Psychomusicology: Music, Mind, and Brain
2012, Vol. 22, No. 2, 97–104

0275-3987/12/$12.00 DOI: 10.1037/a0030858

The High-Voice Superiority Effect in Polyphonic Music Is Influenced by Experience: A Comparison of Musicians Who Play Soprano-Range Compared With Bass-Range Instruments

Céline Marie
McMaster University

Takako Fujioka
McMaster University and Rotman Research Institute

Leland Herrington
McMaster University

Laurel J. Trainor
McMaster University and Rotman Research Institute

Western polyphonic music is typically composed of multiple simultaneous melodic lines of equal importance, referred to as "voices." Previous studies have shown that adult nonmusicians are able to encode each voice in separate parallel sensory memory traces during passive listening. Specifically, when presented with sequences composed of two simultaneous voices (melodies), listeners show mismatch negativity (MMN) responses to pitch changes in each voice, although only 50% of trials are unchanged. Interestingly, MMN is larger for the change in the higher compared to lower voice in both musicians and nonmusicians. This high-voice superiority effect has also been found in nonmusician adults and 7-month-old infants presented with two simultaneous tones, suggesting that a more robust memory trace for the higher-pitched voice might be an innate or early-acquired characteristic of human auditory processing. The present study tested whether musicians with experience playing a bass-range instrument (e.g., cello, double bass) would show a similar high-voice superiority effect as musicians with experience playing a soprano-range instrument (e.g., violin, flute). We found that musicians playing soprano-range instruments showed a high-voice superiority effect in line with previous studies, but musicians playing bass-range instruments showed similar MMN responses for both voices. These results suggest that with years of experience playing a lower-voiced instrument, cortical encoding of the lower of two simultaneous voices can be enhanced to some extent despite the early developing bias for better encoding of the higher voice.

Keywords: musical expertise, brain plasticity, auditory scene analysis (ASA), mismatch negativity (MMN), polyphonic music

Supplemental materials: http://dx.doi.org/10.1037/a0030858.supp

Auditory environments typically contain multiple overlapping sounds. For example, it is not unusual in a social situation to experience simultaneously several conversations, music, animal noises, and various environmental sounds. Separating these different sound sources is referred to as auditory scene analysis (Bregman, 1990). Music itself often contains multiple simultaneous tones or melodies. For example, Western polyphonic music is composed of two or more simultaneous melodic lines (often referred to as "voices"), which can carry equal importance in the music. To process such music, it is crucial for individuals to be able to separate and simultaneously analyze the individual melodies as they unfold over time. Previous research has shown that for both adults (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005; Fujioka, Trainor, & Ross, 2008) and infants (Marie & Trainor, 2012), separate memory traces are formed in auditory cortex for each of two simultaneous voices. Furthermore, these studies have shown a high-voice superiority effect in that the memory trace for the higher-pitched voice is more robust than that for the lower-pitched voice. The purpose of the present experiment is to investigate whether experience playing a higher-pitched (soprano-range) compared with lower-pitched (bass-range) musical

Céline Marie, Department of Psychology, Neuroscience and Behaviour, and McMaster Institute for Music and the Mind, McMaster University, Hamilton, Ontario, Canada; Takako Fujioka, McMaster Institute for Music and the Mind, McMaster University, and Rotman Research Institute, Baycrest, Toronto, Ontario, Canada; Leland Herrington, Department of Psychology, Neuroscience and Behaviour, McMaster University; Laurel J. Trainor, Department of Psychology, Neuroscience and Behaviour, and McMaster Institute for Music and the Mind, McMaster University, and Rotman Research Institute.

This research was supported by a grant from the Canadian Institutes of Health Research (to Laurel J. Trainor). Céline Marie was supported by a postdoctoral fellowship from the Natural Sciences and Engineering Research Council—CREATE Grant in Auditory Cognitive Neuroscience. We thank Elaine Whiskin for her help with testing and Andrea Unrau for proof-reading.

Photos and brief biographies of each author are available at http://dx.doi.org/10.1037/a0030858.supp

Correspondence concerning this article should be addressed to Laurel J. Trainor, Department of Psychology, Neuroscience and Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada. E-mail: LJT@mcmaster.ca

instrument affects the encoding of polyphonic music, in particular, whether the high-voice superiority effect is modified by intensive experience playing the lowest line in a musical ensemble.

