity was calibrated using a sound pressure level (SPL) meter (Model LxT, Larson-Davis, Depew, NY) measured in a 2-cc coupler (IEC 60126, G.R.A.S., Twinsburg, OH).

2.2.2 Forward masked PTCs

PTCs were measured in each listener using a standard forward masking paradigm presented in a three-interval, forced-choice task.[13] The masker was a 300 ms pure tone gated with 5 ms $\cos^2$ ramps. Tonal maskers had normalized frequencies of 0.12, 0.25, 0.50, 0.62, 0.75, 0.87, 1.00, 1.05, 1.12, 1.25, and 1.50 relative to the probe's CF. The masker was followed immediately by a contiguous probe tone (0-ms masker-probe offset). Probe signals were brief (35 ms, 10 ms ramps) sinusoidal tones (1 or 4 kHz), presented at a fixed low-level intensity (20 dB sensation level (SL) re listeners' threshold at CF). Absolute sound intensity was roughly 22 and 15 dB SPL for the 1 and 4 kHz probes, respectively. The two CF conditions and masker frequency order were randomized both within and between participants.

Each trial consisted of a given masker-probe tone combination. A high-pass noise (cutoff frequency: 1.2*CF Hz) was presented at a low intensity (−50 dB spectrum level re probe) concurrent with the masker-probe stimuli to limit off-frequency listening effects.[17] Masked thresholds were measured adaptively by varying the level of the masker with the probe fixed at a low presentation level. A fixed-signal method is preferable to a fixed-masker level (i.e., adaptively varying probe level) as it provides a more accurate depiction of auditory filter shape.[11] The masker was initially set at a level −10 dB below that of the probe. On each trial, participants heard three sequential intervals, two which contained only the masker and one which contained the masker and probe (assigned randomly). Listeners were required to identify the interval containing the probe. Intervals were indicated in the GUI presented on the computer screen; responses were made via the computer keyboard or mouse and visual feedback was provided after each trial. Each interval was separated by a 400 ms interstimulus interval (ISI). Masked thresholds were measured using a 2-down, 1-up adaptive procedure (71% performance). Following two correct responses, masker level was increased for the subsequent trial and decreased following a single incorrect response. The geometric mean of the last 8/12 reversals were used to compute each listener's masked threshold. A single masked threshold was obtained for each of the 11 masker-probe combinations and used to construct a listener's PTC at a given CF. Intense training can improve NMs' performance on perceptual pitch tasks to the level of Ms.[1] Thus, only brief task familiarization was provided to obtain forward-masked PTCs with minimal learning effects.

Filter "sharpness" was quantified from PTCs by measuring the quality (Q_{10}) factor of the auditory filter. Q_{10} is a normalized measure of filter sharpness and was used to quantify frequency selectivity (i.e., tuning) for each listener. PTCs were interpolated (i.e., up-sampled × 1000) and Q_{10} computed as $Q_{10} = f_c/\mathrm{BW}$, where f_c is the filter's center frequency and BW is its + 10 dB BW.

2.2.3 Simultaneous masked PTCs

Simultaneous masked PTCs were mapped using the "Fast PTC" method.[8] In this procedure, listeners monitored a low intensity probe tone (1 and 4 kHz; 20 dB SL) concurrent with a masking noise. A narrowband noise masker (1 kHz probe: 200 Hz BW; 4 kHz probe: 320 Hz)[8] was used to reduce the detection of beats between the masker and probe. Probes were 500 ms pure tones (20-ms ramps), continuously pulsed on/off at a regular rate (ISI: 200 ms) to help subjects maintain attention to the target. The center frequency of the masker swept upward from f_{min} to f_{max} over a time span of 4 min, where $f_{\mathrm{min}}/f_{\mathrm{max}}$ are frequencies 1.5 and 0.6 octaves below and above the CF, respectively. Masker level was continuously varied according to a Békésy track at a rate of 2 dB/s. The run began with the masker set at 50 dB SPL. Subjects were asked to press and hold a button so long as the probe tone remained audible and release it when it became inaudible. Using this procedure, the masker level needed to just mask the probe frequency was obtained as a function of masker center frequency. Fast PTCs were measured twice for each listener at each CF (i.e., 1 and 4 kHz). Only the second PTC was analyzed to ensure subjects had acclimatized to the task. As in forward masked PTCs, filter sharpness was quantified from simultaneous masked PTCs by measuring Q_{10} of the auditory filters. A 2-point moving average was applied to raw fast PTCs prior to quantification to smooth the continuous threshold obtained in the

