

The Influence of Lifelong Musicianship on Neurophysiological Measures of Concurrent Sound Segregation

Benjamin Rich Zendel^{1,2} and Claude Alain^{1,2}

Abstract

■ The ability to separate concurrent sounds based on periodicity cues is critical for parsing complex auditory scenes. This ability is enhanced in young adult musicians and reduced in older adults. Here, we investigated the impact of lifelong musicianship on concurrent sound segregation and perception using scalp-recorded ERPs. Older and younger musicians and nonmusicians were presented with periodic harmonic complexes where the second harmonic could be tuned or mistuned by 1–16% of its original value. The likelihood of perceiving two simultaneous sounds increased with mistuning, and musicians, both older and younger, were more likely to detect and report hearing two sounds when the second harmonic was mistuned

at or above 2%. The perception of a mistuned harmonic as a separate sound was paralleled by an object-related negativity that was larger and earlier in younger musicians compared with the other three groups. When listeners made a judgment about the harmonic stimuli, the perception of the mistuned harmonic as a separate sound was paralleled by a positive wave at about 400 msec poststimulus (P400), which was enhanced in both older and younger musicians. These findings suggest attention-dependent processing of a mistuned harmonic is enhanced in older musicians and provides further evidence that age-related decline in hearing abilities are mitigated by musical training. ■

INTRODUCTION

Most auditory scenes are complex, in that there are multiple active sound sources at any given time. Accordingly, one critical process in auditory scene analysis is the ability to segregate concurrent sounds (Alain, 2007; Bregman, 1990). Sounds that are perceptually segregated can then be tracked as separate auditory streams over time to form a dynamic perceptual auditory scene (Alain & Bernstein, 2008; Carlyon, 2004). There are multiple cues the auditory system can use to detect the presence of concurrent sound objects, including onset asynchrony, spatial location, differences in fundamental frequency (f_0), and periodicity (Hautus & Johnson, 2005; McDonald & Alain, 2005; Assmann & Summerfield, 1994; Bregman, 1990). When acoustic energy is periodic across the frequency spectrum (i.e., bands of acoustic energy [harmonics or overtones] are multiples of a f_0), the separate bands of energy are perceived as a single sound object whereas acoustic energy that is not related to the same f_0 is segregated into a second auditory precept. This is because natural sound sources (i.e., vibrating bodies) normally produce periodic patterns of acoustic energy.

Given the importance of periodicity in auditory perception, it is not surprising that humans are very sensitive to mistuned (i.e., nonperiodic) harmonics (Moore, Peters, &

Glasberg, 1985). Mistuning higher frequency harmonic components adds a roughness to overall timbre of the sound (Moore et al., 1985), whereas mistuning lower harmonics results in the perception of two simultaneous sounds, one with a “buzz-like” quality and the other with a pure-tone “beep-like” quality (Alain, 2007; Moore, Glasberg, & Peters, 1986). The differential effect of mistuning higher or lower harmonics is thought to be related to the ability of auditory nerve fibers to phase lock to acoustic energy (Hartmann, McAdams, & Smith, 1990).

Using auditory ERPs, Alain, Arnott, & Picton (2001) found that the perception of concurrently occurring sounds is paralleled by an increase in negativity, known as an “object related negativity” (ORN) that peaks between the N1 and P2 waves, around 150 msec poststimulus onset. Importantly, the ORN reflects the perception of concurrent sound objects, as it has been observed when the perception of the second sound is due to a mistuned harmonic (Alain, Schuler, & McDonald, 2002), a dichotic pitch produced by interaural time differences (Hautus & Johnson, 2005), the spatial location of a harmonic component (McDonald & Alain, 2005), the onset asynchrony of a harmonic component (Lipp, Kitterick, Summerfield, Baily, & Paul-Jordanov, 2010), or a difference in f_0 between concurrent vowels (Alain, Reinke, He, Wang, & Lobaugh, 2005). Interestingly, ORN amplitude was reduced when listeners were given sequential cues that could aid in segregating the mistuned component, such as increased

¹Rotman Research Institute, Toronto, Canada, ²University of Toronto

stimulus probability (Bendixen, Jones, Klump, & Winkler, 2010) or an onset asynchrony between the mistuned harmonic and the harmonic complex (Weise, Schröger, & Bendixen, 2012), which provides further support for the hypothesis that the ORN is related to the perception of concurrent sounds. Moreover, the ORN is thought to index the automatic detection of the mistuned harmonic as a separate sound object because it can be observed even when the stimuli are not task relevant (e.g., participants reading a book or watching a silent, subtitled movie; Alain, 2007) and is little influenced by task demands or selective attention (Alain & Izenberg, 2003). Critically, when listeners were asked to make a perceptual judgment about the incoming acoustic stimulus, the amplitude of the ORN correlated with the likelihood of reporting the perception of concurrently occurring sounds (Alain, Arnott, et al., 2001).

In addition to the ORN, a P400 can also be observed when a listener consciously detects the presence of concurrently occurring sounds. The P400 is a positive wave that peaks around 400 msec poststimulus onset, is correlated with the likelihood of perceiving concurrent sounds, and is only observed when listeners are asked to make a judgment about an incoming acoustic stimulus (Alain, 2007; Hautus & Johnson, 2005; Alain, Arnott, et al., 2001). Given that the P400 was present only when participants were required to make a response, it is thought to index the conscious registration of concurrent sound objects and reflects the transfer of the automatic detection of a second auditory object to a working memory process where the second object can be identified. It is therefore likely that concurrent sound segregation occurs in two stages. In the first stage, acoustic features are organized automatically, regardless of a listeners' attentional state; this stage is reflected in the ORN. In the second stage, there is a conscious registration of the automatically segregated mistuned harmonic as a second auditory object. This stage of processing requires a listeners' focused attention and is reflected in the P400.

Aging and Concurrent Sound Segregation

Older adults often have difficulty segregating speech from background noise (e.g., Pichora-Fuller, Schneider, & Daneman, 1995; Duquesnoy, 1983), a problem that may be partly related to deficits in parsing concurrent sounds. Indeed, older adults have more difficulty detecting inharmonicity within a harmonic complex (Zendel & Alain, 2012; Grube, von Cramon, & Rübsamen, 2003; Alain, McDonald, Ostroff, & Schneider, 2001), and when passively presented with mistuned harmonic stimuli, the ORN is reduced in older adults (Alain & McDonald, 2007). When concurrent vowel sounds were presented, the ORN associated with segregating and identifying two vowels presented simultaneously was smaller in older adults; however, later activity (reported as an N2b) related to the conscious detection of concurrently occurring vowels was comparable between older and younger adults (Snyder & Alain, 2005).

This pattern of results suggests that aging negatively impacts automatic processing of acoustic features, whereas attention-dependent, endogenous processing of acoustic information is relatively spared. Further support for this theory comes from a gap detection paradigm where older, middle-aged, and younger adults were asked to detect a stimulus that contained a near threshold silent gap (i.e., the gap was longer for older adults; Alain, McDonald, Ostroff, & Schneider, 2004). Despite the gaps being equally detectable by all age groups, neural activity related to the automatic processing of acoustic information was reduced in older adults whereas the ERP wave (i.e., P3b) related to conscious detection of the gap was preserved (Alain et al., 2004). Behaviorally, when listening to speech in noisy environments, older adults use contextual cues within the sentence to overcome age-related decline in hearing abilities (Pichora-Fuller et al., 1995). These studies demonstrate that older adults likely rely on attention-dependent cognitive mechanisms to overcome presbycusis and age-related decline in automatic auditory processing. Accordingly, one important question is how aging influences the P400 wave when using a mistuned harmonic paradigm. Alain and McDonald (2007) recorded ERPs to mistuned harmonic stimuli only during passive listening, and thus no P400 was evoked. The N2b reported in Snyder and Alain (2005) was likely related to the conscious perception of simultaneous vowels; however, the use of vowel sounds likely engaged schema-driven processes because of the overlearned nature of speech stimuli; thus, this negativity is likely different from the P400 wave observed in young adults while using a mistuned harmonic paradigm.

Musicians and Concurrent Sound Segregation

Although aging has a deleterious effect on the ability to detect and segregate a mistuned harmonic, young musicians have an enhanced ability to detect a mistuned harmonic component (Koelsch, Schroger, & Tervaniemi, 1999), an advantage that remains throughout the lifespan (Zendel & Alain, 2012). More importantly, musicians are more likely to hear a mistuned harmonic as a separate auditory object, and accordingly, the ORN and P400 are enhanced in younger musicians (Zendel & Alain, 2009). Furthermore, participants trained to segregate simultaneous vowels showed significant improvement in their ability to correctly identify both vowels (Reinke, He, Wang, & Alain, 2003), which suggests that the advantage musicians have in segregating concurrently occurring sounds is at least partially due to training and not inborn genetic predispositions. Finally, the benefits of musical training extend beyond the ability to separate concurrently occurring sounds into other domains of auditory processing (e.g., Schellenberg & Moreno, 2010; Micheyl, Delhommeau, Perrot, & Oxenham, 2006; Rammsayer & Altenmüller, 2006; Koelsch et al., 1999; Jeon & Fricke, 1997).