For both musical and nonmusical sounds, to determine what auditory objects are present, the auditory system must perform a spectrotemporal analysis of the incoming sound wave to determine which components belong together (e.g., the harmonics of a single sound source such as a musical instrument or a talker, or the successive sounds of an instrument or talker) and which groups of components belong to separate objects (e.g., two different instruments or two different talkers). These processes are known as auditory stream integration and segregation, respectively, and together they constitute auditory scene analysis. Bregman (1990) proposed that much of auditory scene analysis occurs automatically and preattentively, and this is corroborated by numerous event-related potential (ERP) studies based on electroencephalograph (EEG) or magnetoencephalograph (MEG) recordings (e.g., Brattico, Winkler, Näätänen, Paavilainen, & Tervaniemi, 2002; Lee, Skoe, Kraus, 2009; Nager, Teder-Sälejärvi, Kunze, & Münte, 2003; Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995; Shinozaki et al., 2000; Sussman, Ritter, & Vaughan, 1999; Winkler, Paavilainen, Näätänen, 1992; Yabe et al., 2001). Many of these studies measure the mismatch negativity (MMN) component of the ERP. MMN is seen between approximately 150 and 250 ms after the onset of a deviant sound in a stream of standard sounds. The deviance can involve a change in any sound feature such as pitch, duration, loudness, or timbre, as well as a change in sound category (e.g., one phoneme to another) or a change in the pattern of the presented sounds, such as a reversal in order (Näätänen, Gaillard, & Mäntysalo, 1978; for reviews, see Näätänen, Paavilainen, Rinne, & Alho, 2007; Näätänen & Winkler, 1999; Picton, Alain, Otten, Ritter, & Achim, 2000). MMN is observed even when listeners are paying no attention to the sounds. Thus, MMN is thought to reflect an automatic process of updating of sensory memory traces when the brain fails to predict the next sound event as in the case of deviant sounds. MMN is only produced when deviants are relatively rare, and the size of the MMN increases as the ratio of deviants to standards decreases.

Fujioka and colleagues investigated preattentive sound processing for simultaneous musical streams or voices using MMN protocol. Measuring MEG, Fujioka et al. (2005) presented adults with trials of two simultaneous five-note melodies, and introduced deviants on the fifth note on 50% of trials, such that 25% contained a wrong note in the higher, and 25% a wrong note in the lower, melody. MMN was elicited by deviants in both voices, although the overall deviance rate was 50%, indicating that expectations and therefore separate memory traces were formed for the higher and lower melodies. At the same time, MMN was larger for deviants in the high than the low voice, reflecting a bias for better encoding of the high voice. Furthermore, Fujioka et al. (2008) found the same pattern of results with simplified stimuli, in which each standard trial (50% of trials) consisted of a dyad of two simultaneous tones rather than melodies, and 25% of trials contained a pitch change (deviant) in the higher-pitched tone and 25% contained a pitch change (deviant) in the lower-pitched tone. Again, MMN was elicited by deviants in both streams, even with an overall deviance rate of 50%, suggesting that separate memory traces were formed for each stream of tones. Moreover, the results showed a high-voice superiority effect as well, with larger MMN elicited by deviants for the higher-pitched than lower-pitched tone. Finally, similar results have recently been found in 7-month-old infants (Marie & Trainor, 2012). The results of this last study indicate that simultaneous sound processing and, more surprisingly, the high-voice superiority effect emerge early in development, suggesting a possible innate basis. However, little is known about whether this bias can be modified by experience-related neuroplastic changes. Over the past 15 years, there has been a great deal of research examining how musical training or expertise affects brain functions and, more specifically, influences auditory processing (for reviews, see Besson, Chobert, & Marie, 2011a; Kujala & Näätänen, 2010; Trainor & Corrigall, 2010). Related to the processing of sound during passive listening, the general conclusion across the literature is that musicians seem to have better processing abilities as compared with nonmusicians. For instance, musicians are able to discriminate pitch changes among unfamiliar and familiar tone patterns faster than nonmusicians, as indexed by a shorter MMN latency (Brattico, Näätänen, & Tervaniemi, 2001). They show larger MMN to changes in melodic contour and interval structure than nonmusicians (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004). When pitch deviations are introduced in polyphonic and single-voice contexts, musicians show a better encoding of these variations, reflected by larger MMN, than nonmusicians (Fujioka et al., 2005). Outside the domains of pitch processing, musical expertise influences duration processing, as well as the ability to detect changes in temporal structure and in numerical regularity (e.g., Marie, Kujala, & Besson, 2012; Rüsseler, Altenmüller, Nager, Kohlmetz, & Münte, 2001; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005; Vuust et al., 2005, Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009). It is even the case that certain aspects of speech processing, such as meter and suprasegmental and segmental pitch features, are also enhanced by musical training (for a review, see Besson, Chobert, & Marie, 2011b). These results suggest that through years of intensive training with a musical instrument, musical expertise is associated with improved sound processing.

