

# Effects of Selective Attention on Syntax Processing in Music and Language

Clemens Maidhof<sup>1,2,3</sup> and Stefan Koelsch<sup>4</sup>

## Abstract

■ The present study investigated the effects of auditory selective attention on the processing of syntactic information in music and speech using event-related potentials. Spoken sentences or musical chord sequences were either presented in isolation, or simultaneously. When presented simultaneously, participants had to focus their attention either on speech, or on music. Final words of sentences and final harmonies of chord sequences were syntactically either correct or incorrect. Irregular chords elicited an early right anterior negativity (ERAN), whose amplitude was decreased when music was simultaneously presented with speech, compared to when only music was presented. However, the amplitude of the ERAN-like waveform elicited when music

was ignored did not differ from the conditions in which participants attended the chord sequences. Irregular sentences elicited an early left anterior negativity (ELAN), regardless of whether speech was presented in isolation, was attended, or was to be ignored. These findings suggest that the neural mechanisms underlying the processing of syntactic structure of music and speech operate partially automatically, and, in the case of music, are influenced by different attentional conditions. Moreover, the ERAN was slightly reduced when irregular sentences were presented, but only when music was ignored. Therefore, these findings provide no clear support for an interaction of neural resources for syntactic processing already at these early stages. ■

## INTRODUCTION

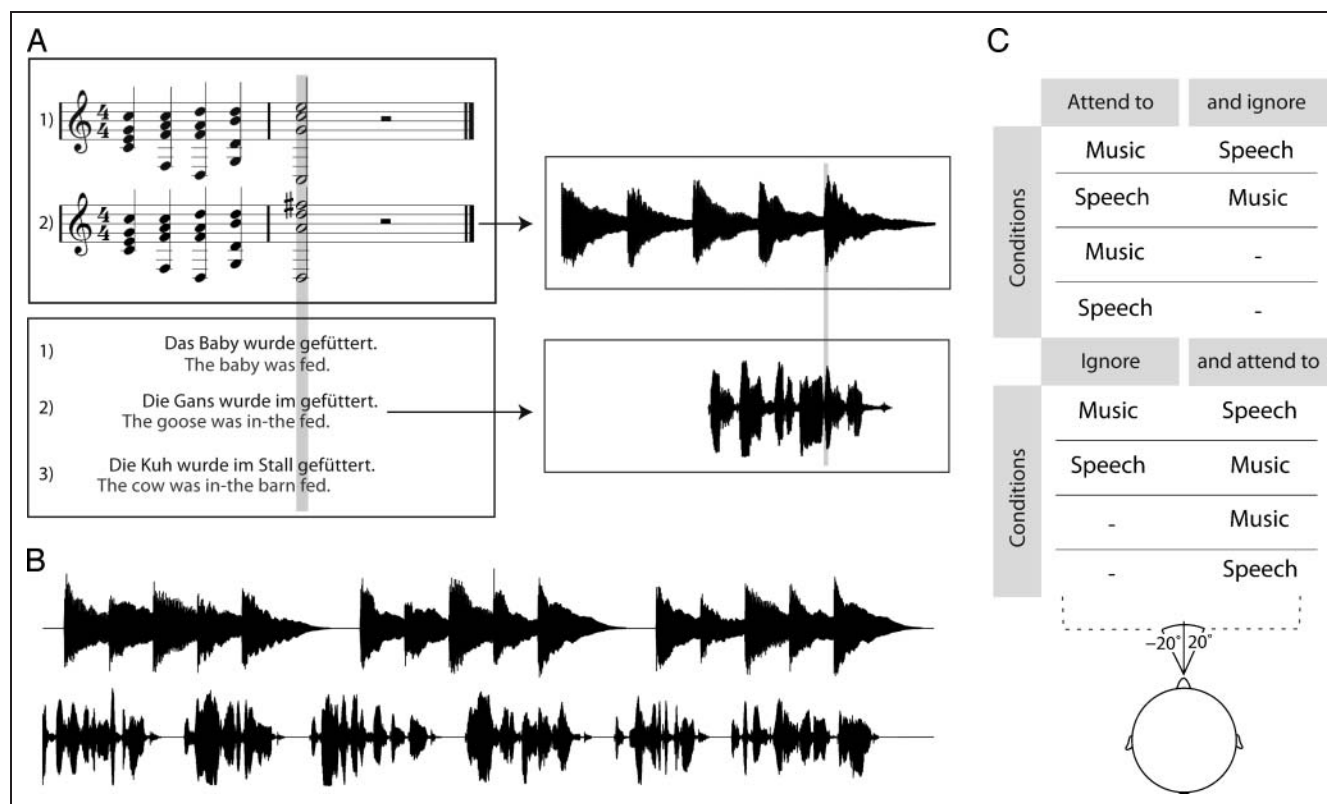
Let's imagine we are at a cocktail party, where many people are talking while some music is playing. This is a highly complex auditory environment, similar to many other everyday situations in which our auditory system is confronted with multiple, simultaneously active sound sources. Even when various sound sources are present (such as speech and music at a cocktail party), we are able to selectively attend to one single sound source, while ignoring other incoming auditory information. However, it is unknown to what extent ignored speech is processed when we selectively attend to music, or to what extent ignored music is processed when we focus our attention on speech. In the present ERP study, we investigated effects of selective attention on the processing of syntactic information in simultaneously presented speech and music (see also Figure 1).

Early psychological theories of attention proposed a filter model with a structural limitation (the attentional bottleneck), and subsequent theories proposed that stimuli are selected for further processing at an early (Broadbent, 1957, 1958), or late stage (Duncan, 1980; Deutsch & Deutsch, 1963). An intermediate theory (Treisman, 1964) proposed that filtering attenuates rather than completely

prevents processing of unattended stimuli. In contrast to filter models, capacity models of attention (e.g., Kahneman, 1973) assume a general limit of cognitive operations, and that the processing capacity can be flexibly allocated to any stage in the processing chain. Early stages require no attention at all and are not under strategic control of participants (preattentive or automatic processing; but see, e.g., Logan, 1992, for a dissociation between automatic and preattentive processes), whereas later stages require increasing amounts of capacity and can be controlled by participants (see, e.g., Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). However, the all-or-none concept of automaticity, that is, the assumption of fully automatic processes that are independent of attention (and do not use limited capacity resources), was extended. Recent views (e.g., Hackley, 1993) hold that processes can vary in their degree of automaticity, and distinguish between strongly automatic (obligatory and not modifiable by attention), partially automatic (obligatory but modifiable by attention), and controlled processes (nonobligatory requiring attentional resources; for a recent review on the concept of automaticity, see Moors & De Houwer, 2006).

In the music domain, the question of the degree of automaticity of syntactic processes has so far been only poorly investigated. In the present study, we used the early right anterior negativity (ERAN) as an electrophysiological marker of syntactic processing. The ERAN can be elicited by harmonically irregular chords presented within a progression of regular chords (e.g., Leino, Brattico, Tervaniemi, & Vuust, 2007;

<sup>1</sup>Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, <sup>2</sup>University of Helsinki, Helsinki, Finland, <sup>3</sup>University of Jyväskylä, Finland, <sup>4</sup>Freie Universität Berlin, Berlin, Germany



**Figure 1.** (A) Examples of chord sequences and sentences. Final chords of sequences were either music-syntactically correct or incorrect. Likewise, participles of sentences were either syntactically correct or incorrect. Onsets of final chords coincided with onsets of participles. (B) Stream of chord-sentence sequences. In the upper part, music is depicted, and in the lower part, speech is depicted. (C) Experimental design. There were eight conditions, in which music, speech, or both were presented. Stimuli were presented from two easily discriminable locations ( $20^\circ$  and  $-20^\circ$  in the azimuthal plane, respectively). Participants had to focus their attention on either music or speech and to perform a timbre detection task on the attended stimulus.

Steinbeis, Koelsch, & Sloboda, 2006; Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005; Koelsch, Schmidt, & Kansok, 2002; Koelsch, Gunter, Friederici, & Schroeger, 2000; for reviews, see Koelsch, 2005, 2009). Three previous studies suggest that the neural mechanisms underlying the generation of the ERAN operate preattentively: The ERAN can be elicited when participants play a video game under the instruction to ignore all acoustic stimuli (Koelsch et al., 2001), when participants read a self-selected book (Koelsch, Schroger, & Gunter, 2002), or when participants perform a reading comprehension task (Loui et al., 2005).

However, the latter study (Loui et al., 2005) found evidence for partially automatic music-syntactic processes, that is, it seems that these processes do not require attention but that they can be enhanced by different attentional demands (similar to other findings in auditory attention research, e.g., Woldorff, Hillyard, Gallen, Hampson, & Bloom, 1998; Hillyard, Hink, Schwent, & Picton, 1973; see also Hackley, 1993; Näätänen, 1992). Importantly, in the study from Loui et al. (2005), attention was diverted by requiring participants to read during the presentation of the musical stimulus. In contrast, we tested in the present study whether the neural mechanisms underlying the processing of music-syntactic information (as reflected in the ERAN) are active even when participants selectively attend to a

speech stimulus (i.e., to a simultaneously presented complex auditory stimulus).

