Processing Tonal Modulations: An ERP Study

Stefan Koelsch[1,2], Thomas Gunter[1], Erich Schröger[3], and
Angela D. Friederici[1]

Abstract

■ A common stylistic element of Western tonal music is the change of key within a musical sequence (known as *modulation* in musical terms). The aim of the present study was to investigate neural correlates of the cognitive processing of modulations with event-related brain potentials. Participants listened to sequences of chords that were infrequently modulating. Modulating chords elicited distinct effects in the event-related brain potentials: an early right anterior negativity reflecting the processing of a violation of musical regularities and a late frontal negativity taken to reflect processes of harmonic integration. Additionally, modulations elicited a tonic negative potential suggested to reflect cognitive processes characteristic for the processing of tonal modulations, namely, the restructuring of the "hierarchy of harmonic stability" (which specifies musical expectations), presumably entailing working memory operations. Participants were "nonmusicians"; results thus support the hypothesis that nonmusicians have a sophisticated (implicit) knowledge about musical regularities. ■

INTRODUCTION

The investigation of the cognitive processing of music has become a substantial aspect of cognitive neuroscience. The present study investigates how a change of a tonal key is electrophysiologically reflected in the brains of "nonmusicians" (i.e., in listeners with no special musical expertise).

Most of the melodies in Western tonal music are normally based on a harmonic structure that refers to one single key (e.g., C major). However, Western tonal music consists of various keys, and music becomes more interesting for listeners when the key of a musical piece sometimes changes, for example, between two melodies or even within one melody. In every first movement of a classical sonata or symphony, the first change of key takes place between the first and the second theme (i.e., often within the first minute of the movement). In music theory, a change of key (e.g., from C major to G major) within a progression of harmonies is termed *modulation*. Modulations have a strong "dynamic aspect in time" (Krumhansl & Kessler, 1982) because they induce the expectancy for a completion of the modulation or even the expectation for a return to the initial key (Schenker, 1956). During the last centuries, composers have found numerous ways to modulate from one key to another. Modulations are usually hardly detectable for a "nonmusician," and even "musicians" often have difficulties in detecting modulations. This holds especially for modulations between closely related keys and when elegantly composed. Thus, interestingly, modulations are musically highly relevant but subjectively often quite unsalient.

A musical key always refers to a tonal center that is ideally represented by the tonic chord (Krumhansl & Kessler, 1982; Schönberg, 1969; Hindemith, 1940). However, there are several other chords that have a function within a key (e.g., subdominant, or dominant). Importantly, the function of each chord in respect to one key can be interpreted as a different function in respect to another key: for example, in a C major context, a G major chord functions as the dominant (of C major), but a G major chord may also function as the subdominant in D major (Figure 1) or as the tonic in G major. Once a key is established (usually within the first few notes or chords), listeners tend to interpret following chords as belonging to this initial key (e.g., Berent & Perfetti, 1993; Bharucha & Krumhansl, 1983; Krumhansl & Kessler, 1982). Nevertheless, because each chord may also belong to another key, there always remains a residual ambiguity for every chord. A modulation from one key to another can, for example, be performed by interpreting an in-key chord of the first key as a different in-key chord function of another key. Such a modulation is termed *diatonic modulation*, and this type of modulation was investigated in the present study. The chord that has one function in the old and another function in the new key is termed *pivot chord* (Figure 1).

Psychologically, a change of key can be detected by realizing that certain notes do not belong to the formerly established key. This requires an exact representation of the major–minor tonal system, especially during the modulation between two closely related keys.

---

[1] Max-Planck-Institute of Cognitive Neuroscience, [2] Harvard Medical School, [3] University of Leipzig
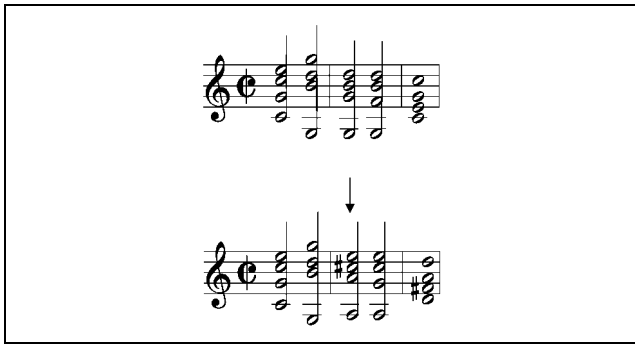
Figure 1. Examples of stimuli. Top: C major chord sequence exclusively consisting of in-key chords. Bottom: chord sequence modulating from C major to D major. The second chord is the pivot chord, functioning as dominant in C major, as well as subdominant in D major. The third chord of the modulating sequence (indicated by the arrow) is the dominant chord of D major, introducing one out-of-key note in respect to C major.

In such a case, only few notes (or even only one single note) belong to only one of the two keys. Within a modulation, at least one chord occurs that does not belong to the old, but to the new key. Usually, this chord directly follows the pivot chord. The fact that a chord belongs to a new key is indicated by notes that are not compatible with the preceding key (but with another key). Only after this moment, listeners know that a modulation might take place (it might also be the case that merely a "harmonic sidestep" without a modulation was performed, followed by an instant return to the initial key, for effects elicited by such events, see Koelsch, Gunter, Friederici, & Schröger, 2000).

In the present study, chord sequences were employed as stimuli, each sequence consisting of five chords (similar to previous studies from Koelsch, Gunter, et al., 2000; Koelsch, Maess, & Friederici, 2000; Koelsch et al., 2001; Maess, Koelsch, Gunter, & Friederici, 2001; Koelsch, Schmidt, & Kansok, 2002; Koelsch, Schröger, & Gunter, 2002). When consisting of in-key chords only, chord sequences were composed in a way that a harmonic context was built up toward the end of each sequence (Krumhansl & Kessler, 1982; Schönberg, 1969; Hindemith, 1940). Infrequently (with $p = .25$), a sequence that modulated two-fifths upward was presented (in the sense of the circle of fifths, Krumhansl & Kessler, 1982; Schönberg, 1969), for example, from C major to D major. Sequences were presented one directly following the other, sounding like a musical piece.[1] Subjects were not informed about the existence of modulations and they did not have a task connected to the modulations (to avoid an overlap of P3 potentials elicited by detectional and decisional processes with potentials reflecting original music processing). To control that participants nevertheless attended the stimulation, they were asked to detect infrequently occurring deviant instrumental sounds (see Methods).