Most related to our concern, different types of musical experience appear to affect the way musicians process musical sounds. For example, Pantev, Roberts, Schulz, Engelien, Ross (2001) demonstrated that brain responses to trumpet sounds are larger in trumpeters compared with violinists, whereas the opposite pattern is true for violin sounds. Musicians who did not use musical scores when practicing or playing an instrument showed greater behavioral discrimination and an enhanced MMN to contour changes compared with those playing with a score and with nonmusicians (Tervaniemi, Rytkonen, Schröger, Ilmoniemi, & Näätänen, 2001). In addition, Seppänen, Brattico, & Näätänen (2007) demonstrated that there can even be differences between musicians, depending on their preferred practice strategies such as reading a musical score, improvising, playing by ear, and rehearsing by listening to recordings. Those who preferred aural practicing were faster at discriminating changes in melodic interval and contour compared with those who preferred other practice strategies (see also Vuust, Brattico, Seppänen, Näätänen, Tervaniemi, 2012). Thus, different musical experiences can result in different cortical reorganizations in the brain. Here we examined how specific musical training can affect the processing of melodies in a polyphonic context.

The origin of the high-voice superiority effect is not yet well understood. The fact that it is present already in young infants suggests that it either has an innate origin or is readily learned through exposure to Western music, in which the most important melody line is typically placed in the highest voice. However, it is not clear to what extent the high-voice superiority effect is malleable by experience beyond the infancy period. If it is modifiable, we expected that musicians with experience playing cello or bass lines in an ensemble would overcome this bias and show larger and/or earlier MMNs to pitch changes in the lower than in the higher voice compared to musicians with experience playing violin or flute lines in an ensemble, who would be expected to show a typical high-voice superiority effect.

Materials and Methods

Participants

Twenty-five adult musicians were tested. Three were excluded owing to excessive artifacts in the EEG data, leaving 22 musicians in the final sample (12 male and 10 female individuals; mean age = 29.8 years). After providing informed written consent to participate, musicians completed a questionnaire for auditory screening purposes and to assess musical and linguistic background. Subjects were required to be trained musicians who were playing regularly in a musical group (e.g., orchestra, ensemble, choral) with more than 5 years of training on their main instrument. Eleven subjects (mean age = 30 years, *SD* = 12) played instruments in the higher voice register (violin, flute, soprano singer; mean training = 21 years, *SD* = 8; see Table 1 for details). Eleven subjects (mean age = 28 years, *SD* = 11) played instruments in the lower voice register (bass, cello, or bass-vocal-range singer; mean training = 15 years, *SD* = 10; see Table 1).

Stimuli

The two five-note melodies (A, B) from Fujioka et al. (2005) were used. They were composed using the first five diatonic scale notes of the Western major scale (e.g., C, D, E, F, and G). In the key of C major, Melody A was the sequence "C-D-F-E-G" and Melody B was "G-F-D-C-E." Melodies A and B were combined in two versions (see Figure 1), one with Melody A in the high voice (C5–G5) and Melody B in the low voice (C4–G4; American notation, High-A/Low-B) and one with the voices reversed (High-B/Low-A). Deviant versions were created by either raising the last note of Melody A by one tone (one-sixth octave) or lowering the last note of Melody B by one tone. This created four deviant versions: High-A, Low-A, High-B, and Low-B. These changes did not alter the pitch contours of the melodies. From one trial to the next, the combined melodies were transposed to one of eight keys sequenced in the order "C-E-C#-F-D-F#-D#-G" to avoid potential priming effects that could arise from absolute and not relative pitch processing. For both melody-voice combinations (High-A/Low-B and High-B/Low-A), the sequences of stimuli were presented in an oddball paradigm. Sixty-four percent of the trials were presented with the standard last note (see Figure 1). Thirty-six percent of the sequences were presented with either the deviant last note in the high voice (18%) or the low voice (18%). Although Melodies A

Table 1
Musical Background of Each Participant

Participants	First instrument	Practice first (yr)	Music theory (yr)	Second instrument	Practice second (yr)
Bass-range players					
1 F	Cello	36	7	Viola da Samba	8
2 M	Bass	6	15	Piano	17
3 M	Bass	16	5	Drums	20
4 M	Bass vocal	34	3	Saxophone	2
5 M	Bass	5	1	Guitar	9
6 M	Upright bass	8	6	Electric bass	11
7 F	Cello	15	4	Piano	10
8 M	Cello	15	1	Piano	21
9 M	Bass vocal	10	4	none	—
10 M	Cello	8	5	Piano	6
11 F	Cello	15	10	Piano	16
Mean		15.3	5.6		12
SD		10.5	4.1		6.3
Soprano-range players					
1 F	Soprano vocal	30	0	Cello	5
2 M	Flute	35	5	none	—
3 F	Flute	15	6	Piano	20
4 M	Violin	29	12	Piano	6
5 F	Flute	27	14	Piano	6
6 M	Violin	27	8	Piano	20
7 F	Violin	16	5	Piano	15
8 F	Violin	10	3	Piano	10
9 F	Flute	11	10	Piano	4
10 F	Violin	15	10	Piano	15
11 M	Violin	14	0	Viola	3
Mean		20.8	6.6		10.4
SD		8.8	4.6		6.6

Note. M = male; F = female.