In the language domain, the question of automaticity of syntactic processes operating during sentence comprehension is an ongoing debate (e.g., Deutsch & Bentin, 1994; Flores d'Arcais, 1988; Frazier, 1987). Previous studies have shown that initial structure building processes are fairly automatic (Hahne & Friederici, 1999; Flores d'Arcais, 1982, 1988; Frazier, 1987; Forster, 1979), but the degree of automaticity of syntactic processes when a second complex stimulus is present remains unclear. We used the early left anterior negativity (ELAN) as a neurophysiological index for syntactic processing (Hahne & Friederici, 1999). The ELAN can be elicited by phrase structure violations and is considered to reflect early syntactic processes (Friederici, 2002).

Only a few studies investigated the degree of automaticity of the ELAN and found that the underlying processes operate quite automatically (Hahne & Friederici, 1999, 2002; see also Friederici, Hahne, & Mecklinger, 1996; Friederici, Pfeifer, & Hahne, 1993). However, in none of the studies investigating the ELAN was the participants' attention directed away from the stimulus material, or were participants confronted with a task that required a high attentional load, which would be necessary to draw stronger conclusions about the automaticity of the ELAN (but see Pulvermüller, Shtyrov, Hastings,

& Carlyon, 2008, for a study investigating the processing of ungrammatical word pairs under a distraction task with high attentional load within the auditory modality).

We used an intramodal design in which speech and music were both presented auditorily while participants focused their attention on either music or speech (see Figure 1), thus allowing to obtain ERPs both when music or language were attended or ignored within the same modality. This enabled us to investigate the processing of musical syntax and linguistic syntax with a distraction task that requires higher attentional demands than the tasks used in previous studies (Loui et al., 2005; Hahne & Friederici, 1999, 2002; Koelsch, Schroger, et al., 2002; Koelsch et al., 2001; Friederici et al., 1993, 1996). The stimulus material was identical for both conditions (“attend to speech and ignore music,” as well as “attend to music and ignore speech”), thus the differences between ERPs elicited in both conditions can only be due to the direction of attention.

The experimental design also enabled us to examine interactions between music- and language-syntactic processing. Based on behavioral evidence (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Slevc, Rosenberg, & Patel, 2009) and data from ERP and functional imaging studies (Steinbeis & Koelsch, 2008; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Tillmann, Janata, & Bharucha, 2003; Koelsch, Gunter, et al., 2002; Platel et al., 1997), it is assumed that both processes rely on overlapping neural resources: The ELAN and the ERAN both show a similar time course and a comparable scalp distribution (although the ELAN often has a more left hemispheric weighting while the ERAN is often reported to be more pronounced over right frontal leads), and studies attempting to localize the neural generators of these ERP components found evidence for overlapping neural resources (mainly in inferior fronto-lateral cortex and the planum polare of the superior temporal gyrus; Maess, Koelsch, Gunter, & Friederici, 2001; Friederici, Wang, Herrmann, Maess, & Oertel, 2000; Knoesche, Maess, & Friederici, 1999; for overviews, see Koelsch, 2005; Patel, 2003; Friederici, 2002).

However, to date, only two ERP studies directly examined the simultaneous processing of syntactic aspects in speech and music (Steinbeis & Koelsch, 2008; Koelsch, Gunter, et al., 2005). These studies reported an interaction between the processing of musical syntax (as reflected in the ERAN) and linguistic syntax (as reflected in the left anterior negativity, LAN), supporting the assumption of an overlap of neural resources involved in syntactic processing of music and language. In the present study, we aimed at investigating if the ERAN also interacts at earlier stages of syntactic processing, as reflected in the ELAN.

We hypothesized that both language- and music-syntactic processes operate even when attention is focused on another complex auditory stimulus. That is, we expected that an ELAN is elicited even when participants focus their attention on the music, and that an ERAN is elicited even when participants focus their attention on the language. Due to

overlapping neural resources for the processing of syntax in language and music, we additionally hypothesized that the processing of syntax in language would influence music-syntactic processing, resulting in an interaction between the ERAN and the ELAN.

## METHODS

### Subjects

Nineteen subjects (10 women, age range = 19–30 years, mean = 23.7 years) participated in the experiment. All participants were German native speakers and had no or moderate musical training: 10 participants had never had any extracurricular music lessons, 9 participants had, on average, 6.2 years of musical practice (range: 3–12 years). All subjects were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and were reported to have normal hearing.

### Stimuli

#### *Musical Stimuli*

There were two types of chord sequences (identical to previous experiments, see Koelsch, Jentschke, Sammler, & Mietchen, 2007; Koelsch, Heinke, Sammler, & Olthoff, 2006), each consisting of five chords (see Figure 1A). The first four chord functions were identical: tonic, subdominant, supertonic, and dominant, each with a presentation time of 600 msec. The final chord was either the tonic (Type 1) or the double dominant (Type 2) and was presented for 1200 msec. Sequences were separated by a 400-msec pause. The sequences were transposed to all 12 tonal keys, resulting in 24 different musical sequences. Each sequence type occurred with a probability of  $p = .5$  and both sequence types were pseudorandomized with the constraint that no sequence was presented in the same tonal key as the one of the preceding sequence. Musical stimuli were generated as MIDI-files and exported as audio-files (WAV, stereo, 44.1 kHz, 16 Bit) with a piano sound (“The Grand,” Steinberg Media Technologies GmbH, Hamburg, Germany). In 10% of the sequences, the timbre of one chord was slightly modified by applying an FFT-Filter (2048 points, Blackman-window) using CoolEdit Pro (Syntrillium Software Company, USA). As a starting point, we used an FFT-Filter suppressing all frequencies below 440 Hz and above 4050 Hz. Then, we adjusted this filter individually for each timbre deviant by excluding lower frequencies from the filter (thereby including these frequencies in the resulting spectrum of the chord), until the percept of the chord was only slightly different to the original, unfiltered version. To keep the attention of the participants focused on the music during the course of a chord sequence, 50% of the timbre deviants occurred on the final chord (and 25% on the fourth chord, 16.6% on the third chord, and 8.3% on the second chord).

## Linguistic Stimuli

There were three types of sentences (identical to previous studies, see Friederici et al., 1993, 2000; Hahne & Friederici, 1999): Syntactically correct sentences (Type 1; see Figure 1A) consisted of a noun phrase, an auxiliary, and a past participle: *Das Baby wurde gefüttert*. [The baby was fed.]. Syntactically incorrect sentences (Type 2) contained a phrase structure violation: In those sentences, a noun phrase and an auxiliary were directly followed by a preposition and a past participle, which is a syntactically illegal phrase structure in German (*Die Gans wurde im gefüttert*. [The goose was in-the fed]). In addition to these types, there were also filler sentences with a complete prepositional phrase (Type 3), for instance, *Die Kuh wurde im Stall gefüttert*. [The cow was in-the barn fed]. These filler sentences were employed to ensure that participants do not anticipate the correctness of a sentence when a preposition was presented. Thus, the critical word on which an error became overt was the participle, which was identical for all three versions of one sentence (for more details, see Hahne & Friederici, 1999). We used 96 different correct, 96 different incorrect, and 48 different filler sentences which were spoken by a female native speaker of German. The order of the sentences was pseudorandomized in a way that no sentences with the same participle directly followed each other. As for the musical stimuli, in 10% of the sentences the timbre of one word was manipulated by applying an FFT-Filter to its spectral shape (in case of short words such as *im*, the following word was additionally manipulated to create a comparable length of musical and linguistic timbre deviants). The procedure was identical to the procedure used for the musical timbre deviants.

## Combination of Musical and Linguistic Stimuli

The onset of the final chord of each sequence coincided with the participle of each sentence (see Figure 1A). Two hundred forty (240) experimental chord–sentence sequence pairs were created. Because the length of the two stimuli was different (average length of a sentence: 1.78 sec; length of a chord sequence: 3.6 sec) and to obtain comparable periods of silence, we presented additional sentences between the sentence–chord pairs. This resulted in two continuous stimulus streams, with periods of silence between chord sequences of 400 msec and between sentences of 250–350 msec (see Figure 1B). The proportion of timbre deviants was 10%, and the proportion of the sentences was 40% correct, 40% incorrect, and 20% filler sentences. To facilitate selective attention (see also Woldorff & Hillyard, 1990) and to create a more ecologically valid situation, the stimuli were filtered with nonindividual head-related transfer functions and presented with a direction angle of 20° and –20° in the azimuthal plane (“Maven 3d professional”; Emersys, USA). That is, the sound source of music was (virtually) spatially separated from the sound source of speech, leading to the impression that music was presented from one side and speech from the other side.

There were eight conditions, in each of which 240 chord sequences, 240 experimental sentences, or 240 chord–sentence sequence pairs were presented (see Figure 1C). Each condition was divided into four subblocks (approximately 4 min), and the order of the resulting 32 subblocks was balanced across participants.

## Procedure

Data were recorded in two experimental sessions (each with a duration of about 65 min; on average, there were 13.8 days between the two sessions, range 6–27 days). Stimuli were presented via headphones (AKG 240 studio) at a comfortable listening level (approx. 55 dB). Participants were informed about irregular sentences and irregular chord sequences. Before the experiment, participants were familiarized with the task and the situation. They were instructed to focus their attention on either music or speech (while ignoring the other stimulus) and to concentrate only on the timbre. They were specifically informed that timbre deviants could occur sometimes rarely (e.g., only two times in a minute), but sometimes also right after each other. Thus, to successfully perform the task, they had to focus their attention strongly on only one stream and to ignore the other one. They were asked to press a button as fast as possible whenever they detected a timbre deviant in the attended stimulus type. Before each block, the instruction appeared on a computer screen and informed about the next condition. During the stimulus presentation, a fixation cross was presented on the left–center or right–center of the screen (depending from which side the to-be-attended stimulus was presented). After each block, there was a short break and participants could continue the experiment by pressing a button whenever they were ready.