As described before, during a sequence of harmonies belonging to one single key, listeners familiar with major–minor tonal music usually tend to expect that subsequent chords are harmonically appropriate. That is, listeners expect that subsequent chords are harmonically closely related and that their chord function is compatible with the regularities of major–minor tonal music (Maess et al., 2001; Koelsch, Gunter, et al., 2000; Tillmann, Bigand, & Madurell, 1998; Tillmann, Bigand, & Pineau, 1998; Bigand & Pineau, 1997; Bigand, Madurell, Tillmann, & Pineau, 1999; Bigand, Tillmann, Poulin, Adamo, & Madurell, 2001; Bharucha & Krumhansl, 1983; Krumhansl & Kessler, 1982). The modulations employed in the present experiment introduced chords that were harmonically less appropriate than in-key chords: They were harmonically less closely related than in-key chords (the modulating chord at the third position contained one note that did not belong to the key established by the first two chords of the sequence) and their chord functions did not directly belong to the previously established key. Harmonically inappropriate chords are perceived as unexpected (e.g., Janata, 1995; Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986; Krumhansl & Kessler, 1982) and the violation of a musical expectancy for harmonically appropriate chords has been shown to be reflected in the event-related brain potentials (ERPs) as an early right anterior negativity (denoted, as a working term, ERAN, Koelsch, Gunter, et al., 2000). Note that the generation of musical expectancies (such as the expectation for a specific chord function) is based on (implicit) knowledge of complex musical regularities, in the present study regularities of major–minor tonal music. These regularities are described by music theory (Maess et al., 2001; Koelsch, Gunter, et al., 2000; Koelsch et al., 2001; Schönberg, 1969; Piston, 1948/1987; Hindemith, 1940) and have been taken as part of a musical syntax (Koelsch, Schmidt, et al., 2002; Maess et al., 2001; Tillmann, Bharucha, & Bigand, 2000; Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Swain, 1997; Sloboda, 1985; Deliège, 1984). Because the ERAN can be elicited by violations of these regularities, the ERAN might be taken as a reflection of music-syntactic processing.

Several studies revealed that this ERP effect shares a number of features with the "mismatch negativity" (MMN, Koelsch, Gunter, et al., 2000; Koelsch et al., 2001; Koelsch, Schmidt, et al., 2002, Koelsch, Schröger, et al., 2002; Schröger, 1998; Näätänen, 1992), but that there are also important differences between the ERAN and both the frequency and the abstract-feature MMN (Koelsch, Gunter et al., 2000; Koelsch et al., 2001; Koelsch & Mulder, 2002). The most important difference is that the ERAN has been shown to be sensitive to musical events that are deviant in respect to a fairly complex system of (music-syntactic) regularities, whereas the MMN is broadly taken to reflect auditory sensory memory operations (Schröger, 1997). However, because
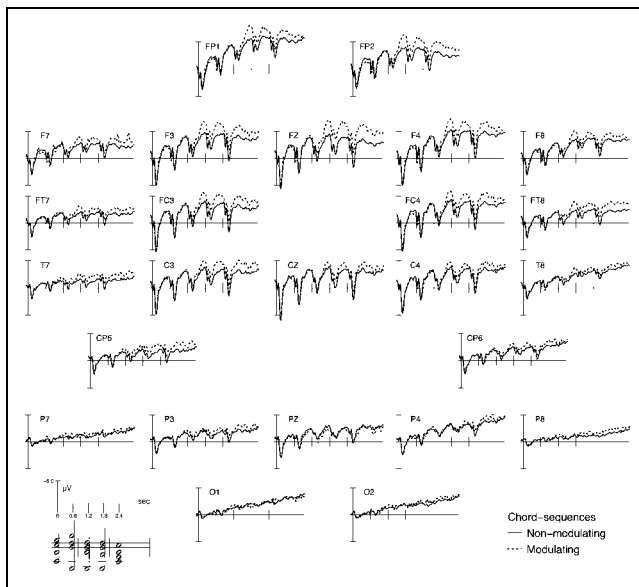
Figure 2. Grand average ERPs of entire chord sequences, separately for in-key and modulating sequences. Each vertical line corresponds to the onset of a chord. Modulations (presented from positions 3 to 5) elicited negative effects with frontal preponderance.

of the similarities between ERAN and MMN, and due to the findings that the ERAN is sensitive to violations of complex musical regularities, it has recently been suggested that the ERAN might best be understood as a music-syntactic MMN (Koelsch et al., 2001; Koelsch, Schröger, et al., 2002).

The ERAN, or music-syntactic MMN, can be elicited preattentively (Koelsch, Schröger, et al., 2002), reflecting that musical expectancies are automatically generated during the perception of a musical context (Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986, 1987; Krumhansl & Kessler, 1982). Moreover, the ERAN is larger in "musicians" (i.e., in subjects with explicit musical long-term training) than in "nonmusicians" (Koelsch, Schmidt, et al., 2002), reflecting that more specific representations of musical regularities lead to more specific musical expectancies. The neural generators of the ERAN have been localized in the inferior fronto-lateral cortex (referred as Broca's area in the left hemisphere), reflecting that musical regularities are processed in brain structures that are also involved in the processing of syntactic rules of language (Maess et al., 2001).

The ERAN is usually followed by a late frontal negativity that is maximal around 500 msec. This effect was termed the N5 (Koelsch, Gunter, et al., 2000) and is taken to reflect processes of harmonic integration. During listening to a sequence of harmonies that build up a musical context, progressing chords are integrated into this context, each chord specifying the tonal schema established by the preceding chords (Koelsch, Gunter, et al., 2000; Krumhansl & Kessler, 1982). Harmonically inappropriate chords require a higher amount

of integration into a musical context compared to harmonically appropriate chords, presumably leading to a larger amplitude of the N5.

In the present study, it was hypothesized that (a) modulating chords elicit an ERAN because they violate harmonic expectancies, (b) modulating chords elicit an N5 because they require a higher amount of musical integration, and (c) the processing of a change of the tonal key is reflected in the ERPs of modulations.

RESULTS

Behaviorally, participants responded on average with 93% hits (range 67–100%), indicating that participants could easily detect the deviant instruments and that they attended to the timbre of the musical stimulus (note that the modulations were task-irrelevant).

Brain responses to entire chord sequences (i.e., from chords 1 to 5) are shown in Figure 2. ERPs elicited by modulating sequences distinctly differed from those elicited by in-key chord sequences. Strong negative ERP effects were elicited during the presentation of all modulating chords (i.e., by chords at the third, fourth, and fifth positions of the modulating sequences). These effects were predominant at frontal electrode sites and larger over the right compared to the left hemisphere.