Although concurrent sound segregation is enhanced in musicians and declines in older adults, the relationship between aging and musical training is less well understood. Previous research has found that age-related decline of gray matter volume in Broca's area is mitigated in older musicians (Sluming et al., 2002). Krampe and Ericsson (1996) found that older musicians experienced less age-related decline on speeded motor tasks related to music performance, but that general processing speed was not influenced by musical training. Andrews, Dowling, Bartlett, and Halpern (1998) found that the ability to recognize speeded or slowed melodies declines with age and that musicians were better than nonmusicians, but that age and musical training did not interact. Meinz (2000) found that memory and perceptual speed in musical situations declined with age in pianists and that more experienced pianists performed better than nonmusicians, but that all levels of pianists declined at the same rate. Finally, subcortical responses to speech sounds are enhanced in older musicians compared with older nonmusicians (Parbery-Clark, Anderson, Hitter, & Kraus, 2012). Although these studies investigated diverse cognitive abilities, they consistently demonstrate an advantage for older musicians.

The goal of the current study was to investigate how aging and musical training interact to influence concurrent sound perception. It is likely that older musicians, compared with older nonmusicians, will have an advantage in segregating a mistuned harmonic as a separate sound object. Given the age-related switch to controlled processing of acoustic information, (Snyder & Alain, 2005; Alain et al., 2004), it is likely that the benefit for older musicians will be reflected in the P400 and not in the ORN.

METHODS

Participants

Fifty-seven participants were recruited for the study and provided formal informed consent in accordance with the joint Baycrest Centre and University of Toronto Research Ethics Committee. These participants were made up of four groups: older musicians (range = 58–91 years, $M = 69$ years, $SD = 9.24$ years), older nonmusicians (range = 61–84 years, $M = 69.2$ years, $SD = 6.69$ years), younger musicians (range: 23–33 years, $M = 28.1$ years, $SD = 3.17$ years), and younger nonmusicians (range = 23–39 years, $M = 29.9$ years, $SD = 5.97$ years). Musicians were defined as having advanced musical training (e.g., university degree, Royal Conservatory Grade 8, college diploma, or equivalent) and continued practice on a regular basis until the day of testing, whereas nonmusicians had no more than 2 years of formal training throughout life and did not currently play a musical instrument. The musicians played a variety of musical instruments; the most common primary instruments played were piano ($n = 8$) and clarinet ($n = 4$). Two participants each played violin, voice, trumpet, trombone, saxophone, or percussion. Finally, the French horn,

guitar, bassoon, tuba, and euphonium were each played by one participant. All participants were screened for neurological or psychiatric illness and hearing loss. Noise-induced hearing loss is a common problem for older musicians because of lifelong exposure to high amplitude sounds (Jansen, Helleman, Dreschler, & de Laat, 2009). Not surprisingly, some participants in the older musician group met the criteria for mild hearing loss based on a pure-tone threshold audiometric assessment (i.e., 25–35 dB HL). To compensate for this, older nonmusicians with mild-hearing loss were recruited so that pure-tone thresholds in older nonmusicians did not differ from older musicians. To confirm this, a 2 (Musical Training: musician, non-musician) \times 6 (Pure Tone Frequency: 250, 500, 1000, 2000, 4000, and 8000 Hz) repeated-measures ANOVA was calculated for the older adults. Neither the main effect of Musical Training nor the interaction between Musical Training and Pure Tone Frequency was significant ($p > .5$ for both). All younger adults had pure thresholds within the normal range (i.e., below 25 dB HL at all frequency octaves). Finally, the majority of participants were monolingual; however, 14 of the participants were bilingual. There were three bilingual participants in the younger musician group, four in the older musician group, seven in younger nonmusician group, and none in the older non-musician group.

Stimuli

Stimuli consisted of six complex sounds that were created by adding together six pure tones of equal intensity (i.e., 220, 440, 660, 880, 1100, and 1320 Hz). The f_0 was 220 Hz, and the third tonal element was either tuned (i.e., 660 Hz) or mistuned by 1% (666.6 Hz), 2% (673.2 Hz), 4% (686.4 Hz), 8% (712.8 Hz), or 16% (675.6 Hz) of its original value, yielding six complex sounds, henceforth referred to as "Stimulus type."

The pure tones were generated at a sampling rate of 22,050 Hz using Sig-Gen software (Tucker-Davis Technology, Alachua, FL) and were combined into a harmonic complex using Cubase SX (Steinberg, V.3.0, Las Vegas, NV). All six harmonic complex tones had durations of 150 msec with 10 msec rise/fall times. They were presented binaurally at 80 decibels sound pressure level (dB SPL) using a GSI 61 Clinical Audiometer via ER-3A transducers (Etymotic Research, Elk Grove, IL, USA). The intensity of the stimuli were measured using a Larson–Davis sound pressure level meter.

Procedure

The same stimuli were used in active and passive listening conditions. In both listening conditions, 720 stimuli were presented (120 exemplars of each Stimulus type). The stimuli were presented at an ISI that was randomly varied according to a rectangular distribution between 1200 and 2000 msec during passive trials and 2000–3000 msec

during active trials to allow time for a response. In the active listening condition, participants were asked to indicate whether the incoming stimulus was perceived as a single complex sound (i.e., a buzz) or two concurrently occurring sounds (i.e., a buzz plus another sound with a pure tone quality; see Alain, Arnott, et al., 2001; Moore et al., 1986). Responses were registered using a multibutton response box, and no feedback related to the responses was given. In the passive condition, participants were instructed to relax and to ignore the sounds while watching a muted subtitled movie of their choice. This design allowed for the examination of the effects of age and musical training on exogenous cortical activity elicited by stimuli while minimizing the influence of top-down processes on ERP amplitudes. The use of muted subtitled movies has been shown to effectively capture attention without interfering with auditory processing (Pettigrew et al., 2004). All participants completed six blocks of trials. The first and last blocks were passive, and each included 360 stimulus presentations (60 exemplars of each stimulus type); the middle four blocks were active, and each included 180 stimulus presentations (30 exemplars of each stimulus type). The experimental procedure lasted about 1 hr.

Recording of Electrical Brain Activity

Neuroelectric brain activity was digitized continuously from 65 scalp locations with a band-pass filter of 0.05–100 Hz and a sampling rate of 500 Hz per channel using SynAmps2 amplifiers (Compumedics Neuroscan, El Paso, TX) and stored for offline analysis. Electrodes on the outer canthi and at the superior and inferior orbit monitored ocular activity (IO1, IO2, LO1, LO2, FP9, and FP10). During recording, all electrodes were referenced to the midline central electrode (i.e., Cz); however, for data analysis, the ERPs were rereferenced to an average reference, and electrode Cz was reinstated. All averages were computed using BESA software (version 5.2). The analysis epoch included 100 msec of prestimulus activity and 1000 msec of poststimulus activity. Trials containing excessive noise ($\pm 130 \mu\text{V}$) at electrodes not adjacent to the eyes (i.e., IO1, IO2, LO1, LO2, FP1, FP2, FPz, FP9, and FP10) were rejected before averaging. ERPs were then averaged separately for each condition, stimulus type, and electrode site.

For each participant, a set of ocular movements was obtained before and after the experiment (Picton et al., 2000). From this recording, averaged eye movements were calculated both for lateral and vertical eye movements as well as for eye blinks. A PCA of these averaged recordings provided a set of components that best explained the eye movements. The scalp projections of these components were then subtracted from the experimental ERPs to minimize ocular contamination such as blinks, saccades, and lateral eye movements for each individual average. ERPs were then digitally low-pass filtered to attenuate frequencies above 30 Hz.

Data Analysis (Behavioral)

For the behavioral task, participants were asked to indicate whether they heard the incoming harmonic complex as either a single buzz or a buzz with an additional pure-tone (beep-like) component by pressing a button on a response box. The behavioral data were analyzed in two ways. The first utilized the percentage of trials that participants reported hearing two sounds as the dependent measure. For the tuned stimulus, this measure approaches zero percent (i.e., most trials were perceived as a single sound) whereas the perceptual judgment of 16% mistuned stimulus approaches 100% (i.e., most trials were reported as two sounds). This analysis was termed “perceptual judgment.”

RT was calculated from the onset of the stimulus to the button press indicating a response and is reported in milliseconds (msec).

d' is the difference in the z -score distribution between hits and false alarms. For the calculation of d' , trials in which participants were presented with a tuned stimulus and reported hearing two sounds were treated as false alarms and trials on which participants were presented with mistuned stimulus and reported hearing two sounds were treated as hits (Moore et al., 1986). Accordingly, d' cannot be calculated for the tuned stimulus. Higher d' indicates greater ability to detect the mistuned harmonic.

All behavioral measures were statistically analyzed with a 6 (Stimulus type [5 levels for d']) \times Age group [2] \times Musical training [2] mixed design repeated-measures ANOVA, and the probability values of all follow-up comparisons were corrected using the Bonferroni procedure. In situations where there was heterogeneity of variance between conditions, the degrees of freedom were adjusted using the Greenhouse–Geisser epsilon. In these cases, the original degrees of freedom were reported, but the p values were adjusted.

