Tonal Expectations Influence Early Pitch Processing

Frédéric Marmel, Fabien Perrin, and Barbara Tillmann

Abstract

■ The present study investigated the ERP correlates of the influence of tonal expectations on pitch processing. Participants performed a pitch discrimination task between penultimate and final tones of melodies. These last two tones were a repetition of the same musical note, but penultimate tones were always in tune whereas final tones were slightly out of tune in half of the trials. The pitch discrimination task allowed us to investigate the influence of tonal expectations in attentive listening and, for penultimate tones, without being confounded by decisional processes (occurring on final tones). Tonal expectations were manipulated by a tone change in the first half of the melodies that changed their tonality, hence changing the tonal expectedness of penulti-
mate and final tones without modifying them acoustically. Manipulating tonal expectations with minimal acoustic changes allowed us to focus on the cognitive expectations based on listeners' knowledge of tonal structures. For penultimate tones, tonal expectations modulated processing within the first 100 msec after onset resulting in an Nb/P1 complex that differed in amplitude between tonally related and less related conditions. For final tones, out-of-tune tones elicited an N2/P3 complex and, on in-tune tones only, tonal manipulation elicited an ERAN/RATN-like negativity overlapping with the N2. Our results suggest that cognitive tonal expectations can influence pitch perception at several steps of processing, starting with early attentional selection of pitch. ■

INTRODUCTION

Music perception requires processing that goes beyond hearing successive sounds. Western tonal music contains structural regularities to which listeners are sensitive (Francès, 1958). When processing music, listeners use their knowledge of these tonal regularities to "organize all the pitch events of a piece into a single coherent structure" (Lerdahl & Jackendoff, 1983, p. 106) and to develop expectations about future events. Although some of these structural regularities concern the time dimension (rhythmic and metric regularities), one of the central organizational domains of music is pitch. Pitch regularities concern patterns like frequency of occurrence of tones, melodic contour, and interval size, but also more abstract structures like tonal structures (also referred to as tonal syntax).[1] Numerous studies have shown the influence of tonal structures on music perception (see Bigand & Poulin-Charronnat, 2006; Tillmann, Bharucha, & Bigand, 2000; Krumhansl, 1990; Francès, 1958, for reviews). The key findings suggest that nonmusician listeners have acquired implicit knowledge of the abstract regularities defining tonal structures by mere exposure to tonal music and that this tonal knowledge drives the formation of expectations for future pitch events. Recent behavioral data suggest that tonal expectations can influence pitch perception even at perceptual processing steps and not only at task-related, decisional processing (Marmel, Tillmann, & Dowling, 2008). Aiming to further understand the processing steps influenced by tonal expectations, our present study investigated the in-

fluence of tonal expectations on the neural correlates of pitch processing.

ERPs Associated with Tonal Expectations

Over the last 20 years, numerous ERP studies have investigated the neural correlates underlying tonal expectations. ERP components that were modulated by tonal expectations include N1 (Krohn, Brattico, Välimäki, & Tervaniemi, 2007; Schön & Besson, 2005), MMN (Brattico, Näätänen, & Tervaniemi, 2002), early right anterior negativity (ERAN; Koelsch & Sammler, 2008; Koelsch, Jentschke, Sammler, & Mietchen, 2007; Koelsch, Gunter, Friederici, & Schröger, 2000), right anterior temporal negativity (RATN; Patel, Gibson, Ratner, Besson, & Holcomb, 1998), N5 (Poulin-Charronnat, Bigand, & Koelsch, 2006; Koelsch et al., 2000), and P3-like components (Regnault, Bigand, & Besson, 2001; Patel et al., 1998; Besson & Faïta, 1995; Janata, 1995).

The diversity of the observed components suggests that tonal expectations modulate perception at several processing steps. Differences in materials, tasks, and attentional load in these ERPs studies might have highlighted different processing steps, partly explaining the diversity of the components observed. The earliest influences of tonal relatedness have been observed under attentive listening conditions. Tonal relatedness modulated the N1 evoked by infrequent tones in an oddball paradigm, only when infrequent tones had to be detected by participants (Krohn et al., 2007). When five-note melodies were presented simultaneously in the visual and auditory modalities and participants judged the congruence of both information for

Université de Lyon

the final note, tonal relatedness modulated the N1 when the auditory note mismatched the visual note (Schön & Besson, 2005). Under preattentive conditions (i.e., with participants instructed to ignore the stimuli), the earliest influence of tonal relatedness was slightly delayed, as reflected in the MMN. For example, infrequent tone changes elicited larger MMNs when the tones defined a tonal context (the first five tones of the A major scale) than when they defined a non-tonal context (five tones that did not correspond to a tonal scale; Brattico et al., 2002).

Violations of tonal regularities have also elicited components similar to MMNs that have been denoted the ERAN (or music-syntactic MMN; Koelsch, 2009; Koelsch et al., 2000; see Koelsch, 2009, for a discussion of similarities and differences between the ERAN and the basic feature MMN). In the studies by Koelsch and colleagues, chord sequences were presented with one target chord being either tonally expected or unexpected. The unexpected chord elicited an ERAN in both attentive (Koelsch et al., 2000, 2007; Koelsch & Mulder, 2002) and preattentive conditions (Koelsch, Schröger, & Gunter, 2002). The ERAN has been interpreted as an indicator of music-syntactic processing (Koelsch et al., 2007). The ERAN has also been reported for melodic contexts (Koelsch & Jentschke, 2010; Miranda & Ullman, 2007). Using short melodies, Miranda and Ullman (2007) contrasted violations of tonal relatedness and of long-term memory representations and reported a double dissociation: Violations of tonal relatedness (but not memory violations) elicited an ERAN, and memory violations (but not violations of tonal relatedness) elicited a later negativity similar to an N400. Koelsch and Jentschke (2010) compared tonal manipulations in chord sequences to tonal manipulations in melodies consisting of the top voice of the chord sequences and observed an ERAN starting with an N125 component for both manipulations.

