Available online at www.sciencedirect.com

SCIENCE DIRECT®

Cognitive Brain Research 25 (2005) 678 – 687

COGNITIVE BRAIN RESEARCH

www.elsevier.com/locate/cogbrainres

Research Report

Effects of attention on the neural processing of harmonic syntax in Western music

Psyche Loui[a,b], Tineke Grent-'t-Jong[a,d], Dana Torpey[a], Marty Woldorff[a,c,*]

[a]Center for Cognitive Neuroscience, Duke University, Durham, NC 27708, USA
[b]Department of Psychology and Center for New Music and Audio Technologies, University of California at Berkeley, Berkeley, CA 94709, USA
[c]Department of Psychiatry, Duke University, Durham, NC 27708, USA
[d]Department of Psychopharmacology, Utrecht University, 3584 CA Utrecht, The Netherlands

Accepted 31 August 2005
Available online 28 October 2005

Abstract

The effects of selective attention on the neural response to the violation of musical syntax were investigated in the present study. Musical chord progressions were played to nonmusicians while Event-Related Potentials (ERPs) were recorded. The five-chord progressions included 61% harmonically expected cadences (I–I^6–IV–V–I), 26% harmonically unexpected cadences (I–I^6–IV–V–N^6), and 13% with one of the five chords having an intensity fadeout across its duration. During the attended condition, subjects responded by pressing a button upon detecting a fadeout in volume; during the unattended condition, subjects were given reading comprehension materials and instructed to ignore all auditory stimuli. In response to the harmonic deviant, an Early Anterior Negativity (EAN) was observed at 150–300 ms in both attention conditions, but it was much larger in amplitude in the attended condition. A second scalp-negative deflection was also identified at 380–600 ms following the harmonic deviants; this Late Negativity onset earlier during the attended condition. These results suggest strong effects of attention on the neural processing of harmonic syntax.

Theme: Neural basis of behavior
Topic: Cognition

Keywords: Selective attention; Music perception; Harmonic expectation; Auditory processing; Event-related potential

1. Introduction

Music we encounter every day is composed according to rules of tonality and harmony. Following repeated cultural exposure, the human brain is thought to implicitly form expectations in accordance with traditional musical styles [13,14,26]. In recent years cognitive neuroscience researchers have explored the neural underpinnings of musical expectation, especially through the use of event-related potentials (ERPs) [1,7–12].

1.1. Music theory and harmonic expectation

The theoretical musical context of these studies relates to a rich set of rules and principles of harmonic music theory. Traditional Western music is generally composed of sequentially and simultaneously occurring pitches organized around a major or minor key, where the key typically serves as the reference and endpoint of a musical piece. Sequentially presented pitches form the melody in music, whereas simultaneously presented pitches give rise to harmony. A chord is comprised of three or more simultaneously occurring pitches, and the root of a chord is the reference pitch within the chord. Chord progressions, or series of chords, illustrate the basic grammatical structures upon which musical harmony

* Corresponding author. Center for Cognitive Neuroscience, Duke University, Box 90999, Durham, NC 27708, USA. Fax: +1 919 681 0815.
E-mail address: woldorff@duke.edu (M. Woldorff).
URL: http://www.mind.duke.edu/woldorfflab (M. Woldorff).

0926-6410/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2005.08.019

is built. A prototypical chord progression begins and ends with a *tonic* chord, which has a root that is the same pitch as the name of the key—for instance, for a piece in the key of C major, the tonic chord has a root of C and is comprised of the pitches C, E, and G. The tonic chord of a progression is notated with the Roman numeral "I". The penultimate chord of the prototypical progression is the *dominant* chord, which has a root that is five increments away from tonic and is thus notated "V". Immediately preceding the dominant is usually the *predominant*, which can be one of several chords, typically a chord with its root four increments away from the root of the tonic (notated "IV"). Between the first tonic and the predominant, composers usually insert other chords to prolong the sense of the initial tonic. One chord frequently used here is a special case of the tonic chord—the *tonic in first inversion* ("I^6"), which is a tonic chord with a different configuration of pitches from the typical tonic chord. The five-chord progression of *tonic–tonic in first inversion–predominant–dominant–tonic* (i.e., I–I^6–IV–V–I) provides a phrase structure that tends to be a prototypical basis of harmonic syntax in Western music (e.g., [24]).

As the above phrase structure has become an accepted norm in music theory over the centuries, vast numbers of musical phrases have been composed according to this underlying structural rule. The chord progression I–I^6–IV–V–I thus reflects a structural rule which can be viewed as an instance of syntax in musical grammar [1]. Following repeated exposure to traditional Western music, most individuals have come to expect the prototypical musical (or harmonic) syntax, and listening to a musical phrase that deviates from this syntactical rule leads to a violation of expectations. These expectations and their violations have been studied from theoretical and empirical perspectives [2,7,17] and have been implicated as the source of emotion and meaning in music [17].

In addition to the usual chords, composers occasionally substitute other chords, such as the Neapolitan sixth, in the predominant position. The Neapolitan sixth (N^6) chord is a major chord where the root pitch is one half of an increment above the tonic pitch [24]; it is derived from the minor scale and is typically used in the first inversion in place of predominant chords, especially in minor keys [24]. Thus, in the key of C, it would be a combination of the pitches F, A-flat, and D-flat (half increment from C). The N^6 chord is conventionally placed in the predominant position; when it is placed in a different position, such as in place of the final tonic, the typical or expected musical syntax is violated. The N^6 chord is not dissonant by itself, but when the context of a particular key has been established, the N^6 chord sounds out of place at the end of the sequence and thus violates harmonic expectation [7,9,11]. Such a violation of harmonic expectation has been examined in a few recent ERP studies.

1.2. ERP studies of musical expectation

An early study involving ERPs and musical expectancy was conducted in order to compare musically-evoked neural activity in musicians and nonmusicians [2]. In an attempt to find an electrophysiological marker of musical expectancy, melodies with congruous (expected) terminal notes were contrasted against melodies with incongruous endings that either obeyed or violated the theoretical norms of melodic expectation. A late positive component (LPC) peaking at around 600 ms over parietal sites was evoked by incongruous endings, especially those that violated harmonic rules. This effect was significant in both musicians and nonmusicians, but was more pronounced in musicians. Based on these results, the LPC was thought to reflect a decision process based on implicit musical knowledge [2].