To test the discriminability of the timbre deviants, an additional behavioral experiment was conducted. Fourteen participants (7 women; aged 18–28 years, mean age = 24.1 years) were presented with one block of 32 sentence pairs and one block of 32 chord sequence pairs (order of blocks was counterbalanced). Thus, the duration of the experiment was about 10 min. Each sentence and each chord sequence could include a timbre deviant (exactly the same timbre deviants as in the ERP study were used). There were four conditions: (a) the first sequence was presented with standard timbre, as well as the second sequence; (b) the first sequence contained a timbre deviant, the second sequence was presented with standard timbre; (c) the first sequence contained a timbre deviant, and the second sequence contained a timbre deviant at the same position; (d) the first sequence was presented with standard timbre, the second sequence contained a timbre deviant. Participants were informed about occasional timbre deviants and were instructed to judge after the presentation (by a button press) whether both sequences had the identical timbre or differed in timbre.

On average, participants correctly classified 81% of the sentence pairs and 65% of the chord sequence pairs (which

was well above chance level, as confirmed by one-sample *t* tests:  $p < .001$  in both conditions). This shows that the timbre deviants were difficult to detect.

### Data Recording and Analysis

Testing was carried out in an acoustically and electrically shielded EEG cabin. The EEG was recorded with Ag/AgCl electrodes from 61 electrode sites of the extended 10–20 system (Fp1, Fp2, Fpz, AF7, AF3, AFz, AF4, AF8, F9, F7, F5, F3, Fz, F4, F6, F8, F10, FT9, FT7, FC5, FC3, FCz, FC4, FC6, FT8, FT10, A1, T7, C5, C3, Cz, C4, C6, T8, A2, TP9, TP7, CP5, CP3, CPz, CP4, CP6, TP8, TP10, P9, P7, P5, P3, Pz, P4, P6, P8, P10, PO7, PO3, POz, PO4, PO8, O1, Oz, O2), referenced to the left mastoid. The ground electrode was located on the sternum. EOGs were recorded bipolarly by electrodes placed on the left and right outer canthus (horizontal EOG), and Fpz and the tip of the nose (vertical EOG). Electrode impedance was kept below 5 k $\Omega$ . Data were digitized with a sampling frequency of 500 Hz. For each subject, EEG and behavioral data were pooled for both sessions to obtain a higher signal-to-noise ratio of the data.

After data acquisition, EEG data were re-referenced to the mean of both mastoid electrodes and band-pass filtered (0.25–25 Hz, 2001 points, finite impulse response). Artifacts caused by eye movements were rejected off-line whenever the standard deviation within a 200-msec window centered around each sampling point exceeded 25  $\mu$ V in the EOG. Artifacts caused by drifts and body movements were eliminated by rejecting sampling points whenever the standard deviation within a 200- or 800-msec window exceeded 25  $\mu$ V at any electrode. Trials with typical eye blinks were marked and corrected by applying EOG correction (xeog, EEP software; ANT, Netherlands). ERPs were computed for 1200 msec time-locked to the onset of the participle or last chord, with a baseline ranging from –100 to 0 msec before the onset of the last chord or word.

For statistical analysis of the ERAN, mean ERP amplitude values were calculated for two anterior regions of interest (ROIs, as in previous studies, Koelsch et al., 2000; Koelsch, Gunter, et al., 2005): left anterior (F7, FT7, F3, FC3) and right anterior (F8, FT8, F4, FC4). Because the ELAN showed a more lateralized topographic distribution, the following ROIs were used for statistical analysis of the ELAN: left anterior (F7, FT7, F9, FT9) and right anterior (F8, FT8, F10, FT10). For the ERPs elicited by the timbre deviants, mean amplitude values were computed for two fronto-central (left: FC5, FC3, C5, C3; right: FC4, FC6, C4, C6) and one posterior–central ROI (PO3, POz, PO4, O1, Oz, and O2).

ERPs were statistically evaluated by repeated measures ANOVAs as univariate tests of hypotheses for within-subjects effects. ANOVAs were computed with factors attention (on music, on speech), distraction (second stimulus present, second stimulus absent), chord type (regular, irregular), sentence type (correct, incorrect), and hemisphere (left, right). Results of the analysis with the additional factor sound direc-

tion (20° azimuth, –20° azimuth) are only reported in footnotes whenever this factor interacted with the factors of chord type or sentence type. Time windows for statistical analyses of ERP data of irregular chord sequences and incorrect sentences were: 90–170 msec (ELAN), 160–220 msec (ERAN), 250–800 msec (late sustained negativity, LSN), and 500–650 msec (N5; for the “attend music” and “ignore music” blocks, a slightly different time window of 550–650 msec was used).

Additionally, we tested for effects of moderate musical expertise to guarantee that the observed effects were not only due to the few individuals with moderate musical training. Results of an ANOVA with an additional between-subject factor group (nonmusicians/amateur musicians) showed that musical expertise had no significant effect on the ERAN in any of the blocks (no main effect or interaction with factor chord type; all  $p > .26$ ); therefore, data of nonmusicians and amateur musicians were pooled.

To test for differences in peak latency of the N100 and P200, peak latencies of these potentials elicited by regular and irregular chords were measured at FC4 for the “only music,” “attend music,” and “ignore music” blocks. Two participants were excluded from this analysis because no unambiguous peak latencies could be determined. Peak latencies were statistically evaluated by repeated measures ANOVAs (Huynh–Feldt corrected  $p$  values where necessary).

The  $d'$  values were calculated by subtracting the Z-scores of the false alarm rates from the Z-scores of the hit rates. A hit was recorded whenever a participant responded within 2000 msec after the onset of a target. To test whether participants performed above chance level, we calculated one-sample *t* tests with a test value of zero, and for the differences between conditions,  $d'$  values, hit rates, and RTs were statistically evaluated by ANOVAs as univariate tests of hypotheses for within-subjects effects with factors attention (on music, on speech), distraction (second stimulus present, second stimulus absent), and sound direction (20° azimuth, –20° azimuth).

## RESULTS

### Behavioral Results

Participants detected the timbre deviants well above chance level ( $p = .001$  in all tests). The results of the  $d'$  values (see Table 1) showed that linguistic and musical target stimuli were both better detected when there was no second stimulus present. The results of the ANOVA showed a main effect of distraction [ $F(1, 18) = 30.23, p = .001$ ] and no interactions. Similarly, the hit rates (see Table 1) were reduced when a second stimulus was present. However, when participants attended to the music stimulus, the hit rate was influenced more strongly by the presence of the second stimulus compared to when participants attended to the linguistic stimulus. The results of the ANOVA showed main effects of distraction [ $F(1, 18) = 54.12, p < .0001$ ] and attention

**Table 1.** Means of Behavioral Results of the Target Detection Task

Condition		<i>d'</i> (SD)	Hit Rate (%)	False Alarms (%)	Hits Ignore (%)
Attended	Distraction				
Music		4.45 (0.88)	72.24	0.68	
Music	Speech	3.68 (0.75)	56.53	0.80	2.57
Speech		4.32 (0.98)	77.10	1.00	
Speech	Music	3.46 (0.86)	69.68	1.61	3.69

[ $F(1, 18) = 9.19, p = .007$ ], and an interaction between these factors [ $F(1, 18) = 12.54, p = .002$ ].

RTs for target words were longer when there was a second stimulus present. The results of the ANOVA showed a main effect of Distraction [ $F(1, 18) = 7.35, p = .014$ ] and an interaction between factors Distraction and Attention [ $F(1, 18) = 14.89, p = .001$ ]. Follow-up ANOVAs conducted separately for the “attend speech” and “attend music” blocks showed only a main effect of Distraction when participants attended the linguistic stimuli [ $F(1, 18) = 38.77, p = .001$ ].

## ERP Results: Music

### *Blocks with Music Only (No Speech)*

This section presents the ERPs recorded while only music was presented to the participants (with the instruction to detect the chords with deviant timbre). Compared to regular chord sequence endings (tonic chords), music-syntactically irregular endings (double dominant chords) elicited an ERAN which was maximal at around 190 msec (see Figure 2A for ERPs and Table 2 for statistical results). With nose reference, the ERAN inverted polarity at mastoid leads (as in all other blocks, see Figure 2B, D, and F). In contrast to previous studies (Koelsch & Jentschke, 2008; Koelsch & Sammler, 2008; Koelsch et al., 2007), the ERAN was not lateralized (see Discussion). The ERAN was followed by a late negativity (the N5) which was maximal around 550 msec.

### *Attend Music Blocks (Ignore Speech)*

This section presents the ERPs recorded while music and language were presented simultaneously, and while participants focused their attention to the music (with the instruction to ignore the speech and detect the musical timbre deviants). Irregular chord sequence endings (compared to regular ones) elicited a marginally significant ERAN (see Figure 2C for ERP and Table 2 for statistical results). However, the ERAN was clearly significant when analyzed over a right central ROI including electrodes F4, Fz, FC4, and FCz [main effect of chord type:  $F(1, 18) = 7.48, p =$

.014]. The N5 was not observed in this condition (for statistical results, see Table 2).