The ERAN was sometimes followed by an N5 (Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005; Koelsch et al., 2000), which has been interpreted as reflecting the integration of musical events into their tonal context (Koelsch et al., 2000). Finally, violations of tonal expectations elicit P3-like components (labeled either LPC, P600, or P3b; Patel et al., 1998; Besson & Faïta, 1995; Janata, 1995). These components are most often observed when participants have to perform a task based on tonal regularities, suggesting that they may be related to an influence of tonal expectations on later task-related decisional processes (but see Besson & Faïta, 1995, Experiment 3, for the observation of an LPC in a passive condition).

Disentangling Cognitive and Sensory Expectations

In addition to differences in tasks and attentional load, the diversity of the components modulated by tonal expectations may be explained by differences in experimental materials. Depending on the experimental material, ERPs may not only reflect processes linked to tonal expectations but also processes linked to pitch incongruities created by the violations of tonal expectations. Some of the observed ERPs may, thus, reflect not the influence of structural knowledge of tonality but rather the processing of acoustic features or deviance (such as the detection of acoustic dissonance). For example, as stated by Koelsch et al. (2007), an ERP evoked by a C# major chord presented in a C major context cannot be attributed solely to top–down processes linked to tonal knowledge, because this chord also creates some acoustical incongruency and is, thus, causing a violation of sensory expectations. In a first attempt to disentangle cognitive and sensory processes of tonal expectations, Regnault et al. (2001) compared ERPs elicited by a tonal violation to those elicited by an acoustic violation. In eight-chord sequences, the first six chords were manipulated in such a way that the final chord was more or less tonally related (thus supposed to be more or less strongly expected), albeit being acoustically the same chord (i.e., the same acoustic signal) in both conditions. These sequences had been used previously in a behavioral priming study to elicit stronger expectations for the final chord when it was strongly tonally related than when it was less tonally related (Bigand & Pineau, 1997). This tonal manipulation was crossed with an acoustical manipulation: the final chord being either kept consonant or rendered dissonant. Different ERP components were elicited by the final chord depending on whether this chord created a tonal violation or an acoustical violation: The acoustical violation elicited an LPC between 300 and 800 msec, whereas the tonal violation elicited an earlier P3 between 200 and 300 msec. This result challenges the interpretation of the LPC (as well as of P600 and P3b) components as reflecting tonal structure processing (Patel et al., 1998; Besson & Faïta, 1995; Janata, 1995). In these three studies, musically unexpected events were acoustically more dissimilar to their context than the expected events; therefore, the observed positivities might be confounded with sensory processing of acoustical dissonance. More recently, Koelsch et al. (Koelsch & Sammler, 2008; Koelsch et al., 2007) took into account that their previously used tonal manipulation (i.e., the unexpected Neapolitan sixth chord, e.g., in Koelsch et al., 2000) confounded cognitive and sensory processes. Aiming to focus on cognitive processes, Koelsch et al. (Koelsch & Sammler, 2008; Koelsch et al., 2007) used chord sequences that were controlled for acoustical incongruencies. Tonally unexpected chords still elicited an ERAN, which pleads for the ERAN reflecting cognitive processing of tonal violations.

Up to now, ERP studies aiming to disentangle cognitive and sensory expectations have only investigated chord processing and not tone processing. Focusing on single tones instead of chords increases the possible control over the experimental material and, thus, allows stronger control of sensory influences. In addition, as sequences of single tones convey less tonal information than chord sequences, observing ERPs associated with tonal relations would highlight the strength of cognitive tonal processes. Finally, sequences of tones allow investigating the

influence of tonal expectations on single pitch processing. Previous ERP studies have shown pitch processing to be sensitive to top–down processes linked to attention (Kauramäki, Jääskeläinen, & Sams, 2007) or musical expertise (Besson, Schön, Moreno, Santos, & Magne, 2007; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005), as reflected in modulations of N1 and N2/P3 components. In the present study, we used the melodic sequences of Marmel et al. (2008, Experiment 3) to investigate the influence of tonal expectations on pitch processing. To focus on cognitive expectations, these melodic sequences manipulated tonal expectations while keeping melodies acoustically as similar as possible. Pitch processing was investigated by repeating the last tone of the melodies; the repetition was identical or out of tune (in half of the trials), and participants had to judge whether the last tone was repeated identically or not. Thus, the task concerned the pitch dimension (pitch discrimination task) and not the tonal relatedness of the tones. Facilitated pitch processing was observed for the expected, tonally related tone in comparison with the less expected, less tonally related tone. In the present study, these melodic sequences and this task allowed us to observe different ERPs on penultimate and final tones: Penultimate tones should be associated with ERP components elicited by the processing of the tonal expectancy violation (i.e., contrasting the less expected subdominant to the expected tonic), whereas final tones should be rather associated with components elicited by processing of pitch deviation (i.e., out-of-tune vs. in-tune tones) and by task-related decisional processes.

METHODS

Participants

Eighteen participants were recruited. Four participants were rejected from the analyses because of excessive artifacts in their EEG. For the remaining 14 participants (eight men and six women between 18 and 30 years old, with a mean age of 25 years), instrumental instruction (i.e., number of years of formal instrumental training) varied between 5 and 25 years, with a mean of 12.2, a standard deviation of 5.5, and a median of 12.