Results from the above study, while being the first ERP data to address musical expectation, investigated the neural response to deviants that violated both melodic and harmonic expectation. From a somewhat different standpoint, ERPs were used in the study of musical expectancy with respect to the violation of harmonic contexts [7]. Participants in this study listened to chord progressions and decided whether they resolved correctly. Results from this study provided evidence for an earlier, centrally distributed P3a component (peaking at approximately 350 ms) that was separate from the later, parietal, and much larger P3b (450 ms) component within the P300 complex. It was proposed that the P3a component was elicited in response to unexpected musical-syntactic events, whereas the P3b component was produced by a shift in voluntary attention in response to a target. In a study testing the language-like specificity of the syntax-related P600 component, a positive component peaking at 600 ms was elicited in response to an out-of-key chord within a harmonic phrase; in addition, a positive component peaking at approximately 300 ms was again elicited, the amplitude of which correlated with the degree of harmonic violation in musical melodies [23].

This positive component at 300 ms was also reported by Koelsch et al., who presented subjects with sequences of various chord progressions [9]. While most of the progressions followed the (I–I^6–IV–V–I) syntax, some of the progressions contained the Neapolitan sixth (N^6) chord: 25% of the trials included the N^6 in the syntactically proper third position (replacing the IV from the prototypical progression), whereas another 25% of the trials included the N^6 in a harmonically improper and unexpected fifth position, replacing the final tonic chord. Relative to the IV and I chords, the Neapolitans in both third and fifth positions elicited a response termed the Early Right Anterior Negativity (ERAN) at around 150–250 ms, the aforementioned P3a component peaking at 350 ms, and a late frontal negative wave (N5) component with an onset at around 380 ms and a peak at 550 ms. The ERAN was thought to reflect neural activity in response to the initial detection of the violation of harmonic expectancy, and was unaffected by task relevance of the unexpected chords. In contrast, the subsequent N5, occurring later at a cognitive rather than sensory level, was proposed to reflect processes of harmonic integration [9,11]. The ERAN, P3a, and N5 were all elicited by N^6 chords, but all these waveforms were smaller in amplitude when the N^6 chord was in the third rather than the fifth position, a result suggesting that the N^6 chord was less unexpected when in third position, in accordance with its predominant harmonic function.

An important aspect to processing activations such as the ERAN and N5 is their degree of automaticity. To investigate whether the ERAN and the N5 could be elicited preattentively, another study manipulated the allocation of attentional resources during harmonic syntax processing [11]. The experiment involved attended and unattended conditions. In the unattended block, subjects read a self-selected book under the instruction to ignore all acoustic stimuli. In the attended block, subjects were asked to detect and respond with a button press to N^6 chords in the third and fifth positions. An early right anterior negativity (ERAN) during 150–210 ms and an N5 (380–600 ms) was found in response to N^6 chords in both the third and fifth positions. In comparing attended and unattended blocks, the authors report that "The amplitude of the ERAN does not significantly differ between both blocks, suggesting that the processes underlying the generation of the ERAN are fairly independent of attention." ([11], p. 45) The N5 in the attended condition was not observed, but it appears that it may have been masked by the large target-related P3b wave in response to the target N^6 chord. [11].

1.3. Other ERP studies on attention

Previously, neural processes underlying other sensory and cognitive features of auditory stimuli were shown to be modulated by attention. One waveform that seems to be particularly comparable to the ERAN is the Mismatch Negativity (MMN), a negative component elicited by a deviant auditory stimulus (deviant in a basic auditory sensory feature such as pitch or intensity) in an ongoing stimulus train [9,20,25]. Prior studies have shown that the MMN, like the ERAN, is automatically elicited [10], but relatively larger in amplitude in attended channels [28,30,32]. In addition, the unattended MMN was also shown to be larger in musicians than nonmusicians [10]. As the MMN has an onset time approximately similar to that of the ERAN, and is sensitive to attentional modulation, it would seem likely that the ERAN could also be modulated by attention factors.

Other more cognitive waveform components that are comparable to the N5 are also modulated by attention. One such component is the N400, which is typically elicited by semantic incongruity for words [15]. The N400 was also found to be strongly modulated in amplitude when in an attended vs. an unattended channel [16]. As the N400 and the N5 are evoked by semantic violations in language and music respectively [9,15], and are elicited at similar latencies, these two waveforms could reflect comparable mental processes, and thus might also be similarly modulated by selective attention.

In the above paradigm, the unattended condition involved subjects reading self-selected material, and unlike most ERP studies, in which trials with eye movements are rejected as containing artifacts (e.g., [1,5,20,29]), only EEG trials that actually contained horizontal eye movements were included in the data analysis, as an indicator that the subjects were actually reading during EEG recording. To account for the increased variance caused by eye movements, more trials were included in the unattended than in the attended condition. While reading had been previously reported as an ignore-auditory task (e.g., [20,25]), there had been few reports of specifically including only trials containing eye movements in EEG data analysis. As eye movements were the only indicator of reading in the original study, it may not be the case that subjects were truly devoting maximal attentional resources to the reading task; subjects could have been scanning the material without truly paying attention. Moreover, although horizontal eye movements during reading are unavoidable, they add non-neural physiological noise into the EEG and ERPs. In order to ensure attentiveness to reading during the unattended condition, the present study employed controlled reading passages and post-run tests of comprehension. Additionally, in order to minimize ocular artifacts from the horizontal eye movements during reading, the eye movements were kept at a minimum by setting the reading material at a substantial distance from the eyes.

During the attended block, subjects in the Koelsch et al. study were instructed to listen for and respond to the N^6 chords; this implies that instead of comparing an attended condition to an unattended one, the study was in fact comparing an attended and *task-relevant-target* condition to an unattended condition. To test for the effects of attention on the processing of harmonic deviants, separately from the effects of those deviants being targets or not, the present study required subjects to listen carefully to the music, but to detect and respond to a feature unrelated to harmony. From the data presented by the previous study it was evident that the task relevance of the N^6 chord led to a very large target-related P3b wave, which seems likely to have masked or otherwise distorted the N5 in the attended (and task-relevant) condition, and possibly also affected the ERAN. By devising another task that ensures continuous attention to the auditory stimuli, but avoids any task relevance of the N^6 chord, the masking effect of the P3b could be avoided.