To analyze the ERP effects of modulations in detail, in the first step, ERPs of in-key chords at the third position were opposed to those elicited by modulating chords at the corresponding position (Figure 3). Modulating chords elicited an early negativity that was present around 180–280 msec (in respect to the onset of the third chord) and had a right anterior preponderance (the ERAN, or "music-syntactic MMN"). This ERP effect was followed by a late frontal negativity that had an onset around 400 msec and peaked
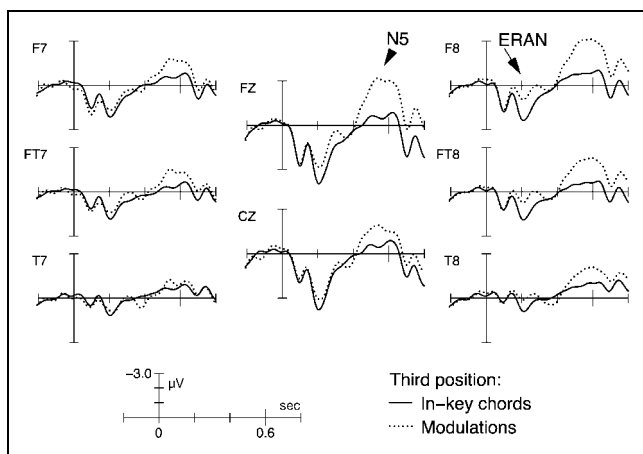

Figure 3. Grand average ERPs elicited at position 3 of the sequences, separately for modulating and in-key chords. Modulating chords elicited an early right anterior negativity (long arrow) and a late right frontal negativity (N5, short arrow).

approximately at 500 msec (the N5). The late negativity was frontally maximal (central electrode sites showed considerably smaller negative potential values) and lateralized to the right.

An ANOVA with factors sequence type (modulating vs. in-key chords, both from the third position of the chord sequences) and hemisphere, conducted for a 180- to 280-msec time interval, revealed an effect of sequence type, $F(1,21) = 11.1$, $p < .005$, and an interaction between the two factors, $F(1,21) = 12.48$, $p < .005$. Likewise, the analogous ANOVA for a late time interval (500–600 msec) showed an effect of sequence type, $F(1,21) = 24.93$, $p < .0001$, and an interaction between the two factors, $F(1,21) = 6.38$, $p < .05$.

ERPs of the last three chords of the chord sequences are shown in Figure 4. As can best be seen in the difference waves of Figure 4, the late negative ERP effect elicited by chords at the third position of the modulating sequences seems to be rather tonic, in contrast to the early negativity that is more phasic and returns to baseline around 400 msec. Thus, the slow negative potential elicited by the modulating chords at the third position might overlap in time with subsequent ERP effects elicited by the modulating chords at the fourth position. Similarly, potentials elicited by modulating chords at the fifth position overlap with potentials elicited by previous chords (especially at right anterior leads).

To discern more phasic processes from those with a more prolonged (or cumulative) time-course, EEGs were 0.5 Hz low-pass filtered before averaging the ERPs (to extract the slow potentials) in one analysis, whereas in another analysis EEGs were 0.5–10 Hz band-pass filtered before averaging (to extract more phasic processes; for a similar procedure employed for the analysis of ERPs elicited in language experiments, see, e.g., Kutas & King, 1996).
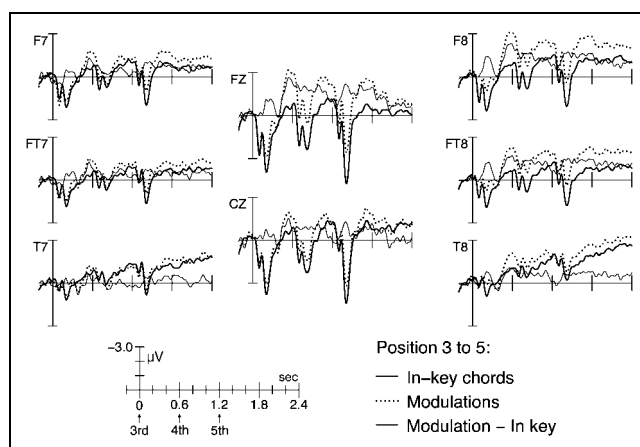

Figure 5. The 0.5-Hz low-pass-filtered EEG data: Grand average ERPs of modulating and in-key chord sequences from position 3 to 5. Thin line: Difference wave (in-key subtracted from modulating sequences). Modulations elicited a distinct tonic negativity that was maximal around 500–1500 msec, frontally predominant, and larger over the right than over the left hemisphere.

Figure 5 shows the ERPs averaged from the 0.5-Hz low-pass-filtered EEGs. Compared to nonmodulating sequences, modulating sequences elicited a tonic negative potential that was maximal around 500–1500 msec after the onset of a modulation (best to be seen in the difference wave of Figure 5, see also Figure 6). This effect was frontally predominant and stronger over the right compared to the left hemisphere. In a statistical analysis of the 0.5-Hz low-pass-filtered data, an ANOVA conducted for the time interval from 500 to 1500 msec with factors sequence type and hemisphere revealed an effect of sequence type, $F(1,21) = 10.73$, $p < .005$, and an interaction between the two factors, $F(1,21) = 5.46$, $p < .05$.

ERPs of band-pass-filtered (0.5–10 Hz) EEGs are shown in Figures 7 and 8. It appears that each modulating chord

Figure 4. Grand average ERPs of modulating and in-key chords elicited from position 3 to 5. Thin line: Difference wave (in-key subtracted from modulating chords). Onsets of the chords are indicated by the arrows in the scale. Compared to in-key chords, modulations elicited overlapping phasic and tonic negative potentials.
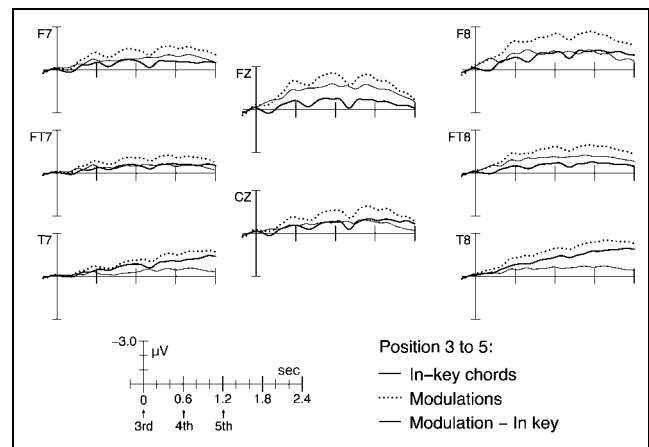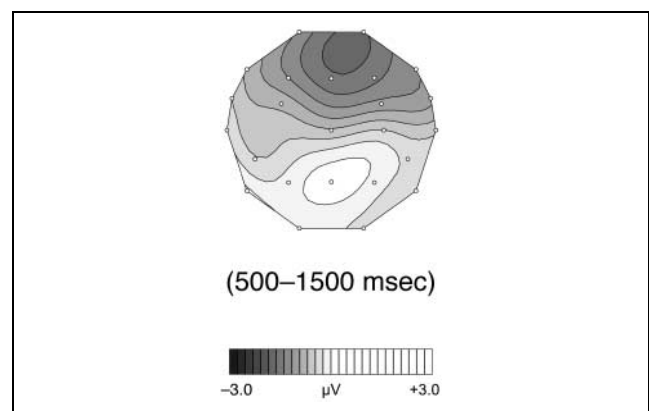

Figure 6. The 0.5-Hz low-pass-filtered data, potential map of the slow negative shift elicited by modulations (difference ERPs: in-key subtracted from modulating sequences, view from top, nose is upward), interpolated over a time window from 500 to 1500 msec (negative potentials are shaded in gray, areas of positive potentials are white).