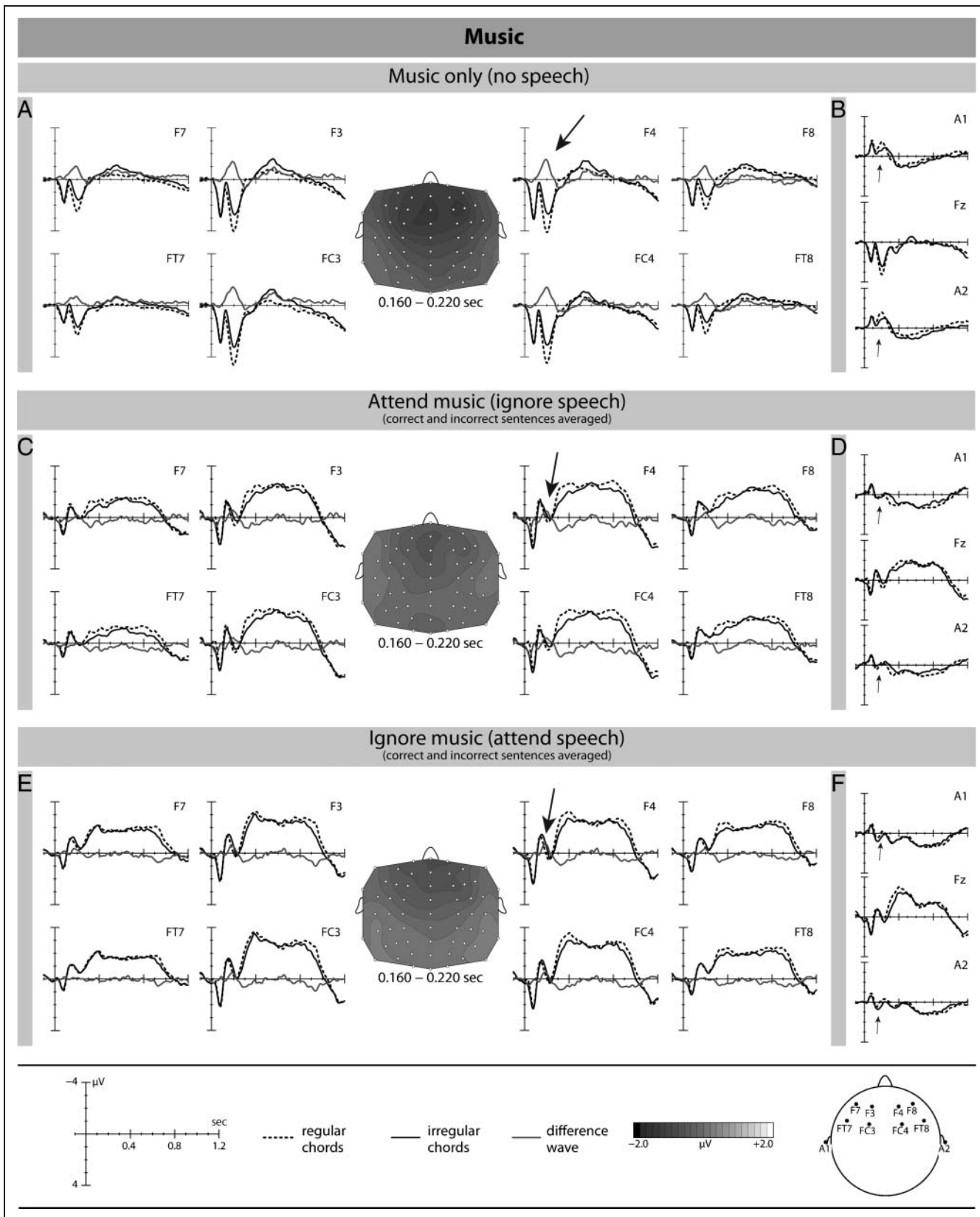
The amplitude of the ERAN was reduced compared to the “only music” blocks: An ANOVA (160 to 220 msec) with factors attention (“attend music,” “only music”), chord type (regular, irregular), and hemisphere (left, right) showed effects of Chord type [ $F(1, 18) = 107.83, p < .0001$ ] and Distraction [ $F(1, 18) = 151.80, p < .0001$ ], as well as an interaction between these two factors [ $F(1, 18) = 21.17, p = .0002$ ; see Table 3 for statistical results].

### *Ignore Music Blocks (Attend Speech)*

This section presents the ERPs to final chords while music and language were presented simultaneously, and while participants ignored the musical stimulus, focusing their attention to the language (under the instruction to detect the words with deviant timbre). Note that the stimulus material presented in this condition was identical as in the “attend music (ignore speech)” blocks, the only difference between blocks was the instruction for the participants. Figure 2E shows that in this condition, the processing of irregular chord sequences also differed from the processing of correct chord sequences. This resulted in a negative difference wave (measured as the difference between irregular and regular chords) peaking at around 190 msec and resembling the ERAN in the other blocks. An N5 was not observed in this condition (see Table 2 for statistical results).

Although the latency of the early negative peak observed in the difference wave was around 190 msec (as in all other blocks), this effect was possibly due to a phase shift of the N100 and the P200 in the “ignore music” condition. To further specify this, we analyzed the peak latencies of both N100 and P200. Results showed that the peak latency of the N100 differed between regular and irregular chords (longer latencies for irregular chords), but independently of the different blocks. An ANOVA with factors Chord type (regular, irregular) and Block (“only music,” “attend music,” “ignore music”) showed main effects of Chord type [ $F(1, 16) = 9.43, p = .007$ ] and Block [ $F(1, 16) = 28.25, p < .0001$ ], but no interaction between these factors [ $F(1, 16) = 1.19, p = .31$ ]. The peak latencies of the P200 did not differ between regular and irregular chords, in none of the blocks. An ANOVA with the same factors showed only a main effect of Block [ $F(1, 16) = 5.31, p = .015$ ], but no effect of Chord type [ $F(1, 16) = 0.85, p = .37$ ], and no interaction between Chord type and Block [ $F(1, 16) = 0.52, p = .6$ ]. This analysis shows, in addition to the separate analysis of the ERAN elicited on correct and incorrect sentences in the “ignore music” blocks (see Figure 4), that the early negative peak observed in the difference wave in the “ignore music” blocks (compared to the “only music” and “attend music” blocks) was not only due to a phase shift of N100 and P200 elicited by irregular (compared to regular) chords.

The amplitude of the negative difference in the ERAN time window was decreased in the “ignore music” blocks



**Figure 2.** Grand-average ERPs (and scalp distributions) elicited by the irregular and regular chords in the blocks with (A) music only, (C) attend music (ignore speech), and (E) ignore music (attend speech). B, D, and F show the polarity inversion of the ERAN in each of these blocks with a nose reference. The difference waves are computed by subtracting regular sequences from irregular sequences. ERPs are averaged across correct and incorrect sentences.

**Table 2.** Summary of ANOVAs for the ERAN and N5 in the Three Music Blocks, and Summary of ANOVAs for the ELAN and LSN in the Three Speech Blocks

	<i>F(1, 18) and p values</i>			
	<i>ERAN</i>		<i>N5</i>	
<i>Only Music Blocks (No Speech)</i>				
Chord type	124.12	<.0001	10.60	.0044
Hemisphere	<2	<i>ns</i>	4.41	.05
Hemisphere × Chord type	<1	<i>ns</i>	5.26	.034
<i>Attend Music Blocks (Ignore Speech)</i>				
Chord type	3.63	.073	<1	<i>ns</i>
Hemisphere	13.19	.0019	6.57	.02
<i>Ignore Music Blocks (Attend Speech)</i>				
Chord type	6.73	.018	<1	<i>ns</i>
<hr/>				
	<i>ELAN</i>		<i>LSN</i>	
<i>Only Speech Blocks (No Music)</i>				
Sentence type	10.43	.0046	<3	<i>ns</i>
Hemisphere	3.39	.082	17.22	.0006
Hemisphere × Sentence type	4.00	.061	5.19	.035
<i>Attend Speech Blocks (Ignore Music)</i>				
Sentence type	13.77	.0016	<3	<i>ns</i>
Hemisphere	3.39	.082	4.71	.044
Hemisphere × Sentence type	<1	<i>ns</i>	3.15	.09
<i>Ignore Speech Blocks (Attend Music)</i>				
Sentence type	10.07	.0053	<1	<i>ns</i>
Hemisphere	2.65	.12	20.17	.0003

Time windows for the ERAN and the N5 were 160–220 msec and 500–650 msec, respectively. Time windows for the ELAN and the LSN were 90–170 msec and 250–800 msec, respectively.

LSN = late sustained negativity.

compared to the “only music” blocks, but the amplitude did not differ between the “ignore music” and “attend music” blocks. Comparing the “ignore music” and “only music” blocks, an ANOVA for the ERAN time window with factors Attention (“ignore music,” “only music”), Chord type (regular, irregular), and Hemisphere (left, right) showed main effects of Chord type [ $F(1, 18) = 72.58, p < .0001$ ], Attention [ $F(1, 18) = 151.75, p < .0001$ ], and an interaction between these factors [ $F(1, 18) = 22.79, p = .0002$ ]. Comparing the “ignore music” and “attend music” blocks, an ANOVA for the same time window with factors

Attention (“ignore music,” “attend music”), Chord type (regular, irregular), and Hemisphere (left, right) showed an effect of Chord type [ $F(1, 18) = 10.01, p = .0054$ ] and an effect of Attention [ $F(1, 18) = 5.16, p = .036$ ], but no interaction between these two factors [ $F(1, 18) = 0.3, p = .59$ ; see Table 3 for complete statistical results].