1.4. Purpose of the present study

The purpose of the present study was to investigate the effects of attention on brain activity associated with violation of harmonic expectancy, using a manipulation of attention in which sequences with the deviant N^6 chords were either attended or ignored, but their deviancy was neither target-defining nor task-relevant. Thus, in the attended condition here, subjects were instructed to detect an occasional decrease in sound intensity of any of the sounds. This use of an orthogonal task was similar to an approach used in some prior studies of music processing [e.g., 9]. In the unattended condition, subjects studied reading comprehension passages in preparation for answering questions concerning these passages. Moreover, their answers to these questions were statistically analyzed, thereby providing behavioral data confirming that the subjects were attending away from the chord progressions during this condition.

2. Materials and methods

2.1. Subjects

Eighteen nonmusicians (10 females, 8 males, mean age 23.3 years, age range 18–49, standard deviation 6.9) participated in this study. No participant had received more than 2 years of education in any instrument or voice outside of normal school education. All subjects were right handed and reported having normal hearing, normal or corrected-to-normal vision, and no history of neurological or psychiatric disorder. All subjects were volunteers recruited by email from the Duke University community; each subject gave written informed consent prior to the experiment, and was paid $10 per hour for their participation. Subjects were excluded from the analysis if more than a third of their trials contained physiological artifacts (such as eye blinks, erratic eye movements, and excessive muscle activity), or if more than half of their answers to the reading-comprehension questions were incorrect. According to these criteria, 5 of the 18 subjects were excluded, leaving 13 in the final analysis.

2.2. Apparatus and stimuli

Subjects were presented with sequences of auditory stimuli. All stimuli were generated electronically on a PC using CoolEdit Pro2.0. The auditory stimuli were series of five-chord progressions, with four chords of 600 ms each followed by a fifth chord of 1200 ms. In order to minimize distortions in overlapping ERPs resulting from short inter-stimulus intervals, especially in the alpha frequency range (see [27]), the inter-stimulus intervals (ISIs) between each five-chord sequence were randomly varied between 0 and 200 ms. These stimuli were presented using a Dell PC with the subject seated 58 cm from a screen at eye level, and 78 cm away from the left and right speakers. The key of each chord progression was randomly varied among eight keys; each chord progression belonged to either the key of C, G, D, or A major or minor. Of the total of 800 trials divided into 8 runs, 61% (488 sequences) were of the standard musical syntax of tonic–tonic in first inversion–subdominant–dominant–tonic (I–I^6–IV–V–I). Among the remaining 39% of chord sequences, two-thirds (208 sequences) were of the improper (i.e. unexpected) musical syntax of tonic–tonic in first inversion–subdominant–dominant–Neapolitan sixth (I–I^6–IV–V–N^6). The remaining one-third (104 chord sequences) were in standard musical syntax, but in one of the five chords in each progression there was a fadeout in sound intensity across the duration of the chord (the intensity-decrement targets) (Fig. 1).

2.3. Procedure

The experiment consisted of 800 trials, divided into eight runs of six min each, with alternating runs of attended and unattended conditions. Half the subjects began the experiment with the attended condition and the other half with the unattended. In the attended condition, subjects were instructed to listen carefully to the chord sequences in order to detect the occasional target chords (those with the intensity fade-outs) by pressing a button as quickly as possible. Since it was anticipated that some subjects would perform better than others, three sets of trials of varying levels of target-discrimination difficulty were made, with the fadeout chords being the most obvious in the easy task and the least obvious in the difficult task. Individual pilot tests were run before EEG was recorded to ensure that the task

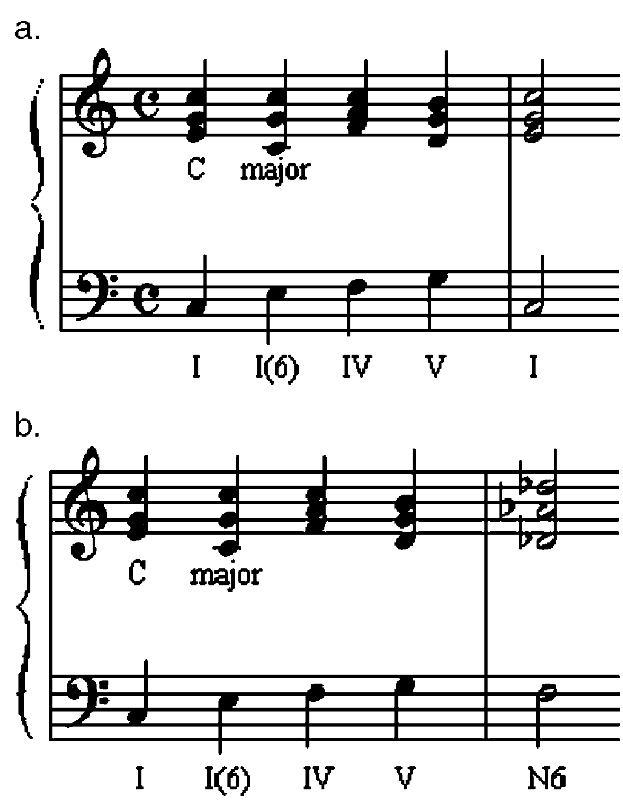

Fig. 1. Examples of stimuli. (a) Standard; (b) Deviant fifth.

was sufficiently attention-demanding, but that the subject could perform the task. During the unattended-condition runs, subjects were given practice reading comprehension passages from the Graduate Record Exam [3] to read, with instructions to ignore all auditory events. To minimize the range of eye movements during reading, these passages were placed 58 cm away from subjects' eye level. Between experimental runs, subjects were asked questions on the reading material, and their answers recorded, to ensure that the subject was attending to the reading.