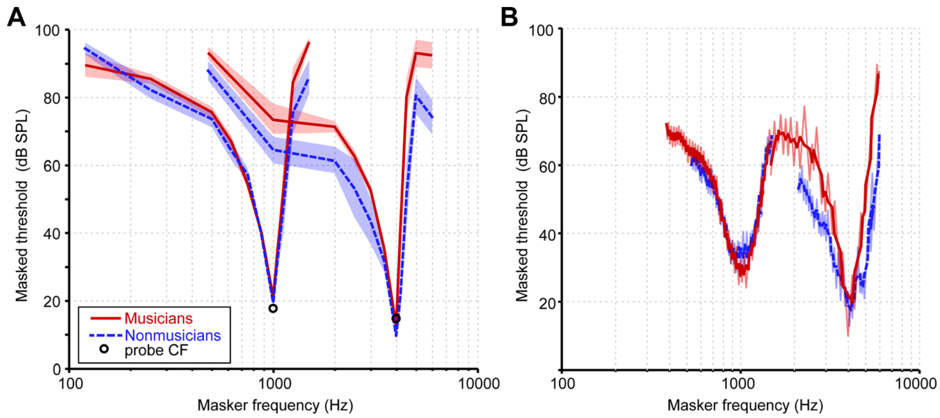

Fig. 1. (Color online) PTCs for Ms and NMs. (A) Forward masked PTCs (group average). Shading = ± 1 standard error of the mean (s.e.m.). (B) Simultaneous masked PTCs (Ref. 8). Thin lines: Raw Békésy tracked thresholds; thick lines: 2-point moving average.

Békésy track[8] [see Fig. 1(B)]. From smoothed PTCs, Q_{10} was then quantified at each probe CF for each listener.

3. Results

Raw forward and simultaneous masked behavioral PTCs are shown in Fig. 1. PTCs showed a typical "V-shape" with a low-frequency tail, highly selective tip, and steep high-frequency skirt characteristic of auditory filters measured via psychophysical paradigms.[7–10]

Auditory filter sharpness, quantified by filter Q_{10}, is shown for Ms and NMs in Fig. 2. A repeated measures (rm)analysis of variance (ANOVA) conducted on forward masked PTC Q_{10} values revealed a significant main effect of group [$F_{1,17} = 4.90$, $p = 0.04$] and probe CF [$F_{1,17} = 6.22$, $p = 0.023$] with a marginal group*CF interaction [$F_{1,\ 17} = 3.35$, $p = 0.08$]. The main effect of group indicates that Ms had sharper tuning across the board. Nevertheless, given the marginal interaction, we used *post hoc* Tukey-Kramer adjusted multiple comparisons to further investigate group differences in tuning at each CF. Results revealed that Ms had higher filter Q_{10} (i.e., more selective tuning) at 4 kHz relative to their NM peers. As with forward masked PTCs, auditory filters were sharper in Ms when measured via simultaneous masking. An rmANOVA conducted on simultaneous masked Q_{10}'s revealed a significant group*CF interaction [$F_{1,17} = 4.19$, $p = 0.05$]. Follow-up contrasts revealed this was again attributable to Ms having larger Q_{10} at 4 kHz relative to NMs. No group differences were observed at 1 kHz for PTCs derived either by simultaneous or forward masking.