(i.e., chords at the third, fourth, and fifth positions) elicited an early negativity with a right anterior predominance. This effect was observable around 180–280 msec after the onset of each chord and consistently over all modulating chords predominant over right anterior leads (Figure 8, note that the topographical maximum of the ERAN effect is very similar for all modulating chords). Although the amplitude of the ERAN is reduced at some electrodes (e.g., F7), other electrodes (e.g., Fz and Cz) show no or only little reduction. At some electrodes (especially at Fz and Cz), the ERAN effect returns to baseline after the presentation of each modulating chord. Thus, it appears unlikely that the negative effects shown in Figure 7 are merely part of one single slow-going negativity.

To evaluate these effects statistically, an ANOVA was conducted with factors sequence type, hemisphere, and position of the chord within the sequence (three levels: third, fourth, and fifth positions). The time windows analyzed, in respect, to the onset of the third chord, were 180–280, 780–880 (i.e., 180–280 msec after the onset of the fourth chord), and 1380–1480 msec (i.e., 180–280 msec after the onset of the fifth chord). Note that chords 4 and 5 were not averaged separately with a prestimulus baseline, because such a baseline would have been diluted by late negativities elicited by the previous modulating chord. The ANOVA indicated an effect of sequence type, $F(1,21) = 8.34$, $p < .01$, and an interaction between factors sequence type and hemisphere, $F(1,21) = 5.74$, $p < .05$, but no interaction between factors sequence type and position ($p > .80$, supporting the observation that the early right anterior negativity elicited by each modulating chord did not clearly differ in amplitude between positions 3 and 5).

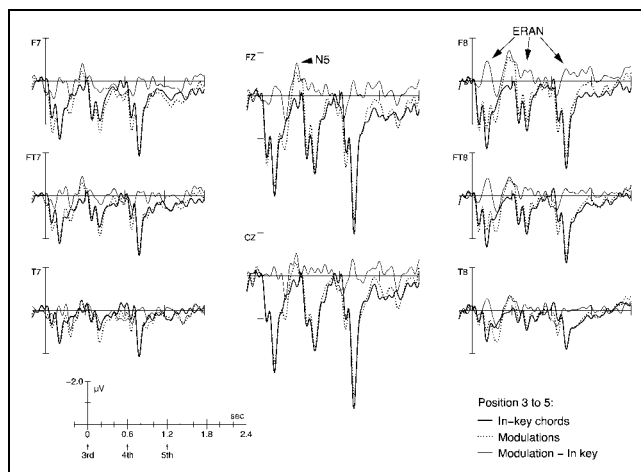In contrast to the early negativity, a clear N5 was elicited only at the third position. Analogously to the

Figure 7. The 0.5- to 10-Hz band-pass-filtered data: Grand average ERPs of modulating and in-key chord sequences from position 3 to 5. Thin line: Difference waves (in-key subtracted from modulating sequences). Each modulating chord elicited an ERAN (long arrows), an N5 was elicited at the third position only (short arrow).
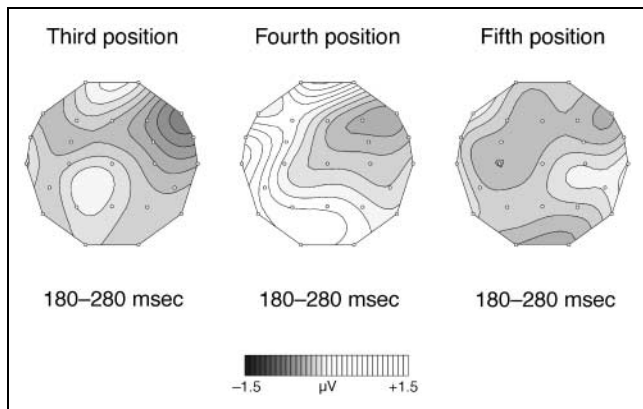

Figure 8. Early right anterior negativity: Potential maps of the early effect elicited by each modulating chord (difference ERPs: in-key subtracted from modulating chords, view from top, nose is upward). From left to right: modulating chords at the third, fourth, and fifth positions. ERPs were interpolated over the time interval from 180 to 280 msec in respect to the onset of each chord (negative potentials are shaded in gray, areas of positive potentials are white).

previous ANOVA, an ANOVA was conducted with factors sequence type, hemisphere, and position of the chord within the sequence (three levels: third, fourth, and fifth positions). Time windows analyzed, in respect to the onset of the third chord, were 500–600, 1100–1200 (i.e., 500–600 msec after the onset of the fourth chord), and 1700–1800 msec (i.e., 500–600 msec after the onset of the fifth chord). The ANOVA indicated an effect of sequence type, $F(1,21) = 7.58$, $p < .02$, and an interaction between factors sequence type and position, $F(1,21) = 8.56$, $p < .0008$, indicating that the N5 differed in amplitude between positions. ANOVAs with factors hemisphere and sequence type conducted for the three time windows separately indicated an effect of sequence type for the modulating chords at the third position, $F(1,21) = 18.66$, $p < .0005$, but no effect for modulating chords at the fourth ($p > .25$) or at the fifth ($p > .15$) position.

DISCUSSION

The brain responses elicited by modulating chord sequences distinctly differed from those elicited by in-key chord sequences. Modulating chords at the third position of a sequence elicited an ERAN (or "music-syntactic MMN") that was present around 180–280 msec. The ERAN was followed by a late frontal negativity (the N5, peaking around 500 msec). The finding of early and late negativities elicited by harmonically, or music-syntactically, inappropriate chords replicates findings of previous experiments, indicating that these effects are neither specific for modulations nor for the inappropriate harmonies employed in the previous experiments (e.g., Neapolitan sixth chords; Koelsch et al., 2001; Koelsch, Schmidt, et al., 2002; Koelsch, Schröger, et al., 2002; Maess et al., 2001; Koelsch, 2000; Koelsch, Gunter, et al., 2000).

As noted in the Introduction, the early negativity is taken to reflect a musical expectancy violation: The modulations introduced chord functions that were harmonically less closely related and functionally less appropriate and it has been shown with behavioral measures that such musical events are perceived as less expected (e.g., Koelsch, Gunter, et al., 2000; Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986; Krumhansl & Kessler, 1982). The distinction between harmonically "appropriate" and "inappropriate" refers in the present study to a processing of harmonic relations and relations of chord functions according to the regularities of major–minor tonal music. These regularities are described by music theory and have been denoted as part of a musical syntax (Koelsch, Schmidt, et al., 2002; Maess et al., 2001; Tillmann et al., 2000; Patel et al., 1998; Swain, 1997; Sloboda, 1985; Delière, 1984). With this respect, the ERAN appears to reflect music-syntactic processing.