Data Analysis (Electrophysiological)

ORN amplitude was quantified as the mean amplitude during the 100–190 msec epoch, over nine fronto-central electrodes (F1, Fz, F2, FC1, FCz, FC2, C1, Cz, and C2). These sites were chosen because previous studies have found that the ORN is largest at frontocentral sites (Alain, 2007; Alain, Arnott, et al., 2001). A visual inspection of the current data confirmed a similar topography in all participants; slight differences in the topography between groups are accounted for by including multiple electrode sites. This epoch was chosen because it captured the peak amplitude of the ORN in each group (see Results). Importantly, the ORN is a difference wave (i.e., ERPs from a tuned stimulus is subtracted from the mistuned stimulus) and is therefore measured statistically as a main effect of Stimulus type. Specifically, the ORN is due to an increase in negativity during the 100–190 msec epoch, related to an increasing amount of mistuning of a single harmonic

in the stimulus. During active listening, this increase in negativity is associated with the likelihood of hearing concurrently occurring sounds; however, the ORN is also observed during passive listening. Therefore, only the main effect of Stimulus type and interactions with Stimulus type are indicative of an ORN. To quantify the change in mean amplitude related to Stimulus type, orthogonal polynomial decompositions were calculated, with a focus on the linear or quadratic trends. Before analysis, activity from each of the nine frontocentral electrodes was rereferenced to the linked mastoid. That is, the average amplitude of electrodes M1 and M2 was subtracted from the amplitude of each of the frontocentral electrodes. The purpose of this rereferencing was to maximize voltage potentials at frontocentral sites. Previously, source analysis of the ORN (and P400) revealed generators along the superior temporal plane that was oriented toward the vertex (i.e., electrode Cz; Alain, Arnott, et al., 2001). This source configuration results in a polarity reversal at mastoid sites. Visual analysis of the ORN scalp topographies from the current data set confirm that the ORN was maximal around electrode Cz (see Figure 1, top views) and was reversed in polarity at mastoid sites (see Figure 1, side views). Thus, by using a linked mastoid reference, the polarity reversal was included in the analysis of the frontocentral electrodes, which increases the ORN amplitude over the frontocentral scalp region. The analysis was carried out using a mixed design ANOVA that included Age group and Musical training as between-subject factors and Listening condition and Stimulus type as within-subject factors.

P400 amplitude was quantified as the mean amplitude during the 250–350 and 350–450 msec epoch over a frontal-right electrode montage (FC2, C2, CP2, C4, FC6, CP6, and

C6). Separate epochs were used because the morphology and time course of the P400 was different in each group, despite having similar peak latency. The early P400 window was chosen to capture the onset of the P400 response, and the late P400 window was chosen to capture the offset of the response. This electrode montage was chosen based on a visual inspection of the data that revealed the P400 peak to have a fronto-right distribution for all participants (see Figure 1). The P400 is best illustrated as a difference wave (i.e., ERPs from a tuned stimulus is subtracted from the mistuned stimulus) and is therefore expressed statistically as a main effect of Stimulus type. The analyses of P400 amplitude and latency were limited to data from the active listening condition, as there was no clear P400 during passive listening. To quantify the change in mean amplitude related to Stimulus type, orthogonal polynomial decompositions were calculated, with a focus on the linear or quadratic trends. Like the ORN data, the P400 data were rereferenced to the linked mastoid. Statistical analyses were the same as the ORN analysis, except they did not include Listening condition as a factor.

Whereas the amplitude of the ORN and P400 was quantified by comparing the mean amplitude between tuned and mistuned conditions, the latencies for the ORN and P400 were determined by calculating a difference wave between the tuned and 16% mistuned condition for each participant. This limits the measure of ORN and P400 latency to the 16% mistuned condition. This was done because the 16% mistuned condition resulted in a clear ORN and P400 in all participants, while the ORN and P400 became increasingly difficult to observe when the stimulus had smaller levels of mistuning. ORN latency was defined as the largest negative value in the difference wave between 100 and 200 msec poststimulus onset at frontocentral electrodes,

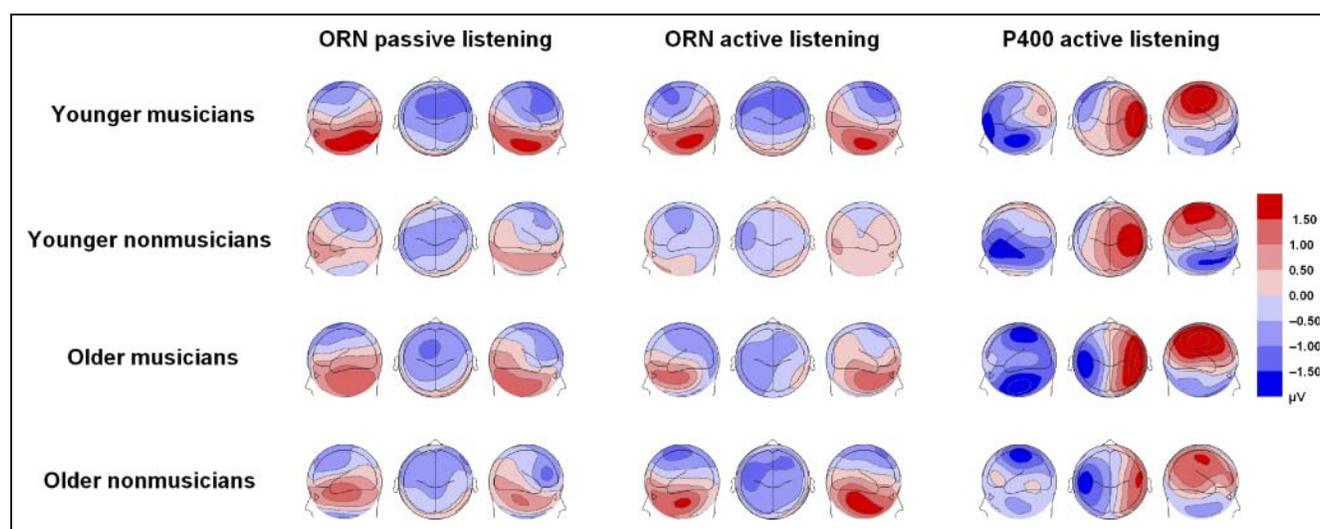


Figure 1. Topographical map of the ORN (during both active and passive listening) and the P400. Each column illustrates the head from the left, top, and right views. Latencies for the ORN were chosen based on peak latency from the frontocentral electrode array, whereas latencies for the P400 were chosen based on peak latency from the frontal-right electrode array. For younger musicians, the ORN latency was 131 msec and the P400 latency was 395 msec. For younger nonmusicians, the ORN latency was 156 msec and the P400 latency was 395 msec. For older musicians, the ORN latency was 151 msec and the P400 latency was 395 msec. For older nonmusicians, the ORN latency was 143 msec and P400 was 395 msec.

whereas P400 latency was calculated as the largest positive value in the difference wave between 250 and 500 msec poststimulus onset at frontal-right electrodes. P400 latency was only calculated for active listening because there was no P400 during passive listening. The peak amplitude of the ORN and P400 was also calculated from the same data. That is, the amplitude of the largest negative deflection between 100 and 200 msec poststimulus onset was the ORN peak amplitude, and the largest positive deflection between 250 and 500 msec poststimulus onset was the P400 peak amplitude. The final analysis for ORN and P400 peak latency and amplitude was a 2 (Musical training) \times 2 (Age Group) \times 2 (Listening condition [ORN only]) ANOVA. All post hoc analyses are corrected for multiple comparisons using the Bonferonni procedure.

One issue related to analyzing ERP components at a specific montage of scalp electrodes is that the underlying neural sources may be different in each group. To determine if there were age- or musical training-related shifts in the sources of the ORN and P400 an analysis of the topography of the each of the four components was calculated (ORN active, ORN passive, P400 250–350 msec, and P400 350–450 msec). First, data were normalized within each subject and for each component using the original, average referenced data. Normalization was done by subtracting the minimum value and dividing by the difference between the minimum and maximum value at all 65 electrodes (McCarthy & Wood, 1985). These values were then compared using an ANOVA that included Age group, Musical training, and Electrode as factors. Significant group by electrode interactions suggests topographical differences between groups and thus differences in the underlying neural sources.

For all electrophysiological measures, in situations where there was heterogeneity of variance between conditions, the degrees of freedom were adjusted using the Greenhouse–Geisser epsilon. In these cases, the original degrees of freedom were reported, but the p values were adjusted.

To determine the relationship between behavioral and electrophysiological measures, within-subject correlations were calculated between the amplitude of ORN (during active listening) and P400 (early and late epoch were included separately) with the three behavioral measures: perceptual judgment, RT, and d' . That is, a correlation coefficient was calculated for each participant between the electrophysiological data and each behavioral measure across the six levels of mistuning. The mean of these correlations is reported and is indicative of the relationship between behavior and electrophysiology in each participant. Significance was assessed using a one-sample t test that compared the value of the correlation coefficient to zero ($\alpha > 0.05$).