2.4. EEG recording

EEGs were recorded from 64-channel tin electrodes mounted in a custom-designed elastic cap (Electro-Cap International, Inc.) and referenced to the right mastoid during recording. This custom-made cap covered the whole head from above the eyebrows back to lower posterior scalp over the inferior occipital lobule, with electrodes equally spaced across the cap [31,33]. Because of this special distribution of electrodes, only some electrode-positions completely matched the International 10–20 system locations (e.g., Cz). Horizontal eye movements were measured by electrodes at the outer canthi of the eyes, and vertical eye movements and blinks were measured by electrodes placed below the eyes. In addition, a video zoom lens camera was used to monitor eye movements during recording. Electrode impedances were kept below 2 kΩ for the mastoids and the ground, below 10 kΩ for the eye channels, and below 5 kΩ for all remaining electrodes. All channels were continuously recorded with a bandpass filter of 0.01–100 Hz, and a gain of 1000 (Synamps amplifiers from Neuroscan, Inc.). The raw signal was digitized with a sampling rate of 500 Hz. Recordings took place in an electrically shielded, sound-attenuated chamber, and the experiment was conducted in moderately dim light.

2.5. Data analysis

A single factor ANOVA was conducted on behavioral responses to the intensity-decrement target chords in the attended condition, with the average hit rate for each subject as a factor relative to a chance level of 2.6% (13 targets over 500 chords in each run). ERPs were averaged separately for tonic and Neapolitan chords at the fifth position, over a time window of 400 ms prestimulus to 1000 ms poststimulus, with a 200-ms prestimulus baselining period. Artifact rejection was applied offline by discarding EEG epochs that were contaminated by eye movements, eye blinks, excessive muscle-related potentials, drifts, or amplifier blocking. Trials with horizontal eye movements were rejected as artifacts in the attended condition but were included in the unattended, as the task of reading during the unattended condition necessitated some horizontal eye movements. The raw EEG data were digitally low-pass filtered with a non-causal, zero-phase running average filter of 9 points, which strongly reduces frequencies at and above approximately 57 Hz at the sampling frequency of 500 Hz. ERPs were grand-averaged across the 13 acceptable subjects, and statistical analyses across subjects were performed on mean amplitude measures across specific latency windows, relative to the 200-ms prestimulus baseline. Repeated-Measure Analyses of Variance (ANOVAs) in a within-subject design were conducted at individual electrode channels over the most activated sites: left anterior (Fp1m[1], F3a, FC1, C1a), right anterior (Fp2m, F4a, FC2, C2a), and midline anterior (Fzp, Cza). ANOVAs were conducted with factors of chord type (standard tonic chords versus Neapolitan sixth chords), attention condition (attended versus unattended), and hemisphere (corresponding left versus right electrode sites).

3. Results

3.1. Behavioral data

For the unattended condition, the reading comprehension questions were multiple-choice with five choices; therefore, chance level was 20%. All but one of the subjects performed above 50% correct in answering the reading comprehension questions. The high level of performance among these subjects was especially significant as the national average of performance on such reading comprehension passages is below 50% [3]. To ensure that subjects were attending closely to the reading (and ignoring the auditory stimuli), only participants who responded correctly to more than half of the multiple-choice reading comprehension questions were included. Based on this criterion, the one subject scoring below 50% was dropped from the final data analysis.

Each of the four runs in the auditory attended condition contained 13 auditory intensity-decrement target trials. The average hit rate was 80.3% (10.4 out of 13 targets, SD = 1.50). A button press was recorded as a hit if it occurred within a latency window of 1000 ms after the onset of the target stimulus. Given this constraint, the average reaction time was 729 ms (SD = 51.7 ms) after the onset of the intensity fadeout target chords. On average, subjects had a mean false alarm rate of 0.75% of the non-target chords.

[1] Due to the use of a custom cap with equidistantly configured electrode locations, channel-labels differed slightly from the standard 10–20 system. The closest standard 10–20 site is given, with a suffix indicating additional localization information. More specifically, channel labels followed by an "a" are located slightly (less than 1 cm) more anterior than the noted standard position, a "p" indicates a slightly more posterior position, an "i" a more inferior position, an "s" a more superior position, and an "m" a more medial position. Finally, a quote (') with no letter indicates approximately equivalent (within 5 mm) to the noted 10–10/20 location.

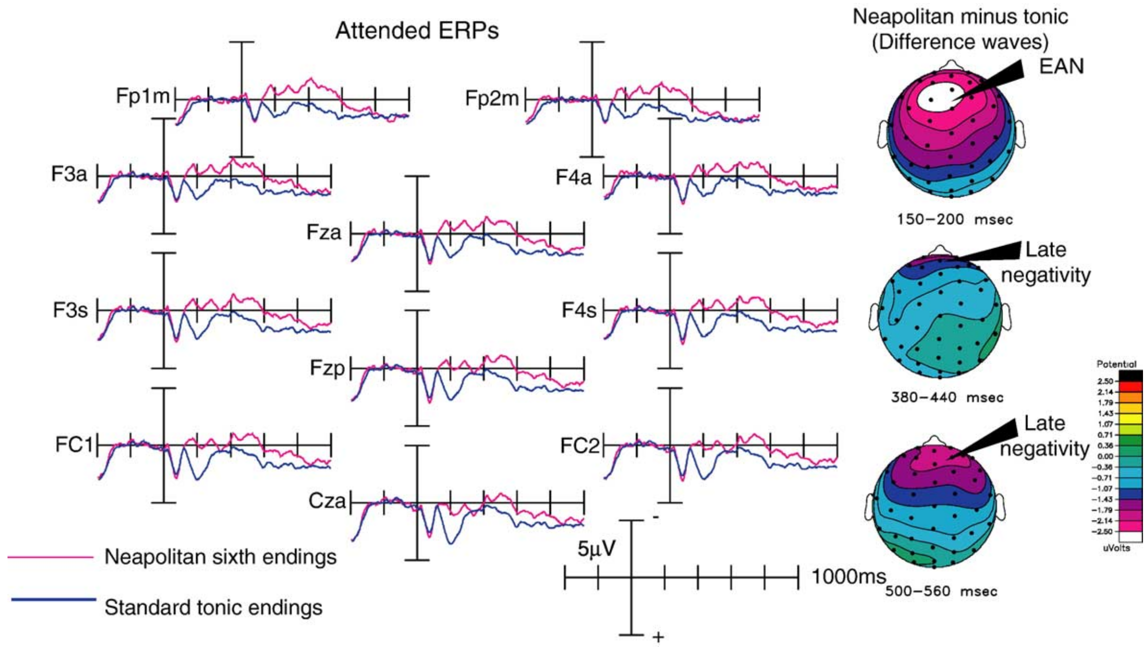

Fig. 2. Attended ERPs: deviant N6 endings versus standard tonic endings in the attended condition. Grand-averaged data from 13 subjects.

