Article

Musicianship Modulates Cortical Effects of Attention on Processing Musical Triads

Jessica MacLean [1,2], Elizabeth Drobny [1], Rose Rizzi [1,2] and Gavin M. Bidelman [1,2,3,*]

1. Department of Speech, Language, and Hearing Sciences, Indiana University, Bloomington, IN 47408, USA
2. Program in Neuroscience, Indiana University, Bloomington, IN 47408, USA
3. Cognitive Science Program, Indiana University, Bloomington, IN 47408, USA
* Correspondence: gbidel@iu.edu; Tel.: +1-(812)-855-9339

Abstract: Background: Many studies have demonstrated the benefits of long-term music training (i.e., musicianship) on the neural processing of sound, including simple tones and speech. However, the effects of musicianship on the encoding of simultaneously presented pitches, in the form of complex musical chords, is less well established. Presumably, musicians' stronger familiarity and active experience with tonal music might enhance harmonic pitch representations, perhaps in an attention-dependent manner. Additionally, attention might influence chordal encoding differently across the auditory system. To this end, we explored the effects of long-term music training and attention on the processing of musical chords at the brainstem and cortical levels. **Method**: Young adult participants were separated into musician and nonmusician groups based on the extent of formal music training. While recording EEG, listeners heard isolated musical triads that differed only in the chordal third: major, minor, and detuned (4% sharper third from major). Participants were asked to correctly identify chords via key press during active stimulus blocks and watched a silent movie during passive blocks. We logged behavioral identification accuracy and reaction times and calculated information transfer based on the behavioral chord confusion patterns. EEG data were analyzed separately to distinguish between cortical (event-related potential, ERP) and subcortical (frequency-following response, FFR) evoked responses. **Results**: We found musicians were (expectedly) more accurate, though not faster, than nonmusicians in chordal identification. For subcortical FFRs, responses showed stimulus chord effects but no group differences. However, for cortical ERPs, whereas musicians displayed P2 (~150 ms) responses that were invariant to attention, nonmusicians displayed reduced P2 during passive listening. Listeners' degree of behavioral information transfer (i.e., success in distinguishing chords) was also better in musicians and correlated with their neural differentiation of chords in the ERPs (but not high-frequency FFRs). **Conclusions**: Our preliminary results suggest long-term music training strengthens even the passive cortical processing of musical sounds, supporting more automated brain processing of musical chords with less reliance on attention. Our results also suggest that the degree to which listeners can behaviorally distinguish chordal triads is directly related to their neural specificity to musical sounds primarily at cortical rather than subcortical levels. FFR attention effects were likely not observed due to the use of high-frequency stimuli (>220 Hz), which restrict FFRs to brainstem sources.

Keywords: music training; musical chords; event-related potential (ERP); frequency-following response (FFR)

1. Introduction

Many studies have demonstrated the benefits of long-term music training (i.e., musicianship) for neural processing of sound, including simple tones [1] and speech [2]. This "musician enhancement" in brain processing has been linked to a variety of auditory perceptual skills, especially those related to spectral processing. For example, musicians show improved abilities to detect violations in musical pitch patterns [3–5], smaller frequency

difference limens [6–8], and reduced susceptibility to timbral influences on pitch perception [9]. Musicians' auditory benefits also confer improvements to non-musical sounds, including speech [10–14]. Such enhancements to behavioral and neural indicators of sound processing are important for understanding how music engagement may support everyday listening skills.

Despite a wealth of cross-domain investigations into the influence of music experience on neural speech encoding, less is known about the effects of music training on neural encoding of *musical* sounds. Given that music performance necessitates the precise manipulation and monitoring of pitch information (e.g., to remain in tune within an ensemble), musicianship may engender superior musical pitch encoding abilities. Bidelman et al. [1] investigated neural processing of arpeggiated triad chords, in which one note was played at a time (i.e., melodically). Major, minor, and detuned ($\pm 4\%$ frequency) triads only differed in the chordal third, the middle note of which determined the perceived quality of the chord. Major and minor chords are often heard in Western music and can be reliably distinguished by individuals without music training as sounding "happy" and "sad", respectively [15]. In contrast, the detuned chord represented a triad less familiar to participants given its absence from common Western scales. Bidelman et al. [1] found musicians' frequency-following responses (FFRs), neurophonic potentials reflecting subcortical (brainstem) phase locking, showed stronger neural onset synchronization and encoding of the chordal third than nonmusicians', regardless of whether the chord was in or out of tune. Musicians' enhanced FFRs were paralleled by their improved performance on pitch discrimination tasks. While musicians showed superior discrimination of three chord variants, nonmusicians were only sensitive to major/minor distinctions and could not detect more subtle chord mistuning. Though this study suggests musicianship strengthens passively evoked brainstem responses to musical triads, it does not elucidate whether such experience-dependent enhancements are uniform at different levels of auditory processing, nor how they change under different states of attention. Here, we investigated the effects of musicianship on the neural processing of harmonic musical chords at multiple levels of the auditory system.

In this vein, several studies have shown musicians' enhancements to speech and music at a cortical level utilizing event-related potentials (ERPs) [16–20]. However, cortical neuroplasticity associated with music training is sometimes observed in the absence of subcortical enhancements [21]. This raises the question of where in the auditory system music training exerts its neuroplastic influences on auditory coding and which level of processing is most predictive of musicians' behavioral benefits for musical pitch. Relatedly, the effects of musicianship may be attention-dependent. Attention can modulate both subcortical FFRs [22,23] and cortical ERPs to speech [24,25]. However, most studies demonstrating FFR enhancements in musicians have utilized only passive listening paradigms where participants watch a silent movie and do not interact with the stimuli [2]. At least for speech sounds, musicians' FFR enhancements are not necessarily observed under active listening even when auditory cortical potentials do show attentional gains [21]. Thus, although musicianship may exert influences on attention at the behavioral level [10,26], it may do so with differing degrees across the auditory brain hierarchy.

Given the equivocal nature of the literature, the current study aimed to test whether brainstem FFRs and cortical ERPs are sensitive to attention and reflect a neural correlate of active music perception. We used musical chord triads [1] in an active identification paradigm [23,27,28] designed to simultaneously record FFRs/ERPs with behavioral responses. Chordal stimuli included the major and minor triads of Western music and a detuned (4% sharp) version. Chords were advantageous to our design because their perceptual identification required listeners to correctly "hear out" only a single middle note (i.e., chordal third) flanked by two other pitches. We hypothesized the chordal third would be selectively enhanced in neural responses when listeners correctly identified the chord. We also measured listeners with a range of musical training, as we hypothesized that musically experienced listeners would be more successful at musical chord identification tasks (e.g., [1]) and thus show stronger attentional changes in their FFRs and/or ERPs.

2. Materials and Methods

2.1. Participants

Our sample included *n* = 15 young adults (age range = 19–31 years, 10 female/ 5 male). All participants were monolingual English-speakers and had normal hearing (pure tone thresholds ≤ 25 dB HL; octave frequencies 250–8000 Hz). Participants had an average of 17.6 ± 2.80 years of education. Participants self-reported musical training of 9.47 ± 8.26 years (range = 0–24 years). For the purposes of visualization, the data are plotted with participants split into two groups: musicians (M; *n* = 9) had greater than 5 years of training and nonmusicians (NM; *n* = 6) had fewer than 5 years of training (see [29] for a similar definition of "musician"). However, the analyses considered music as a continuous predictor rather than discrete grouping variable (see Section 2.8). Each participant provided written, informed consent in compliance with protocol approved by the Institutional Review Board of Indiana University and was paid for their time.