RESULTS

Perceptual Judgment

Figure 2A shows the group mean perceptual judgment in younger and older musicians and nonmusicians. As

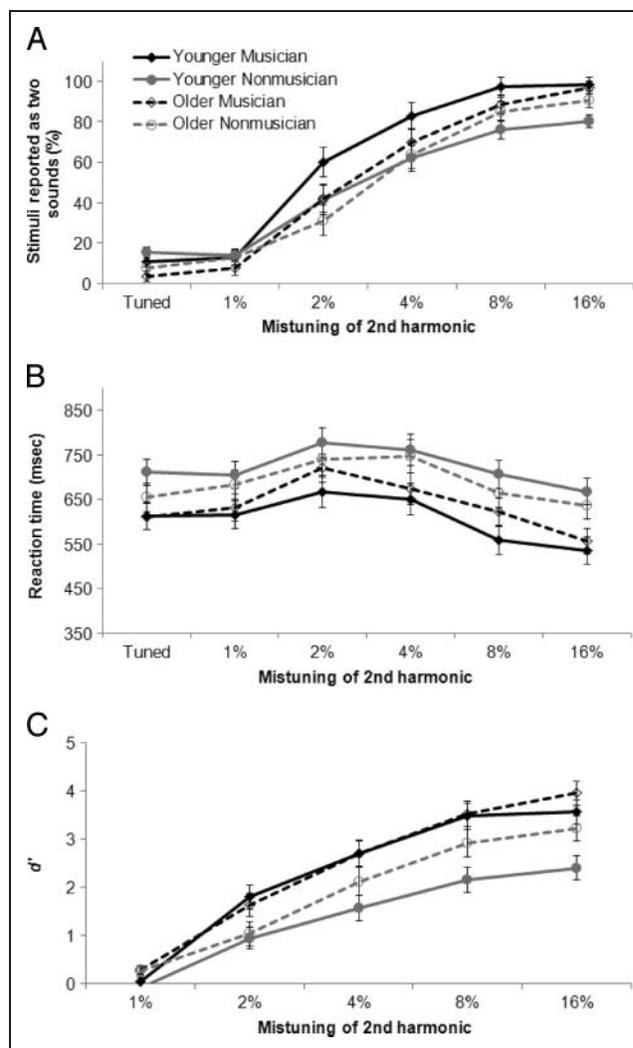


Figure 2. (A) Perceptual judgment: Percentage of responses perceived as two sounds as a function of stimulus type. (B) RT to a harmonic complex in milliseconds (msec) as a function of stimulus type. (C) d' : Ability to detect the mistuned harmonic as a function of stimulus type.

expected, the likelihood of reporting the perception of two concurrent sound objects increased as the Stimulus contained greater mistuning in the second harmonic [$F(5, 265) = 282.75, p < .001$; linear trend $F(1, 53) = 837.55, p < .001$]. The main effect of Musical training was significant [$F(1, 53) = 5.86, p < .05$]. Moreover, the interaction between Stimulus type and Musical training was also significant [$F(5, 265) = 4.21, p < .01$; linear trend $F(1, 53) = 7.98, p < .01$]. Follow-up pairwise comparisons indicated that musicians were more likely to report hearing two sounds when the harmonic was mistuned by 4%, 8%, and 16% [$t(55) = 2.05, 2.56, \text{ and } 3.65, \text{ respectively}, p < .05$ in all cases]. In addition, there was a trend for musicians to report hearing two sounds more often than nonmusicians when the harmonic was mistuned by 2% [$t(55) = 1.87, p = .066$]. The main effect of Age group was not significant ($p = .17$), whereas the interaction between Age group and Stimulus type was marginally significant ($p = .06$).

Although the influence of musical training appears to be smaller in older adults compared with younger adults, the interaction between Age group and Musical training was not significant ($p = .14$) nor was the three-way interaction between Age group, Musical training, and Stimulus type ($p = .67$).

RT

Figure 2B shows the group mean RTs. There was a main effect of Stimulus type on RT [$F(5, 265) = 38.52, p < .01$; quadratic trend $F(1, 53) = 94.1, p < .001$], where participants had the longest RTs to the 2% and 4% mistuned stimuli compared with the tuned 1%, 8%, and 16% stimuli ($p < .001$ in both cases). Moreover, musicians responded more quickly than nonmusicians [$F(1, 53) = 8.25, p < .01$]. The main effect of Age group was not significant ($p = .95$) nor was the interaction between Age group, Musical training, and Stimulus type ($p = .49$).

Signal Detection (d')

Figure 2C shows d' for each group. There was a main effect of Stimulus type on d' [$F(4, 212) = 231.61, p < .001$; linear trend $F(1, 53) = 481.00, p < .001$]. The main effect of Musical training was significant [$F(1, 53) = 15.56, p < .001$]. Moreover, the interaction between Musical training and Stimulus type was also significant [$F(4, 212) = 4.76, p < .01$; linear trend $F(1, 53) = 7.43, p < .01$]. Follow-up pairwise comparisons revealed a higher d' for musicians compared with nonmusicians in the 2%, 4%, 8%, and 16% stimulus conditions ($p > .01$ combined). The main effect of Age group was not significant ($p = .09$). The interaction between Age group and Stimulus type was not

significant ($p = .11$) nor was the three-way interaction between Age group, Musical training, and Stimulus type ($p = .75$).

Electrophysiological Data

Figure 3 shows the group mean ERPs elicited by the tuned and the 16% mistuned stimuli for young and older musicians and nonmusicians. A clear ORN can be seen overlapping the N1-P2 complex during active (Figure 3A) and passive (Figure 3B) listening, whereas the P400 was present only during active listening. The ORN and P400 are labeled on the plot for Younger Musicians. The scalp topographies for these responses are illustrated in Figure 1 before being rereferenced to the linked mastoid, separately at three angles (top, left, and right) for each wave and each group of participants. The ORN had a frontocentral distribution, whereas the P400 was lateralized slightly to the right central scalp region. The inversion of the ORN and P400 activity can be seen around mastoid sites on both the left and right sides.

ORN

The main effect of Listening condition on ORN latency was not significant ($p = .23$). Accordingly, group mean ORN latencies across both listening conditions were 135 msec ($SE = 4.07$) for younger musicians, 158 msec ($SE = 3.93$) for younger nonmusicians, 150 msec ($SE = 3.93$) for older musicians, and 145 msec ($SE = 4.22$) for older nonmusicians. The main effect of Musical training was significant [$F(1, 53) = 4.64, p < .05$], whereas the main effect of Age group was not ($p = .99$). However, the interaction between Age group and Musical training was significant

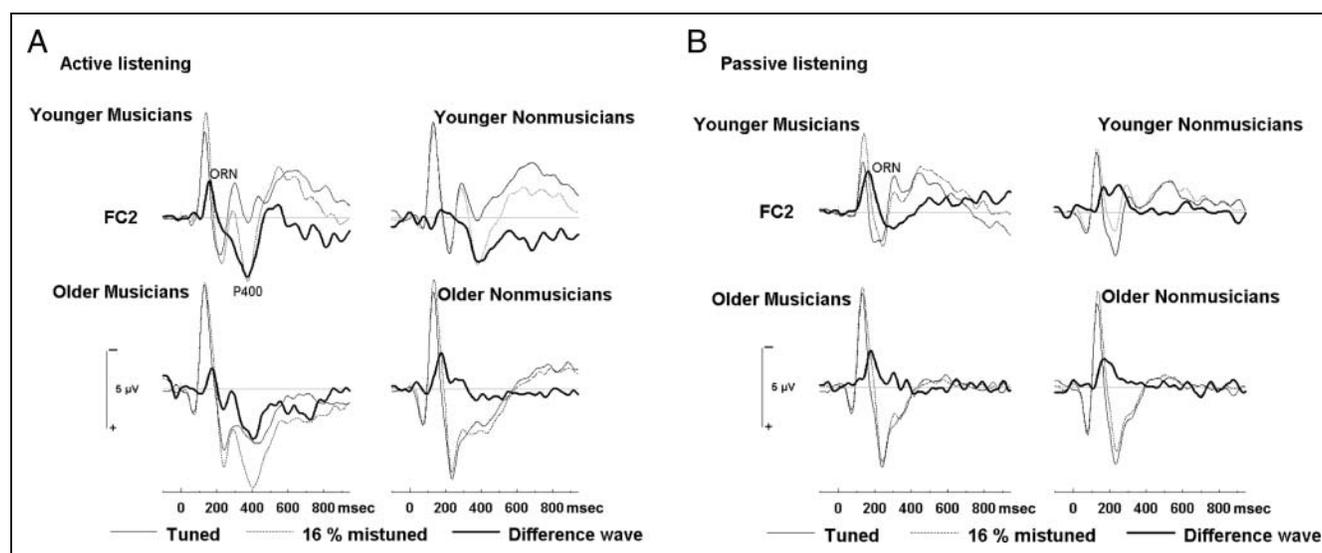


Figure 3. Auditory-evoked response to the tuned stimulus (solid line) and the 16% mistuned stimulus (dashed line) at electrode F2. The thick black line is the difference between the two responses. Responses from active listening are shown in (A) and passive listening in (B). The ORN can be seen as a negative peak ~150 msec in both active and passive listening, whereas the P400 can be seen as a positive peak around ~400 msec only in active listening.