Materials

Twelve pairs of melodies were composed like the example pair shown in Figure 1. All melodies had a length of two bars of four beats each, followed by one additional beat. The two melodies of a pair had the same rhythmic patterns. The 12 melodic pairs had very similar rhythmic patterns: the first three beats of the two first bars consisted of eighth and/or sixteenth notes, the fourth beats of the two first bars and the final beat were quarter notes, and the same rhythmic pattern was used for the two bars of each melody. The two melodies of a pair, in addition to having the same rhythms, contained the same tones except for one—possibly repeated—tone in the first bar. The serial position of this changed note was varied across melodic pairs from the second eighth note of the first beat to the fourth beat. This note change modified the tonal function of the last two tones of the melodies: these last tones, thus, functioned as either the tonic (tonally expected) or the subdominant (tonally less expected). In 50% of the experimental sequences, the last tone was in tune (thus being strictly identical to the penultimate tone). In the other 50%, the last tone was slightly out of tune (its pitch being lowered by 13 cents, a 0.75% deviation in frequency). Twelve melodic pairs were composed so that each of the 12 major keys was represented. The 12 melodic pairs (12 melodies ending on the tonic and 12 melodies ending on the subdominant) and the two pitch conditions (in tune or out of tune) resulted in a total of 48 experimental melodies. To train participants with the task, eight additional example melodies (composed in the same way as the experimental melodies) and 16 shorter melodies (1.5 four-beat bars long, with half ending on the tonic and half on the subdominant) were constructed.

All melodies were created in MIDI with Cubase SX2 software (Steinberg, Hamburg, Germany) and were transformed into audio files using The Grand (a VST piano instrument by Steinberg, Hamburg, Germany). MIDI velocity was constant for all pitches. Melodies were recorded at a tempo of 789.5 msec per beat without any expressive or stylistic timing variations. This tempo represents a duration of 789.5 msec for a quarter note, 394 msec for an eighth note, and 197 msec for a sixteenth note. The overall duration of a melody was 7600 msec (including instrumental resonance

Figure 1. One example of the 12 pairs of melodies used in the present study. The tone differing between related and less related melodies can be visually identified by alteration marks (in the first bar: on the first note of the second beat and on the second note of the third beat). This example can be heard on olfac.univ-lyon1.fr/unite/equipe-02/tillmann/sound_examples.html.

for the final target tone that increased its duration). Cubase's microtuner was used to create the last tone lowered in pitch by 13 cents (a cent is 1/100 of a semitone in logarithmic units). Presentation of stimuli and collection of participants' responses was done with the software Presentation (Neurobehavioral Systems, Albany, CA).

Procedure

Participants were seated in a soundproof booth and performed the pitch discrimination task while their EEG was being recorded. Sound stimuli were presented through headphones, instructions were presented both orally and on a computer screen, and participants responded by pressing keys on a computer keyboard. In a training phase, participants were first familiarized with the pitch deviation by listening to tone pairs, with the tone being either repeated identically or out of tune. Participants were then trained for the task with the 16 short melodies and the eight example melodies. The rhythmic pattern of the melodies was explained to the participants so that they could anticipate the temporal moment of occurrence of the last two tones. Participants judged whether the last two tones were identical or not by using a 4-point scale (1 = sure different; 2 = not sure different; 3 = not sure same; 4 = sure same). No time limit was imposed for responses. The experimental phase contained six blocks of the 48 melodies presented in a pseudorandom order (i.e., random orders with two constraints: the two melodies of a pair had to be separated by at least four melodies and a given pitch deviation condition was not repeated more than five times in succession). Thus, participants judged a total of 288 melodies in the experimental phase. Participants received auditory feedback on errors one beat (789 msec) after their response. To minimize contamination by motor processes, participants were instructed to avoid moving and blinking during the melody and to wait one imaginary beat after the end of the melody to respond. Participants had to press a key to proceed to the next trial and were told that they could relax and blink before proceeding to the next melody.

EEG Recording and Analysis

The EEG was recorded with 64 Ag–AgCl electrodes positioned on a 64-channel electrode cap following the 10–20 system (Electro-Cap Systems). The reference electrode was placed on the tip of the nose, and the ground was FPz. Voltage changes in the EEG caused by the horizontal eye movements were monitored bipolarly with electrodes positioned at the outer canthi of the two eyes, and voltage changes due to vertical eye movements were monitored with electrodes below and above the left eye. The recordings were done in an acoustically and electrically shielded booth. The signal was recorded with a Brain Quick SD64 amplifier and the System Plus software (Micromed, Treviso, Italy) at a resolution of 16 bits and a sampling rate

of 512 Hz and was stored on a hard disk for off-line analysis. All impedances were kept below 10 kΩ. EEG analysis was done with EEGLAB 6.01b (Delorme & Makeig, 2004). The EEG was bandpass-filtered between 1 and 100 Hz. Artifact rejection was done automatically by rejecting epochs with EEG exceeding ±50 μV on Fz, Cz, or Pz and manually by discarding epochs seemingly contaminated by eye movements or muscle activity on any of the electrodes. For the remaining trials, the penultimate and final tones were epoched separately. Epochs ranged from 50 msec before the tone onset to 789 msec after the tone onset and the 50 msec before the tone onset were used as a baseline.[2] For illustration purposes only, a low-pass filter with a cutoff frequency of 15 Hz was applied. ERPs for the penultimate tones and for the final tones were analyzed separately. ERPs were averaged over trials for each of the two conditions for penultimate tones (Related/Less Related) and for each of the four conditions for the final tones (In Tune, Related/In Tune, Less Related/Out of Tune, Related/Out of Tune, Less Related).

For statistical analysis, time windows were defined by visual inspection of grand averages, and mean amplitudes were calculated over each time window. The time windows were centered on the peak amplitudes of Cz on the basis of the grand average across participants. These time windows included 0–100 msec (labeled Nb/P1), 100–200 msec (labeled N1), 200–260 msec (P2 for penultimate tones), 190–250 msec (P2 for final tones), 280–380 msec (labeled N2), 430–570 msec (final tones only, labeled P3), and 610–720 msec (penultimate tones only, labeled P6). Only the analyses on time windows where significant differences were found are reported. These time windows correspond to an early Nb/P1 component between 0 and 100 msec for penultimate tones and to an N2/P3 complex for final tones (280–380 and 430–570 msec).[3] As the main difference between tonal conditions, namely, the Nb/P1 component, was consistent over the scalp, we focused our statistical analyses on the three midline electrodes Fz, FCz, and Cz, where the amplitude of this component was maximal. In addition, as a laterality effect was noticeable on the N2 component, we also performed statistical analyses on three Left (F1, FC1, C1) and three Right (F2, FC2, C2) electrodes. The statistical analyses were performed on mean amplitudes using repeated measure ANOVAs. For penultimate tones, the ANOVA on midline electrodes used Electrode (Fz/FCz/Cz) and Tonal Relatedness (Related/Less Related) as within-participant factors, and the ANOVA on left and right electrodes used Laterality (Left/Right), Electrode (F1 or 2/FC1 or 2/C1 or 2), and Tonal Relatedness (Related/Less Related). For final tones, the same factors were used in the ANOVAs with the addition of the within-participant factor Pitch (In-tune/Out-of-tune). To assess potential repetition effects over the experimental sessions, we also ran analyses including Blocks (first three blocks/second three blocks) as an additional factor in the ANOVAs.