3.2. ERP data

Relative to the expected tonic chords, the deviant N^6 chords elicited an early anterior negativity that onset at around 150 ms and was larger in the attended condition (Figs. 2 and 3). This component was maximal between 150 and 300 ms after stimulus onset, with the peak at ~180 ms (see difference waves in Fig. 4). This early frontal negativity did not differ between conditions in latency, but its amplitude was much larger in the attended than in the unattended condition (see Figs. 2 and 3 and difference waves in Fig. 4). This was statistically confirmed by a one-way analysis of variance, which revealed highly significant main effects of chord type (Neapolitan vs. tonic, $F(1,12) = 31.1$, $P < 0.0001$) and a significant interaction of chord type and

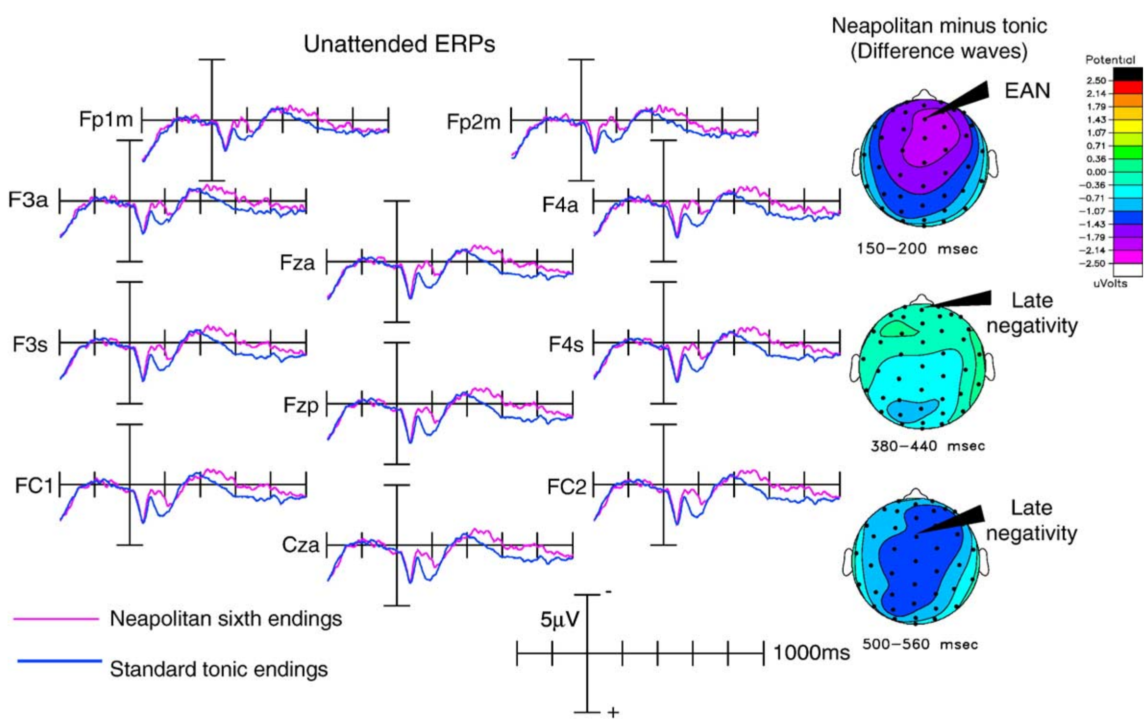

Fig. 3. Unattended ERPs: deviant N6 endings versus standard tonic endings in the unattended condition. Grand-averaged data from 13 subjects.

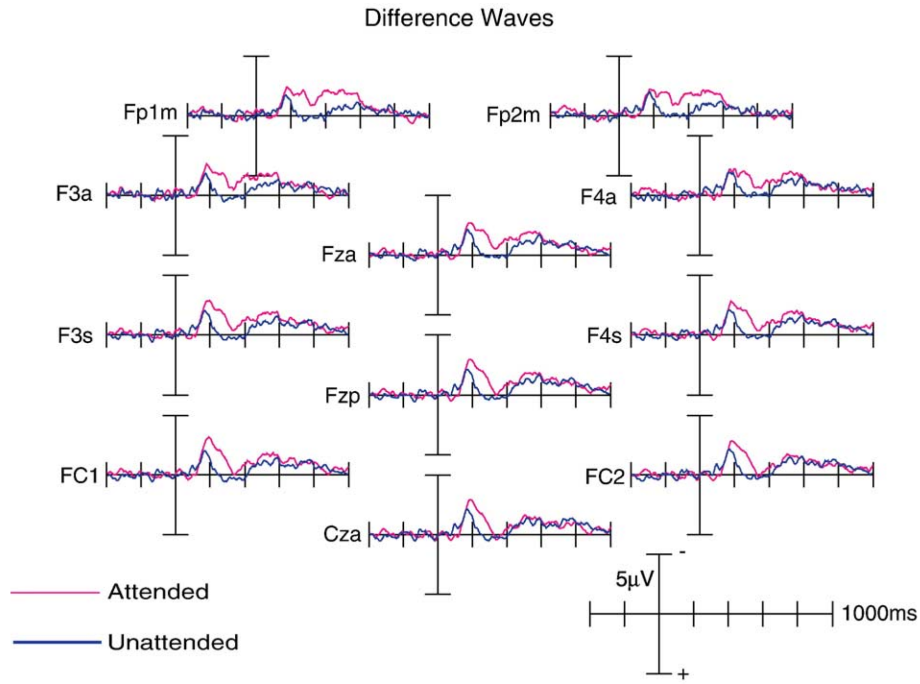

Fig. 4. Difference waves: deviant N^6 chords minus standard tonic chords in the attended versus unattended conditions. Deviance-related negativities are significantly larger and/or earlier when attended. The effect is frontal and bilaterally symmetrical.

attention condition ($F(1,12)$ = 6.72, $P<0.03$) in repeated measurement ANOVAs on anterior channels between 150 and 250 ms. Although the amplitude and latency of the frontal negativity appear to correspond to the ERAN reported in prior studies [9,11], this study revealed no significant effect of hemisphere ($F(1,12)=0.51$, P = 0.49), which differs from the right-anterior dominant results in some previous studies. Thus, this component will be referred to as the EAN (Early Anterior Negativity) for the remainder of this article.