Correlational analyses tested whether an individual's degree of musical training, as measured in years of engagement, predicted their auditory filter sharpness. Years of formal music training was positively correlated with filter Q_{10} at 4 kHz when measured via simultaneous masking [$r = 0.44$, $p = 0.02$] [Fig. 2(B)] in that longer training predicted sharper cochlear tuning. Correlations with forward masked PTC Q_{10} and results at 1 kHz were insignificant.

Q_{10} values provide a normalized measure of frequency selectivity and a means to directly compare filter tuning estimates obtained under the two disparate masking paradigms. Pooling across musician and nonmusician listeners, a rmANOVA conducted on filter Q_{10}'s revealed significant main effects of probe CF [$F_{1,18} = 6.98$, $p = 0.017$] and masking paradigm [$F_{1,18} = 47.73$, $p < 0.0001$] with no interaction [$F_{1,\ 18} = 0.26$, $p = 0.62$]. The simple main effect of masking type suggests that forward masking provides higher estimates of filter tuning, i.e., larger PTC Q values, relative to

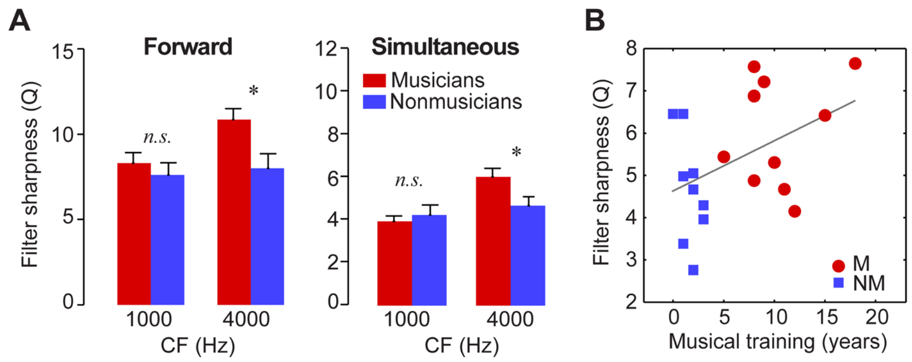

Fig. 2. (Color online) Group comparison of auditory filter tuning (Q-factor) in Ms and NMs. (A) Estimates of tuning are 2 times sharper for forward compared to simultaneous masking. Relative to NMs, Ms demonstrate more selective (i.e., narrower) auditory filters, particularly at higher CFs (4 kHz). (B) Years of formal musical training predict increased filter sharpness at 4 kHz measured via simultaneous masking; longer music experience is associated with higher Q_{10}. Error bars = ± 1 s.e.m.; $*p < 0.05$.

simultaneous masking across the board. Similarly, the main effect of CF indicates that tuning was better at the higher (4 kHz) relative to the lower CF (1 kHz).

4. Discussion

Previous studies have demonstrated superior spectral acuity in musicians relative to nonmusician listeners[1,2,13,14] implying that musical training shapes cochlear processing and increases the resolution of peripheral auditory filters. By measuring PTCs in musicians and nonmusicians—a measure widely believed to reflect peripheral cochlear filtering[5,6,10,11]—our results provide strong evidence for sharper auditory filtering in musically trained listeners (cf. Ref. 16). Importantly, this effect was predicted by an individual's years of musical training [Fig. 2(B)]. These findings are consistent with the notion that musicianship improves peripheral cochlear filtering, increasing peripheral spectral resolution in an experience-dependent manner.[12] Superior cochlear tuning in musicians may account for their enhanced auditory performance in a wide variety of auditory behavioral studies, including speech/language and musical tasks.[1,2,4,14]

In the current study, PTCs were obtained via both forward and simultaneous masking. Pooled across listeners, estimates of tuning were generally sharper (i.e., higher Q) by nearly a factor of 2 under forward compared to simultaneous masking. This is consistent with previous psychoacoustic studies and the notion that suppression plays a stronger role in the latter approach.[10,11,18] Nevertheless, larger Q_{10} in musicians derived under both methods suggests increased peripheral auditory filter selectivity than musically naive listeners even in the presence of strong cochlear nonlinearities (e.g., suppression).