Behavioral Pretest

A behavioral pretest was first conducted with the aim to replicate the improved pitch discrimination for tonally related tonic tones reported by Marmel et al. (2008, Experiment 3) with the here used 13-cent deviation (instead of the previously used 9-cent deviation). Nine participants (instrumental instruction between 0 and 10 years, with a mean of 5.1) had to judge whether penultimate and final tones were identical or different on a 4-point rating scale (1 = sure different; 2 = not sure different; 3 = not sure same; 4 = sure same). One block of the 48 experimental melodies was presented. Participants' ability to discriminate between in-tune and out-of-tune final tones was analyzed by calculating areas under the receiver operating characteristic (ROC) curves (Swets, 1973) and by performing an ANOVA on the area scores, with Tonal Relatedness (Related/Less Related) as a within-participant factor. This ANOVA revealed a significant effect of Tonal Relatedness: Pitch discrimination was better when the two final tones were tonics (ROC areas with standard errors: 78.8 ± 6.2) than when they were subdominants (72.1 ± 7.5) ($F(1, 8) = 6.37, p < .05$). This result replicates the finding of Marmel et al. (2008, Experiment 3).

RESULTS

Behavioral Data

Area scores averaged over the 14 musician participants were high overall and did not differ between the related tonic tones (.96) and the less related subdominant tones (.95), as shown by a two-sided paired t test ($t(13) = 2.16$, $p = .37$). The absence of a significant difference may be because of a ceiling effect, reflected by the very high performance scores, and caused by the participants being more experienced musicians than the participants of our pretest (12.2 years of musical instruction, on average, vs. 5.1). Informal comments of participants gathered after the experimental session suggested that participants were not aware of the tonal manipulation.

ERP Results

Penultimate Tones

On midline electrodes, an early effect of Tonal Relatedness on amplitudes was found: An early Nb/P1 component was observed with more positive amplitudes for related tones than for less related tones ($F(1, 13) = 8.04, p < .05$). In addition, the effect of Electrode was significant, with the more frontal electrode showing more positive amplitudes ($F(1, 13) = 5.60, p < .05$) (Figure 2). The ANOVA on left and right electrodes confirmed these two effects (effect of Tonal Relatedness: $F(1, 13) = 8.61, p < .05$; effect of Electrode: $F(1, 13) = 10.43, p < .01$), but no effect of Laterality was found. The magnitude of the effect of tonal relatedness (i.e., the difference between the two conditions) did not correlate with instrumental instruction. The ANOVA with Blocks as an additional factor did not reveal any effect of repetition (i.e., no main effect of Block nor interactions involving Block).