[$F(1, 53) = 13.09, p < .01$]. Follow-up t tests were calculated to compare the influence of aging in musicians and nonmusicians. The ORN latency was shorter in younger musicians compared with older musicians [$t(27) = 2.88, p < .01$] but was also shorter in older nonmusicians compared with younger nonmusicians [$t(26) = 2.31, p < .05$]. For the ORN peak amplitude, the main effect of condition was significant, with the ORN being larger in passive listening compared with active listening [$-1.39 \mu\text{V}$ vs. $-1.11 \mu\text{V}$; $F(1, 53) = 6.78, p < .05$]. The main effects of Age group ($p = .81$) and Musical training ($p = .19$) were not significant; however, the interaction between Age group, Musical training and Listening condition was significant [$F(1, 53) = 4.75, p < .05$]. Follow-up t tests revealed a larger ORN amplitude in young nonmusicians during the passive compared with the active listening condition [$t(14) = 2.72, p < .05$]. The ORN amplitude was not different between the active and passive listening conditions in young and older musicians ($p = .21$ and $.10$) as well as in older nonmusicians ($p = .23$).

For the mean amplitude during 100–190 msec interval, the main effect of Stimulus type was significant, which was indicative of an ORN, as the amount of negativity increased from the tuned to the 16% mistuned condition [$F(5, 265) = 45.76, p < .001$; linear trend $F(1, 53) = 88.98, p < .001$; Figures 3 and 4]. The main effects of Age group and Musical training were not significant ($p = .43$ and $.52$) nor were the Stimulus type \times Age group and the Stimulus type \times Musical training interactions ($p = .50$ and $.14$). However, the interaction between Stimulus type, Age group, and Musical training was significant [$F(5, 265) = 2.89, p < .05$; linear trend; $F(1, 53) = 5.55, p < .05$]. The four-way interaction involving Listening condition, Stimulus type, Age group, and Musical training was not significant ($p = .48$). Therefore, follow-up tests for the Stimulus type \times Age group by Musical training interaction were based on the average ORN amplitude during active and passive listening. To determine the influence of Age group on the ORN, follow-up simple two-way interactions were calculated separately for musicians and nonmusicians. These analyses revealed a greater influence of Stimulus type in older nonmusicians compared with younger nonmusicians [$F(5, 130) = 3.37, p < .05$; linear trend $F(1, 26) = 7.65, p < .01$] and only a marginal age-related difference on the effect of Stimulus type for musicians ($p = .06$). In a second follow-up analysis, to determine the influence of Musical training, simple two-way interactions confirmed that the effect of Stimulus type was larger in younger musicians compared with younger nonmusicians [$F(5, 135) = 3.94, p < .01$; linear trend $F(1, 27) = 7.35, p < .05$], but that the effect of Stimulus type was similar between older musicians and nonmusicians ($p = .11$). The main effect of Listening condition was significant, as overall there was greater negativity in the active listening condition [$-1.52 \mu\text{V}$ vs. $-0.78 \mu\text{V}$; $F(1, 53) = 16.87, p < .001$]. In addition, the interaction between Listening condition and Stimulus type was significant as the increase in negativity as a func-

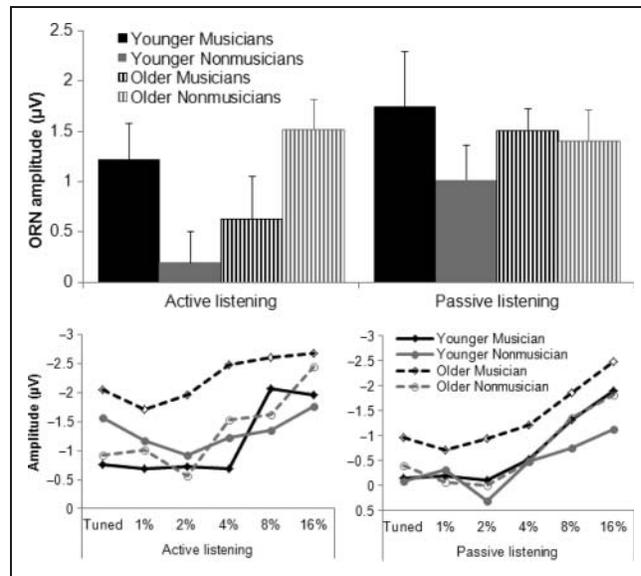


Figure 4. Top: The ORN amplitude, shown during active and passive listening at frontocentral electrodes. ORN amplitude was calculated as the difference in mean amplitude between the tuned and 16% mistuned conditions during the 100–190 msec epoch. Error bars represent 1 SE. To calculate standard error for this graph, the difference in mean amplitude across the nine frontocentral electrodes was calculated between the tuned and 16% mistuned condition. The standard deviation of the difference was then calculated for each group. The standard deviation was then divided by the square root of N , separately for each group, yielding a standard error. Bottom left: Mean amplitude during the 100–190 msec epoch as a function of Stimulus type during active listening. Bottom right: Mean amplitude during the 100–190 msec epoch as a function of Stimulus type during passive listening. ORN amplitude was multiplied by (-1) so the ORN amplitude was positive on all three graphs.

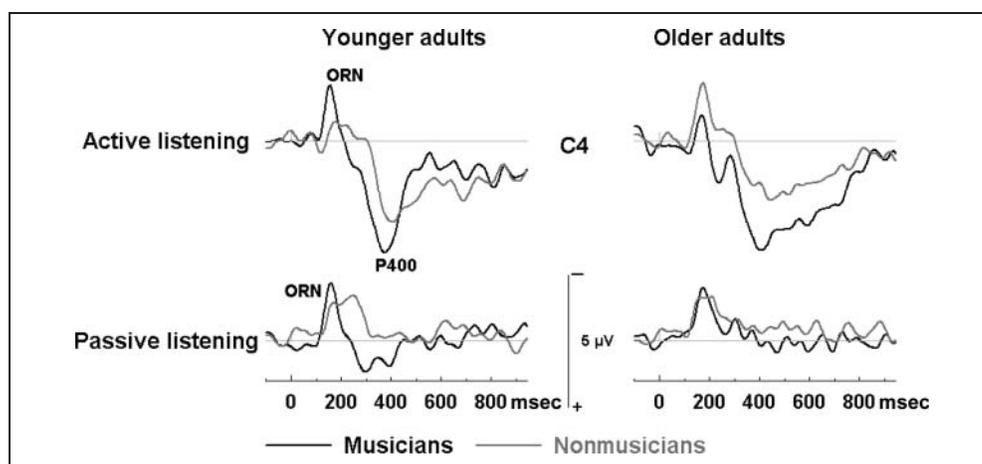
tion of mistuning was different during active and passive listening [$F(5, 265) = 2.45, p < .05$ linear trend; $F(1, 53) = 6.17, p < .05$], but as mentioned above, the Listening condition by Stimulus type interaction was not influenced by group factors.

Topography of the ORN was marginally different between musicians and nonmusicians in active [$F(64, 3392) = 2.00, p = .07$] and passive listening [$F(64, 3392) = 2.29, p = .04$]. ORN topography was similar in older and younger adults during both active [$F(64, 3392) = 1.68, p < .13$] and passive listening [$F(64, 3392) = 0.72, p = .9$]. The three-way Electrode \times Musical training \times Age group interaction was not significant during both active [$F(64, 3392) = 1.09, p = .37$] and passive listening [$F(64, 3392) = 0.34, p = .89$].

P400

The P400 peaked around 395 msec in all four groups of participants (Figure 5). The main effects of Age group, Musical training, and their interactions were not significant ($p = .99, .16, \text{ and } .88$, respectively). The P400 peak amplitude was larger in musicians compared with nonmusicians [$F(1, 53) = 4.10, p < .05$]. The main effect of

Figure 5. The difference wave (tuned – 16% mistuned) shown at electrode C4. The ORN can be seen in all conditions as a negative wave that peaks around 500 msec. The P400 can be seen only in the active listening condition as positive wave that peaks around 400 msec.



Age group and the interaction between Age group and Musical training were not significant ($p = .95$ and $.64$, respectively).

During the 250–350 msec epoch, the main effect of Stimulus type was significant, which was indicative of the P400 [$F(5, 265) = 19.82, p < .001$; linear trend $F(1, 53) = 31.67, p < .001$; Figures 5 and 6]. The interaction between Musical training and Stimulus type was also significant, indicating a larger influence of Stimulus type in musicians compared with nonmusicians [$F(1, 53) = 5.48, p < .01$; linear trend $F(1, 53) = 6.66, p < .05$]. The Age group by Stimulus type and the Age group \times Musical training \times Stimulus type interactions were not significant, indicating that the influence of Stimulus type was not influenced by age ($p = .10$ and $.60$); however, the main effect of Age group was significant [$F(1, 53) = 10.21, p < .01$], as the mean amplitude during this epoch was larger in older adults.

During the 350–450 msec epoch over the right frontal sites, the main effect of Stimulus type was significant, which was indicative of a P400 [$F(5, 265) = 35.84, p < .001$; linear trend $F(1, 53) = 57.24, p < .001$; Figures 5 and 6]. The effect of Stimulus type was larger for musicians compared with nonmusicians but was only significant at a trend level [$F(5, 265) = 2.05, p = .07$; quadratic trend $F(1, 53) = 2.04, p > .05$]. The Stimulus type \times Age group \times Musical training interaction and the Stimulus type \times Age group interactions were not significant ($p = .33$ and $.67$, respectively); however, the main effect of Age group was significant [$F(1, 53) = 6.04, p < .05$], and the interaction between Age group and Musical training approached significance [$F(1, 53) = 3.11, p = .08$], indicating that during this epoch the mean amplitude was larger in older adults, and this difference was driven mainly by the older musicians.