**ERP Results: Speech**

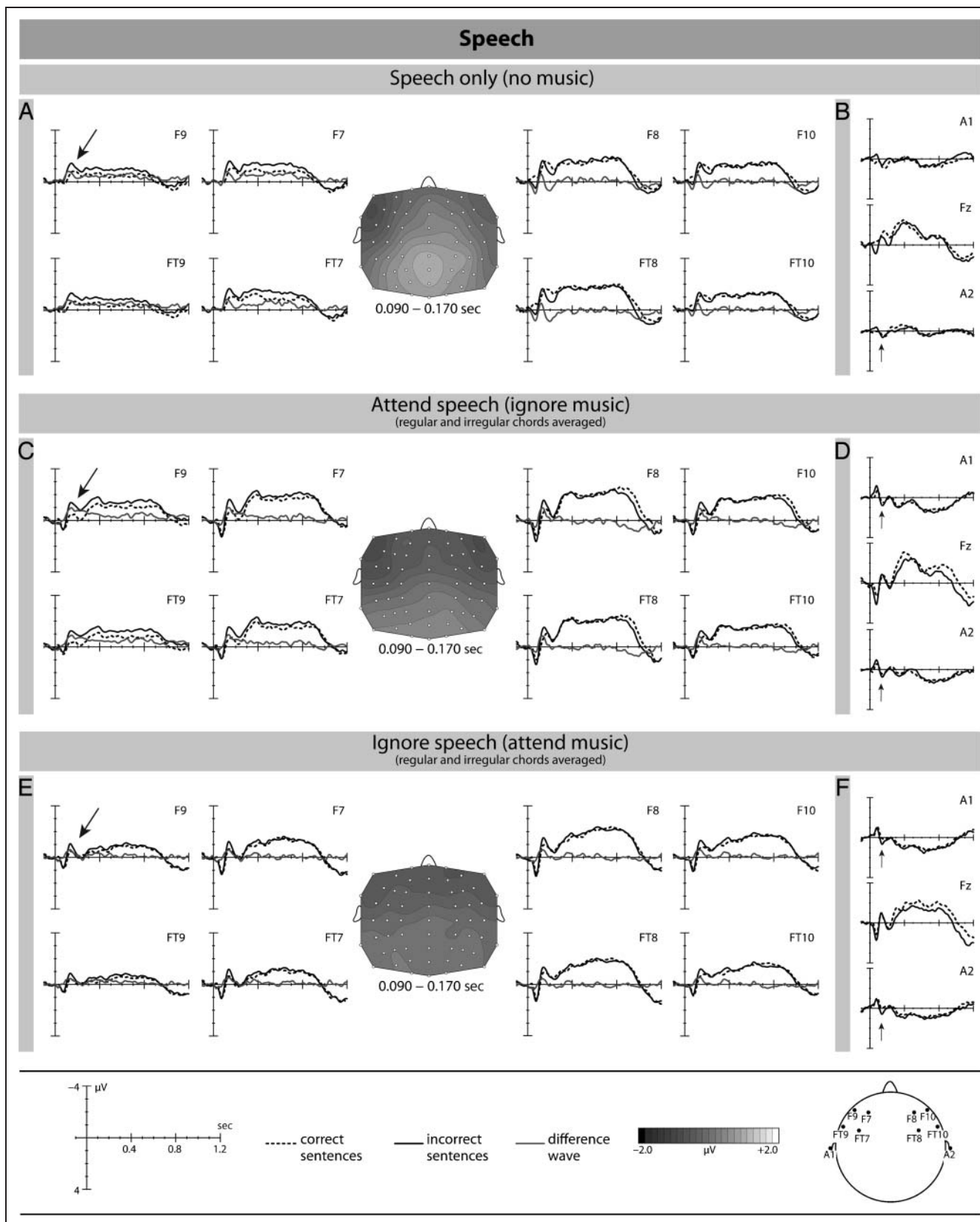
*Blocks with Speech Only (No Music)*

This section presents the ERPs recorded while only speech was presented to the participants (with the instruction to detect the words with deviant timbre). Compared to correct sentences, syntactic violations elicited an ELAN with a peak latency of around 130 msec, which showed slightly increased amplitude values over left-frontal electrodes (see Figure 3A for ERPs and Table 2 for statistical results). As in all other blocks, the ELAN inverted polarity at mastoid leads with a nose reference (see Figure 3B, D, and F). The ELAN was followed by an LSN ranging from around 250 to

**Table 3.** Results of ANOVAs for the ERAN Compared in the Different Music Blocks, and Results of ANOVAs for the ELAN Compared in the Different Speech Blocks

	<i>F(1, 8) and p values</i>			
	<i>ERAN</i>		<i>ELAN</i>	
<i>Music (or Speech) Only Blocks vs. Attend Music (or Speech) Blocks</i>				
Sequence type	107.83	<.0001	17.43	.0006
Hemisphere	6.27	.022	<1	<i>ns</i>
Distraction	151.80	<.0001	15.39	.001
Hemisphere × Sequence type	<1	<i>ns</i>	2.88	.107
Distraction × Sequence type	21.17	.0002	<3	<i>ns</i>
Distraction × Hemisphere	2.95	.103	6.21	.023
<i>Attend Music (or Speech) Blocks vs. Ignore Music (or Speech) Blocks</i>				
Sequence type	10.01	.0054	15.61	.0009
Attention	5.16	.036	<1	<i>ns</i>
Attention × Sequence type	<1	<i>ns</i>	<3	<i>ns</i>
Attention × Hemisphere	13.21	.0019	3.81	.067
<i>Music (or Speech) Only Blocks vs. Ignore Music (or Speech) Blocks</i>				
Sequence type	72.58	<.0001	18.65	.0004
Hemisphere	<1	<i>ns</i>	8.26	.01
Attention	151.75	<.0001	19.19	.0004
Attention × Sequence type	22.79	.0002	<1	<i>ns</i>





**Figure 3.** Grand-average ERPs (and scalp distributions) elicited by the incorrect and correct sentences in the blocks with (A) speech only, (C) attend speech (ignore music), and (E) ignore speech (attend music). B, D, and F show the polarity inversion of the ELAN in each of these blocks with a nose reference. The difference waves are computed by subtracting correct sentences from incorrect sentences. ERPs are averaged across regular and irregular chord sequences.

800 msec, showing also larger amplitude values over left-frontal electrodes.<sup>1</sup>

### *Attend Speech Blocks (Ignore Music)*

This section presents the ERPs recorded while speech and music were simultaneously presented, and while participants focused their attention to the speech (with the instruction to ignore the music and detect the words with deviant timbre). Compared to correct sentences, incorrect sentences elicited an ELAN with a maximal amplitude around 130 msec, but amplitude values of the ELAN were not larger over left-frontal than over right-frontal electrodes (see Figure 3C and Table 2).

The amplitude of the ELAN did not differ between the “attend speech (ignore music)” compared to the “only speech” blocks. An ANOVA (from 90 to 170 msec) with factors Attention (“attend speech,” “only speech”), Sentence type (correct, incorrect), and Hemisphere (left, right) showed an effect of Sentence type [ $F(1, 18) = 17.43, p = .0006$ ], an effect of Attention [ $F(1, 18) = 15.39, p = .001$ ], but no interaction between Attention and Sentence type [ $F(1, 18) = 2.38, p = .14$ ; see Table 3 for complete statistical results].

### *Ignore Speech Blocks (Attend Music)*

This section presents the ERPs elicited by correct and incorrect sentences while speech and music were presented simultaneously, and while participants ignored the linguistic stimulus, focusing their attention to the music (with the instruction to detect the chords with deviant timbre). The stimulus material was identical with the material in the “attend speech (ignore music)” blocks, with the only difference between blocks being the instruction. As depicted in Figure 3E, incorrect sentences containing a phrase structure violation elicited an ELAN, even though the sentences were ignored. The ELAN had a maximal amplitude around 130 msec. The LSN was not observed in this condition (for statistical results, see Table 2).

The amplitude of the ELAN did not differ between the “ignore speech” and “only speech” blocks: An ANOVA (from 90 to 170 msec) with factors Attention (“ignore speech,” “only speech”), Sentence type (correct, incorrect), and Hemisphere (left, right) showed main effects of Sentence type [ $F(1, 18) = 15.61, p = .0009$ ] and Attention [ $F(1, 18) = 19.19, p = .0004$ ], but no interaction between these factors [ $F(1, 18) < 1, p > .81$ ]. Although the amplitude of the ELAN was nominally smaller in the “ignore speech” blocks (mean:  $-0.46 \mu\text{V}$ ,  $SEM: 0.19$ ) than in the “attend speech” blocks (mean:  $-0.77 \mu\text{V}$ ,  $SEM: 0.22$ ), this difference was not statistically significant. An ANOVA for the same time window with factors Attention (“attend speech,” “ignore speech”), Sentence type (correct, incorrect), and Hemisphere (left, right) revealed a main effect of Sentence type [ $F(1, 18) = 15.61, p = .0009$ ], but no interaction between factors Attention and Sentence type [ $F(1, 18) = 2.04, p = .17$ ; see Table 3 for complete results].