Interestingly, the ERAN is reminiscent (although with the opposite distribution over the scalp) of the early left anterior negativity (ELAN), an ERP component reflecting the processing of syntactic relations during the processing of language (phrase structure violations, cf. Hahne & Friederici, 1999; Friederici, 1998; Friederici, Wang, Hermann, Maess, & Oertel, 2000). Both ERAN and ELAN receive major contributions from neural generators located in the inferior pars opercularis (referred as Broca's area in the left hemisphere, Maess et al., 2001; Friederici et al., 2000, the generators of the ELAN show a slight left hemispheric, those of the ERAN a slight right hemispheric weighting). Note that the early negativity was maximal around 200 msec, that is, the latency of this effect reflects that participants processed musical irregularities surprisingly fast.

The N5 is taken to reflect processes of harmonic integration. Modulating chords at the third position contained an out-of-key note (in respect to the old key) and had a chord function that was not closely related to the preceding in-key chords. Thus, modulating chords could hardly be integrated into the old, but into a new key, leading to a higher amount of musical integration required for the modulating than for the in-key chords (see also Koelsch, Gunter, et al., 2000; Krumhansl & Kessler, 1982).

This interpretation of the N5 is supported by experimental evidence from a previous ERP study (Koelsch, Gunter, et al., 2000) in which both harmonically appropriate and inappropriate chords elicited an N5 effect, although the amplitude of the N5 was considerably larger when elicited by harmonically unusual chords. This phenomenon is reminiscent of the N400 that can be elicited by both contextually appropriate and inappropriate words, although with a considerably larger amplitude when elicited by contextually unusual words (Van Petten & Kutas, 1990). Importantly, the amplitude of the N5 elicited by (harmonically appropriate) in-key chords de-

creased toward the end of the sequence (Koelsch, Gunter, et al., 2000), reflecting that the integration of an in-key chord into a preceding musical context is easier with progressing musical context buildup. Previously, it had been shown with behavioral data that a music-theoretically derived establishment of a musical context is cognitively represented in listeners and that listeners integrate each new chord into a preceding musical context; the processes of integration are accompanied by a specification of a hierarchy of harmonic stability and, thus, by the perception of an increasing stability of the musical context (Bharucha & Krumhansl, 1983; Krumhansl & Kessler, 1982). A similar amplitude reduction can be observed for the N400 elicited by semantically appropriate words across a sentence. This reduction of the N400 is interpreted as the reflection of semantic context buildup during sentence comprehension (Van Petten & Kutas, 1990).

Note that, compared to the N400, the N5 usually shows (a) a more anterior scalp distribution, (b) a longer latency, and (c) polarity inversion at mastoidal sites when nose reference is used. These differences indicate that N5 and N400 do not reflect identical cortical processes. However, the N400 may, on a more abstract level, be interpreted as a reflection of the processing of meaning information (Kutas & Federmeier, 2000). Because the N5 seems to be (at least partly) connected to the processing of meaning information in music (Koelsch, Gunter, et al., 2002), it is possible that the N5 entails processes that also contribute to the N400 (or vice versa); this issue remains to be specified.

It is unlikely that the N5 is related to processes reflected in the O wave of the contingent negative variation (CNV), or the reorienting negativity (RON, Schröger & Wolff, 1998): The O wave of the CNV is a long-latency, long-duration ERP component that usually (a) occurs in experimental settings in which motor responses are required, (b) has a frontally negative and parietally positive scalp distribution (with mastoid reference), and (c) is elicited under attend conditions. The N5, in contrast, can be elicited under preattentive listening conditions (where no responses are required, see Koelsch, Schröger, et al., 2002) and does usually not show a parietally positive scalp distribution (with mastoid reference, Koelsch, Gunter, et al., 2000). The RON (which also occurs around 500 msec after stimulus onset) is elicited when participants turn attention back to a primary task after being oriented away. Thus, the RON is thought to reflect the reorienting to task-relevant information. In contrast to the N5, the RON (a) is confined to the condition in which the deviant is task-irrelevant (the N5 is also present under attend conditions in which harmonically inappropriate chords are task-relevant, Koelsch & Friederici, in press; Koelsch, Gunter, et al., 2000) and (b) has not been reported to show a clear polarity inversion at mastoidal sites with nose reference.

Modulating sequences elicited an additional effect, namely a slow-going negativity that was maximal around 500–1500 msec and right frontally predominant. The time-course of this negativity correlates specifically with the time-course of the modulation: It declined with the establishment of the new tonic, and returned to baseline approximately with the end of the last modulating chord. Up to now, no such effect has to our knowledge been described to be elicited by music; other musical regularity violations like unexpected but nonmodulating Neapolitan chords, secondary dominants (Koelsch, 2000), or tone clusters do not elicit such an effect (Koelsch, Gunter, et al., 2000). The slow negativity elicited by modulations is therefore suggested to reflect cognitive operations characteristic for the processing of a change of key.

These cognitive operations presumably reflect integration processes in respect to the change of key, namely the restructuring of the "hierarchy of harmonic stability" (Bharucha & Krumhansl, 1983; Krumhansl & Kessler, 1982; Krumhansl, Bharucha, & Kessler, 1982): During a sequence in which harmonies belong to one key (as the in-key chord sequences of the present study), listeners establish a "hierarchy of harmonic stability" that specifies notions (a) of the key membership of chords, (b) of the central chord functions, or the "harmonic core," of this key (consisting of tonic, dominant, and subdominant), and (c) of the relations between the central chord functions and other functions of this key (for a detailed description of these principles, see Bharucha & Krumhansl, 1983). The notion of a stable harmonic hierarchy specifies expectancies for harmonically appropriate chords to follow (Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986; Krumhansl, Bharucha, & Castellano, 1982). Because modulations introduced a new key, the hierarchy of stability established by the chord sequences preceding a modulation had to be restructured, this restructuring was presumably reflected in the slow negativity.

A decrease of stability is connected to an increase in demand of resolution (Bharucha, 1984; Bharucha & Krumhansl, 1983), that is, increase of expectancy for a return to a stable harmonic structure (either in the old or new key). The restructuring of harmonic expectancies demands the resolution to a more stable harmonic hierarchy (Bharucha, 1984; Bharucha & Krumhansl, 1983; Bharucha & Stoeckig, 1986); this demand reflects what Krumhansl and Kessler (1982) termed the "strong dynamic aspect of modulations in time." Because time is involved in the restructuring of the tonal hierarchy, working memory operations are most presumably strongly entailed in this process. The slow shift found to be elicited by modulations had a frontal maximum. Thus, the notion about the relation of this potential to working memory processes is in accordance with findings that indicate the involvement of frontal brain areas in central executive processes (Kandell, Schwartz, & Jessell, 2000). This interpretation receives further support from findings that indicate interactions between temporal and frontal cortices to be entailed in working memory for the pitch of single tones (Zatorre & Samson, 1991; Zatorre, Evans, & Meyer, 1994). Moreover, slow negative potential shifts have also been observed during the processing of language and associated with working memory operations (e.g., King & Kutas, 1995, in that experiment "object vs. subject relative sentences," elicited a slow frontal negativity). In which respect the neural resources of working memory overlap between music and language processing remains to be specified. However, shared neural resources for the processing of language and music have been reported by Koelsch, Gunter et al. (2002), Maess et al. (2001), and Patel et al. (1998). Moreover, some studies found that brain structures important for the processing of language are also crucially involved in the processing of music (Liegeois-Chauvel, Peretz, Babaie, Laguitton, & Chauvel, 1998; Platel et al., 1997; Peretz, Kolinsky, Tramo, & Labrecque, 1994; Zatorre et al., 1994).