Topography of the P400 was different in musicians and nonmusicians during the 250–350 msec epoch [$F(64, 3392) = 2.93, p < .01$] but not during the 350–450 msec epoch [$F(64, 3392) = 0.74, p = .94$]. P400 topography was different between older and younger adults during

both the 250–350 msec epoch [$F(64, 3392) = 7.03, p < .01$] and the 350–450 msec epoch [$F(64, 3392) = 4.60, p < .01$]. The three-way Electrode \times Musical training \times Age group interaction was not significant during both the 250–350 msec epoch [$F(64, 3392) = 0.71, p = .65$] and the 350–450 msec epoch [$F(64, 3392) = 1.62, p = .15$], indicating that the age-related changes in P400 topography were not influenced by Musical training.

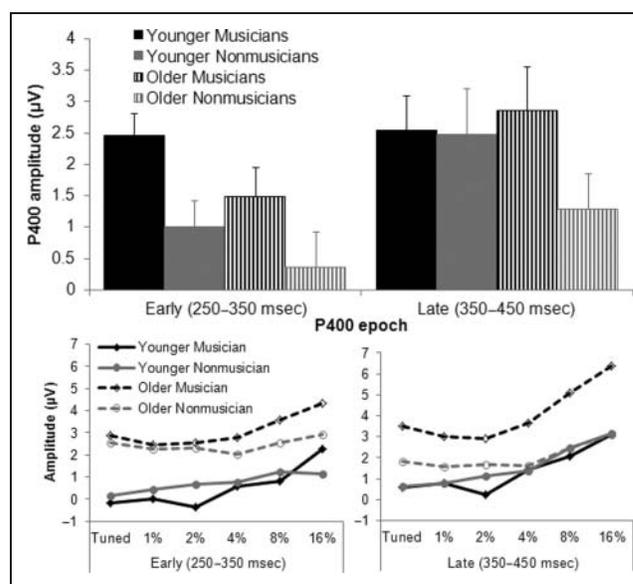


Figure 6. The P400 amplitude shown during active listening at right frontal electrodes during the 250–350 msec epoch and the 350–450 msec epoch. P400 amplitude was calculated as the difference in mean amplitude between the tuned and 16% mistuned stimulus. Error bars represent 1 SE. To calculate standard error for this graph, the difference in mean amplitude across the seven right frontal electrodes was calculated between the tuned and 16% mistuned conditions. The standard deviation of the difference was then calculated for each group. The standard deviation was then divided by the square root of N , separately for each group, yielding a standard error. Bottom left: Mean amplitude during the 250–350 msec epoch as a function of Stimulus type. Bottom right: Mean amplitude during the 350–450 msec epoch as a function of Stimulus type.

Correlations between Behavior and Electrophysiology

Within-subject correlation values are presented in Table 1, along with the within-subject correlation values for each group. That is, a within-subject correlation was calculated for each participant between the mean amplitude of the ERP component for each Stimulus type with the behavioral performance for the same stimulus type. Accordingly, this correlation represents the relationship between brain activity and performance for each participant. Importantly, d' was only available for the mistuned stimuli (performance for the tuned stimuli was compared with each level of mistuning); therefore, the mean ERP amplitude from the tuned stimulus was subtracted from the ERP amplitude for each of the mistuned stimulus to calculate brain–behavior correlations for d' . The ORN and P400 amplitude were correlated with perceptual judgment, RT, and d' . Interestingly, the ORN was most strongly correlated with d' , the early portion of the P400 was most strongly correlated with RT, and the late portion of the P400 was most strongly correlated with the response judgment. Group differences in the size of each correlation were assessed with a between-subject ANOVA. The relationship between the P400 and RT was significantly higher in musicians compared with nonmusicians. Finally, to understand the relationship between each of the electrophysiological measures, correlations were calculated between the ORN (in active listening) and the early and late portions of the P400. For the brain–brain correlations, the values used were the difference between the 16% mistuned condition and the Tuned condition (i.e., the ORN and P400). The ORN was correlated with the early portion of the P400 [$r(57) = .29, p < .05$], and the early portion of the P400 correlated with the late portion of the P400 [$r(57) = .72, p < .01$]. The

ORN was not correlated with the later portion of the P400 ($p > .1$).

DISCUSSION

There were four main findings in this study. First, musicians were better able to detect the mistuned harmonic when the stimulus was mistuned by 2% or more. Accordingly, musicians had faster RTs and were more likely to report hearing two sounds when mistuning was above 2%. Second, there was little age-related difference in the likelihood of reporting the perception of concurrent sound objects and the effects of age on perception was comparable in musicians and nonmusicians. Third, the ORN was larger in younger musicians compared with the other three groups. Fourth, the P400 started earlier and was larger in musicians compared with nonmusicians. The next section will consider each of the results in more detail, which will be followed by a broader interpretation of the overall pattern of results in terms of how they relate to previous research.

Although one study found a reduced ORN amplitude in older adults (Alain & McDonald, 2007), a more recent study found that this age-related difference was due to the length of the stimulus (Alain, McDonald, & Van Roon, 2012). When the harmonic complex was short (i.e., 40 msec), older adults were less likely to hear the mistuned harmonic as a separate sound object and this age-related decline coincided with a reduction in ORN amplitude recorded during passive listening (Alain & McDonald, 2007). On the other hand, when the stimulus was longer (i.e., 200 msec), there was no age-related differences in the likelihood of hearing the mistuned harmonic as a separate sound nor were age-related differences observed in the ORN amplitude (Alain et al., 2012). Furthermore, it

Table 1. Brain–Behavior Correlations

ERP	Behavior	Pearson r	t Test	YM (r)	YN (r)	OM (r)	ON (r)
ORN	RJ	-.38	-6.23**	-.43	-.14	-.44	-.51
	RT	.23	4.22**	.43	.15	.14	.22
	DP	-.45	-7.62**	-.50	-.33	-.50	-.47
P400a	RJ	.30	4.28**	.46	.23	.37	.11
	RT	-.35*	-6.37**	-.54	-.17	-.49	-.19
	DP	.33*	4.95**	.45	.21	.48	.18
P400b	RJ	.48	8.05**	.55	.41	.49	.49
	RT	-.44*	-7.91**	-.65	-.33	-.55	-.23
	DP	.55	9.78**	.61	.44	.60	.54

Within-subject correlations between an ERP component (ORN, P400a [250–350 msec], P400b [350–450 msec]) and a behavioral measure (response judgment [RJ], RT, and d' [DP]). Significance of the relationship was assessed by a one-sample t test. Within-subject correlations are also displayed separately for each group: younger musicians (YM), younger nonmusicians (YN), older musicians (OM), and older nonmusicians (ON).

*Musicians > Nonmusicians ($p > .05$).

** $p < .0001$.

has been shown that age-related decline in detection of a mistuned harmonic is smaller when the stimulus length is longer (Alain, McDonald, et al., 2001). These findings suggest that older adults may require a longer time frame to resolve and segregate a mistuned component as a separate auditory object, which would impede concurrent sound segregation when stimuli are short (Alain et al., 2012). In the current study, the length of the stimulus was 150 msec—shorter than 200 msec (Alain et al., 2012), but longer than 40 msec used in Alain and McDonald's (2007) study. The lack of age-related difference in ORN amplitude during passive listening is consistent with Alain et al. (2012) and suggests that 150-msec sound duration is sufficient for older adults to process the mistuned harmonic.

In older adults, the effect of musical training on the ORN amplitude was not significant, whereas the ORN was enhanced in younger musicians. This finding suggests that in older adults concurrent sound segregation, as indexed by the ORN, is little affected by musical training. One possible reason for the similarity of the ORN in older musicians and nonmusicians is that older adults may increasingly rely on more cognitive and attention-dependent processes to make acoustic judgments (Snyder & Alain, 2005; Alain et al., 2004). This is especially possible, considering that physical changes in the cochlea (Gates & Mills, 2005) and functional changes in subcortical auditory structures (Clinard, Tremblay, & Krishnan, 2010; Poth, Boettcher, Mills, & Dubno, 2001) make the encoding of incoming sensory information more variable for older adults. In the current study, it is likely that the effects of age on the cochlea were similar in older musicians and nonmusicians, as there were no differences in their pure-tone thresholds. Although we did not investigate subcortical responses, it is likely that the ORN is related to processing of the mistuned harmonic in subcortical structures (Sinex, 2008; Sinex, Guzik, Li, & Sabes, 2003; Sinex, Sabes, & Li, 2002). Thus, although the ORN is generated along the superior temporal plane, ORN generation likely depends on earlier processing of the mistuned harmonic in subcortical structures. The studies that have compared subcortical responses in musicians and nonmusicians have found numerous enhancements in younger musicians compared with younger nonmusicians, including pitch tracking (e.g., Wong, Skoe, Russo, Dees, & Kraus, 2007), faster onset responses, stronger stimulus–response correlations, and more robust tracking of upper harmonics in noise (Parbery-Clark, Skoe, & Kraus, 2009). At the same time, enhancements to subcortical responses in older musicians, compared with older nonmusicians, were limited to tracking a speech formant transition (Parbery-Clark et al., 2012). On the basis of these studies, it seems as if the benefit of musical training at the subcortical level may be reduced in older adults. This reduced influence of musicianship in subcortical, and hence automatic, stages of auditory processing may explain why the ORN is similar in older musicians and older nonmusicians. The age-related difference on the impact of musical train-

ing may be related to an age-related decline in inhibition of noise in the auditory system (Casparly, Ling, Turner, & Hughes, 2008). Weakened inhibitory function in subcortical structures would reduce the ability of higher cortical structures (i.e., auditory cortex, frontal lobes) to fine-tune subcortical structures via the efferent cortico-fugal pathway, which is a proposed mechanism for the influence of musical training on subcortical structures (Parbery-Clark et al., 2009; Wong et al., 2007). Accordingly, the influence of musical training on early automatic processing of acoustic information would be reduced in older adults.