A later negative deflection was identified in the ERPs evoked by the deviant Neapolitan in both the attended and unattended conditions. This component was observed from around 380 to 600 ms, with a peak at ~550 ms, appearing to correspond to the N5 component reported in previous studies [9,11]. The greater negativity for Neapolitan sixth chords in comparison to standard chords was confirmed by a significant main effect of chord type ($F(1,12)$ = 24.69, $P <$ 0.01) in an ANOVA over the time window of 400–600 ms poststimulus. This Late Negativity appeared to be larger and

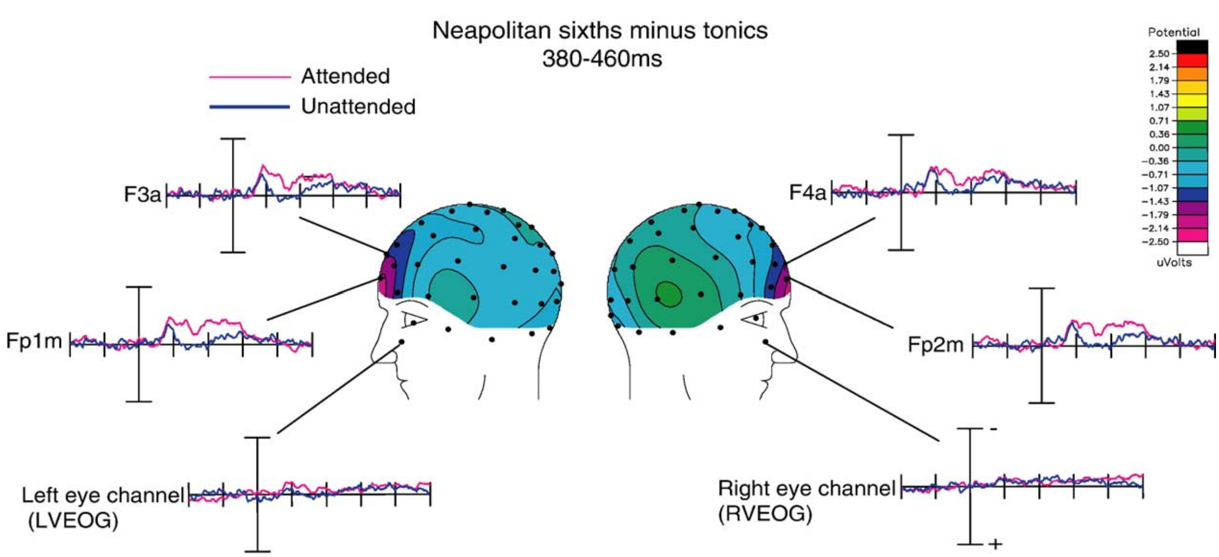

Fig. 5. Difference waves at prefrontal sites and eye electrodes. Note that the late frontal negativity effect (Late Negativity) is seen clearly at the prefrontal channels Fp1m, Fp2m, F3a, and F4a and do not invert at the left and right EOG channels located below the eyes, indicating that the Late Negativity is neural in origin and not an eye artifact.

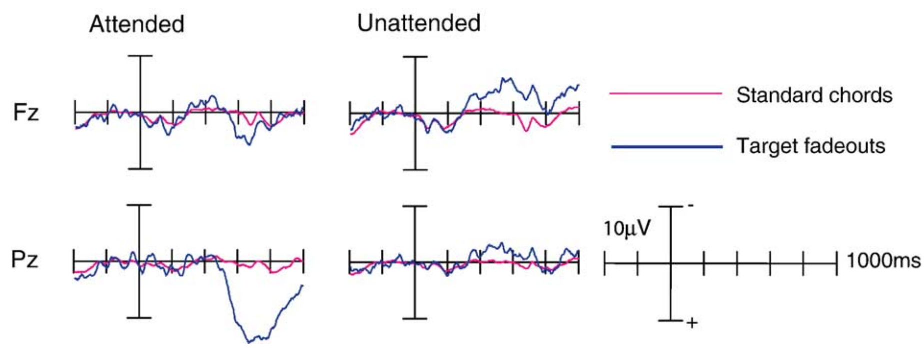

Fig. 6. Frontal and parietal ERPs to target fadeout chords. A large parietal P3b wave is shown in the attended condition, and a mismatch negativity is in the unattended condition. Note the larger scale of this figure compared to previous ERP figures. The P3b would have masked any Late Negativity effect if the Neapolitan sixth endings had been task relevant.

to onset earlier in the attended condition. The effect of hemisphere was not significant, but showed a left-dominant trend ($F(1,12) = 4.12$, $P = 0.065$), in contrast to bilateral and right-dominant findings in some prior studies [9,11]. An ANOVA over the time window of 380–450 ms revealed a significant interaction between condition and chord type ($F(1,12) = 5.06$, $P < 0.05$), with the waveform being larger in the attended condition. The interaction at this early epoch of the Late Negativity suggests that the onset of the Late Negativity is earlier in the attended condition than in the unattended condition. It is also important to note that the later part of the waveform was observed over both frontal and prefrontal channels, whereas the earliest part was largest mainly over the most anterior prefrontal sites Fp1, Fp2, F3a, and F4a. A comparison of this waveform to the eye electrodes confirms that these results cannot be explained by eye artifacts and is a robust effect detected maximally over the anterior prefrontal cortex (Fig. 5).