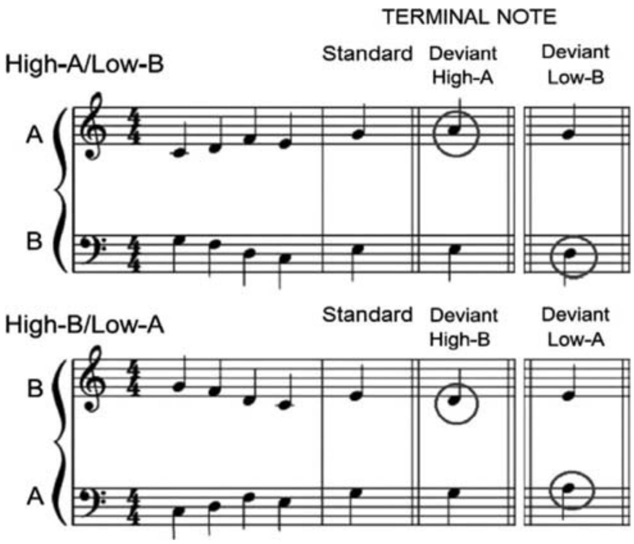

Figure 1. Description of the stimulus sequences illustrated in musical notation. Two different melodies (A and B) are played simultaneously in high and low voice corresponding to two lines in musical notation. Both melodies consist of five notes with a standard or a deviant terminal note. In the High-A/Low-B case, the high voice is Melody A and the low voice is Melody B (top), whereas in the High-B/Low-A case, the melody–voice combination was reversed (bottom).

and B used the same group of five notes for standards, they had different notes at each time point in the melodies. The harmony at each time point was always musically consonant, even in the deviant versions. The sound files were created from digitally recorded piano timbres for each note at a sampling rate of 44,100 Hz. The duration of each note was 300 ms for a total melody length of 1,500 ms. Successive trials were separated by a 900-ms silent interval.

Procedure

All participants were tested individually. The procedure was explained to each participant, who gave consent to participate. The procedures were cleared by the McMaster Research Ethics Board. Each participant sat in the sound-attenuated room (Industrial Acoustics Company) facing a loudspeaker and a screen placed 1 m in front of their head. During the experiment, the participant watched a silent movie with subtitles and was instructed to pay attention to the movie and not to the sounds that were coming from the loudspeaker. They were also asked to minimize their movement, including blinking and facial movements, so as to obtain the best signal-to-noise ratio. Four blocks of 384 trials each were presented using E-prime software in pseudorandom order such that trials of the same deviant never followed successively. Each block lasted approximately 12 min. Two blocks contained Melody A in the higher voice and Melody B in the lower voice, and two blocks contained the reverse combination.

EEG Recording and Processing

EEG data were recorded at a sampling rate of 1,000 Hz from 128-channel HydroCel GSN nets (Electrical Geodesics, Eugene, OR) referenced to Cz. The impedance of all electrodes was < 50 kΩ during the recording. EEG data were band-pass filtered between 0.5 and 20 Hz (roll-off $=$ 12 dB/oct) using EEprobe software. Recordings were re-referenced off-line using an average reference and then segmented into 500-ms epochs (-100 to 400 ms relative to the onset of the last note of the melody). EEG responses exceeding ± 70 μV in any epoch were considered artifacts and excluded from the averaging.

ERP Data Analysis

For each participant for each condition (High-A/Low-B, High-B/Low-A), responses to standards and each of the two deviants were averaged separately, and difference waveforms were computed for each condition and participant by subtracting ERPs elicited by the standards from those elicited by each of the deviants. Thus, four different difference waves were created, one for each of the deviants High-A, Low-A, High-B and Low-B in both the bass-range and soprano-range musicians. Subsequently, for statistical analysis, 70 electrodes were selected and divided into four groups for each hemisphere (left and right) representing frontal, central, occipital, and temporal regions (FL, FR, CL, CR, OR, OL, TL, TR; see Figure 2). Grouping the electrodes in this way enhances signal-to-noise ratios and enables examination of the average response across classic scalp regions. Fifty-eight electrodes were excluded from the groupings owing to the following considerations: electrodes on the forehead near the eyes to further reduce the contamination of eye movement artifacts, electrodes at the edge of the geodesic net to reduce contamination of face and neck muscle movement, electrodes in the midline to enable com-