2.2. Stimuli and Task

We used harmonic chord stimuli to evoke FFRs [1]. Each chord consisted of three concurrent pitches (root, 3rd, 5th) that differed only in the chordal third (i.e., middle pitch). Consequently, category labeling could only be accomplished if listeners were able to perceptually "hear out" the chord-defining pitch. Two chords were exemplary of Western music practice (major and minor chords); the third represented a detuned version in which the chord's third was made sharp (4%) from its major version. The minor and major third intervals connote the valence of "sadness" and "happiness" even to non-musicians [30] and are easily described to listeners unfamiliar with categorical labels derived from music theory [31]. A 4% deviation is greater than the just-noticeable difference for frequency (<1%) [32] but smaller than a full semitone (6%). This amount of deviation is similar to previously published reports examining musicians' and non-musicians' EEG responses to detuned triads [1,4,30]. Individual chord notes were synthesized using complex-tones (6 iso-amplitude harmonics added in sine phase; 100 ms duration; 5 ms ramp). The fundamental frequency (F0) of each of the three notes (i.e., root, 3rd, 5th) per triad was as follows: major = 220, 277, 330 Hz; minor = 220, 262, 330 Hz; detuned = 220, 287, 330 Hz. Critically, the F0 (indeed all spectral cues) in our stimuli exceeded 220 Hz, which is substantially higher than the phase-locking limits of cortical neurons observed in any animal or human studies [33–35]. This ensured our FFRs were of brainstem origin [21,28,35]. Before initiating the task proper, participants were allowed to freely replay the three stimuli for familiarization.

To efficiently record FFRs during an online behavioral task while obtaining the high (i.e., several thousand) trial counts needed for response visualization, we used a clustered interstimulus interval (ISI) presentation paradigm (Figure 1) [23,27,28]. Single tokens were presented in blocks of 15 repetitions with a rapid ISI (10 ms). After the clustered block of tokens ended, the ISI was increased to 1500 ms and a single token was presented to cue the behavioral response and allow for the recording of an un-adapted cortical ERP [27]. Participants then indicated their percept (major, minor, detuned) as quickly and accurately as possible via the keyboard. Following the behavioral response and a period of silence (250 ms), the next trial cluster commenced. Per token, this paradigm allowed 1980 presentations for input to FFR analysis and 66 ERP presentations and behavioral responses. We calculated percent correct identification per stimulus condition.

In addition to the active condition, we ran a passive condition that did not require an overt task. For passive presentation, participants did not actively engage with the sounds, and instead passively watched a self-selected movie with subtitles to maintain a calm and wakeful state [19]. Because there was no behavioral response in passive blocks, we included an ISI of 600 ms between token clusters (i.e., comparable to the estimated reaction time (RT) in the active block) to ensure that the overall pacing of stimulus delivery was comparable between listening conditions. The order of active and passive blocks was randomized.
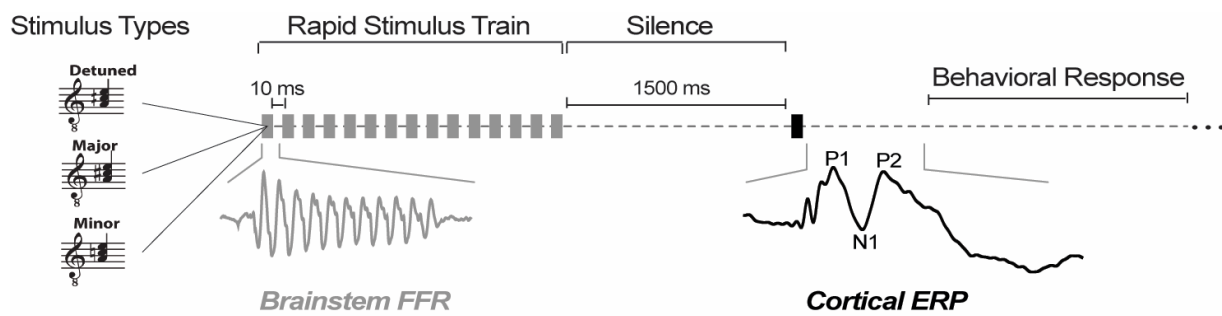
Figure 1. Stimulus paradigm to elicit neural and behavioral responses. We used a clustered stimulus paradigm [23,27,28] to elicit FFRs and ERPs within a single trial. For a given trial, the stimulus (a major, minor, or detuned chord) was presented in a rapid stimulus train used to elicit the FFR, followed by a 1500 ms gap of silence and then a single presentation of the chord which elicits an ERP. Following the ERP-evoking stimulus, participants were asked to behaviorally identify the chord via keyboard press.

Stimulus presentation was controlled via MATLAB (The MathWorks, Natick, MA, USA) routed to a TDT RZ6 (Tucker-Davis Technologies, Alachua, FL, USA) signal processor. Stimuli were presented binaurally at 80 dB SPL with rarefaction polarity. We used shielded insert headphones (ER-2; Etymotic Research) to prevent pickup of electromagnetic artifacts from contaminating neural recordings [36,37].

2.3. EEG Recording Procedures

We used Curry 9 software (Compumedics Neuroscan, Charlotte, NC, USA) and Neuroscan Synamps RT amplifiers to record the EEG data. Continuous EEGs were recorded differentially between scalp Ag/AgCl disk electrodes placed on the high forehead at the hairline (~Fpz) referenced to linked mastoids (M1/M2); a mid-forehead electrode served as ground. This montage is optimal for pickup of the vertically oriented FFR dipole in the midbrain [38]. Electrode impedances remained ≤ 5 kΩ throughout the duration of recording. EEGs were digitized at 10 kHz to capture the fast activity of FFR. Trials exceeding a ± 120 μV threshold were automatically rejected from the average. We separately band-pass-filtered responses from 200 to 2500 Hz and 1 to 30 Hz (zero-phase Butterworth filters; slope = 48 dB/octave) to isolate FFRs and ERPs, respectively [39,40]. Data were then epoched relative to the time-locking trigger for each response class (i.e., FFR: 0–105 ms, ERP: -200–1000 ms), pre-stimulus baselined, and averaged for each stimulus token. Data preprocessing was performed in BESA Research v7.1 (BESA, GmbH, Gräfelfing, Germany).

2.4. Brainstem FFR Analysis

From each FFR waveform, we generated Fast Fourier Transforms (FFTs) to quantify the spectral information in each response. Amplitudes of the spectral peak corresponding to the chordal 3rd (227–287 Hz) were identified in all conditions by two independent observers. Inter-rater reliability across all measurements was exceptional [$r = 0.99$, $p < 0.0001$]. F0—related to pitch encoding—represents the dominant energy in the FFR and can be modulated by attention and listeners' trial-by-trial categorical hearing [22,23,37,41]. We focused on the F0 of the chordal 3rd as this was the category-defining pitch of each stimulus. Onset latency and amplitude were measured as the first positive peak in the FFR waveform in the 7–15 ms time window, the expected onset latency of brainstem responses [42,43].

2.5. Cortical ERP Analysis

From ERP waveforms to each stimulus, we measured the amplitude and latency of the N1 and P2 deflections between 100 and 140 ms and 135 and 175 ms, respectively. The analysis window was guided by visual inspection of the grand-averaged data. We selected

N1 due to its well-documented susceptibility to attention effects and musicianship [44–50] and P2 for its sensitivity to long-term plasticity from musicianship [19,21].