For the purposes of comparison, ERPs elicited by task-relevant fadeout targets are shown in Fig. 6, showing the large positive P3b with onset around 450 ms. This target-related component is largest parietally, in accordance to previous studies (see [6] for a review). The P3b to the fadeout chords is only elicited in the attended condition, when these stimuli were relevant targets; when unattended, only a frontal negativity is evoked by the fadeout chords. This frontal negativity is presumably a form of mismatch negativity, as the volume decrement compared to other chords constitutes a stimulus mismatch [18,19,22,29]. The onset of this MMN appears to be somewhat later here than in the typical oddball paradigm, presumably due to the relatively gradual temporal envelope of the intensity fadeouts (~100 ms). The amplitudes of these waveforms (P3b and MMN) in response to the fadeouts are larger than the EAN and Late Negativity in both the attended and unattended conditions. This indicates that if the harmonic deviant chords had been the task-relevant targets in this experiment, the EAN and the Late Negativity would have been masked or otherwise heavily distorted by the P3b and other activity related to target processing that have no specificity for music processing.

4. Discussion

From previous research, the prominent waveforms identified as reflecting violation of musical expectancy were the Early Right Anterior Negativity (ERAN), Late Positive Complex (LPC), and N5 [2,7,9,11,21,23]. In the present study, we observed an Early Anterior Negativity (EAN), similar to the ERAN and a Late Negativity similar to the N5.

The EAN was evoked in response to the unexpected Neapolitan chord in the latency window between 150 and 300 ms (peaking at around 180 ms) over frontal sites; this corresponds in latency and prefrontal distribution with previous findings of an ERAN maximal at 150–250 ms [9,11], although the previously observed right hemispheric dominance was not observed in the present study. The EAN component observed here was significantly larger in the attended condition than the unattended condition; this contrasts with previous reports of no interaction between attention and the ERAN [11]. As Figs. 2 and 3 show, the effect of attention was especially pronounced in the EAN at 150–300 ms in the present study. Several possible explanations can account for these divergent results. First, the difference between the attended and unattended condition may be bigger in our experiment as a result of a more demanding cognitive task in the unattended condition, as the reading comprehension task we used here may have more fully directed cognitive resources away from the auditory stimuli. Second, our subjects' task of detecting a dropoff in loudness during the attended condition ensured that subjects would need to pay attention not only to the beginning of each chord, but uniformly to the entire duration of all auditory stimuli, thus enhancing any attentional effects evoked by the auditory stream.

More importantly, the attentional effect in the Koelsch et al. (2002) study was probably influenced by the elicitation of a large P3b wave in the attended condition for the harmonically deviant targets. The P3b wave is generally associated with task relevance and target detection, and is not specific for the violation of musical expectation. By eliminating the task relevance of the N^6 chord, the present study specifically tested for the attentional modulation of music syntactically related components, instead of target-related ones.

A large P3b indicating task relevance was elicited in the present study in response to the target fadeout chords as shown in Fig. 6. This component is large, parietally centered, and maximal around 600 ms poststimulus onset. The delayed onset of this component was probably due to the physical feature of the task-relevant targets: the chords did not drop off in intensity until later in the chord duration, and the target was no different from standard chords at its moment of onset. Our behavioral data, with an average hit rate of ~80% and false alarm rate of 0.75%, suggests that while the task in the attended condition was not overly demanding, it did require considerable recruitment of attentional resources on the part of the subjects.

In the present study, the Late Negativity was defined as the significant effect of chord type in the 380–600 ms time window. This effect corresponds in latency to the N5, which has been proposed as the musically specific index of an attempt to integrate an unexpected chord into its harmonic context [9,11]. The amplitude of the Late Negativity in the present study was larger in the attended condition, with the interaction of chord type and attention condition being statistically significant at 380–450 ms. This interaction pattern suggests the possibility that the onset of the Late Negativity was earlier in the attended condition. However, another possibility to consider is that the early and later portions of the Late Negativity may be comprised of somewhat different sets of neural generators, with the earlier ones being more sensitive to attentional resource allocation. Lastly, considering the prefrontal distribution of these effects, it is important to note that the data pattern indicates that it is very unlikely that this component was derived from any eye movements or other ocular artifacts (see Fig. 5).

Previous studies have associated the Early Right Anterior Negativity with the violation of harmonic expectancy, and the slower and later negativity to a cognitive attempt at harmonic integration [9,11]. This was taken as a musical analog to the Early Left Anterior Negativity (ELAN), an index of syntactic deviance in language comprehension [5]. Although an anterior negative component was identified at a similar latency in the present study, the EAN was strongly bilateral and the Late Negativity was slightly left-dominant in our data, in contrast to the right-anterior topography reported by several studies on musical syntactic expectation. The observation of a bilateral EAN is comparable to results from a prior study investigating harmonic processing in children, which showed a gender-dependent effect in which the EAN was left-dominant in boys and bilateral in girls [12]. Although we cannot draw conclusions regarding hemispheric localization from this study, the present data calls to question the clear-cut model of language processing in the left hemisphere versus music processing in the right hemisphere. Future studies can be conducted to further explore the role of nonlinguistic syntax processing in the left hemisphere.

Lastly, it is important to note that, although the deviance-related responses in the present study (i.e., the EAN and the Late Negativity) were larger and/or earlier in the attended than in the unattended condition, some of this activity was significantly present in the unattended condition as well. This pattern of effects would thus be similar to that previously reported for the mismatch negativity (MMN) for auditory sensory mismatches, which has been shown to be larger within an attended channel, but still present in an unattended one [30,32]. Thus, this deviance-related activity would fit into the category of being automatically elicited (i.e., attention not being required for its elicitation), but capable of being modulated by attention [4,29].

In summary, the results of the present study provide evidence that the nonmusician brain can detect deviations in harmonic context even when the stimuli are not attended, but that attention can strongly enhance the neural activity related to the processing and detection of these deviations. Effects of attention on both early and late ERP components suggest that both the earlier sensory and later cognitive processing of music can be modulated by the top-down allocation of resources. These findings provide evidence that endogenous resource allocation influences neural processing of complex stimuli such as musical harmonic syntax.

Acknowledgments

This work was supported by NIMH RO1-MH60415 to M.G.W. We thank A. Landau, A. Finn, C. Lucas for suggestions on earlier versions of the manuscript, along with S. Koelsch and an anonymous referee for their helpful review comments.