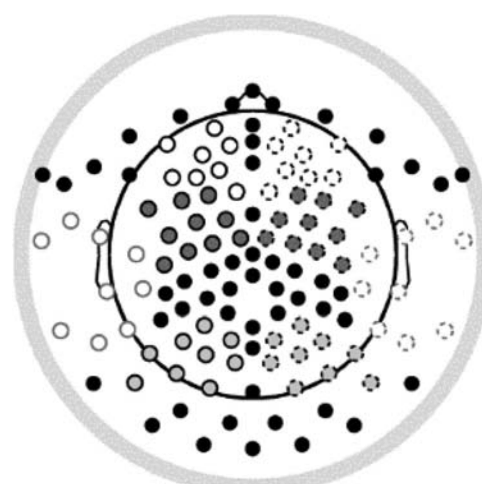

Figure 2. The grouping of electrodes in the geodesic net. Of 128 electrodes, 70 were selected to be divided into four groups (frontal, central, occipital, and temporal) for each hemisphere. The waveforms for all channels in each region were averaged together to represent EEG responses from that scalp region. Open circle with black solid line, frontal left; open circle with black dotted line, frontal right; dark gray filled circle with black solid line, central left; dark gray filled circle with black dotted line, central right; light gray filled circle with black solid line, occipital left; light gray filled circle with black dotted line, occipital right; open circle with gray solid line, temporal left; and open circle with gray dotted line, temporal right.

parison of the EEG response across hemispheres; and parietal electrodes where MMN amplitude is close to 0 μV.

To analyze MMN amplitude, first the most negative peak in the right frontal region (FR) between 150 and 250 ms after stimulus onset was determined from the grand average difference waves for each of the four conditions (High-A, Low-A, High-B, Low-B), and a 50-ms time window was constructed centered at this latency. For each subject and each region, the average amplitude in this 50-ms time window for each condition was used as the measure of MMN amplitude. Finally, the latency of the MMN was measured as the time of the most negative peak between 150 and 250 ms at the FR region for each subject, for each condition, as visual inspection showed the largest MMN amplitude at this region. Analyses of variance (ANOVAs) were conducted on amplitude and latency data. Greenhouse–Geisser corrections were applied where appropriate, and Tukey post hoc tests were conducted to determine the source of significant interactions.

Results

MMN Amplitude

Amplitude. A five-way ANOVA was conducted with Group (soprano, bass) as a between-subjects factor; Voice (high, low), Melody (A, B), Hemisphere (left, right), and Region (frontal: FL and FR; central: CL and CR; temporal: TL and TR; occipital: OL and OR) as within-subject factors; and MMN amplitude as the dependent measure. The main effects of Group ($F(1, 20) = 0.33$, $p = .57$) and Voice ($F(1, 20) = 1.94$, $p = .17$) were not significant. However, there was a significant interaction between Group and Voice ($F(1, 20) = 4.79$, $p = .04$). Post hoc tests showed a slight trend for a larger MMN difference between soprano-range and bass-range players on the Low voice ($p = .23$) but no trend for a group difference for the High voice ($p = .82$, see Table 2 for details). As can be seen in Table 2, differences appear to be largest at FR and CR regions. Moreover, an interaction between Group by Voice by Melody by Region was significant ($F(3, 60) = 4.52$, $p = .03$). No other interactions were significant.

Table 2
Detailed Mean MMN Amplitude (Value in μV) for Each Group, for Each Voice and Each Region by Hemisphere

	Soprano-range players		Bass-range players	
Regions	High voice	Low voice	High voice	Low voice
Frontal left	−0.81	−0.37	−0.34	−0.39
Frontal right	−0.33	−0.18	−0.55	−0.44
Central left	−0.58	−0.30	−0.27	−0.33
Central right	−0.18	−0.14	−0.54	−0.35
Occipital left	0.23	0.13	0.41	0.28
Occipital right	0.43	0.24	0.25	0.13
Temporal left	0.11	0.10	0.47	0.45
Temporal right	0.56	0.36	0.17	0.18

Note. An inversion of MMN polarity in the occipital and temporal regions.

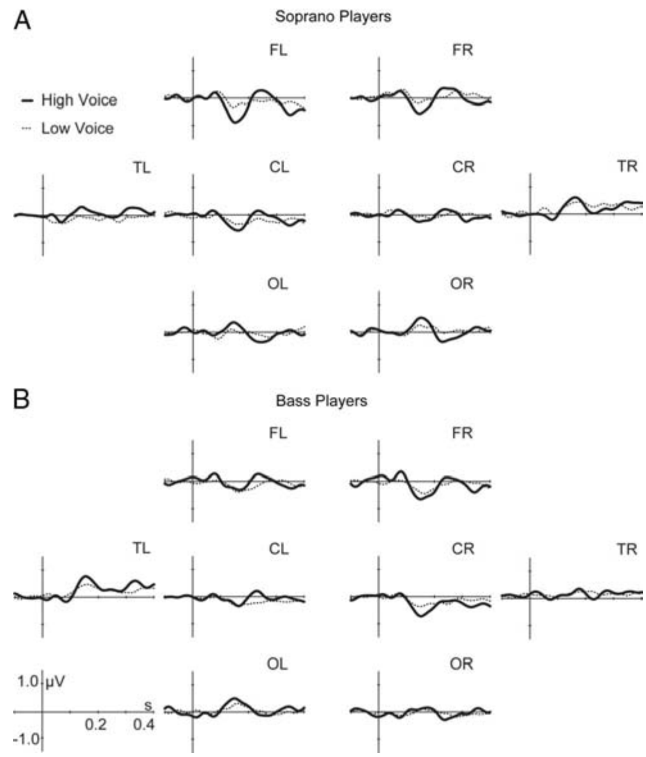

Figure 3. Difference waveforms (deviant−standard) in (A), the group of soprano-range players ($n = 11$) and (B), and the group of bass-range players ($n = 11$) for the High voice (black line) and the Low voice (dashed line). Deviants in Melody A and B were combined for each voice.