The ERPs of band-pass-filtered electroencephalogric (EEG) data (i.e., ERPs that do not contain the slow negative potential) suggest that each modulating chord (i.e., of chords at the third, the fourth, and the fifth positions of the modulating sequences) elicited an early right anterior negativity in the time interval from 180 to 280 msec after the onset of each chord (cf. Figures 7 and 8). That is, the ERAN (or "music-syntactic" MMN) was elicited by three modulating chords in a row. Note that the frequency ratios between third and fourth as well as between fourth and fifth chords were identical when comparing modulating with nonmodulating sequences, because in both modulating and nonmodulating sequences, chord functions at positions 3–5 were (in respect to the fifth chord) dominant–dominant seventh chord–tonic. That is, modulating chords from positions 3 to 5 were on average physically identical with the chords from positions 3 to 5 of the nonmodulating sequences (see Methods). Hence, the degree of "sensory dissonance," that is, the degree to which the chords shared less component tones (or overtones) with the preceding chords (Parncutt, 1989; Schmuckler, 1989), was identical for modulating and nonmodulating sequences between chords 3 and 4, as well as between chords 4 and 5.

Because the degree of "sensory dissonance" was balanced, the presence of the ERAN elicited at positions 4 and 5 indicates a complex regularity-based processing, rather than a merely sensory processing, suggesting that participants processed the chord functions with reference to (implicit) knowledge about the structural regularities of major–minor tonal music (the knowledge about these regularities most presumably being highly related to the cultural background of listeners). Notably, these ERP effects were present although the harmonic dimension of the stimulation was task-irrelevant. Results thus support the hypothesis that the ability to acquire and

apply (implicit) knowledge about complex musical regularities is a general human ability (Koelsch, Gunter, et al., 2000; Bigand & Pineau, 1997; Bigand et al., 2001, 1999; Tillmann, Bigand, & Madurell, 1998; Papousek, 1996; Sloboda et al., 1996; Sloboda, Davidson, & Howe, 1994).

The phenomenon that three deviant events in a row elicit an MMN is not to be expected for the "classical" MMN (e.g., the frequency MMN). The amplitude of the "classical" MMN is known to decrease already in response to the second of two directly succeeding deviants (cf. Giese-Davis, Miller, & Knight, 1993; Näätänen, 1992; Sams, Alho, & Näätänen, 1984), a third deviant does virtually not elicit any MMN (in the studies from Sams et al., 1984; Giese-Davis et al., 1993, all deviants were physically identical; to our knowledge, no experiment has so far been conducted with three abstract-feature deviants occurring directly succeeding in a row, but note that chords at the third and fourth positions were both physically and functionally very similar). Thus, the finding that the early right anterior negativity was elicited by three directly succeeding chords supports the hypothesis that the ERAN, or "music-syntactic" MMN, is specifically connected to the processing of highly complex rule-based musical information (rather than due to sensory memory operations, see also Koelsch, Gunter, et al., 2000), and thus to be differentiated from the physical and abstract-feature MMN.

This notion is corroborated by the findings that the physical MMN is mainly generated within or in the close vicinity of the primary auditory cortex (i.e., in supra-temporal areas, Tervaniemi et al., 2000; Alho, 1995; Alho et al., 1996, 1998; Giard, Perrin, & Pernier, 1990; Giard et al., 1995). In contrast, the ERAN has been reported to receive its main contributions from areas located in the inferior fronto-lateral cortex (inferior pars opercularis, Maess et al., 2001), probably with additional contributions from anterior supratemporal areas (planum polare, Koelsch & Friederici, in press). This difference between physical MMN and ERAN is supported by a recent fMRI study (carried out by our group) with a musical stimulus similar to that used in the present study: In that study, the processing of harmonically inappropriate chords activated inferior fronto-lateral, but not primary (or peri-primary) auditory areas. Thus, it appears that the main generators of the ERAN are located in different cortical regions than the main generators of the physical MMN.

However, it is important to note that the MMN also receives contributions from a frontal component (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000; Chao & Knight, 1997; Alho, Woods, Algazi, Knight, & Näätänen, 1994; Giard et al., 1990). This frontal MMN component presumably reflects an involuntary attention switch and seems to be generated in the dorsolateral prefrontal cortex. Patients with lesions in this region show a marked impairment in attentional control of irrelevant sensory input (Chao & Knight, 1997), as well as a reduction of the MMN, especially over the lesioned hemisphere (Alho et al., 1994). It is possible that, like the physical MMN, the ERAN partly entails processes of attention switching (because the ERAN is usually followed by subsequent processes of musical integration that presumably require increased attention); this issue remains to be specified.

Another ERP effect elicited by unexpected auditory events is the phonological mismatch negativity (PMN, Connolly & Phillips, 1994). The PMN is related to expectancies for the sound of a phoneme: The PMN can be observed when a word's initial phoneme is not expected in relation to a preceding semantic context. As the ERAN, the PMN often shows a right anterior scalp distribution (the main generators of the PMN seem to be located in the dorsolateral prefrontal cortex of the left hemisphere, Connolly, Service, D'Arcy, Kujala, & Alho, 2000). However, whereas the ERAN is linked to the processing of structure (like the ELAN), the phonological processing function of the PMN is sensitive to lexical (and thereby phonological) expectancies developed by semantic constraints. Furthermore, the latency of the ERAN is slightly shorter than that of the PMN (which usually peaks in the 250- to 280-msec region). Also note that the paradigms used to study the PMN have involved active attention on the part of the subjects; physical MMN and ERAN, in turn, are elicited even by changes in unattended auditory inputs (Koelsch et al., 2001; Koelsch, Schroger, et al., 2002; Nääatänen, 1992).

Given the similarities among physical MMN, ERAN, ELAN, and PMN, these ERP effects may reflect operations of a peri-sylvian system that can mediate auditory information processing with respect to single tones (Schröger, 1998; Näätänen, 1992), acoustic patterns (Schröger, 1994), phonemes (Näätänen et al., 1997; Connolly & Phillips, 1994), tonal music (Koelsch, Gunter, et al., 2000), and speech (Friederici, 2002).