Figure 2. Grand average ERPs for the penultimate tones, presented as a function of tonal relatedness (Related: tonic; Less Related: subdominant), recorded on three left electrodes (F1, FC1, C1), three midline electrodes (Fz, FCz, Cz), and three right electrodes (F2, FC2, C2). Arrows indicate the Nb/P1 component (labels shown on FCz only). Negativity is up.
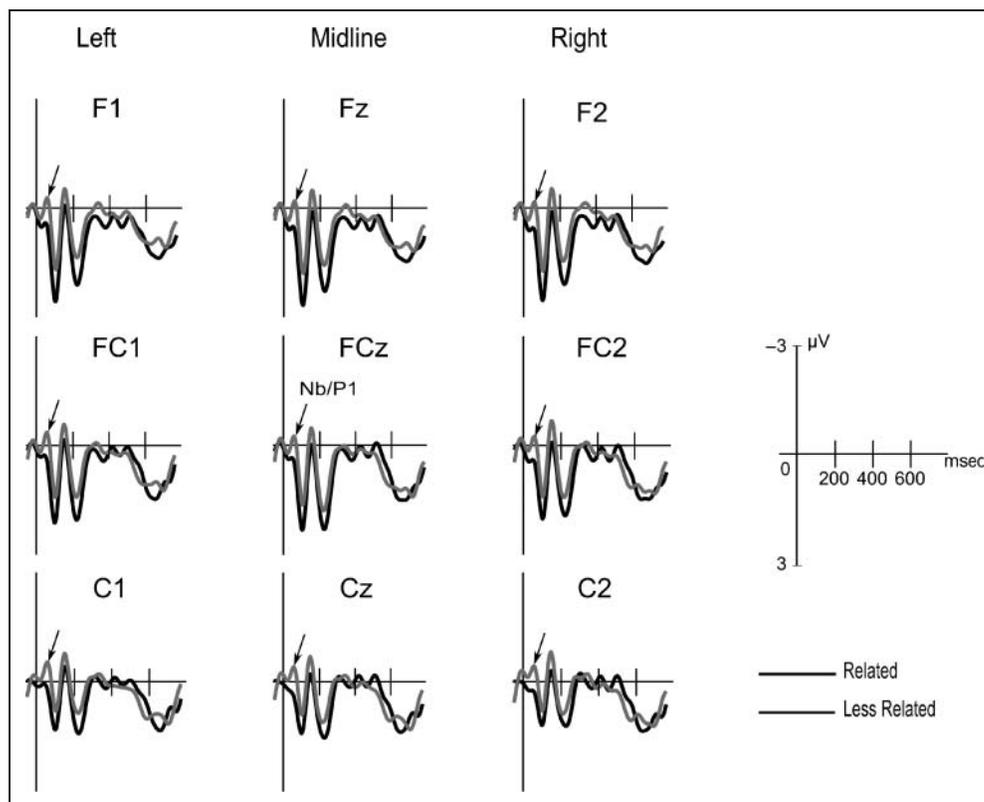
Figure 3. Grand average ERPs for the final tones, presented as a function of tonal relatedness (Related: tonic; Less Related: subdominant) and pitch deviation (In tune; Out of tune), recorded on three left electrodes (F1, FC1, C1), three midline electrodes (Fz, FCz, Cz), and three right electrodes (F2, FC2, C2). Arrows indicate the N2 and P3 components (labels shown on FCz only). Negativity is up.
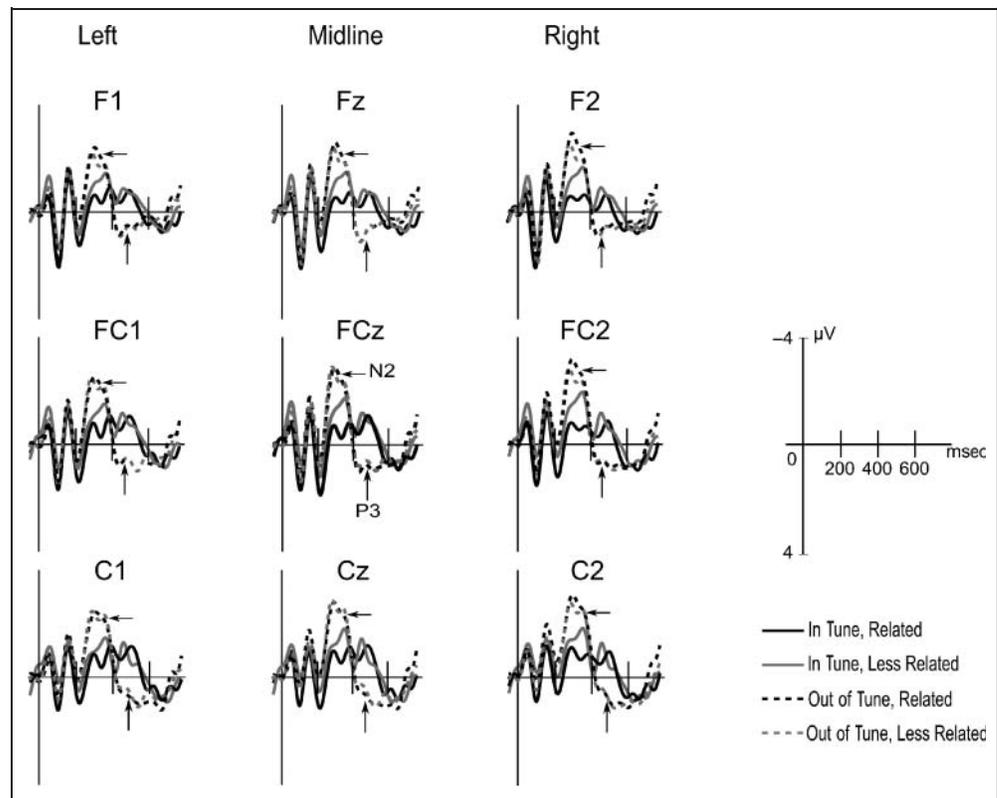
Final Tones

On midline electrodes, the introduction of a pitch deviation on final tones elicited a N2/P3 complex: A larger N2 was observed for out-of-tune tones than for in-tune tones ($F(1, 13) = 19.39$, $p < .001$), followed by a larger P3 for out-of-tune tones than for in-tune tones ($F(1, 13) = 8.14$, $p < .05$) (Figure 3). The ANOVA on left and right electrodes confirmed this effect of Pitch on N2 ($F(1, 13) = 20.36$, $p < .001$) and P3 ($F(1, 13) = 8.23$, $p < .05$).

A main effect of Laterality was observed on N2, with larger negativities on the right electrodes than on the left electrodes ($F(1, 13) = 8.16$, $p < .05$). Laterality also interacted with Tonal Relatedness and Pitch (three-way interaction: $F(1, 13) = 5.61$, $p < .05$), reflecting a pattern of larger N2 differences between tonally related and less related tones for in-tune tones than for out-of-tune tones, this being more pronounced on the right electrodes than on the left electrodes (Figure 3). Visual inspection of the results displayed in Figure 3 suggests that the difference between related and less related in-tune tones emerged slightly later than did the main N2 difference between out-of-tune and in-tune tones, and topographies suggested that this difference between related and less related in-tune tones was larger on fronto-right electrodes. This led us to perform the same ANOVA analysis, but on a time window starting later, that is, between 320 and 380 msec: The three-way interaction between Laterality, Tonal Relatedness, and Pitch was found again ($F(1, 13) = 5.94$, $p < .05$), and contrast analyses showed that amplitudes were more

negative for less related than for tonally related tones on F2 (two-tailed t test: $t(1, 13) = 1.77$, $p < .05$) and FC2 ($t(1, 13) = 1.77$, $p < .05$). The ANOVA with Blocks as an additional factor did not reveal any effect of repetition.

DISCUSSION

The present study investigated the influence of tonal expectations on ERPs elicited in a pitch discrimination task. To focus on cognitive expectations linked to listeners' implicit knowledge of tonal regularities, listeners' tonal expectations were manipulated while keeping the melodies acoustically as similar as possible. Target tones (both penultimate and final tones) as well as almost all context tones were kept identical in related and less related conditions, thus controlling bottom–up expectations linked to contour, intervals, or tone repetition (i.e., sensory expectations). Also, to ensure that the ERPs elicited by tonal relatedness were not confounded with ERPs elicited by acoustic features (such as consonance/dissonance), the tonal manipulation involved diatonic tones only (i.e., tones belonging to the tonality). This design ensured that the ERPs elicited by the penultimate tones were not because of the processing of acoustic incongruities. The "repeated tone" design further allowed us to study ERPs elicited on the penultimate tones that were not confounded with task-related decisional processes. For penultimate tones, this allowed us to focus on the perceptual processing steps of pitch processing (without the decisional

processes that occurred on the final tones, together with the perceptual processes). Finally, ERPs associated with pitch discrimination were investigated on final tones in interaction with tonal relatedness as final tones could be played either in tune or out of tune in addition to being tonally related or less related.