### *Different Effects of Attention on the ERAN and the ELAN*

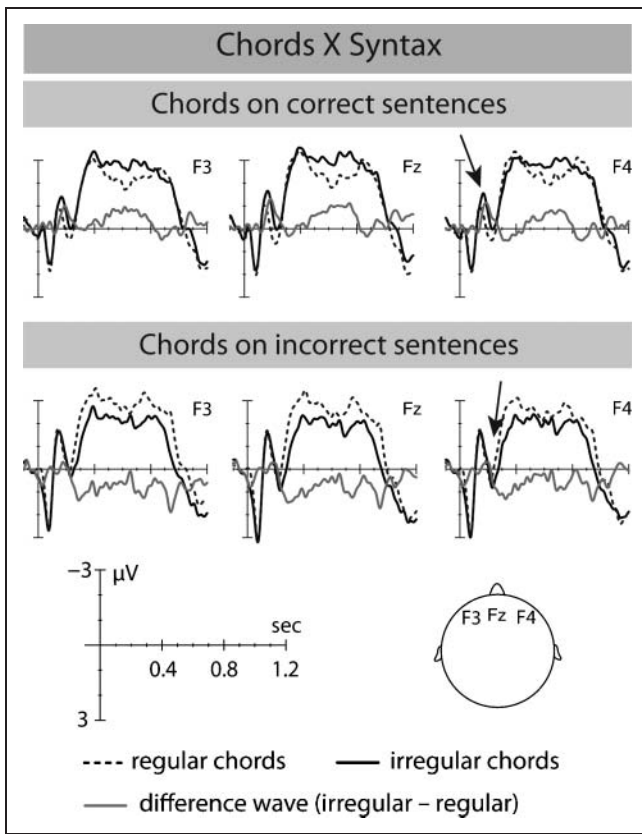
The analysis of the ERAN showed that the ERAN amplitude decreased in the “ignore music” blocks compared to the “only music” blocks (and the ERAN amplitude did not differ between the “ignore music” and “attend music” blocks). By contrast, the ELAN amplitude did not significantly differ between blocks. To test statistically whether the ERAN and the ELAN were differently influenced by the presence/absence of another stimulus and the direction of attention, we computed an ANOVA with factors Regularity (regular, irregular), Block (one stimulus, stimulus attended with additional second stimulus, stimulus ignored with additional second stimulus), and Domain (music, speech). Results showed a three-way interaction between these factors [ $F(1, 18) = 5.11, p = .015$ ], indicating that attentional demands influenced the ERAN more strongly than the ELAN.

### **ERP Results: Interaction between Speech and Music**

We hypothesized that the processing of the linguistic syntax would interact with the processing of the musical syntax (due to overlapping neural resources, see Introduction), potentially resulting in a decreased amplitude of the ERAN when the linguistic syntax is violated (a decrease of the ELAN was not expected, because the ELAN occurs earlier in time than the ERAN). Figure 4 shows the ERAN (in the “ignore music” blocks) for the conditions in which chords were presented on correct sentences and incorrect sentences. The amplitude of the ERAN was smaller when elicited during the presentation of incorrect sentences ( $-0.03 \mu\text{V}$ ,  $SEM: 0.26$ ) compared to when elicited during correct sentences ( $-0.65 \mu\text{V}$ ,  $SEM: 0.24$ ). An ANOVA with factors Sentence type (correct, incorrect), Chord type (correct, incorrect), and Hemisphere (left, right) showed an effect of Chord type [ $F(1, 18) = 6.73, p = .018$ ] and a marginally significant interaction between Sentence type and Chord type [ $F(1, 18) = 3.57, p = .075$ ], indicating that the correctness of the sentences slightly influenced the amplitude of the ERAN. However, the amplitude of the ERAN in the “attend music” blocks was not influenced by the correctness of the sentences ( $F < 1$ ).

### **ERP Results for Target Stimuli**

This section presents the ERPs elicited by chords or words with a slightly deviant timbre (i.e., the target stimuli). Figure 5 shows the difference waveforms of target stimuli minus standard stimuli. When participants focused on the music, chords with deviant timbre were task-relevant (and words with deviant timbre were to be ignored). When participants focused on the speech, on the other hand, words with deviant timbre were task-relevant (and chords with deviant timbre were to be ignored).



**Figure 4.** Interaction between music- and language-syntactic processing. In “ignore music” blocks the amplitude of the ERAN (difference wave: irregular sequences minus regular sequences; see arrows) is slightly reduced when chords were presented on incorrect sentences compared to when chords were presented on correct sentences.

Chords with deviant timbre elicited an early negative deflection, presumably consisting of an MMN that was partly overlapped by an N2b (the latter one being due to the controlled and conscious detection of task-relevant deviants; Novak, Ritter, Vaughan, & Wiznitzer, 1990). This negative deflection was followed by a P3b, reflecting the conscious detection of target stimuli (Comerchero & Polich, 1999; Mecklinger & Ullsperger, 1995; Donchin & Coles, 1988). Importantly, the amplitude of the P3b (as well as the amplitude of the negative deflection) was considerably smaller in the “ignore music” blocks than in the “attend music” blocks. Similar to the chords with deviant timbre, target words with deviant timbre elicited an early negative deflection, followed by a P3b. The amplitude of the P3b (as well as the negative deflection) was smaller in the “ignore speech” blocks than in the “attend speech” blocks (see Figure 5 and Table 4).

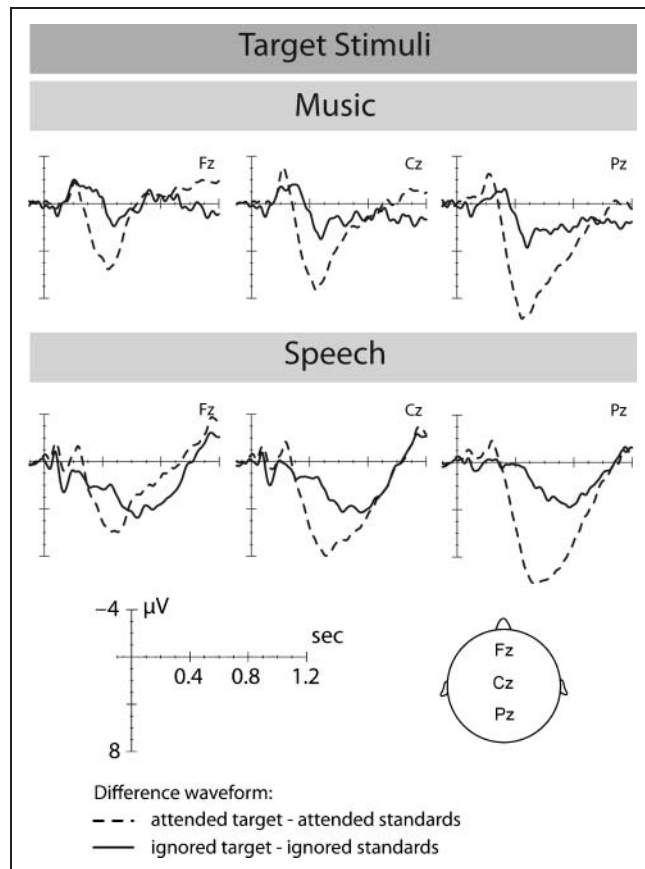
## DISCUSSION

### Automaticity of the ERAN and the ELAN

The present study investigated the effects of selective attention on two neurophysiological indices of musical and linguistic syntax processing (ERAN and ELAN). Under all three different attentional and stimulus conditions (only

music, attend music and ignore speech, ignore music and attend speech), irregular chords elicited an ERAN. However, the amplitude of the ERAN was largest when only music was presented, and was significantly decreased in conditions in which a second complex auditory stimulus (speech) was additionally presented. Importantly, a small ERAN-like waveform was present when participants ignored the chord sequences and focused on the linguistic stimuli (and the amplitude of the ERAN did not differ between the “attend music” and “ignore music” blocks). This indicates that the syntactic structure of music is processed even when attention is focused onto another auditory stimulus such as speech. In this regard, the music-syntactic processes investigated in this study fall under the category of “partially automatic processes” (Hackley, 1993), meaning that they operate obligatory (i.e., without the participant’s intention) but can be influenced by attention.

Irregular sentences elicited an ELAN, independently of whether another stimulus was presented, and independently of whether speech was attended or to be ignored.



**Figure 5.** Grand-average ERPs of the target detection task in attend and ignore conditions. The upper part shows the difference waveforms of ERPs elicited by attended chords (with deviant timbre minus standard timbre) and of ERPs elicited by ignored chords (with deviant timbre minus standard timbre). The lower part shows the difference waveforms of ERPs elicited by attended words (with deviant timbre minus standard timbre) and of ERPs elicited by ignored words (with deviant timbre minus standard timbre).