2.6. Confusion Matrices and Information Transfer (IT) Analysis

To examine the pattern of behavioral responses to chordal stimuli (in the active task), we first computed confusion matrices [51]. Confusion matrices plot the percentage of times an actual stimulus was presented vs. the listener's percentage of response labels. Correct responses appear along the diagonal of the matrix, where 100% corresponds to perfect perceptual identification of the stimuli. Deviations in the off-diagonal cells show incorrect responses and thus patterns of confusion between the chords that were actually presented vs. what was instead perceived. We then used information transfer (IT) [52] to directly assess the correspondence between stimulus input and behavioral output. IT is defined as the ratio of transmitted information between x and y [i.e., $T(x;y)$] to the input entropy (H_x), expressed as a percentage. $T(x;y)$ represents the transmission of information (in an information theoretic sense) from x to y, measured in bits per stimulus, and was computed from confusion matrices via Equation (1):

$$T(x;y) = -\sum_{ij} p_{ij} \log_2 \frac{p_i p_j}{p_{ij}} \tag{1}$$

where p_i and p_j are the probabilities of the observed input and output variables, respectively, and p_{ij} is the joint probability of occurrence for observing input i with output j. These probabilities were computed from the confusion matrices as $p_i = n_i/N$, $p_j = n_j/N$, and $p_{ij} = n_{ij}/N$, where n_i is the frequency of stimulus i, n_j is the frequency of response j, and n_{ij} is the frequency of the joint occurrence of stimulus i and response j (i.e., diagonal elements of the confusion matrix) in the sample of N total observations [52]. The input entropy H_x is given by Equation (2):

$$Hx = -\sum_{x} p_x \log_2 p_x \tag{2}$$

In the current study, all stimuli occurred with equal probability (i.e., $p_x = 0.33$). IT was then computed as (Equation (3)):

$$IT = \frac{T(x;y)}{Hx} \tag{3}$$

This metric varies from 0 to 100%. Intuitively, if the transmission is poor and a listener's response does not closely correlate with the stimulus, then *IT* will approach zero; alternatively, if the response can be accurately predicted from the stimulus, then *IT* will approach unity (i.e., 100% information transfer). *IT* was computed from the behavioral chordal confusion matrices for each participant and then group-averaged. Comparisons of *IT* allowed us to assess the degree to which stimulus information was faithfully transmitted to perception in each listener.

2.7. Neural Differentiation of Musical Chords

We reasoned that listeners with more robust perceptual identification and fewer behavioral confusions would show maximally differentiable cortical ERPs across chordal stimuli [51]. To this end, we quantified the dissimilarity in listeners' ERPs across the three chordal triads via the pairwise difference in ERP N1-P2 amplitude between chords (i.e., major–minor, major–detuned, minor–detuned). We focused on the N1-P2 for this analysis because this complex represents the overall magnitude of the auditory cortical ERPs and previous work has shown robust neural decoding of auditory stimulus properties in the time window of the N1-P2 deflections [51,53]. Differences were computed via the Euclidean distance between peak-to-peak amplitude values (MATLAB *pdist2* function). We then averaged the three distances (per participant) to quantify the overall differentiation of stimuli in their ERPs.

2.8. Statistical Analysis

Unless otherwise noted, we analyzed dependent variables using mixed-model ANOVAs in the R (version 4.2.2) [54] lme4 package [55]. To analyze the behavioral outcomes of accuracy (percent correct, PC) and information transfer [IT; 52], we modeled fixed effects of group (2 levels; M vs. NM) and stimulus (3 levels; major vs. minor vs. detuned), with a random effect of subject. To analyze neural outcomes (FFR onset latency and amplitude, ERP P2 latency and amplitude), we added an additional fixed effect of condition (2 levels; active vs. passive). Effect sizes are reported as partial eta squared (η_p^2) and degrees of freedom (*d.f.*) using Satterthwaite's method. A priori significance level was set at $\alpha = 0.05$. We examined brain–behavior correspondences via Pearson correlations between behavioral IT and ERP distances. Unless otherwise noted, statistical analyses are reported using years of music training as a continuous variable in the ANOVA models. Discrete musician and nonmusician groups are used in some instances for visualization purposes. However, results were qualitatively similar whether music was treated as a discrete or continuous variable.

3. Results

3.1. Behavioral Chord Identification

Figure 2 displays behavioral results for all participants separated by musician (M) and nonmusician (NM) groups across all three chordal stimuli (major, minor, detuned). An ANOVA on behavioral accuracy revealed an interaction between music training and stimulus [$F(2, 26) = 3.44$, $p = 0.047$, $\eta_p^2 = 0.21$]. Post hoc Tukey-adjusted pairwise comparisons revealed that whereas Ms performed equally well identifying all three chords (all pairwise comparisons: $p > 0.77$), NMs showed a gradient performance with better identification of the minor relative to other chords (Figure 2A). These data suggest that musicians were more accurate (but not faster) in identifying musical chords. Behavioral confusion matrices (Figure 2B) for all stimuli were used to calculate information transfer (IT) [52] for both groups (Figure 2C). A *t*-test between groups indicated significantly higher IT for Ms relative to NMs [$t(13) = 6.15$, $p < 0.001$].
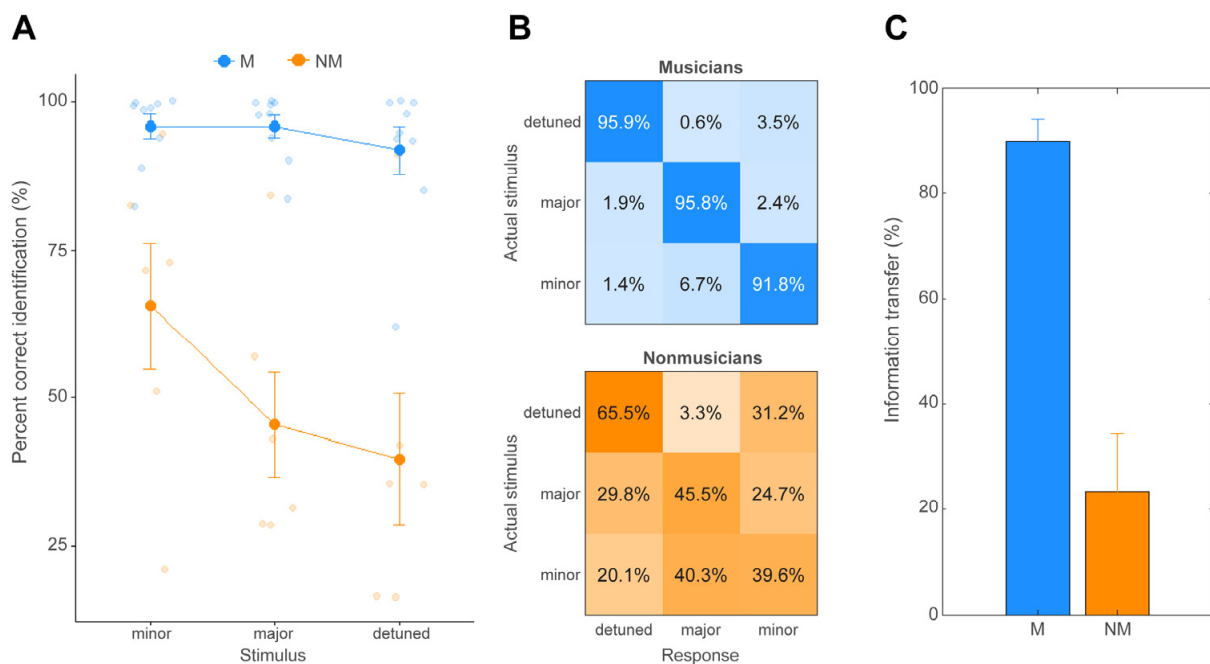

Figure 2. Behavioral performance in chord identification for musicians and nonmusicians (active condition). (**A**) Musicians outperformed nonmusicians in identification accuracy across the board. Nonmusicians showed more gradient performance and were better at identifying minor chords relative to the major and detuned chords. (**B**) Mean behavioral confusion matrices per group. Cells

denote the proportion of responses labeled as a given chord relative to the actual stimulus category presented. Diagonals show correct responses. Chance level = 33%. NMs show more perceptual confusions than Ms, especially between major and detuned chords. (**C**) Information transfer (IT) derived from the perceptual confusion matrices. IT represents the degree to which listeners' responses can be accurately predicted given the known input stimulus [52]. Values approaching 100% indicate perfect prediction of the response given the input; values approaching 0 indicate total independence of the stimulus and response. IT is near the ceiling and much larger in Ms, indicating higher fidelity differentiation of musical stimuli. Error bars = ±1 S.E.M.