Processing of Tonal Relatedness

The main result of the present study is the observation of top–down influences of tonal expectations on ERP components within the first 100 msec after tone onset. On penultimate tones, the Nb/P1 complex differed in amplitude between the tonally related tones and the less related tones. To our knowledge, modulations of components as early as this Nb/P1 have not been previously reported in studies investigating musical expectations. The closest components in previous auditory research to which we can compare our finding might be the early components observed in auditory attention studies using the dichotic listening paradigm (Woldorff et al., 1993; Woldorff & Hillyard, 1991). In these studies, participants focused on tones presented to one ear while ignoring tones presented to the opposite ear at another pitch. In both ears, tones were either frequent "standard" tones or infrequent (9%) tones of weaker intensity. Frequent standard tones elicited an enhanced positivity between 20 and 50 msec when presented in the attended ear than when presented in the unattended ear (named the P20-50 attention effect). This early response was reported to originate from the primary auditory cortex (Woldorff et al., 1993). Woldorff and Hillyard suggested that the occurrence of this early attentional effect was created by the high attentional load (because of the fast presentation rate of the tones and the difficulty of the task) and by the use of different pitches for both ears, which helped listeners to focus their attention on the requested ear. This effect was interpreted as strong evidence for the theory of early selection, which postulates the existence of a filtering (or gain) mechanism that allows sensory input to be selected before the completion of perceptual analysis (Woldorff, 1999). There were some differences between the early P20-50 in the studies of Woldorff and collaborators (Woldorff et al., 1993; Woldorff & Hillyard, 1991) and the Nb/P1 in our study, which may be attributed to the differences in experimental materials and task. Notably, the Nb/P1 was delayed in comparison with the usual latency of P1 components (about 50 msec) as it peaked at 100 msec, and this delay also affected the subsequent components with the N1 peaking around 150 msec. This delay could be explained by the long durations of the tones of the prime context (i.e., thus preceding the target tone), as suggested by previous studies that have reported delayed P1s and N1s in experimental contexts using long stimuli durations (Proverbio, Esposito, & Zani, 2002). The differences between the P20-50 and Nb/P1 do not rule out the proposed attentional interpretation as the research by Woldorff and Hillyard integrates into a larger set of re-

search showing early effects of attention. The data of Woldorff and Hillyard were observed for spatial auditory attention, but other studies have shown that early effects can arise from attention being allocated by features other than location (e.g., color, orientation, pitch; see Astheimer & Sanders, 2009). For example, a modulation of an early P1 component was reported in a study on temporal attention in the visual modality (Correa, Lupiáñez, Madrid, & Tudela, 2006). These findings suggest that early selection may be an amodal attentional mechanism that can be triggered by top–down processes in different modalities. Hence, the observed Nb/P1 in our study may be interpreted in terms of attention and suggest that top–down mechanisms can modulate pitch processing at early attentional levels. Attention and musical structures have been linked in the dynamic attending theory (Large & Jones, 1999; Jones & Boltz, 1989). Musical structures guide and modulate attention over time so that more attentional resources are available for processing musically related events than less related events. Facilitated processing for tonally related (expected) events in comparison with less-related (less expected) events and to neutral events has been interpreted in this attentional framework (Escoffier & Tillmann, 2008; Tillmann, Janata, Birk, & Bharucha, 2008). In our study, it is worth noting that the observed effect cannot be explained by different frequencies of occurrence that might have been orienting attention toward one target type in particular (related tonic or less related subdominant) as the two tonal relatedness conditions had the same frequency of occurrence in the experimental session (50%). Instead, the observed effect could be explained by listeners' tonal knowledge and tonal expectations, with the two tonal conditions orienting attention differently. The orientation of attention would, thus, not result from differences in probabilities of occurrence or in experimental instructions to participants but from tonal expectations based on musical structures (as suggested by the dynamic attending theory of Jones; Large & Jones, 1999; Jones & Boltz, 1989). In the tonally related condition, listeners' tonal knowledge would orient attention toward the pitch of the actual penultimate tone (because it is tonally related to the melodic context), whereas in the less tonally related condition listeners' attention would have been oriented toward a pitch different from the one of the actually sounding penultimate tone (i.e., toward the tone that would be tonally related in this melodic context).

Alternatively to the interpretation in terms of more attentional resources allocated for related tones, the tonal effect on Nb/P1 in our study could be interpreted as resulting from a mismatch detection on the less related tones. It is worth noting that both interpretations suggest that listeners build up tonal expectations for the penultimate tone while listening to the melodic context, such that by the time the penultimate tone is presented they would strongly expect the pitch of the tonally related tone. In the attentional interpretation, the Nb/P1 would emerge from more resources being allocated to the processing of related

tones whereas, if mismatch detection is involved, the Nb/P1 would result from the participants' expectations being deceived in the less related condition.