A second possible explanation is that the age-related increase in ORN for nonmusicians could be partly accounted for by the superimposition of another wave such as the MMN. The MMN is an electrophysiological response to an oddball sound in a stream of otherwise similar stimuli and is observed during a similar epoch and at a similar scalp location to the ORN (Näätänen, Pakarinen, Rinne, & Takegata, 2004). The MMN may have been selectively evoked in the older nonmusicians because as a group they may have only automatically detected the 8% and 16% mistuned harmonic (thereby making these stimuli more salient, i.e., deviant), whereas the other groups were more likely to automatically detect the 4%, 8%, and 16% mistuned harmonics. Support for this proposal comes from a previous study of ours that demonstrated the thresholds for detecting a mistuned harmonic were below 4% for younger adults and older musicians, but above 4% for older nonmusicians (Zendel & Alain, 2012). Therefore, because the threshold to automatically detect a mistuned harmonic was higher in nonmusicians, they may not have distinguished the 8% and 16% mistuned stimuli and the tuned 1%, 2%, and 4% stimuli, thus perceiving them both categorically (i.e., concurrent sound and single sound, respectively). Accordingly, the 8% and 16% stimuli would be perceived as “oddballs” selectively in the older nonmusicians. In addition, the ORN tended to be larger in nonmusicians during active listening. This is important because the ORN and MMN are functionally separable (Bendixen et al., 2010), and increased attention levels are related to increased MMN amplitude but not ORN amplitude (Alain & Izenberg, 2003). Accordingly, increased attention may have selectively enhanced this hypothetical MMN response in the older nonmusicians.

Finally, small differences in ORN topography were observed between musicians and nonmusicians during both active and passive listening. This suggests that musicians may automatically engage additional brain areas to process a mistuned harmonic and is consistent with previous reports of topographic differences in the N1 response between musicians and nonmusicians when processing speech material (Ott, Langer, Oechslin, Meyer, & Jäncke, 2011). Given that the ORN and N1 overlap in time, the difference in ORN topography may be related to differences in N1 topography. Critically, no shift in ORN topography was observed in older adults, which suggests that the differences in ORN topography

between musicians and nonmusicians are little influenced by age.

This study extends prior research on aging and concurrent sound segregation based on inharmonicity by measuring neuroelectric brain activity during active listening. As expected, we found a P400 response whose amplitude was related to the perceptual judgment in all participants. That is, a larger P400 was related to an increased likelihood of reporting the perception of concurrent sounds. This response differs from other late positivities associated with response generation or target detection as it reflects perceptual differences based on stimulus characteristics. Furthermore, the P400 was lateralized to the right hemisphere, consistent with the hypothesis that the right auditory cortex is more specialized for processing spectral relationships (Warrier et al., 2009; Zatorre, 1988). In the current context, the P400 could index the transfer of concurrent sound stimuli to conscious awareness.

In this study, the P400 latency and amplitude were little affected by age. The lack of age-related difference in P400 mimics that of behavioral data and suggests that the perceptual decision regarding the presence of a secondary auditory object is little affected by age. At the same time, the topography of the P400 was different in older adults compared with younger adults, suggesting that older adults may recruit additional brain areas when parsing concurrent sounds. This finding is consistent with the idea that older adults will recruit additional brain areas to complete a variety of cognitive tasks (e.g., Reuter-Lorenz, 2002).

The P400 was larger in both older and younger musicians. Moreover, P400 activity started much earlier in musicians compared with nonmusicians, despite the findings that the peak of the P400 occurred at the a similar time in all participants. In musicians, the early stage of the P400 was more strongly correlated with both RT and signal detection, whereas these behavioral measures were more strongly related to the later portion of the P400 in nonmusicians. The effect of musical training on the P400 and behavioral measures is consistent with the hypothesis that musicians are better able to actively detect and quickly organize the auditory scene into separate auditory objects. Previous work that examined how musicians spatially organize the auditory environment demonstrated that musicians and conductors, specifically, were more sensitive to peripheral components of the auditory scene, and this was reflected in late positive electrophysiological activity (Nager, Kohlmetz, Altenmuller, Rodriguez-Fornells, & Munte, 2003). In both the current study and in Nager et al. (2003), participants had to assign acoustic input to at least two perceptual objects (or streams). Thus, musicians are likely better at organizing the auditory environment, and in the current study, this was related to an enhanced P400. Enhanced organizational abilities may be related to topographical differences between musicians and nonmusicians during the early portion of the P400,

suggesting possible recruitment of additional brain areas. On the other hand, later portions of the P400 were similar between musicians and nonmusicians, suggesting that this later stage of processing was similar between musicians and nonmusicians. Critically, the differences in topography for the P400 did not show Age group by Musical training interactions, demonstrating that the enhancement to the P400 in older musicians was not due to different topography.

The pattern of results for the ORN suggests a differential enhancement to the P400 in older musicians. The ORN and early portion of the P400 were correlated, suggesting that some of the variance in P400 amplitude was related to the ORN amplitude. In younger musicians, the ORN and P400 were larger than in the younger nonmusicians. This difference in P400 amplitude could be related to the difference in ORN amplitude. In older musicians, the P400 was larger than in the older nonmusicians, whereas the ORN was similar. Accordingly, the larger P400 in older musicians compared with older nonmusicians cannot be explained by a larger ORN. This suggests that there was either enhanced attention-dependent activity in the older musicians or a more robust connection between the early automatic processing of a mistuned harmonic and later conscious recognition of a second auditory object. Therefore, this overall pattern of results suggests that older musicians preferentially engage attention-dependent neural activity that may be related to using a cognitive strategy to help organize the auditory environment. The cognitive strategy could also be related to an enhanced ability to interpret an impoverished afferent auditory signal or to enhance transfer of the afferent auditory signal from early cortical processing to cognitive processing in working memory. This possibility is further supported by the observation that the P400 was similar in amplitude between the younger adults and older musicians and appeared to be reduced in older nonmusicians, although this effect did not reach statistical significance.

The current study demonstrated that lifelong musicianship can improve the ability to segregate concurrently occurring sounds. This finding is consistent with previous work that demonstrated that lifelong musicianship is associated with preservation of brain structure and some cognitive abilities (Parbery-Clark et al., 2012; Sluming et al., 2002; Meinz, 2000; Andrews et al., 1998; Krampe & Ericsson, 1996). In addition to musical training, other lifestyle choices can also influence age-related decline in cognition. For example, in a meta-analysis, Middleton and Yaffe (2009) reported that engaging in cognitive activities, such as reading, learning, or game playing, and being physically or socially active could delay or prevent the onset of dementia. Other studies have demonstrated that older adults who engage in cognitively stimulating activities show slower rates of cognitive decline, independent of early educational levels (Ghisletta, Bickel, & Lovden, 2006; Valenzuela & Sachdev, 2006). In addition, high educational and occupational achievements can create a “cognitive reserve” that can delay the onset of age-related cognitive

decline (Qiu, Backman, Winblad, Aguero-Torres, & Fratiglioni, 2001; Stern et al., 1994). It is therefore likely that early musical training contributes to a cognitive reserve and that lifelong musicianship can maintain or enhance this reserve. This enhanced cognitive reserve would likely influence attention-dependent processing of complex auditory stimuli, which was one of the major findings of the current research.

Concluding Remarks

Concurrent sound segregation is enhanced in musicians, and this ability is preserved with age. The most likely reason for this advantage is an enhanced ability to recognize a mistuned harmonic as a separate auditory object. Part of this advantage is likely related to musicians having an increased ability to detect a mistuned harmonic. Consistent with previous work, this advantage remains constant throughout the lifespan (Zendel & Alain, 2012). Furthermore, neural activity related to the conscious detection of a mistuned harmonic was enhanced in both older and younger musicians, whereas the neural activity related to the automatic detection of a mistuned harmonic was only enhanced in younger adults. Thus, the advantage for older musicians is most likely related to enhanced endogenous processing of acoustic features related to the segregation of simultaneous sound objects.

Acknowledgments

We would like to thank Yu He for technical assistance. This research was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada.