3.2. Brainstem FFRs

Figure 3 shows subcortical FFRs to all three chordal stimuli for both groups and conditions. FFR response metrics are shown in Figure 4. An ANOVA on FFR onset amplitude revealed a significant interaction between stimulus and condition [$F_{(2, 65)} = 3.71$, $p = 0.030$, $\eta_p^2 = 0.10$] (Figure 4A). Post hoc, Tukey-adjusted comparisons revealed that the interaction was driven by stronger onset amplitudes for the major chord in the active condition relative to the passive condition ($p = 0.037$), despite similar onset amplitudes between conditions for the minor ($p = 0.48$) and detuned ($p = 0.37$) stimuli. A main effect of stimulus [$F_{(2, 65)} = 3.71$, $p = 0.030$, $\eta_p^2 = 0.10$] was driven by stronger onset amplitudes in the detuned condition relative to the minor condition ($p = 0.0015$). There were no other main or interaction effects (all $ps > 0.2$). FFR onset latency did not vary with condition [$F_{(1, 65)} = 0.39$, $p = 0.53$, $\eta_p^2 < 0.01$], stimulus [$F_{(2, 65)} = 1.48$, $p = 0.23$, $\eta_p^2 = 0.04$], music training [$F_{(1, 13)} = 0.70$, $p = 0.42$, $\eta_p^2 = 0.05$], or their two- and three-way interactions (all $ps > 0.60$; Figure 4B).
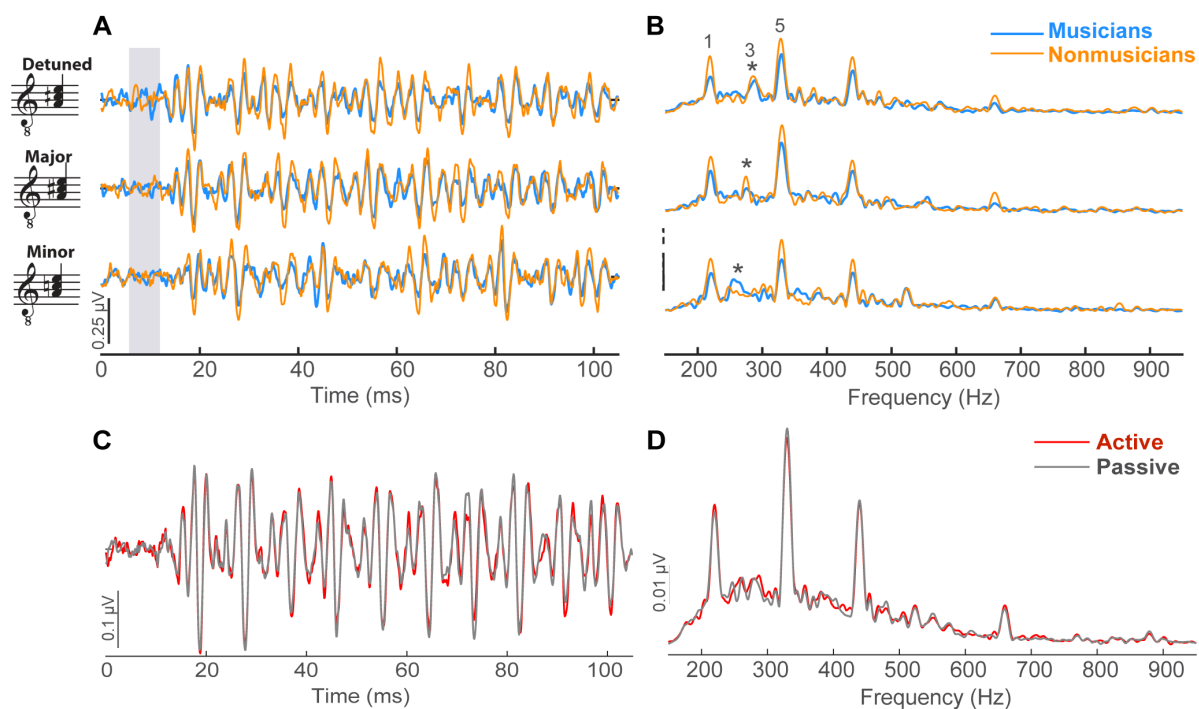

Figure 3. Group-averaged FFRs to chordal stimuli. (**A**) Time waveforms to chordal stimuli in musicians and nonmusicians (active condition only). (**B**) Response spectra (FFTs). FFRs were not influenced by music training. Peaks for the chordal root (1; 220 Hz), third (3*; 262–287 Hz), and fifth (5; 330 Hz) are demarcated. Only the chordal third (*) differed across stimuli. FFRs (collapsed across groups) did not differ between active and passive conditions in the time (**C**) or frequency (**D**) domains.
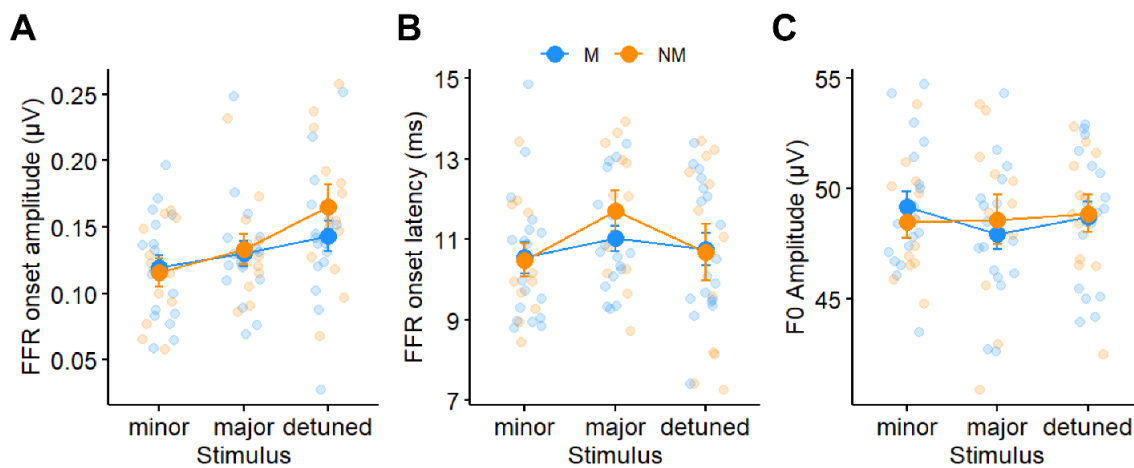
Figure 4. FFR latency and amplitude characteristics. (**A**) FFR onset amplitude varied with chord stimulus but not music training. An interaction between condition and stimulus showed stronger amplitudes in the active relative to the passive condition for the major stimulus only (not shown). (**B**) FFR onset latency and (**C**) FFR F0 amplitude were invariant both within and between participants. Error bars = ±1 S.E.M.

In the spectral domain, F0 amplitudes to the chordal third were isolated from FFR waveforms (Figure 3B). F0 amplitudes did not vary with attention (Figure 3D) [$F(1, 65)$ = 0.075, p = 0.78, η_p^2 < 0.01], stimulus (Figure 4C) [$F(2, 65)$ = 0.039, p = 0.96, η_p^2 < 0.01], music training [$F(1, 13)$ = 0.15, p = 0.70, η_p^2 = 0.01], or their interactions (all ps > 0.1).

3.3. Cortical ERP Responses

Figure 5 depicts cortical ERPs for musicians and nonmusicians in both active and passive conditions. An ANOVA on P2 latencies did not reveal any main effects or interactions (all ps > 0.1). In contrast, P2 amplitudes were influenced by an interaction between condition and music training (Figure 6B) [$F(1, 65)$ = 5.20, p = 0.026, η_p^2 = 0.07]. The interaction was driven by higher amplitudes for the passive relative to the active condition with less music training (p = 0.029). Paralleling latency, response amplitudes were invariant to attention at higher levels of music training (p = 0.14). All other main and interaction effects were non-significant (all ps > 0.1).