To better understand these interactions, four-way ANOVAs were conducted separately for each group, with Voice, Melody, Hemisphere, and Region as within-subject factors. For soprano-range players, the main effect of voice was significant ($F(1, 10) = 6.08$, $p = .03$, partial eta squared [η_p^2] = .38), revealing larger MMN responses for the higher than for the lower voice (see Table 2 and Figure 3). No other main effects or interactions were significant. For the bass-range players, the main effect of voice was not significant ($F(1, 10) = 0.33$, $p = .57$, $\eta_p^2 = .03$; see Table 2 and Figure 3), revealing that MMN responses were not larger for the higher than for the lower voice. No other main effects or interactions were significant (see Figure 4).

MMN Latency

A three-way ANOVA with Group as a between-subjects factor, Voice and Melody as within-subject factors, and latency at region FR as the dependent measure revealed a main effect of melody ($F(1, 20) = 19.46$, $p < .001$], with significantly shorter MMN latency to deviants in Melody A (148 ms) than to deviants in Melody B (163 ms, see Figure 5). No other main effects or interactions were significant.

Discussion

In Western polyphonic music, the melody is mostly commonly placed in the highest voice, but it is not known to what extent this

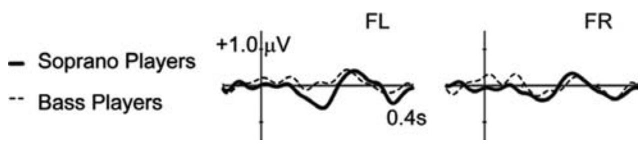

Figure 4. High-voice superiority effect in both groups. Difference between the High-voice difference waves minus the Low-voice difference waves.

is a cultural convention or a result of innate biological constraints on perception. Previous research using EEG and MEG indicates that a high-voice superiority effect is found both in adults (Fujioka et al., 2005, 2008) and in infants (Marie & Trainor, 2012). The infant findings suggest that innate constraints might play a role. Here we investigated whether musical experience might also play a role. Specifically, we tested whether the experience of playing a soprano-range compared to a bass-range instrument modifies the high-voice superiority effect. The results indicated that this effect is indeed malleable by experience beyond the infancy period. As expected, soprano-range players showed a high-voice superiority effect as reflected by larger MMN amplitude for deviants in the high voice than in the low voice. Of most interest, bass-range players did not show a high- or low-voice superiority effect. Post hoc analyses indicated that there was a slight trend for bass-range musicians to show higher-amplitude MMN for changes in the low voice than soprano-range musicians, but no trend for a difference between groups for the high voice.

Across both groups of musicians, we found that MMN latency was shorter in response to deviants in Melody A than in response to deviants in Melody B, regardless of the voice assignment. This effect is not related to the high-voice superiority effect, as Melody A was in the high voice on half of the conditions, whereas Melody B was in the high voice on the other half. The most likely explanation is that in the case of Melody A, the deviant note did not occur in the previous melodic context, whereas in the case of Melody B, the deviant note was identical to the third note of the melody (see Figure 1). Thus, the deviant note in Melody B may have been less novel than the deviant note in Melody A, causing a slight delay in processing the change. Another interpretation of the different MMN latencies for the two melodies is that they resulted from the fact that the deviant last note in Melody A was higher in pitch than the last note of the standard melody, causing an increase in the size of the final interval (minor third to perfect fourth), whereas in Melody B, the deviant last note was lower in pitch, causing a decrease in the size of the final interval (major third to major fourth). However, note that the size of the deviance was identical in both cases (a major second), so if this explanation is correct, it implies that increases in pitch are easier to detect than decreases. A third possibility is related to tonal structure. Melody A might be considered to be more stable than Melody B, as it begins on the tonic (first and most stable note of the scale) and ends on the dominant (fifth and second most stable note). However, previous work indicates that MMN is little affected by tonal structure; rather, it is primarily affected by the physical size of the change rather than the meaning in terms of musical scale and key (e.g., Näätänen, Pakarinen, Rinne, & Takegata, 2004; Trainor, McDonald, & Alain, 2002; Trainor & Zatorre, 2009), making this explanation unlikely. In any case, it would be interesting for future studies to explore the perceptual and developmental origins of this effect as well.