For music cognition, previous electrophysiological evidence for early attentional effects may be found in Krohn et al. (2007). Their study used an oddball paradigm where infrequent tones were more or less tonally related. When participants had to detect these tones (i.e., attentive condition), larger N1s were observed for the most tonally related of the infrequent tones. Although the authors did not discuss their effect as being attentional but rather as reflecting a more accurate representation of the most hierarchically important pitches that would result in a larger N1, the fact that this effect occurred only in the attentive listening condition suggests that attentional processes were involved in the observed N1. The authors suggested that allocation of attention might have been needed for the N1 effect, because their study focused on small tonal contrasts (i.e., within-key tonal hierarchies) as ours also did. In contrast to the oddball paradigm, in which participants cannot predict the specific moment of occurrence of infrequent tones within the stimulus sequence, the rhythmic construction of our melodies allowed participants to anticipate the temporal moment of occurrence of the penultimate tones within a melody (thus making them temporally predictable). This implementation of the pitch comparison task might have allowed participants to attribute more attentional resources to the processing of the pitch of the penultimate tones than would be possible with random occurrences of infrequent tones, thus resulting in an attentional effect earlier than the N1. Another factor that could have favored early components is that the tonal manipulation in our melodies not only changed the tonal function of penultimate and final tones but theoretically changed the tonal function of all the tones following the note change, which occurred at the beginning of the melodies. Thus, the processes eliciting the tonal relatedness effect within the first 100 msec of penultimate tones were probably already at stake before and may have favored an early effect.[4]

In addition, for the final in-tune tones, tonally less related tones elicited more negative ERPs than tonally related tones between 320 and 380 msec on two right electrodes. This ERP component is reminiscent of the ERAN observed by Koelsch et al. (Koelsch & Sammler, 2008; Koelsch et al., 2000, 2002, 2007), even if the ERAN is usually observed earlier (around 200–250 msec). It is also reminiscent of the RATN observed by Patel et al. (1998), a right anterior negativity between 300 and 400 msec, which was larger for out-of-key chords than for in-key chords. Both the ERAN and the RATN have been interpreted as reflecting syntactic processes, and our result is congruent with this interpretation. Two processes might be reflected in the negativities observed for the final tones in our study: an N2 related to the detection of pitch deviations (larger amplitudes for out-of-tune tones than for in-tune tones between 280 and 400 msec) and a smaller and later ERAN/RATN-like component related to the processing of tonal relatedness (larger

negativity for in-tune less related than for in-tune related tones between 320 and 380 msec). The fact that the ERAN/RATN-like component was observed only on the final tones and not on the penultimate tones might be linked to the use of tone sequences with subtle acoustic differences between the two tonal conditions (implemented in melodies). Robust ERANs have been previously observed in melodies with strong tonal violations (Miranda & Ullman, 2007) and when melodies were simultaneously presented in the auditory and visual (musical score) modalities, which may have enhanced tonal expectations in comparison with a purely auditory presentation (Schön & Besson, 2005). Interestingly, Koelsch and Jentschke (2010) reported a latency of the ERAN that was even earlier than the latency of 200–250 msec found in most studies reporting ERANs. This study compared chord sequences to melodic lines that were the extracted top voice of the chord sequences and reported an early negativity at 125 msec in both melodic and chord conditions and an additional, later, negative component at 180 msec in the chord condition. In our study, the use of more complex melodies and of subtle acoustic differences between the two tonal conditions (i.e., controlling for sensory influences and also ending on the same interval patterns before the target) might have weakened the processes indexed by the ERAN so that the ERAN was delayed in comparison with previous studies and emerged only on the final tones, that is, when the tone was repeated and participants focused on detecting a pitch incongruity.

One limitation of our study is that we have investigated implicit tonal knowledge, which is thought to be shared between musicians and nonmusicians, while testing only musician participants. Although numerous studies have shown that musical training leads to changes throughout the neural system (see Kraus & Chandrasekaran, 2010, for a review), ERP studies focusing on the influence of musical (tonal) structures tend to report the same effects for musicians and nonmusicians (e.g., Koelsch & Sammler, 2008; Regnault et al., 2001; see Bigand & Poulin-Charronnat, 2006, for a review including also behavioral studies), even if the amplitudes of the observed ERPs might be smaller for nonmusicians in some cases (e.g., Koelsch et al., 2007). These findings, together with our observation that the amplitudes of ERPs did not correlate with years of instrumental instruction for the early tonal effect on penultimate tones, suggest that the pattern of results should be similar for nonmusician participants, although this remains to be shown in a future study.

Processing of Pitch Deviations

On final tones, out-of-tune tones elicited a larger N2 and a larger P3 than in-tune tones. This N2/P3 complex is in agreement with previous results observed for the detection of pitch deviations. Pitch deviations of 20 cents (slightly larger than the 13-cent pitch deviation used in the present study) at the end of melodies elicited a negativity around 200 msec followed by a positivity between

200 and 800 msec (Schön, Magne, & Besson, 2004). Pitch discrimination was associated with N300-P600 components in 8-year-old children who were given 6 months of musical training (Besson et al., 2007). These components are similar to the results of Tervaniemi et al. (2005), who investigated the detection of mistunings as small as ours (also 0.76%, i.e., 13 cents) in an oddball paradigm and found that the mistuning was associated with an N2/P3 complex. More generally, the N2/P3 complex has been associated with attentional decision-making processes (Woldorff, 1999): In our task, it may reflect the detection of the pitch violations and their categorization as "different" responses.

Conclusion

The main finding of our study is the influence of tonal expectations on an early Nb/P1 component. This suggests that top–down processes linked to tonal relatedness (based on listeners' knowledge of the tonal system) influence pitch processing at perceptual—as opposed to decisional—levels (see Marmel et al., 2008, for behavioral data with nonmusician participants). In the present study, participants' attention was directed to the pitch dimension with a pitch discrimination task. Tonal relatedness modulated early pitch processing of the penultimate tones, possibly by enhancing early attentive selection for expected tones. The design with the repeated last tone allowed us to investigate these influences independently of task-related decisional processes, which were requested on final tones only. These decisional processes elicited an N2/P3 complex for deviant out-of-tune tones, and the negative component of this complex was modulated by tonal relatedness for in-tune tones. This suggests that an ERAN-like component, linked to processing of tonal relatedness, may have overlapped with the N2 component in this time window.