N1 latencies were invariant to all experimental manipulations (Figure 6C). However, N1 amplitudes varied with attention (Figure 6D) [$F(1, 65)$ = 4.82, p = 0.032, η_p^2 = 0.07], whereby N1 amplitudes were stronger (i.e., more negative) in the active than in the passive condition. All other main and interaction effects were non-significant (all ps > 0.08).
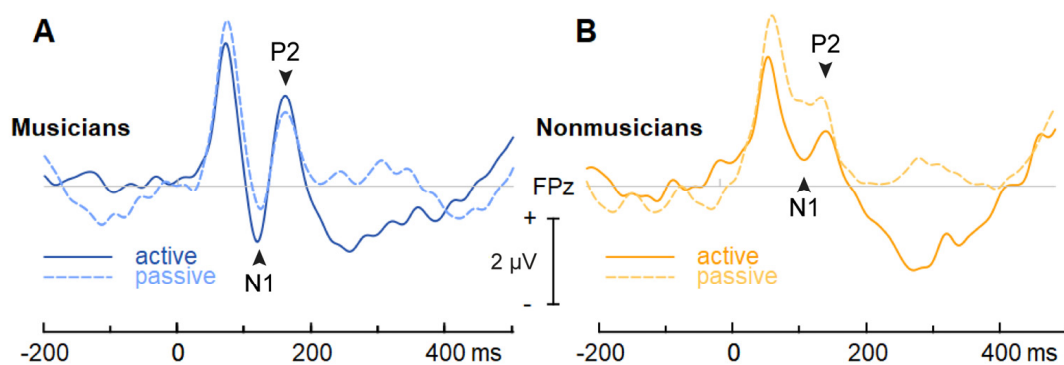

Figure 5. Group-averaged ERP waveforms during active and passive conditions. (**A**) Musicians' P2 latencies did not differ with attention. (**B**) Nonmusicians showed stronger P2 responses in the passive compared to active condition.
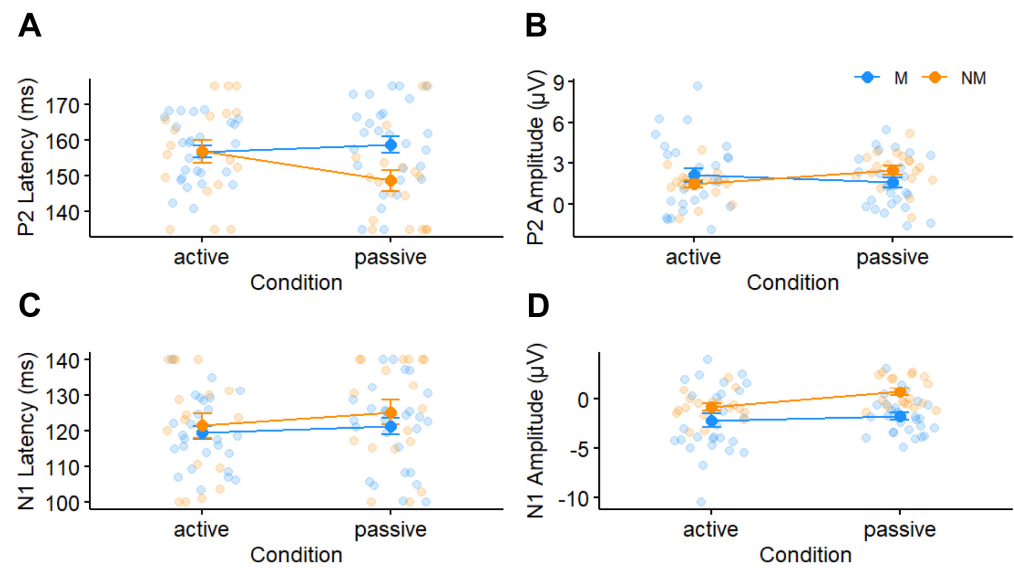
Figure 6. ERP latency and amplitude characteristics. (**A**) P2 latency and (**B**) amplitudes varied with attention and group. (**C**) N1 latency was invariant. (**D**) N1 amplitudes showed a stronger negativity with attention.

3.4. Brain–Behavior Relationships

Figure 7A displays the mean Euclidian distances between all pairwise ERPs (N1-P2 amplitudes) for musicians and nonmusicians. This measure reflects the degree to which the neural responses differentiated musical chords. Musicians demonstrated larger differences in ERP amplitudes across chords than non-musicians, suggesting more distinct neural responses [$t(13) = 2.59$, $p = 0.022$]. Figure 7B shows the correlation between neural ERP distance and behavioral information transfer. The positive association between brain and behavioral responses suggests that greater cortical neural differentiation of chordal stimuli (as in musicians) was related to better behavioral IT ($r = 0.59$, $p = 0.02$) and thus less confusability of the chords. In contrast to the cortical ERPs, none of the FFR measures, including onset and spectral amplitudes/latencies, showed this brain–behavioral correspondence (all ps > 0.05; data not shown). These effects are preserved when treating music training continuously.
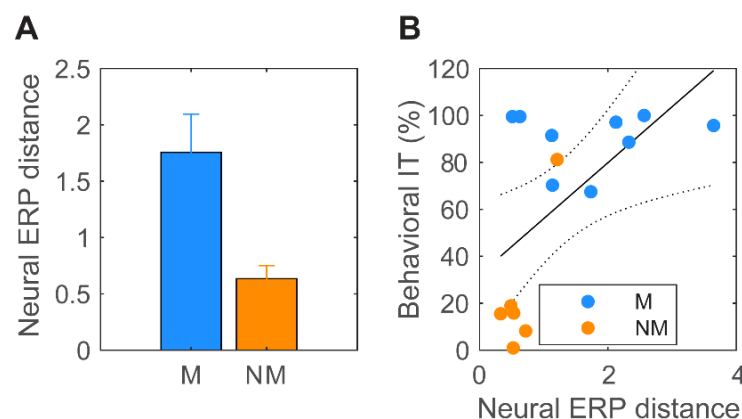

Figure 7. Brain–behavior relation in the differentiation of musical chords. (**A**) ERP distance between chords (Euclidian distance between all pairwise N1-P2 amplitudes) is larger in musicians, indicating more salient neural responses across triads. (**B**) ERP distance between chords correlates with behavioral IT, indicating a brain–behavior correspondence between perceptual and neural (cortical) differentiation. No such brain–behavior correspondence was observed for brainstem FFRs. Error bars: ±1 S.E.M.

4. Discussion

By recording subcortical FFRs and cortical ERPs to harmonic chordal stimuli during active and passive conditions in musicians and nonmusicians, we found that (i) subcortical responses did not vary with musicianship or attention; (ii) musicianship and attention strongly modulated cortical responses; and (iii) better cortical differentiation of chord stimuli in musicians corresponded with their better identification (e.g., fewer confusions) of musical sounds behaviorally.

4.1. (High-Frequency) Subcortical Responses Are Largely Invariant to Musicianship and Attention

As expected, we found FFRs faithfully encoded stimulus acoustics, representing differences in frequency of the chordal third. Similarly, FFR onset amplitude changed with stimuli, suggesting this feature is sensitive to changes in the chordal third in otherwise identical simultaneous chords.

Prior work using musical stimuli to elicit FFRs has been conducted during passive listening [1,2,39,56–60]. Yet, attention has been shown to modulate speech-evoked FFRs [22,23,37,61–63]. To determine the role of attention in subcortical music encoding, we recorded FFRs in both active and passive listening conditions. However, we did not observe an overall difference in FFRs with attention (though attention did interact with stimuli, as discussed below). The lack of a global attention effect could be due in part to our stimulus parameters. FFRs are generated by several cortical and subcortical sources whose engagement varies in a stimulus-dependent manner. Stimuli with high F0s elicit FFRs that are dominated by brainstem contributions [35,64–66]. The high F0 of our stimuli ($\geq$220 Hz), used to intentionally minimize cortical contributions to the FFR, attenuates attentional modulations from the cortex, which may drive previously reported FFR enhancements [67,68]. Though attention may also enhance subcortical FFRs, this modulation is likely dependent on top-down feedback from the cortex [37], the strength of which varies between individuals [69]. Consequently, it is also possible that listeners in our sample had weak cortical feedback, subduing any attentional gain to the FFR that is sometimes measurable [37].