The presence of the high-voice superiority effect in young infants suggests that it either has an innate origin or is readily learned through exposure to Western polyphonic music, in which the most important melody line is typically placed in the highest voice. The present results indicate that, at the same time, the high-voice superiority effect is malleable to some extent by extensive experience playing the lowest voice in musical groups. Interestingly, despite the many years of practice with a low-voice instrument, we did not observe a complete reversal of this bias in musicians who play bass-range instruments, in which case, we would have seen a low-voice superiority effect. Rather, in these musicians, the high and low voices appeared to be encoded equally well, suggesting an interplay between an innate tendency for a high-voice superiority effect and greater experience with lower-

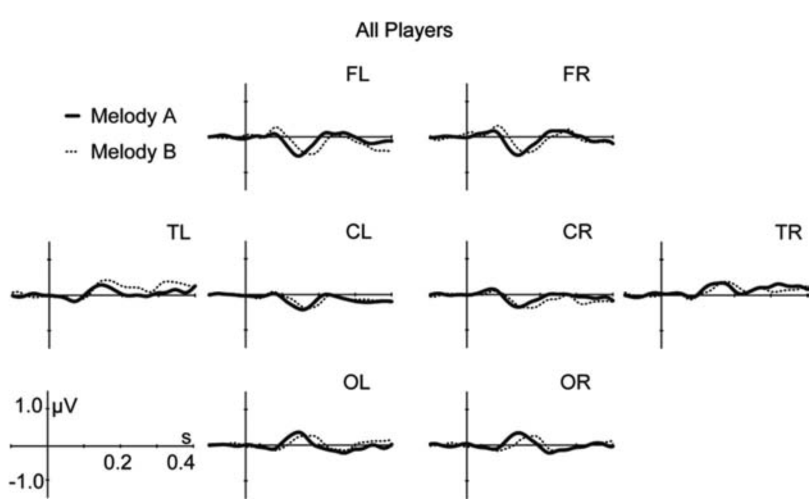

Figure 5. Average difference waveforms for Melody A (black line) and for Melody B (dashed line) with both groups combined. Shorter MMN latency for Melody A than for Melody B.

than higher-voice parts. However, it is possible that we have underestimated the effects of experience. As can be seen in Table 1, it is difficult to recruit musicians who play only one instrument, and many musicians play the piano as a second instrument. Because pianists need to play all melody lines, and have much experience playing the most important melody in the highest voice, as a group, the bass-range instrument players likely had significant experience playing high-voice melody parts. It is also possible that bass instrument players with more extensive training than those in our sample might have shown a low-voice superiority effect. Finally, melodic context clearly affects how memory traces are formed because the high-voice superiority effect appears to be smaller with our polyphonic melody context than in the context of a repeating dyad of complex tones (e.g., Fujioka et al., 2008; Marie & Trainor, 2012). The present results thus suggest that musical experience in playing a bass-range instrument can specifically modify the degree of high-voice dominance even though it is already present in young infants.

Conclusion

This study provides evidence that experience in the form of extensive musical training can influence the high-voice superiority effect in polyphonic music. Years of practice with a bass-range instrument appears to enhance the encoding of the lower voice and to modify the bias for better encoding of the higher melody. Whereas musicians playing a soprano-range instrument demonstrated a typical high-voice superiority effect, those playing a bass-range instrument showed equal encoding of the high and low voices. The presence of the high-voice superiority effect in young infants and the fact that extensive experience playing a bass-range instrument did not result in a low-voice superiority effect suggest an innate bias toward superior encoding of the higher voice. At the same time, modification of the high-voice superiority effect in bass-range musicians indicates that the high-voice superiority effect is modifiable to some extent by experience.

References

Besson, M., Chobert, J., & Marie, C. (2011a). Language and music in the musician brain. *Language and Linguistics Compass, 5,* 617–634. doi:10.1111/j.1749-818X.2011.00302.x

Besson, M., Chobert, J., & Marie, C. (2011b). Transfer of training between music and speech: Common processing, attention and memory. *Frontiers in Psychology, 2, 94.* doi:10.3389/fpsyg.2011.00094

Brattico, E., Näätänen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related potential recordings. *Music Perception: An Interdisciplinary Journal, 19,* 199–222. doi:10.1525/mp.2001.19.2.199

Brattico, E., Winkler, I., Näätänen, R., Paavilainen, P., & Tervaniemi, M. (2002). Simultaneous storage of two complex temporal sound patterns in auditory sensory memory. *Neuroreport, 13,* 1747–1751.