**Table 4.** Summary of ANOVAs for the ERPs Elicited by the Timbre Deviants

	<i>F(1, 18) and p values</i>			
	<i>MMN/N2b</i>		<i>P3b</i>	
<i>Attend Music Blocks vs. Ignore Music Blocks</i>				
Timbre	36.58	<.0001	52.55	<.0001
Attention	<1	<i>ns</i>	105.11	<.0001
Hemisphere	20.32	.0003	–	–
Timbre × Hemisphere	23.44	.0001	–	–
Timbre × Attention	3.4	.081	64.66	<.0001
Attention × Hemisphere	18.42	.0004	–	–
Attention × Hemisphere × Timbre	4.51	.048	–	–
<i>Attend Speech Blocks vs. Ignore Speech Blocks</i>				
Timbre	<1	<i>ns</i>	72.65	<.0001
Attention	40.53	<.0001	26.95	<.0001
Hemisphere	13.56	.0017	–	–
Timbre × Hemisphere	13.04	.002	–	–
Timbre × Attention	15.45	.001	54.31	<.0001
Attention × Hemisphere × Timbre	18.56	.0004	–	–

In attend and ignore music blocks, the time window for MMN/N2b was 150–250 msec and in attend and ignore speech blocks, the time window was 180–250 msec. In attend and ignore music blocks, the time window for the P3 was 350–600 msec and in attend and ignore speech blocks, the time window was 450–850 msec. Statistics were computed for left and right frontal ROIs (MMN/N2b) and a centroparietal ROI (P3).

This shows that the syntactic processes reflected by the ELAN operate (at least) partially automatically. Because the ELAN was less influenced by attention than the ERAN, it appears that the mechanisms underlying the generation of the ELAN operate with a higher degree of automaticity than music-syntactic processes. This might be due to the fact that, for nonmusicians (which were investigated in this study), language is a stimulus that is more common than music, and that nonmusicians are used to process sentences, even in noisy environments, in contrast to chord sequences. Whether attentional influences on the ERAN differ between nonmusicians and musicians remains to be specified.

Previous studies examining attentional effects on the ERAN used cross-modal designs, that is, during the presentation of music, attention was directed to stimuli presented in the visual domain (Loui et al., 2005; Koelsch, Schroger, et al., 2002; Koelsch et al., 2001). However, such tasks may not be optimal for directing attention away from stimuli because some aspects of attention seem to be modality-specific (Duncan, Martens, & Ward, 1997, e.g., reported that when participants had to identify targets presented in two different sensory modalities, performance was not affected,

whereas when participants had to identify targets in two input streams within the same modality, performance was disturbed).

Previous studies investigating attentional effects on the ELAN elicited during sentence processing directed attention to stimuli features other than syntax (Hahne & Friederici, 1999, 2002; Friederici et al., 1993, 1996), using no overly attention-demanding tasks. In the present study, processing of attended and ignored stimuli within the same (auditory) modality could be directly compared.

To answer the question of the level of automatic processing, it is necessary to use demanding diversion tasks and to provide evidence that attention was directed away from the stimulus material. Here, the assumption that participants followed the instructions and selectively focused their attention on one stimulus is evidenced by the modulation of the amplitudes of ERPs elicited by target stimuli: The negative deflection, presumably consisting of an MMN in part overlapped by an N2b, was decreased in the ignore conditions compared to attend conditions. More importantly, the P3b, a component modulated by task relevance (e.g., Gaeta, Friedman, & Hunt, 2003), was significantly reduced in ignore conditions compared to attend conditions. In addition, a previous study suggested that the amplitude of the P3b reflects the amount of information extracted from a stimulus: The larger the amplitude, the more information was extracted (Sussmann, Winkler, & Schröger, 2003). Thus, although it seems that participants partially automatically processed the timbre deviants in the unattended stream (as reflected by the P3b elicited by to-be-ignored stimuli), we assume that participants selectively focused their attention on one stimulus, and that attended targets were more relevant for the participants. Note that the increased P3b amplitude might also reflect additional response or decision-related processes in the attend-conditions, due to the button presses. However, timbre deviants are expected to elicit attention-capturing mechanisms, which include the P3b reflecting that participants made the decision not to press the button. Thus, it seems unlikely that only these additional processes associated with button presses are reflected in the increased P3b amplitude.

The behavioral data further support the assumption that participants were attending to the timbre deviants in the cued stream. First, the results of the discriminability experiment (see Methods), in which participants correctly classified 65% of the musical sequences and 81% of the sentences, showed that the timbre manipulations were difficult to detect. More importantly, these results, in combination with the similar hit rates during the ERP experiment (around 70% during the “blocks with music only,” and around 75% during the “blocks with speech only”), provide compelling evidence that participants were attending to the cued stream in the ERP experiment. Moreover, the hit rates (and  $d'$  values) significantly dropped during blocks in which speech and music were presented simultaneously (compared to when only one stimulus was present). This

shows that attentional demands increased during these more complex conditions. To accomplish the detection task, participants were therefore required to shift their focus of attention to the to-be-attended stimulus and to ignore the other stream (which participants confirmed after the experiment in a questionnaire to be able to do very well).

### *Interaction between Music- and Language-syntactic Processing*

With regard to the interactions between music- and language-syntactic processing, two previous studies (Steinbeis & Koelsch, 2008; Koelsch, Fritz, et al., 2005; Koelsch, Gunter, et al., 2005) indicated that neural resources for syntactic processing in speech and music are not only shared on the level of syntactic integration (as supposed by the “shared syntactic integration hypothesis,” SSIRH, e.g., Patel, 2003), but already at earlier processing stages. More specifically, these studies showed that the ERAN interacts with the LAN, elicited by morphosyntactic violations in speech. The LAN is associated with syntactical-relational processes, and is elicited around 300–500 msec after morphosyntactic violation of tense, number or gender agreement, and verb inflection errors (Gunter, Friederici, & Schriefers, 2000; Osterhout & Mobley, 1995; Friederici et al., 1993; Kutas & Hillyard, 1983). In contrast, the ELAN is an index of initial syntactic structure building (based on word category information; see Friederici, 2002), usually preceding syntactic-relational processes (but see also Hastings & Kotz, 2008, showing that in two-word utterances, phrase structure and morphosyntactic processes do not necessarily operate sequentially). The present study tested directly whether the processing of musical syntax (as reflected in the ERAN) interacts also with these early stages of syntactic language processing (as reflected in the ELAN). Results showed that, when the musical stimulus was ignored, the amplitude of the ERAN was slightly reduced when an irregular chord was simultaneously presented with an irregular word. Surprisingly, the ERAN was not decreased on irregular sentences when participants focused on the music. Thus, these findings provide no clear support for an interaction of neural resources for syntactic processing already at these early stages (i.e., initial structure building). Perhaps the overlap of neural resources is larger for the ERAN and the LAN (resulting in clear interactions; Steinbeis & Koelsch, 2008; Koelsch, Fritz, et al., 2005; Koelsch, Gunter, et al., 2005), than for the ERAN and the ELAN (resulting only in a marginal interaction in the present study). This issue remains to be specified in future studies.

### *Scalp Distribution of the ERAN and the ELAN and Relations to Other ERP Components*

Both the ERAN and the ELAN were maximal over frontal electrode leads, but neither the ELAN nor the ERAN were significantly lateralized, similar to some previous studies on music-syntactic (Miranda & Ullman, 2007; Steinbeis et al., 2006; Loui et al., 2005; Koelsch, Maess, Grossmann,

& Friederici, 2003) and language-syntactic processing (Kubota, Ferrari, & Roberts, 2004; Hahne & Friederici, 2002; Knoesche et al., 1999). With regard to the lateralization of the ERAN and the ELAN, it is important to note that previous studies reported that musical (Koelsch et al., 2003) and linguistic (e.g., Pugh et al., 1996) syntax processing is rather bilateral in some women (but usually not in men). Hence, a relatively large number of subjects are required until the lateralization of the ERAN or the ELAN reaches statistical significance; consequently, studies with larger numbers of participants ( $n \geq 20$ ) usually report a lateralization of the ERAN (e.g., Koelsch & Jentschke, 2008; Koelsch & Sammler, 2008; Koelsch et al., 2007). Additional factors that modulate the lateralization of the ERAN might include the salience of irregular chords, attentional factors, and the signal-to-noise ratio of ERP data. Notably, functional neuroimaging studies consistently showed a right hemispheric weighting of the ERAN (e.g., Tillmann et al., 2003; Koelsch, Gunter, et al., 2002; Maess et al., 2001) and a left hemispheric weighting of the ELAN (e.g., Friederici et al., 2000; Knoesche et al., 1999). Thus, even if the EEG effect is sometimes not clearly lateralized, it is reasonable to assume that the neural generators of the ERAN and the ELAN are, on average, activated with a hemispheric weighting. Although ERAN and ELAN effects were not lateralized in the present study, we use the terms ERAN and ELAN here because these terms have been established for the functional significance of these ERP components, rather than for their scalp distribution. Note that similar conflicts exist for most (if not all) endogenous ERP components (see Koelsch, 2009, for further details).

Furthermore, the ELAN and the ERAN appear roughly in the same time windows as the N100 and the P200, leading to the phenomenon that ELAN potentials overlap in part with N100 potentials, and ERAN potentials overlap in part with P200 potentials. Therefore, one might argue that the ELAN and the ERAN could also be described as effects of syntactical correctness on the N100 and P200 components, respectively. However, previous studies showed that the scalp topography and the neural generators (as defining features of an ERP component, e.g., Picton et al., 2000) of the ELAN differ from those of the N100 (Friederici et al., 2000; Hahne & Friederici, 1999; Knoesche et al., 1999; for a review on N100 and its generators, see Näätänen & Picton, 1987). Similarly, the scalp topography and neural generators of the ERAN significantly differ from those of the P200 (Maess et al., 2001; see also Koelsch, Fritz, et al., 2005; Papanicolaou, Rogers, Baumann, Saydjari, & Eisenberg, 1990; Picton, Hillyard, & Galambos, 1974). Thus, it seems implausible that the ELAN and the ERAN simply reflect modulations of the auditory-evoked N100 and P200 potentials.