Interestingly, we did observe an interaction between attention and stimulus on FFR onset amplitude. We found larger differences in onset amplitude with attention for the major vs. minor and detuned chords. This suggests some degree of stimulus-dependent attention effect at the brainstem level. Prior work has demonstrated that FFR pitch salience corresponds with perceptual ratings of consonance, with major chords having the highest pitch salience and consonance ratings [70]. The interaction we observed could reflect stimulus-dependent attentional enhancements corresponding to consonance or some salient feature of major, but not minor or detuned, chords (e.g., perceptual brightness).

Surprisingly, we did not observe changes in FFRs with musicianship. Several studies have reported enhanced FFRs to both speech [71–75] and musical (usually isolated notes) stimuli [1,5,57] in musicians. However, many of these studies used stimuli with low F0, potentially allowing for greater cortical contributions to the FFR. Indeed, musician's enhanced FFR might be attributable to larger cortical rather than subcortical phase locking [75]. Our use of higher F0 stimuli here intentionally elicits FFRs dominated by brainstem generators, which would render any cortical effects of long-term musical training unobservable. The findings here echo those of MacLean et al. [21], who also reported similar FFRs between musicians and nonmusicians for high-frequency speech stimuli. Our results support the notion that FFR enhancements traditionally reported for musicians might be influenced more by cortical rather than subcortical neuroplasticity.

4.2. Cortical Responses Vary with Attention and Musicianship

In stark contrast to brainstem FFRs, attentional effects were much stronger in the cortical ERPs. Consistent with prior literature on ERP attentional effects, we observed a large enhancement of N1 amplitudes in the active compared to the passive condition [24,44,45,76]. Comparisons between groups showed stronger P2 responses in nonmusicians during pas-

sive listening, whereas musicians' P2 was invariant to attentional manipulation. P2 is a wave that occurs relatively early after stimulus onset (~150 ms) and is often associated with perceptual auditory object formation [77]. Musicians' similar P2 profiles in passive and active conditions suggest automaticity in early sound processing. Contrastively, attention seemed to have a stronger effect on sensory coding in nonmusicians. Though we may have expected stronger P2 during active listening [78], nonmusicians' P2 was reduced in the active compared to the passive condition. Nonmusicians' P2 reductions in the active condition could result from habituation with short-term learning of the musical sounds. This parallels similar habituation effects observed during speech learning [21,79,80].

4.3. Cortical Stimulus Differentiation Mirrored Behavioral Discrimination

We observed a relationship between neural stimulus differentiation (assessed through pairwise N1-P2 amplitude differences) and behavioral information transfer. In other words, the degree to which listeners *behaviorally* distinguished triads was directly related to their *neural specificity* to stimuli primarily at a cortical (rather than subcortical) level. These results suggest that cortical levels of processing more closely support the behavioral distinction of chordal stimuli. Moreover, greater predictive power of ERPs vs. FFRs we find in predicting listeners' behavioral chord confusions implies that cortical neural representations convey a readout closer to the ultimate percept.

Our cortical findings support previous ERP studies examining stimulus differentiation in musicians vs. nonmusicians using mismatch negativity (MMN), a pre-attentive potential indexing the brain's automatic change detection for sounds [3,30,81,82]. Musicians' enhanced behavioral pitch discrimination abilities are mirrored in their enhanced MMNs to speech pitch and timbre [83], detuned chords [30], and changes in the chordal third of musical triads [3]. These findings suggest long-term music training with complex pitched stimuli facilitates the development of superior auditory discrimination at both the neural and behavioral levels [82]. Our results also mirror recent evidence demonstrating changes to cortical encoding of musical sounds following music training [84–86]. Here, we demonstrate that musicians' enhanced behavioral identification of musical chords corresponded with greater cortical discrimination in the ~150 ms time window. This provides further evidence that music training strengthens cortical specificity for musically relevant sounds.

4.4. Limitations

The small sample size of our study limits some conclusions and might warrant caution in interpretating some of the findings. Replication of this study with a larger sample size would be useful to draw definitive conclusions on the roles of attention and music training on cortical vs. subcortical processing of musical chords. Importantly, our findings were nearly identical whether treating music training as a group or continuous variable and so are unlikely to depend on the exact definitional split of a "musician". Additionally, we replicate previous work with much larger samples that similarly demonstrate enhanced behavioral musical chord processing [1] and cortical responses in musicians [21]. Thus, a larger sample would probably only reaffirm the strong cortical and behavioral effects we observe in the present study (Figures 6 and 7). It could be argued that the lack of a musician-related effect in FFRs is underpowered due to a smaller effect size of brainstem vs. cortical responses. However, we find this explanation unlikely as this replicates much larger studies reporting similar null group differences in brainstem responses [21]. Thus, the most parsimonious account of our FFR data is that the lack of FFR effects is due to the use of high-frequency stimuli (>220 Hz) that restrict FFRs to subcortical sources—which are less malleable to attention and neuroplastic effects of music [21,87].

5. Conclusions

Our results suggest that long-term music training strengthens the brain's encoding of musical sounds, particularly at a cortical, rather than subcortical, level of processing. We also find more attention-related changes in the ERPs of nonmusicians vs. musicians. We

propose that these findings add to existing evidence for more automated neural processing of musical chords at higher levels, which do not depend on attention. Additionally, our findings demonstrate that *cortical* specificity to musical triads is better in musicians and predicts their enhanced *behavioral* identification. This argues for a close relationship between high-level neural and behavioral distinction of musical sounds.

Author Contributions: G.M.B. designed the experiment, J.M., E.D. and R.R. collected the data, J.M., E.D. and G.M.B. analyzed the data, and all authors wrote the paper. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Indiana University Institutional Review Board (IRB) (protocol #14860, approval date 14 June 2024).

Informed Consent Statement: Informed, written consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy reasons.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bidelman, G.M.; Krishnan, A.; Gandour, J.T. Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *Eur. J. Neurosci.* **2011**, *33*, 530–538. [CrossRef] [PubMed]
2. Musacchia, G.; Sams, M.; Skoe, E.; Kraus, N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15894–15898. [CrossRef] [PubMed]
3. Koelsch, S.; Schroger, E.; Tervaniemi, M. Superior pre-attentive auditory processing in musicians. *Neuroreport* **1999**, *10*, 1309–1313. [CrossRef]
4. Tervaniemi, M.; Just, V.; Koelsch, S.; Widmann, A.; Schroger, E. Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Exp. Brain Res.* **2005**, *161*, 1–10. [CrossRef]
5. Bidelman, G.M.; Gandour, J.T.; Krishnan, A. Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neurosci. Lett.* **2011**, *503*, 203–207. [CrossRef]
6. Micheyl, C.; Delhommeau, K.; Perrot, X.; Oxenham, A.J. Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.* **2006**, *219*, 36–47. [CrossRef]
7. Bidelman, G.M.; Gandour, J.T.; Krishnan, A. Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *J. Cogn. Neurosci.* **2011**, *23*, 425–434. [CrossRef]
8. Kishon-Rabin, L.; Amir, O.; Vexler, Y.; Zaltz, Y. Pitch discrimination: Are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.* **2001**, *12*, 125–143. [CrossRef]
9. Pitt, M.A. Perception of pitch and timbre by musically trained and untrained listeners. *J. Exp. Psychol. Hum. Percept. Perform.* **1994**, *20*, 976–986. [CrossRef]
10. Yoo, J.; Bidelman, G.M. Linguistic, perceptual, and cognitive factors underlying musicians' benefits in noise-degraded speech perception. *Hear. Res.* **2019**, *377*, 189–195. [CrossRef]
11. Kraus, N.; Chandrasekaran, B. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* **2010**, *11*, 599–605. [CrossRef] [PubMed]
12. Alain, C.; Zendel, B.R.; Hutka, S.; Bidelman, G.M. Turning down the noise: The benefit of musical training on the aging auditory brain. *Hear. Res.* **2014**, *308*, 162–173. [CrossRef] [PubMed]
13. Mankel, K.; Bidelman, G.M. Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 13129–13134. [CrossRef] [PubMed]
14. Bidelman, G.M.; Yoo, J. Musicians show improved speech segregation in competitive, multi-talker cocktail party scenarios. *Front. Psychol.* **2020**, *11*, 1927. [CrossRef]
15. Brattico, E.; Jacobsen, T. Subjective appraisal of music. *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 308–317. [CrossRef]
16. Shahin, A.; Bosnyak, D.J.; Trainor, L.J.; Roberts, L.E. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *J. Neurosci.* **2003**, *23*, 5545–5552. [CrossRef]
17. Schneider, P.; Sluming, V.; Roberts, N.; Scherg, M.; Goebel, R.; Specht, H.J.; Dosch, H.G.; Bleeck, S.; Stippich, C.; Rupp, A. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat. Neurosci.* **2005**, *8*, 1241–1247. [CrossRef]