Reprint requests should be sent to Benjamin Rich Zendel, BRAMS, Suite 0-120, Pavillon 1420 boul. Mont Royal, Université de Montréal, C.P. 6128, Station Centre ville, Montreal, Quebec, Canada, H3C 3J7, or via e-mail: benjamin.rich.zendel@umontreal.ca.

REFERENCES

- Alain, C. (2007). Breaking the wave: Effects of attention and learning on concurrent sound perception. *Hearing Research, 229*, 225–236.
- Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1072–1089.
- Alain, C., & Bernstein, L. J. (2008). From sounds to meaning: The role of attention during auditory scene analysis. *Current Opinion in Otolaryngology & Head and Neck Surgery, 16*, 485–489.
- Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience, 15*, 1063–1073.
- Alain, C., & McDonald, K. L. (2007). Age-related differences in neuromagnetic brain activity underlying concurrent sound perception. *Journal of Neuroscience, 27*, 1308–1314.
- Alain, C., McDonald, K. L., Ostroff, J. M., & Schneider, B. (2001). Age-related changes in detecting a mistuned harmonic. *Journal of the Acoustical Society of America, 109*, 2211–2216.
- Alain, C., McDonald, K. L., Ostroff, J. M., & Schneider, B. A. (2004). Aging: A switch from automatic to controlled processing of sounds? *Psychology and Aging, 19*, 125–133.
- Alain, C., McDonald, K. L., & Van Roon, P. (2012). Effects of age and background noise on processing a mistuned harmonic in an otherwise periodic complex sound. *Hearing Research, 283*, 126–135.
- Alain, C., Reinke, K., He, Y., Wang, C., & Lobaugh, N. (2005). Hearing two things at once: Neurophysiological indices of speech segregation and identification. *Journal of Cognitive Neuroscience, 17*, 811–818.
- Alain, C., Schuler, B. M., & McDonald, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *Journal of the Acoustical Society of America, 111*, 990–995.
- Andrews, M. W., Dowling, W. J., Bartlett, J. C., & Halpern, A. R. (1998). Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychology and Aging, 13*, 462–471.
- Assmann, P. F., & Summerfield, Q. (1994). The contribution of waveform interactions to the perception of concurrent vowels. *Journal of Acoustical Society of America, 95*, 471–484.
- Bendixen, A., Jones, S. J., Klump, G., & Winkler, I. (2010). Probability dependence and functional separation of the object-related and mismatch negativity event-related potential components. *Neuroimage, 50*, 285–290.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sounds*. London: MIT Press.
- Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Science, 8*, 465–471.
- Caspary, D. M., Ling, L., Turner, J. G., & Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *The Journal of Experimental Biology, 211*, 1781–1791.
- Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: Evidence from human frequency-following response recordings. *Hearing Research, 264*, 48–55.
- Duquesnoy, A. J. (1983). The intelligibility of sentences in quiet and in noise in aged listeners. *Journal of Acoustical Society of America, 74*, 1136–1144.
- Gates, G. A., & Mills, J. H. (2005). Presbycusis. *Lancet, 366*, 1111–1120.
- Ghisletta, P., Bickel, J. F., & Lovden, M. (2006). Does activity engagement protect against cognitive decline in old age? Methodological and analytical considerations. *Journals of Gerontology, Series B, Psychological Sciences and Social Sciences, 61*, P253–P261.
- Grube, M., von Cramon, D. Y., & Rübnsamen, R. (2003). Inharmonicity detection: Effects of age and contralateral distractor sounds. *Experimental Brain Research, 221*, 637–642.
- Hartmann, W. M., McAdams, S., & Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *Journal of Acoustical Society of America, 88*, 1712–1724.
- Hautus, M. J., & Johnson, B. W. (2005). Object-related brain potentials associated with the perceptual segregation of a dichotically embedded pitch. *Journal of Acoustical Society of America, 117*, 275–280.
- Jansen, E. J. M., Helleman, H. W., Dreschler, W. A., & de Laat, J. (2009). Noise induced hearing loss and other hearing complaints among musicians of symphony orchestras. *International Archives of Occupational and Environmental Health, 82*, 153–164.
- Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music, 25*, 70–83.

- Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*, 1309–1313.
- Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, *125*, 331–359.
- Lipp, R., Kitterick, P., Summerfield, Q., Baily, P. J., & Paul-Jordanov, I. (2010). Concurrent sound segregation based on inharmonicity and onset asynchrony. *Neuropsychologia*, *48*, 1417–1425.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *62*, 203–208.
- McDonald, K. L., & Alain, C. (2005). Contribution of harmonicity and location to auditory object formation in free field: Evidence from event-related brain potentials. *Journal of Acoustical Society of America*, *118*, 1593–1604.
- Meinz, E. J. (2000). Experience-based attenuation of age-related differences in music cognition tasks. *Psychology and Aging*, *15*, 297–312.
- Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*, 36–47.
- Middleton, L. E., & Yaffe, K. (2009). Promising strategies for the prevention of dementia. *Archives of Neurology*, *66*, 1210–1215.
- Moore, B. C., Glasberg, B. R., & Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *Journal of Acoustical Society of America*, *80*, 479–483.
- Moore, B. C., Peters, R. W., & Glasberg, B. R. (1985). Thresholds for the detection of inharmonicity in complex tones. *Journal of Acoustical Society of America*, *77*, 1861–1867.
- Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology*, *115*, 140–144.
- Nager, W., Kohlmetz, C., Altenmuller, E., Rodriguez-Fornells, A., & Munte, T. F. (2003). The fate of sounds in conductors' brains: An ERP study. *Cognitive Brain Research*, *17*, 83–93.
- Ott, C. G. M., Langer, N., Oechslin, M. S., Meyer, M., & Jäncke, L. (2011). Processing of voiced and unvoiced acoustic stimuli in musicians. *Frontiers in Psychology*, *2*, 1–10.
- Parbery-Clark, A., Anderson, S., Hitter, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging*, *33*, 1483.e1–1483.e4.
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, *29*, 14100–14107.
- Pettigrew, C. M., Murdoch, B. E., Ponton, C. W., Kei, J., Chenery, H. J., & Alku, P. (2004). Subtitled videos and mismatch negativity (MMN) investigations of spoken word processing. *Journal of the American Academy of Audiology*, *15*, 469–485.
- Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *Journal of Acoustical Society of America*, *97*, 593–608.
- Picton, T. W., van Roon, P., Armiljo, M. L., Berg, P., Ille, N., & Scherg, M. (2000). The correction of ocular artifacts: A topographic perspective. *Clinical Neurophysiology*, *111*, 53–65.
- Poth, E. A., Boettcher, F. A., Mills, J. H., & Dubno, J. R. (2001). Auditory brainstem responses in younger and older adults for broadband noises separated by a silent gap. *Hearing Research*, *161*, 81–86.
- Qiu, C. X., Backman, L., Winblad, B., Aguero-Torres, H., & Fratiglioni, L. (2001). The influence of education on clinically diagnosed dementia incidence and mortality data from the Kungsholmen project. *Archives of Neurology*, *58*, 2034–2039.
- Rammsayer, T., & Altenmuller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception*, *24*, 37–47.
- Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*, 781–791.
- Reuter-Lorenz, P. A. (2002). New visions of the aging mind and brain. *Trends in Cognitive Sciences*, *6*, 394–400.
- Schellenberg, E. G., & Moreno, S. (2010). Music lessons, pitch processing, and g. *Psychology of Music*, *38*, 209–221.
- Sinex, D. G. (2008). Responses of cochlear nucleus neurons to harmonic and mistuned complex tones. *Hearing Research*, *238*, 39–48.
- Sinex, D. G., Guzik, H., Li, H., & Sabes, J. H. (2003). Responses of auditory nerve fibers to harmonic and mistuned complex tones. *Hearing Research*, *182*, 130–139.
- Sinex, D. G., Sabes, J. H., & Li, H. (2002). Responses of inferior colliculus neurons to harmonic and mistuned complex tones. *Hearing Research*, *168*, 150–162.
- Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage*, *17*, 1613–1622.
- Snyder, J. S., & Alain, C. (2005). Age-related changes in neural activity associated with concurrent vowel segregation. *Cognitive Brain Research*, *24*, 492–499.
- Stern, Y., Gurland, B., Tatemichi, T. K., Tang, M. X., Wilder, D., & Mayeux, R. (1994). Influence of education and occupation on the incidence of Alzheimers-disease. *Journal of the American Medical Association*, *271*, 1004–1010.
- Valenzuela, M. J., & Sachdev, P. (2006). Brain reserve and cognitive decline: A nonparametric systematic review. *Psychological Medicine*, *36*, 1065–1073.
- Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., et al. (2009). Relating structure to function: Heschl's gyrus and acoustic processing. *Journal of Neuroscience*, *29*, 61–69.
- Weise, A., Schröger, E., & Bendixen, A. (2012). The processing of concurrent sounds based inharmonicity and asynchronous onsets: An object-related negativity (ORN) study. *Brain Research*, *1439*, 73–81.
- Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*, 420–422.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function. *Journal of Acoustical Society of America*, *84*, 566–572.
- Zendel, B. R., & Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience*, *21*, 1488–1498.
- Zendel, B. R., & Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychology and Aging*, *27*, 410–417.