The ERAN inverted polarity at mastoid leads with a nose reference, similar to the MMN (despite different neural generators; Koelsch, 2009). The generation of the ERAN as well as of the MMN involves predicting subsequent acoustic events, and comparing new acoustic information with the predicted sound (Koelsch, 2009). However, the generation

of the MMN is based on representations of regularities of local intersound relationships that are extracted on-line from the acoustic environment, whereas the generation of the ERAN relies on representations of music-syntactic regularities that already exist in a long-term memory format (and often refer to long-distance dependencies involving hierarchical syntactic organization; Koelsch, 2009; see also Steinbeis & Koelsch, 2008; Koelsch, Gunter, et al., 2005; Koelsch et al., 2001). Because the music-syntactically irregular chords used in the present study did not represent a frank acoustic irregularity (Koelsch et al., 2007), it is rather unlikely that the ERAN elicited in the present study overlapped with an MMN.

## Conclusion

In summary, the results of the present study indicate that in complex auditory environments, syntactic features of music and speech are partially automatically processed, even when attention is shifted to another stimulus in the same modality. However, our data suggest that music-syntactic processes can, nevertheless, be influenced by attention. Therefore, when we are at a cocktail party, trying to follow a conversation while there is some music coming from another direction, our brain still keeps track of music-syntactic features. Vice versa, when we want to enjoy the music and try to ignore other people's utterances (for whatever reasons...), our brain, nevertheless, monitors the syntax of these utterances.

## Acknowledgments

We thank Anja Hahne for providing the language stimuli, Sebastian Jentschke and Daniela Sammler for help in data analysis and valuable discussions, Kerstin Flake for help with the images, Nikolaus Steinbeis and Arvid Herwig for helpful comments on earlier versions of this manuscript. In addition, we also would like to thank Uwe Seifert for his support.

Reprint requests should be sent to Clemens Maidhof, Cognitive Brain Research Unit, Institute of Behavioural Sciences, University of Helsinki, Siltavuorenpenger 1 B, 00014 Helsinki, Finland, or via e-mail: clemens.maidhof@helsinki.fi, or Stephan Koelsch: koelsch@fu-berlin.de.

## Note

1. The LSN was more lateralized when speech was presented from the left side: An ANOVA with the additional factor Sound direction (20° azimuth, -20° azimuth) showed an interaction between Sentence type, Hemisphere, and Sound direction [ $F(1, 18) = 7.67, p = .013$ ]. Separate ANOVAs with speech only presented from the left side and presented from the right side showed only an interaction between Sentence type and Hemisphere when speech was presented from the left side [ $F(1, 18) = 11.79, p = .003$ ], but not when presented from the right side ( $F < 1$ ).

## REFERENCES

Broadbent, D. E. (1957). A mechanical model for human attention and immediate memory. *Psychological Review*, *64*, 205–215.

- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon.
- Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*, 23–30.
- Deutsch, A., & Bentin, S. (1994). Attention mechanisms mediate the syntactic priming effect in auditory word identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 595–607.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*, 80–90.
- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–427.
- Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, *87*, 272–300.
- Duncan, J., Martens, S., & Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature*, *387*, 808–810.
- Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integrations in language and music: Evidence for a shared system. *Memory & Cognition*, *37*, 1–9.
- Flores d'Arcais, G. B. (1982). Automatic syntactic computation and use of semantic information during sentence comparison. *Psychological Research*, *44*, 231–242.
- Flores d'Arcais, G. B. (1988). Automatic processes in language comprehension. In G. Denes, C. Semenza, & P. Bisiachi (Eds.), *Perspectives on cognitive neuropsychology* (pp. 92–114). London: LEA.
- Forster, K. I. (1979). Levels of processing and the structure of the language processor. In W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing* (pp. 27–85). Hillsdale, NJ: Erlbaum.
- Frazier, L. (1987). Sentence processing. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 559–586). Hillsdale, NJ: Erlbaum.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.
- Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1219–1248.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*, 183–192.
- Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping*, *11*, 1–11.
- Gaeta, H., Friedman, D., & Hunt, G. (2003). Stimulus characteristics and task category dissociate the anterior and posterior aspects of the novelty P3. *Psychophysiology*, *40*, 198–208.
- Gunter, T. C., Friederici, A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12*, 556–569.
- Hackley, S. A. (1993). An evaluation of the automaticity of sensory processing using event-related potentials and brain-stem reflexes. *Psychophysiology*, *30*, 415–428.
- Hahne, A., & Friederici, A. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11*, 194–205.
- Hahne, A., & Friederici, A. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Brain Research, Cognitive Brain Research*, *13*, 339–356.

- Hastings, A., & Kotz, S. (2008). Speeding up syntax: On the relative timing and automaticity of local phrase structure and morphosyntactic processing as reflected in event-related brain potentials. *Journal of Cognitive Neuroscience*, *20*, 1207–1219.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- Knoesche, T. R., Maess, B., & Friederici, A. D. (1999). Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topography*, *12*, 75–87.
- Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*, 207–212.
- Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, *46*, 179–190.
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage*, *25*, 1068–1076.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schroeger, E. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, *12*, 520–541.
- Koelsch, S., Gunter, T. C., von Cramon, D. Y., 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. C., Schröger, E., Tervaniemi, M., Sammler, D., & Friederici, A. D. (2001). Differentiating ERAN and MMN: An ERP study. *NeuroReport*, *12*, 1385–1389.
- Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, *17*, 1565–1577.
- Koelsch, S., Heinke, W., Sammler, D., & Olthoff, D. (2006). Auditory processing during deep propofol sedation and recovery from unconsciousness. *Clinical Neurophysiology*, *117*, 1746–1759.
- Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: An ERP study. *Brain Research*, *1212*, 55–62.
- Koelsch, S., Jentschke, S., Sammler, D., & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*, 476–490.
- Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport*, *14*, 709–713.
- Koelsch, S., & Sammler, D. (2008). Cognitive components of regularity processing in the auditory domain. *PLoS ONE*, *3*, e2650.
- Koelsch, S., Schmidt, B.-H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*, 657–663.
- Koelsch, S., Schroger, E., & Gunter, T. C. (2002). Music matters: Preattentive musicality of the human brain. *Psychophysiology*, *39*, 38–48.
- Kubota, M., Ferrari, P., & Roberts, T. P. L. (2004). Human neuronal encoding of English syntactic violations as revealed by both L1 and L2 speakers. *Neuroscience Letters*, *368*, 235–240.
- Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition*, *11*, 539–550.
- Leino, S., Brattico, E., Tervaniemi, M., & Vuust, P. (2007). Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, *1142*, 169–177.
- Logan, G. D. (1992). Attention and preattention in theories of automaticity. *American Journal of Psychology*, *105*, 317–339.
- Loui, P., Grent-’t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Brain Research, Cognitive Brain Research*, *25*, 678–687.
- Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: An MEG study. *Nature Neuroscience*, *4*, 540–545.
- Mecklinger, A., & Ullsperger, P. (1995). The P300 to novel and target events: A spatiotemporal dipole model analysis. *NeuroReport*, *7*, 241–245.
- Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage*, *38*, 331–345.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Novak, G. P., Ritter, W., Vaughan, H. G., & Wiznitzer, M. L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, *75*, 255–275.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97–113.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *24*, 739–773.
- Papanicolaou, A. C., Rogers, R. L., Baumann, S., Saydjari, C., & Eisenberg, H. M. (1990). Source localization of two evoked magnetic field components using two alternative procedures. *Experimental Brain Research*, *80*, 44–48.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674–681.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, J. R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.
- Picton, T. W., Hillyard, S. A., & Galambos, R. (1974). Human auditory evoked potentials I. Evaluation of components. *Electroencephalography and Clinical Neurophysiology*, *36*, 179–190.
- Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, J., Frackowiak, R. S. V., et al. (1997). The structural components of music perception. A functional anatomical study. *Brain*, *120*, 229–243.
- Pugh, K., Shaywitz, B., Shaywitz, S., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1996). Cerebral organization of component processes in reading. *Brain*, *119*, 1221–1238.
- Pulvermüller, F., Shtyrov, Y., Hastings, A. S., & Carlyon, R. P. (2008). Syntax as a reflex: Neurophysiological evidence for early automaticity of grammatical processing. *Brain and Language*, *104*, 244–253.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*, 1–66.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual

- learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Slevc, L., Rosenberg, J., & Patel, A. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16, 374–381.
- Steinbeis, N., & Koelsch, S. (2008). Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, 18, 1169–1178.
- Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, 18, 1380–1393.
- Sussmann, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin & Review*, 10, 630–637.
- Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, 16, 145–161.
- Treisman, A. M. (1964). Selective attention in man. *British Medical Bulletin*, 20, 12–16.
- Woldorff, M. G., & Hillyard, S. A. (1990). Attentional influence on the mismatch negativity [commentary]. *Behavioral and Brain Sciences*, 13, 258–260.
- Woldorff, M. G., Hillyard, S. A., Gallen, C. C., Hampson, S. R., & Bloom, F. E. (1998). Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. *Psychophysiology*, 35, 283–292.