18. Virtala, P.; Huotilainen, M.; Partanen, E.; Tervaniemi, M. Musicianship facilitates the processing of Western music chords—An ERP and behavioral study. *Neuropsychologia* **2014**, *61*, 247–258. [CrossRef]
19. Bidelman, G.M.; Weiss, M.W.; Moreno, S.; Alain, C. Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur. J. Neurosci.* **2014**, *40*, 2662–2673. [CrossRef]
20. Bidelman, G.M.; Alain, C. Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J. Neurosci.* **2015**, *35*, 1240–1249. [CrossRef]
21. MacLean, J.; Stirn, J.; Sisson, A.; Bidelman, G.M. Short- and long-term neuroplasticity interact during the perceptual learning of concurrent speech. *Cereb. Cortex* **2024**, *34*, bhad543. [CrossRef] [PubMed]
22. Lai, J.; Price, C.N.; Bidelman, G.M. Brainstem speech encoding is dynamically shaped online by fluctuations in cortical α state. *NeuroImage* **2022**, *263*, 119627. [CrossRef] [PubMed]
23. Carter, J.A.; Bidelman, G.M. Perceptual warping exposes categorical representations for speech in human brainstem responses. *NeuroImage* **2023**, *269*, 119899. [CrossRef] [PubMed]
24. Picton, T.W.; Hillyard, S.A. Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr. Clin. Neurophysiol.* **1974**, *36*, 191–199. [CrossRef]
25. Brown, J.A.; Bidelman, G.M. Attention, musicality, and familiarity shape cortical speech tracking at the musical cocktail party. *bioRxiv* **2023**. [CrossRef]
26. Strait, D.L.; Kraus, N.; Parbery-Clark, A.; Ashley, R. Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hear. Res.* **2010**, *261*, 22–29. [CrossRef]
27. Bidelman, G.M. Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials. *J. Neurosci. Methods* **2015**, *241*, 94–100. [CrossRef]
28. Rizzi, R.; Bidelman, G.M. Duplex perception reveals brainstem auditory representations are modulated by listeners' ongoing percept for speech. *Cereb. Cortex* **2023**, *33*, 10076–10086. [CrossRef]
29. Zhang, J.D.; Susino, M.; McPherson, G.E.; Schubert, E. The definition of a musician in music psychology: A literature review and the six-year rule. *Psychol. Music* **2020**, *48*, 389–409. [CrossRef]
30. Brattico, E.; Pallesen, K.J.; Varyagina, O.; Bailey, C.; Anourova, I.; Jarvenpaa, M.; Eerola, T.; Tervaniemi, M. Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *J. Cogn. Neurosci.* **2009**, *21*, 2230–2244. [CrossRef]
31. Bidelman, G.M.; Walker, B. Attentional modulation and domain specificity underlying the neural organization of auditory categorical perception. *Eur. J. Neurosci.* **2017**, *45*, 690–699. [CrossRef] [PubMed]
32. Moore, B.C.J. *Introduction to the Psychology of Hearing*, 5th ed.; Academic Press: San Diego, CA, USA, 2003.
33. Joris, P.X.; Schreiner, C.E.; Rees, A. Neural processing of amplitude-modulated sounds. *Physiol. Rev.* **2004**, *84*, 541–577. [CrossRef] [PubMed]
34. Brugge, J.F.; Nourski, K.V.; Oya, H.; Reale, R.A.; Kawasaki, H.; Steinschneider, M.; Howard, M.A., III. Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J. Neurophysiol.* **2009**, *102*, 2358–2374. [CrossRef] [PubMed]
35. Bidelman, G.M. Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *Neuroimage* **2018**, *175*, 56–69. [CrossRef] [PubMed]
36. Campbell, T.; Kerlin, J.R.; Bishop, C.W.; Miller, L.M. Methods to eliminate stimulus transduction artifact from insert earphones during electroencephalography. *Ear Hear.* **2012**, *33*, 144–150. [CrossRef] [PubMed]
37. Price, C.N.; Bidelman, G.M. Attention reinforces human corticofugal system to aid speech perception in noise. *NeuroImage* **2021**, *235*, 118014. [CrossRef]
38. Bidelman, G.M. Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient abr. *Hear. Res.* **2015**, *323*, 68–80. [CrossRef]
39. Musacchia, G.; Strait, D.; Kraus, N. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.* **2008**, *241*, 34–42. [CrossRef]
40. Bidelman, G.M.; Moreno, S.; Alain, C. Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage* **2013**, *79*, 201–212. [CrossRef]
41. Saiz-Alía, M.; Forte, A.E.; Reichenbach, T. Individual differences in the attentional modulation of the human auditory brainstem response to speech inform on speech-in-noise deficits. *Sci. Rep.* **2019**, *9*, 14131. [CrossRef]
42. Galbraith, G.C.; Brown, W.S. Cross-correlation and latency compensation analysis of click-evoked and frequency-following brain-stem responses in man. *Electroencephalogr. Clin. Neurophysiol.* **1990**, *77*, 295–308. [CrossRef] [PubMed]
43. Bidelman, G.M.; Momtaz, S. Subcortical rather than cortical sources of the frequency-following response (FFR) relate to speech-in-noise perception in normal-hearing listeners. *Neurosci. Lett.* **2021**, *746*, 135664. [CrossRef] [PubMed]
44. Hillyard, S.A.; Hink, R.F.; Schwent, V.L.; Picton, T.W. Electrical signs of selective attention in the human brain. *Science* **1973**, *182*, 177–180. [CrossRef]
45. Schwent, V.L.; Hillyard, S.A. Evoked potential correlates of selective attention with multi-channel auditory inputs. *Electroencephalogr. Clin. Neurophysiol.* **1975**, *28*, 131–138. [CrossRef] [PubMed]
46. Woldorff, M.G.; Gallen, C.C.; Hampson, S.A.; Hillyard, S.A.; Pantev, C.; Sobel, D.; Bloom, F.E. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 8722–8726. [CrossRef] [PubMed]

47. Neelon, M.F.; Williams, J.; Garell, P.C. The effects of auditory attention measured from human electrocorticograms. *Clin. Neurophysiol.* **2006**, *117*, 504–521. [CrossRef]

48. Ding, N.; Simon, J.Z. Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11854–11859. [CrossRef]

49. Kaganovich, N.; Kim, J.; Herring, C.; Schumaker, J.; Macpherson, M.; Weber-Fox, C. Musicians show general enhancement of complex sound encoding and better inhibition of irrelevant auditory change in music: An ERP study. *Eur. J. Neurosci.* **2013**, *37*, 1295–1307. [CrossRef]

50. Brown, J.A.; Bidelman, G.M. Familiarity of background music modulates the cortical tracking of target speech at the "cocktail party". *Brain Sci.* **2022**, *12*, 1320. [CrossRef]

51. Lee, S.; Bidelman, G.M. Objective identification of simulated cochlear implant settings in normal-hearing listeners via auditory cortical evoked potentials. *Ear Hear.* **2017**, *38*, e215–e226. [CrossRef]

52. Miller, G.A.; Nicely, P.E. An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* **1955**, *27*, 338–352. [CrossRef]

53. Bidelman, G.M.; Yellamsetty, A. Noise and pitch interact during the cortical segregation of concurrent speech. *Hear. Res.* **2017**, *351*, 34–44. [CrossRef]

54. R-Core-Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020; Available online: https://www.R-project.org/ (accessed on 1 August 2024).

55. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [CrossRef]

56. Bidelman, G.M.; Krishnan, A. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.* **2009**, *29*, 13165–13171. [CrossRef] [PubMed]

57. Lee, K.M.; Skoe, E.; Kraus, N.; Ashley, R. Selective subcortical enhancement of musical intervals in musicians. *J. Neurosci.* **2009**, *29*, 5832–5840. [CrossRef] [PubMed]

58. Skoe, E.; Kraus, N. Hearing it again and again: On-line subcortical plasticity in humans. *PLoS ONE* **2010**, *5*, e13645. [CrossRef]

59. Bones, O.; Hopkins, K.; Krishnan, A.; Plack, C.J. Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia* **2014**, *58*, 23–32. [CrossRef]

60. Losorelli, S.; Kaneshiro, B.; Musacchia, G.A.; Blevins, N.H.; Fitzgerald, M.B. Factors influencing classification of frequency following responses to speech and music stimuli. *Hear. Res.* **2020**, *398*, 108101. [CrossRef]

61. Galbraith, G.; Olfman, D.M.; Huffman, T.M. Selective attention affects human brain stem frequency-following response. *Neuroreport* **2003**, *14*, 735–738. [CrossRef]

62. Lehmann, A.; Schonwiesner, M. Selective attention modulates human auditory brainstem responses: Relative contributions of frequency and spatial cues. *PLoS ONE* **2014**, *9*, e85442. [CrossRef]

63. Forte, A.E.; Etard, O.; Reichenbach, T. The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *eLife* **2017**, *6*, e27203. [CrossRef] [PubMed]

64. Tichko, P.; Skoe, E. Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. *Hear. Res.* **2017**, *348*, 1–15. [CrossRef] [PubMed]

65. Coffey, E.B.J.; Nicol, T.; White-Schwoch, T.; Chandrasekaran, B.; Krizman, J.; Skoe, E.; Zatorre, R.J.; Kraus, N. Evolving perspectives on the sources of the frequency-following response. *Nat. Commun.* **2019**, *10*, 5036. [CrossRef]

66. Gorina-Careta, N.; Kurkela, J.L.O.; Hämäläinen, J.; Astikainen, P.; Escera, C. Neural generators of the frequency-following response elicited to stimuli of low and high frequency: A magnetoencephalographic (MEG) study. *NeuroImage* **2021**, *231*, 117866. [CrossRef]

67. Hartmann, T.; Weisz, N. Auditory cortical generators of the frequency following response are modulated by intermodal attention. *NeuroImage* **2019**, *203*, 116185. [CrossRef]

68. Schüller, A.; Schilling, A.; Krauss, P.; Rampp, S.; Reichenbach, T. Attentional modulation of the cortical contribution to the frequency-following response evoked by continuous speech. *J. Neurosci.* **2023**, *43*, 7429–7440. [CrossRef]

69. Lai, J.; Alain, C.; Bidelman, G.M. Cortical-brainstem interplay during speech perception in older adults with and without hearing loss. *Front. Neurosci.* **2023**, *17*, 1075368. [CrossRef]

70. Bidelman, G.M.; Krishnan, A. Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport* **2011**, *22*, 212–216. [CrossRef]

71. Parbery-Clark, A.; Strait, D.L.; Kraus, N. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* **2011**, *49*, 3338–3345. [CrossRef]

72. Strait, D.L.; Kraus, N.; Skoe, E.; Ashley, R. Musical experience and neural efficiency: Effects of training on subcortical processing of vocal expressions of emotion. *Eur. J. Neurosci.* **2009**, *29*, 661–668. [CrossRef]

73. Wong, P.C.; Skoe, E.; Russo, N.M.; Dees, T.; Kraus, N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* **2007**, *10*, 420–422. [CrossRef] [PubMed]

74. Parbery-Clark, A.; Skoe, E.; Lam, C.; Kraus, N. Musician enhancement for speech-in-noise. *Ear Hear.* **2009**, *30*, 653–661. [CrossRef] [PubMed]

75. Coffey, E.B.; Herholz, S.C.; Chepesiuk, A.M.; Baillet, S.; Zatorre, R.J. Cortical contributions to the auditory frequency-following response revealed by MEG. *Nat. Commun.* **2016**, *7*, 11070. [CrossRef] [PubMed]

76. Hansen, J.C.; Hillyard, S.A. Endogenous brain potentials associated with selective auditory attention. *Electroencephalogr. Clin. Neurophysiol.* **1980**, *49*, 277–290. [CrossRef]

77. Ross, B.; Jamali, S.; Tremblay, K.L. Plasticity in neuromagnetic cortical responses suggests enhanced auditory object representation. *BMC Neurosci.* **2013**, *14*, 151. [CrossRef]

78. Crowley, K.E.; Colrain, I.M. A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clin. Neurophysiol.* **2004**, *115*, 732–744. [CrossRef]

79. Alain, C.; Campeanu, S.; Tremblay, K.L. Changes in sensory evoked responses coincide with rapid improvement in speech identification performance. *J. Cogn. Neurosci.* **2010**, *22*, 392–403. [CrossRef]

80. Mankel, K.; Shrestha, U.; Tipirneni-Sajja, A.; Bidelman, G.M. Functional plasticity coupled with structural predispositions in auditory cortex shape successful music category learning. *Front. Neurosci.* **2022**, *16*, 897239. [CrossRef]

81. Tervaniemi, M. Musicians—Same or different? *Ann. N. Y. Acad. Sci.* **2009**, *1169*, 151–156. [CrossRef]

82. Putkinen, V.; Tervaniemi, M.; Saarikivi, K.; de Vent, N.; Huotilainen, M. Investigating the effects of musical training on functional brain development with a novel melodic MMN paradigm. *Neurobiol. Learn. Mem.* **2014**, *110*, 8–15. [CrossRef]

83. Hutka, S.; Bidelman, G.M.; Moreno, S. Pitch expertise is not created equal: Cross-domain effects of musicianship and tone language experience on neural and behavioural discrimination of speech and music. *Neuropsychologia* **2015**, *71*, 52–63. [CrossRef] [PubMed]

84. Jo, H.S.; Hsieh, T.H.; Chien, W.C.; Shaw, F.Z.; Liang, S.F.; Kung, C.C. Probing the neural dynamics of musicians and non-musicians' consonant/dissonant perception: Joint analyses of electrical encephalogram (EEG) and functional magnetic resonance imaging (fMRI). *NeuroImage* **2024**, *298*, 120784. [CrossRef] [PubMed]

85. Panda, S.; Shivakumar, D.; Majumder, Y.; Gupta, C.N.; Hazra, B. Real-time dynamic analysis of EEG response for live Indian classical vocal stimulus with therapeutic indications. *Smart Health* **2024**, *32*, 100461. [CrossRef]

86. Jiang, Y.; Zheng, M. EEG microstates are associated with music training experience. *Front. Hum. Neurosci.* **2024**, *18*, 1434110. [CrossRef]

87. Holmes, E.; Purcell, D.W.; Carlyon, R.P.; Gockel, H.E.; Johnsrude, I.S. Attentional modulation of envelope-following responses at lower (93–109 Hz) but not higher (217–233 Hz) modulation rates. *J. Assoc. Res. Otolaryngol.* **2018**, *19*, 83–97. [CrossRef]