Interestingly, we found that group differences in tuning were more pronounced at higher (4 kHz) relative to lower CFs (1 kHz). Additionally, musical training was correlated with filter Q_{10} sharpness at 4- but not 1-kHz (under simultaneous masking). These findings imply that the experience-dependent effects of musicianship on cochlear processing might act more strongly at higher- relative to lower-frequencies along the cochlear partition. Both psychophysical[19] and neurophysiological[18] data indicate that nonlinearities including suppression and cochlear amplification are larger at higher relative to lower CFs. Indeed, musicians are highly sensitive to manipulations in spectral timbre[2] and show benefits in noise-degraded listening,[2,4] tasks that rely heavily on high frequency spectral coding.[20] We infer that musicians' superiority in exploiting high-frequency spectral cues in these tasks may arise due to their sharper tuning in more basal cochlear channels (e.g., Fig. 1). Such differential effects could arise based on the well-known neuroanatomical configuration and putative physiological effects of the medial olivocochlear (MOC) efferent system, whose innervation density is greater at higher- relative to lower-frequencies.[21] Given these inherent structural asymmetries, the modulatory gain supplied by MOC fibers to the cochlea is likely to be larger at more basal portions of the basilar membrane,[22] consistent with our observations.

It is possible that musicians' tuning benefits might be at least partially driven by enhanced attention or efficiency in exploiting auditory cues.[14,23] However, if this were the case, we would have expected musicians to show superior performance across the board, whereas group differences were observed only at higher CFs. A more likely possibility involves converging evidence from otoacoustic emissions, which suggests that musical training can strengthen cochlear processing and MOC modulatory activity.[24] These studies provide evidence for music-induced plasticity at initial stages of auditory sensory processing mediated by strengthened "top-down" feedback from the caudal brainstem to the cochlea. While the functional role of the MOC in human hearing is still debated,[22] it is speculated that it may provide an "antimasking" effect, helping to improve signal detection in noise[25] and/or discrimination sensitivity.[22] Musicians excel at both of these skills.[2,12,13] By this account then, musicians' stronger MOC activity may provide more antimasking at the probe signal frequency, enhancing its contrast from the noise masker, and consequently providing sharper estimates of filtering in that cochlear channel. Alternatively, stronger top-down efferent control in musicians[24] may allow them to preemptively inhibit MOC gain reduction (which can alter neural tuning curves[22]) and thus maintain a higher degree of frequency resolution compared to nonmusicians. Therefore, it is possible that at least some of musicians' behavioral auditory advantages observed previously, as well as the sharper auditory tuning observed in the current study, might result from enhancements to the MOC-cochlear efferent pathway tuned through rigorous, long-term musical engagement.[24]

Acknowledgments

This work was supported in part by a grant from the GRAMMY Foundation® awarded to G.M.B. and S.P.B.