Our study is in line with recent electrophysiological studies that have highlighted the influence of musical expertise on pitch processing by showing that musicians have better pitch encoding than nonmusicians at the brainstem level (Musacchia, Sams, Skoe, & Kraus, 2007; Wong, Skoe, Russo, Dees, & Kraus, 2007). In contrast to these studies that focused on musical expertise based on explicit musical training (musicians vs. nonmusicians), our study focused on the implicit tonal knowledge shared by both musicians and nonmusicians (i.e., their implicit expertise of music perception; see Bigand & Poulin-Charronnat, 2006). Our study showed an influence of this tonal knowledge on early cortical processes and, thus, raises the question of its possible influences on even earlier subcortical processes, as has been observed for explicit expertise.

Notes

1.   Tonal structures are based on combinations of musical events that follow the organizational principles of the tonal system. In this system, musical events are organized at the levels of tones, chords, and keys. At the tone level, subsets of seven pitch classes form scales (e.g., C, D, E, F, G, A, B form the C-major scale). The combination of three or more tones from a scale form a chord when played simultaneously (e.g., C-E-G form a C-major chord). Musical pieces that use tones and chords mainly built from one scale are said to be in one key or tonality (e.g., the C-major key). Tones belonging to the scale are called diatonic tones or within-key tones; those outside the scale are nondiatonic, out-of-key tones. Music events are organized hierarchically according to their tonal function within the key. The most referential event (for tones and chords) is the tonic, which gives its name to the key. The dominant (fifth degree) and subdominant (fourth degree) are next in tonal hierarchy, followed by the other within-key tones and chords (including supertonic and subdominant). Out-of-key (i.e., nondiatonic) chords and tones are incongruous with the instilled key and have the lowest hierarchical rank.

2.   Two alternative baselines have also been used in additional analyses: one consisting of 200 msec taken before the onset of the penultimate tones and one consisting of 600 msec taken before the onset of the melodies. Analyzing the data with these baselines led to the same results. Also, an ANOVA performed on the 50-msec baseline (for the nine electrodes kept in the presented analyses) did not reveal a difference between the two tonal conditions, thus suggesting that the observed differences on the penultimate tone here were not affected by the ERPs to the previous tone.

3.   For the Nb/P1, we also did peak amplitude and latency analyses on the Nb peak and on the P1 peak. These analyses showed effects of tonal relatedness similar to the one we report with the time window analysis on both amplitudes and latencies for both the Nb and the P1 peaks.

4.   However, statistical analyses of the 100 msec preceding the onset of the penultimate tones did not reveal significant differences. This argues that, even if participants' expectations build up over time, it is their fulfilment or violation on penultimate tones that is reflected in the Nb/P1 effect.

REFERENCES

Astheimer, L. B., & Sanders, L. D. (2009). Listeners modulate temporally selective attention during natural speech processing. *Biological Psychology, 80,* 23–34.

Besson, M., & Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1278–1296.

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience, 25,* 399–410.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics, 59,* 1098–1107.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100,* 100–130.

Brattico, E., Näätänen, R., & Tervaniemi, M. (2002). Context effects on pitch perception in musicians and nonmusicians: Evidence form event-related-potential recordings. *Music Perception, 19,* 199–222.

Correa, Á., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing: A review and new evidence from event-related potentials. *Brain Research, 1076,* 116–128.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods, 134,* 9–21.

Escoffier, N., & Tillmann, B. (2008). The tonal function of a task-irrelevant chord modulates speed of visual processing. *Cognition, 107,* 1070–1083.

Francès, R. (1958). *La perception de la musique* [*The perception of music*] (J. W. Dowling, Trans.). Hillsdale, NJ: Erlbaum, 1988. Paris: Vrin.

Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience, 7,* 153–164.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Kauramäki, J., Jääskeläinen, I. P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. *PLoS ONE, 2,* e909.

Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology, 46,* 179–190.

Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Non-musicians" are musical. *Journal of Cognitive Neuroscience, 12,* 520–541.

Koelsch, S., & Jentschke, S. (2010). Differences in electric brain responses to melodies and chords. *Journal of Cognitive Neuroscience, 22,* 2251–2262.

Koelsch, S., Jentschke, S., Sammler, D., & Mietchen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology, 44,* 476–490.

Koelsch, S., & Mulder, J. (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology, 113,* 862–869.

Koelsch, S., & Sammler, D. (2008). Cognitive components of regularity processing in the auditory domain. *PLoS ONE, 3,* e2650.

Koelsch, S., Schröger, E., & Gunter, T. C. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology, 39,* 38–48.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11,* 599–605.

Krohn, K. I., Brattico, E., Välimäki, V., & Tervaniemi, M. (2007). Neural representations of the hierarchical scale pitch structure. *Music Perception, 24,* 281–296.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch.* New York: Oxford University Press.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106,* 119–159.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research, 25,* 678–687.

Marmel, F., Tillmann, B., & Dowling, W. J. (2008). Tonal expectations influence pitch perception. *Perception & Psychophysics, 70,* 841–852.

Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage, 38,* 331–345.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences, U.S.A., 104,* 15894–15898.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10,* 717–733.

Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience, 18,* 1545–1554.

Proverbio, A. M., Esposito, P., & Zani, A. (2002). Early involvement of the temporal area in attentional selection of grating orientation: An ERP study. *Cognitive Brain Research, 13,* 139–151.

Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience, 13,* 241–255.

Schön, D., & Besson, M. (2005). Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 17,* 694–705.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41,* 341–349.

Swets, J. A. (1973). The relative operating characteristic in psychology. *Science, 182,* 990–1000.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research, 161,* 1–10.

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review, 107,* 885–913.

Tillmann, B., Janata, P., Birk, J., & Bharucha, J. J. (2008). Tonal centers and expectancy: Facilitation or inhibition of chords at the top of the harmonic hierarchy? *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1031–1043.

Woldorff, M. G. (1999). Auditory attention. In R. Wilson & F. Keil (Eds.), *The MIT encyclopedia of the cognitive sciences* (pp. 50–52). Cambridge, MA: MIT Press.

Woldorff, M. G., Gallen, G. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., et al. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences, U.S.A., 90,* 8722–8726.

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. *Electroencephalography and Clinical Neurophysiology, 79,* 170–191.

Wong, P. C. M., 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.