References

[1] L. Bernstein, The Unanswered Question: Six Talks at Harvard (Charles Eliot Norton Lectures), Harvard University Press, 1973.
[2] M. Besson, F. Faïta, An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians, J. Exp. Psychol. Hum. Percept. Perform. 21 (1995) 1278–1296.
[3] Educational Testing Systems, The Graduate Record Exam, 2003 (http://www.gre.org).
[4] S.A. Hackley, An evaluation of the automaticity of sensory processing using event-related potentials and brain–stem reflexes, Psychophysiology 30 (1993) 415–428.
[5] A. Hahne, A.D. Friederici, Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials, Bilingualism: Lang. Cogn. 4 (2) (2001) 123–141.
[6] S.A. Hillyard, T.W. Picton, Electrophysiology of cognition, in: V.B. Mountcastle, F. Plum, S.R. Geiger (Eds.), Handbook of Physiology. Section I. The Nervous System. Vol. V. Higher Functions of the Brain, vol. 2, American Physiological society, Bethesda, 1987, pp. 519–584.
[7] P. Janata, ERP measures assay the degree of expectancy violation of harmonic contexts in music, J. Cogn. Neurosci. 7 (1995) 153–164.
[8] S. Koelsch, E. Schröger, M. Tervaniemi, Superior pre-attentive

auditory processing in musicians, NeuroReport 10 (6) (1999) 1309–1313.
[9] S. Koelsch, T.C. Gunter, A.D. Friederici, E. Schröger, Brain indices of music processing: 'Non-musicians' are musical, J. Cogn. Neurosci. 12 (2000) 520–541.
[10] S. Koelsch, E. Gunter, M. Tervaniemi, D. Sammler, A.D. Friederici, Differentiating ERAN and MMN: an ERP study, NeuroReport 12 (7) (2001) 1385–1389.
[11] S. Koelsch, E. Schröger, T.C. Gunter, Music matters: preattentive musicality of the human brain, Psychophysiology 39 (2002) 38–48.
[12] S. Koelsch, T. Grossmann, T. Gunter, A. Hahne, E. Schröger, A. Friederici, Children processing music: electric brain responses reveal musical competence and gender differences, J. Cogn. Neurosci. 15 (5) (2003) 683–693.
[13] C.L. Krumhansl, Cognitive Foundations of Musical Pitch, Oxford University Press, 1990.
[14] C.L. Krumhansl, E.J. Kessler, Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys, Psychol. Rev. 89 (1982) 334–368.
[15] M. Kutas, S.A. Hillyard, Reading senseless sentences: brain potentials reflect semantic incongruity, Science 207 (1980) 203–205.
[16] G. McCarthy, A.C. Nobre, Modulation of semantic processing by spatial selective attention, Electroencephalogr. Clin. Neurophysiol. 88 (1993) 210–219.
[17] L. Meyer, Emotion and Meaning in Music, University of Chicago Press, 1956.
[18] R. Näätänen, Implications of ERP data for psychological theories of attention, Biol. Psychol. 26 (1988) 117–163.
[19] R. Näätänen, The mismatch negativity: a powerful tool for cognitive neuroscience, Ear Hear. 16 (1) (1995) 6–18.
[20] R. Näätänen, M. Simpson, N.E. Loveless, Stimulus deviance and evoked potentials, Biol. Psychol. 14 (1982) 53–98.
[21] H. Nittono, T. Bito, M. Hayashi, S. Sakata, T. Hori, Event-related potentials elicited by wrong terminal notes: effects of temporal disruption, Biol. Psychol. 52 (2000) 1–16.
[22] P. Paavilainen, H. Tiitinen, K. Alho, R. Näätänen, Mismatch negativity to slight pitch changes outside strong attentional focus, Biol. Psychol. 37 (1993) 25–41.
[23] A.D. Patel, E. Gibson, J. Ratner, M. Besson, P.J. Holcomb, Processing syntactic relations in language and music: an event-related potential study, J. Cogn. Neurosci. 10 (1998) 717–733.
[24] W. Piston, M. Devoto, A. Jannery, Harmony, 3rd ed., W.W. Norton and Co, 1987.
[25] M. Sams, P. Paavilainen, K. Alho, R. Näätänen, Auditory frequency discrimination and event-related potentials, Electroencephalogr. Clin. Neurophysiol., Evoked Potentials 62 (6) (1985) 437–448.
[26] M. Schmuckler, Expectation in music: Investigation of melodic and harmonic processes, Music Percept. 7 (1989) 109–150.
[27] M.G. Woldorff, Distortion of ERP averages due to overlap from temporally adjacent ERPs: analysis and correction, Psychophysiology 30 (1) (1993) 98–119.
[28] M.G. Woldorff, Selective listening at fast stimulus rates: so much to hear, so little time, Perspect. Event-Relat. Potentials Res. (EEG Suppl. 44) (1995) 32–51.
[29] M.G. Woldorff, S.A. Hillyard, Modulation of early auditory processing during selective listening to rapidly presented tones, Electroencephalogr. Clin. Neurophysiol. 79 (1991) 170–191.
[30] M.G. Woldorff, S.A. Hackley, S.A. Hillyard, The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones, Psychophysiology 28 (1991) 30–42.
[31] M.G. Woldorff, P.T. Fox, M. Matzke, J.L. Lancaster, S. Veeraswamy, F. Zamarripa, M. Seabolt, T. Glass, J.H. Gao, C.C. Martin, P. Jerabek, Retinotopic organization of early visual spatial attention effects as revealed by PET and ERPs, Hum. Brain Mapp. 5 (1997) 280–286.
[32] M.G. Woldorff, S.A. Hillyard, C.C. Gallen, S.R. Hampson, F.E. Bloom, Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex, Psychophysiology 35 (1998) 283–292.
[33] M.G. Woldorff, M. Liotti, M. Seabolt, L. Busse, J.L. Lancaster, P.T. Fox, The temporal dynamics of the effects in occipital cortex of visual–spatial selective attention, Cogn. Brain Res. 15 (1) (2002) 1–15.