References and links

[1]C. Micheyl, K. Delhommeau, X. Perrot, and A. J. Oxenham, "Influence of musical and psychoacoustical training on pitch discrimination," Hear. Res. **219**, 36–47 (2006).
[2]G. M. Bidelman and A. Krishnan, "Effects of reverberation on brainstem representation of speech in musicians and non-musicians," Brain Res. **1355**, 112–125 (2010).
[3]G. M. Bidelman, M. W. Weiss, S. Moreno, and C. Alain, "Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians," Eur. J. Neurosci. 1–12 (in press).
[4]A. Parberry-Clark, E. Skoe, C. Lam, and N. Kraus, "Musician enhancement for speech-in-noise," Ear Hear. **30**, 653–661 (2009).
[5]C. A. Shera, J. J. Guinan, Jr., and A. J. Oxenham, "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. U. S. A. **99**, 3318–3323 (2002).
[6]E. F. Evans, "Auditory processing of complex sounds: An overview," Philos.Trans. R. Soc., B **336**, 295–306 (1992).
[7]G. M. Bidelman and A. Syed Khaja, "Spectrotemporal resolution tradeoff in auditory processing as revealed by human auditory brainstem responses and psychophysical indices," Neurosci. Lett. **572**, 53–57 (2014).
[8]A. Sek, J. Alcántara, B. C. Moore, K. Kluk, and A. Wicher, "Development of a fast method for determining psychophysical tuning curves," Int. J. Audiol. **44**, 408–420 (2005).
[9]S. G. Jennings and E. A. Strickland, "Auditory filter tuning inferred with short sinusoidal and notched-noise maskers," J. Acoust. Soc. Am. **134**, 2497–2513 (2012).
[10]B. C. J. Moore, "Psychophysical tuning curves measured in simultaneous and forward masking," J. Acoust. Soc. Am. **63**, 524–532 (1978).
[11]A. J. Oxenham and C. A. Shera, "Estimates of human cochlear tuning at low levels using forward and simultaneous masking," J. Assoc. Res. Oto. **4**, 541–554 (2003).
[12]G. M. Bidelman, S. Hutka, and S. Moreno, "Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music," PloS One **8**, e60676 (2013).

[13]G. M. Bidelman, J. T. Gandour, and A. Krishnan, "Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch," Brain Cogn. **77**, 1–10 (2011).
[14]D. L. Strait, N. Kraus, A. Parberry-Clark, and R. Ashley, "Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance," Hear. Res. **261**, 22–29 (2010).
[15]D. R. Soderquist, "Frequency analysis and the critical band," Psych. Sci. **21**, 117–119 (1970).
[16]P. A. Fine and B. C. J. Moore, "Frequency analysis and musical ability," Music Percept. **11**, 39–53 (1993).
[17]B. J. O'Loughlin and B. C. J. Moore, "Off-frequency listening: Effects on psychoacoustical tuning curves obtained in simultaneous and forward masking," J. Acoust. Soc. Am. **69**, 1119–1125 (1981).
[18]P. J. Abbas and M. B. Sachs, "Two-tone suppression in auditory-nerve fibers: Extension of a stimulus-response relationship," J. Acoust. Soc. Am. **59**, 112–122 (1976).
[19]R. V. Shannon, "Two-tone unmasking and suppression in a forward-masking situation," J. Acoust. Soc. Am. **59**, 1460–1470 (1976).
[20]N. Amos and L. E. Humes, "Contribution of high frequencies to speech recognition in quiet and noise in listeners with varying degrees of high-frequency sensorineural hearing loss," J. Speech Lang. Hear. Res. **50**, 819–834 (2007).
[21]M. C. Liberman, L. W. Dodds, and S. Pierce, "Afferent and efferent innervation of the cat cochlea: Quantitative analysis with light and electron microscopy," J. Comp. Neurol. **301**, 443–460 (1990).
[22]J. J. Guinan, Jr., "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," Ear Hear. **27**, 589–607 (2006).
[23]A. J. Oxenham, B. J. Fliqor, C. R. Mason, and G. Kidd, Jr., "Informational masking and musical training," J. Acoust. Soc. Am. **114**, 1543–1549 (2003).
[24]X. Perrot and L. Collet, "Function and plasticity of the medial olivocochlear system in musicians: A review," Hear. Res. **308**, 27–40 (2014).
[25]J. de Boer, A. R. Thornton, and K. Krumbholz, "What is the role of the medial olivocochlear system in speech-in-noise processing?," J. Neurophysiol. **107**, 1301–1312 (2012).