Conclusions

In the present study, unexpected harmonies elicited both early and late negativities, the ERAN and the N5. Additionally, modulations elicited a slow-going negativity whose amplitude specifically correlated with the time-course of the modulations. This tonic negativity is taken to reflect the restructuring of the tonal hierarchy required by the change of the tonal key (presumably entailing an increased expectancy for a stable harmonic structure). These processes presumably involve working memory operations. Notably, ERPs indicate that listeners processed the modulations with reference to an exact representation of harmonic and functional relations of the major–minor tonal system. Participants did not have formal musical training, but musical expectancies are highly related to the cultural background of listeners. Thus, the implicit knowledge about musical regularities has presumably been acquired during the exposure to this type of music in everyday life (Tillmann et al., 2000).

Results thus strengthen the hypothesis that the ability to acquire knowledge about musical regularities and to process musical information according to this knowledge is a general ability of the human brain.

METHODS

Subjects

Twenty-two subjects (right-handed and with normal hearing, aged 20–30 years, mean 23.3, 11 women) participated in the experiment. All participants were "nonmusicians," that is, they had no special musical education or expertise (besides normal school education). No participant had ever had instrumental or singing lessons.

Stimuli

Stimuli were sequences of chords, each consisting of five chords. The first chord was always a tonic chord. Chords at the second position were dominant or mediant. In nonmodulating sequences, chords at the third position were dominant chords, at the fourth position dominant seventh chords, and at the fifth position tonic. In modulating sequences, dominant chords at the second position were subdominant of the new key (mediants were the supertonic of the new key, respectively), followed by a dominant chord of the new key at the third position, a dominant seventh chord of the new key at the fourth position, and the new tonic at the fifth position. Modulations thus stepped two-fifths upward (in the sense of the circle of fifths) and had a duration of three chords.[2] Note that chords at positions 3, 4, and 5 (i.e., dominant–dominant seventh chord–tonic) were on average physically identical in modulating and nonmodulating sequences. That is, chords at the third position did on average physically not differ between modulating and nonmodulating sequences, and neither did the chords at the fourth or at the fifth position. Presentation time of chords 1–4 was 600 msec and chord 5 was presented for 1200 msec. Chords and chord sequences were presented in direct succession, there was no silent interval between chords or chord sequences.

Chords were presented with different melodic outlines (e.g., beginning with the root, the third, or the fifth in the top voice). All chord sequences were composed in a way that maximally two identical notes occurred in succession in the top voice. To prevent the stimulation from becoming monotonous, both chords at the third and fourth positions were presented equiprobably either in root position or as sixth chords (i.e., with the third as base tone) or as six–four chords (i.e., with the fifth as base tone). Twenty-seven in-key chord sequences and nine modulating sequences were composed, resulting in a pool of 36 chord sequences. From this pool, 172 sequences were randomly chosen in a way that modulations occurred with a probability of 25%.

All chords had the same decay of loudness and were played under computerized control via MIDI on a synthesizer. In 10% of the in-key chord sequences, a chord at the second, third, fourth, or fifth position was played on an instrument other than piano (e.g., trumpet, celesta, strings).

Procedure

Participants were only informed about the presence of the deviant instruments, not about the modulations or their nature. An example of a chord sequence played on a piano and of a sequence in which one chord was played by a deviant instrument (organ) was presented to each participant before starting the EEG measurement. Participants were instructed to look at a fixation cross, to ignore the harmonies, and to count the deviant instruments. They were informed that they would be asked approximately every 2 min about the number of instances of deviant instruments (the question appeared on a computer screen and subjects had to report their answer by pressing a response button). The duration of an experimental session was approximately 12 min. Because of the low number of sequences with deviant instruments (which were only used to control whether participants attended to the timbre of the musical stimulus), these sequences were excluded from further data analysis. Effects of deviant instruments will not be discussed in this article; for very similar effects, see Koelsch, Gunter et al. (2000).

EEG Measurements

Measurements were performed in an acoustically and electrically shielded room. The EEG was recorded with Ag–AgCl electrodes from 26 scalp locations of the 10–20 system, referenced to the left mastoid (Pivik et al., 1993). After the measurements, EEG data were referenced offline to the algebraically mean of both mastoid electrodes to guarantee that a lateralization of effects could not be due to the placement of the reference electrode. The horizontal electrooculogram (EOG) was recorded bipolarly between electrodes situated at the outer right and outer left canthus; the vertical EOG was recorded bipolarly between electrodes situated above and below the right eye. Sampling rate was 250 Hz (30 Hz low-pass).

Data Analysis

For elimination of artifacts caused by eye movements, EEG data were rejected off-line from the raw EEG whenever the standard deviation within a gliding window of 200 msec exceeded 35 μV in the vertical and 15 μV in the horizontal EOG. For elimination of artifacts caused by drifts or movements, EEG data were rejected off-line from the raw EEG whenever the standard deviation within a gliding window of 500 msec

exceeded 20 μV at any electrode. On average, 8.1% of all trials were rejected from further data analysis. Trials were averaged off-line from the artifact-corrected raw EEGs with a 200-msec prestimulus baseline, except when averaging whole chord sequences (as presented in Figure 2), where the baseline was 0–600 msec of each sequence (i.e., the first chord was the baseline for the whole chord sequence).

To evaluate ERPs statistically, variances of ERPs were analyzed by repeated measures ANOVAs (univariate tests of hypotheses for within subjects effects). If not separately indicated, ANOVAs were conducted with factors sequence type (in-key chords vs. modulating chords) and hemisphere (left [mean of F3, FC3, F7, FT7] vs. right [mean of F4, FC4, F8, FT8] frontal electrodes). To discern fast and slow potentials, raw EEGs were filtered off-line (a) with a 0.5-Hz low-pass filter and (b) in another analysis with a 0.5- to 10-Hz band-pass filter (both filters 1001 points, FIR).

Acknowledgments

Notes

1. A sound example of the stimulation is available at www.stefan-koelsch.de.
2. According to music theory, the last chord of a modulation (i.e., the new tonic) is not denoted as a modulating chord. For the sake of simplicity, the last chord of a modulating sequence will be referred to in the present study as "modulating chord."

REFERENCES

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing, 16,* 38–51.

Alho, K., Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R., Knuutila, J., & Näätänen, R. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology, 33,* 369–375.

Alho, K., Winkler, I., Escera, C., Huotilainen, M., Virtanen, J., Jääskelïnen, I., Pekkonen, E., & Ilmoniemi, R. (1998). Processing of novel sounds and frequency changes in the human auditory cortex: Magnetoencephalographic recordings. *Psychophysiology, 35,* 211–224.

Alho, K., Woods, D., Algazi, A., Knight, R., & Näätänen, R. (1994). Lesions of frontal cortex diminish the auditory mismatch negativity. *Electroencephalography and Clinical Neurophysiology, 91,* 353–362.

Berent, I., & Perfetti, C. (1993). An on-line method in studying music parsing. *Cognition, 46,* 203–222.

Bharucha, J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology, 16,* 485–518.

Bharucha, J., & Krumhansl, C. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition, 13,* 63–102.