Bregman, A. S. (1990). *Auditory scene analysis; The perceptual organization of sound.* Cambridge, MA: MIT Press. doi:10.1121/1.408434

Fujioka, T., Trainor, L. J., & Ross, B. (2008). Simultaneous pitches are encoded separately in auditory cortex: An MMNm study. *NeuroReport, 19,* 361–366. doi:10.1097/WNR.0b013e3282f51d91

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience, 16,* 1010–1021. doi:10.1162/0898929041502706

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience, 17,* 1578–1592. doi:10.1162/089892905774597263

Kujala, T., & Näätänen, R. (2010). The adaptive brain: A neurophysiological perspective. *Progress in Neurobiology, 91,* 55–67. doi:10.1016/j.pneurobio.2010.01.006

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience, 29,* 5832–5840. doi:10.1523/JNEUROSCI.6133-08.2009

Marie, C., Kujala, T., & Besson, M. (2012). Musical and linguistic expertise influence pre-attentive and attentive processing of non-speech sounds. *Cortex, 48,* 447–457. doi:10.1016/j.cortex.2010.11.006

Marie, C., & Trainor, L. J. (2012). Development of simultaneous pitch encoding: Infants show a high voice superiority effect. *Cerebral Cortex*. Advance online publication. doi:10.1093/cercor/bhs050

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpretes. *Acta Psychological, 42,* 313–329. doi:10.1016/0001-6918(78)90006-9

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118,* 2544–2590. doi:10.1016/j.clinph.2007.04.026

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115,* 140–144. doi:10.1016/j.clinph.2003.04.001

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125,* 826–859. doi:10.1037/0033-2909.125.6.826

Nager, W., Teder-Sälejärvi, W., Kunze, S., & Münte, T. F. (2003). Preattentive evaluation of multiple perceptual streams in human audition. *NeuroReport: For Rapid Communication of Neuroscience Research, 14,* 871–874. doi:10.1097/00001756-200305060-00019

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport, 12,* 169–174.

Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology and Neuro-Otology, 5,* 111–139.

Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., & Vaughan, H. G. Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review. *Ear and Hearing, 16,* 52–67. doi:10.1097/00003446-199502000-00005

Rüsseler, J., Altenmüller, E., Nager, W., Kohlmetz, C., & Münte, T. F. (2001). Event-related brain potentials to sound omissions differ in musicians and non-musicians. *Neuroscience Letters, 308,* 33–36. doi:10.1016/S0304-3940(01)01977-2

Seppänen, M., Brattico, E., & Näätänen, R. (2007). Practice strategies of musicians modulate neural processing and the learning of sound patterns. *Neurobiology of Learning and Memory, 87,* 236–247. doi:10.1016/j.nlm.2006.08.011

Shinozaki, N., Yabe, H., Sato, Y., Sutoh, T., Hiruma, T., Nashida, T., & Kaneko, S. (2000). Mismatch negativity (MMN) reveals sound grouping in the human brain. *Neuroreport, 11,* 1597–1601. doi:10.1097/00001756-200006050-00001

Sussman, E., Ritter, W., & Vaughan, H. G. Jr. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology, 36,* 22–34. doi:10.1017/S0048577299971056

Tervaniemi, M., Rytkonen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory, 8,* 295–300. doi:10.1101/lm.39501

Trainor, L. J., & Corrigall, K. A. (2010). Music acquisition and effects of musical experience. In M. Riess-Jones & R. R. Fay (Eds.), *Springer handbook of auditory research: Music perception* (pp. 89–128). Heidelberg: Springer. doi:10.1007/978-1-4419-6114-3_4

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience, 14,* 430–442. doi:10.1162/089892902317361949

Trainor, L. J., & Zatorre, R. (2009). The neurobiological basis of musical expectations: From probabilities to emotional meaning. In S. Hallen, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 171–182). Oxford: Oxford University Press.

van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity – a mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research, 23,* 270–276. doi:10.1016/j.cogbrainres.2004.10.007

Vuust, P., Brattico, E., Seppänen, M., Näätänen, R., Tervaniemi, M. (2012). Practiced musical style shapes auditory skills. *Annals of the New York Academy of Science, 1252,* 139–146. doi:10.1111/j.1749-6632.2011.06409.x

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music. *Cortex, 45,* 80–92. doi:10.1016/j.cortex.2008.05.014

Vuust, P., Pallesen, K. J., Bailey, C., Van Zuijen, T. L., Gjedde, A., Roepstor, V. A., Østergaard, L. (2005). To musicians, the message is in the meter—Preattentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage, 24,* 560–564. doi:10.1016/j.neuroimage.2004.08.039

Winkler, I., Paavilainen, P., & Näätänen, R. (1992). Can echoic memory store two traces simultaneously? A study of event-related brain potentials. *Psychophysiology, 29,* 337–349. doi:10.1111/j.1469-8986.1992.tb01707.x

Yabe, H., Winkler, I., Czigler, I., Koyama, S., Kakigi, R., Suto, T., Hiruma, T., & Kaneko, S. (2001). Organizing sound sequences in the human brain: The interplay of auditory streaming and temporal integration. *Brain Research, 897,* 222–227. doi:10.1016/S0006-8993(01)02224-7

Received April 30, 2012
Revision received July 25, 2012
Accepted October 2, 2012 ■