Bharucha, J., & Stoeckig, K. (1986). Reaction time and musical expectancy: Priming of chords. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 403–410.

Bharucha, J., & Stoeckig, K. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics, 41,* 519–524.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 184–197.

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

Bigand, E., Tillmann, B., Poulin, B., Adamo, D., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition, 81,* B11–B20.

Chao, L., & Knight, R. (1997). Human prefrontal lesions increase distractability to irrelevant sensory inputs. *NeuroReport, 6,* 1605–1610.

Connolly, J., & Phillips, N. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience, 6,* 256–266.

Connolly, J., Service, E., D'Arcy, R., Kujala, A., & Alho, K. (2000). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *NeuroReport, 12,* 237–243.

Deliège, C. (1984). *Les fondements de la musique tonale.* Paris: J. C. Lattès.

Friederici, A. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science, 6.*

Friederici, A. D. (Ed.) (1998). *Language comprehension: A biological perspective.* Berlin: Springer-Verlag.

Friederici, A. D., Wang, Y., Herrmann, C., Maess, B., & Oertel, U. (2000). Localisation of early syntactic processes in frontal and temporal cortical areas: An MEG study. *Human Brain Mapping, 11,* 1–11.

Giard, M., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience, 7,* 133–143.

Giard, M., Perrin, F., & Pernier, J. (1990). Brain generators implicated in processing of auditory stimulus deviance. a topographic ERP study. *Psychophysiology, 27,* 627–640.

Giese-Davis, J., Miller, G., & Knight, R. (1993). Memory template comparison processes in anhedonia and dysthymia. *Psychophysiology, 30,* 646–656.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11,* 194–205.

Hindemith, P. (1940). *Unterweisung im Tonsatz, 1. Theoretischer Teil.* Mainz: Schott.

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

Kandell, E., Schwartz, J., & Jessell, T. (2000). *Principles of neural science.* CT: Appleton & Lange.

King, J., & Kutas, M. (1995). Who did what and when—Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7,* 376–395.

Koelsch, S. (2000). *Brain and music—A contribution to the*

investigation of central auditory processing with a new electrophysiological approach. Leipzig: Risse.

Koelsch, S., & Friederici, A. D. (in press). Towards the neural basis of processing structure in music: Comparative results of different neurophysiological investigation methods (EEG, MEG, fMRI). *Annals of the New York Academy of Sciences.*

Koelsch, S., Gunter, T., Cramon, D. Y. V., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *Neuroimage, 17,* 956–966.

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

Koelsch, S., Gunter, T., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP-study. *NeuroReport, 12,* 1385–1389.

Koelsch, S., Maess, B., & Friederici, A. D. (2000). Musical syntax is processed in the area of Broca: An MEG study. *Neuroimage, 11,* 56.

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

Koelsch, S., Schmidt, B., & Kansok, J. (2002). Influences of musical expertise on the ERAN: An ERP-study. *Psychophysiology, 39,* 657–663.

Koelsch, S., Schröger, E., & Gunter, T. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology, 39,* 1–11.

Krumhansl, C., Bharucha, J., & Castellano, M. (1982). Key distance effects on perceived harmonic structure in music. *Perception & Psychophysics, 32,* 96–108.

Krumhansl, C., Bharucha, J., & Kessler, E. (1982). Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 24–36.

Krumhansl, C., & Kessler, E. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89,* 334–368.

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *TICS, 4,* 463–470.

Kutas, M., & King, J. (1996). The potentials for basic sentence processing: Differentiating integrative processes. In I. Ikeda & J. McClelland (Eds.), *Attention and Performance XVI* (pp. 501–546). Cambridge: MIT Press.

Liegeois-Chauvel, C., Peretz, I., Babaie, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain, 121,* 1853–1867.

Maess, B., Koelsch, S., Gunter, T., & Friederici, A. D. (2001). "Musical syntax" is processed in the area of Broca: An MEG-study. *Nature Neuroscience, 4,* 540–545.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, NJ: Erlbaum.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by magnetic brain responses. *Nature, 385,* 432–434.

Papoušek, H. (1996). Musicality in infancy research. In J. Sloboda & I. Deliege (Eds.), *Musical beginnings.* Oxford: Oxford University Press.

Parncutt, R. (1989). *Harmony: A psychoacoustical approach.* Berlin: Springer-Verlag.

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

Peretz, I., Kolinsky, R., Tramo, M., & Labrecque, R. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117,* 1283–1301.

Piston, W. (1948/1987). *Harmony.* New York: Norton.

Pivik, R., Broughton, R., Coppola, R., Davidson, R., Fox, N., & Nuwer, M. (1993). Guide-lines for the recording and quantitative analysis of electroencephalographic activity in research contexts. *Psychophysiology, 30,* 547–558.

Platel, H., Price, C., Baron, J., Wise, R., Lambert, J., Frackowiak, R., Lechevalier, B., & Eustache, F. (1997). The structural components of music perception, a functional anatomical study. *Brain, 120,* 229–243.

Rinne, T., Alho, K., Ilmoniemi, R., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage, 12,* 14–19.

Sams, M., Alho, K., & Näätänen, N. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology, 21,* 434–441.

Schenker, H. (1956). *Neue musikalische Theorien und Phantasien* (2nd ed.). Wien: J.G. Cottäsche Buchhandlung Nach folge.

Schmuckler, M. (1989). Expectations in music: Investigation of melodic and harmonic processes. *Music Perception, 7,* 109–150.

Schönberg, A. (1969). *Structural functions of harmony* (rev. ed.). New York: Norton.

Schröger, E. (1994). An event-related potential study on sensory representations of unfamiliar tonal patterns. *Psychophysiology, 31,* 175–181.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology, 34,* 245–257.

Schröger, E. (1998). Measurement and interpretation of the mismatch negativity (MMN). *Behavior Research Methods, Instruments, & Computers, 30,* 131–145.

Schröger, E., &Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *NeuroReport, 9,* 3355–3358.

Sloboda, J. (1985). *The musical mind: The cognitive psychology of music.* New York: Oxford University Press.

Sloboda, J., Davidson, J., & Howe, M. (1994). Is everyone musical? *Psychologist, 7,* 349–356.

Sloboda, J. A., Davidson, J. W., Have, M. J. A., Moore D. G. (1996). The role of practice in the development of performing mucisians. *British Journal of Psychology,* 87, 287–309.

Swain, J. (1997). *Musical languages.* UK: Norton.

Tervaniemi, M., Medvedev, S., Alho, K., Pakhomov, S., Roudas, M., v.Zuijen, T., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping, 10,* 74–79.

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

Tillmann, B., Bigand, E., & Madurell, F. (1998). Local versus global processing of harmonic cadences in the solution of musical puzzles. *Psychological Research, 61,* 157–174.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception, 16,* 99–117.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18,* 380–393.

Zatorre, R., Evans, A., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience, 14,* 1908–1919.

Zatorre, R., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain, 114,* 2403–2417.