Effects of Musical Expertise and Boundary Markers on Phrase Perception in Music

Christiane Neuhaus[1,2], Thomas R. Knösche[1], and Angela D. Friederici[1]

Abstract

■ A neural correlate for phrase boundary perception in music has recently been identified in musicians. It is called music closure positive shift ("music CPS") and has an equivalent in the perception of speech ("language CPS"). The aim of the present study was to investigate the influence of musical expertise and different phrase boundary markers on the music CPS, using event-related brain potentials (ERPs) and event-related magnetic fields (ERFs). Musicians and nonmusicians were tested while listening to binary phrased melodies. ERPs and ERFs of both subject groups differed considerably from each other. Phrased melody versions evoked an electric CPS and a magnetic CPSm in musicians, but an early negativity and a less pronounced CPSm in nonmusicians, suggesting different perceptual strategies for both subject groups. Musicians seem to process musical phrases in a structured manner similar to language. Non-

musicians, in contrast, are thought to detect primarily discontinuity in the melodic input. Variations of acoustic cues in the vicinity of the phrase boundary reveal that the CPS is influenced by a number of parameters that are considered to indicate phrasing in melodies: pause length, length of the last tone preceding the pause, and harmonic function of this last tone. This is taken as evidence that the CPS mainly reflects higher cognitive processing of phrasing, rather than mere perception of pauses. Furthermore, results suggest that the ERP and MEG methods are sensitive to different aspects within phrase perception. For both subject groups, qualitatively different ERP components (CPS and early negativity) seem to reflect a top-down activation of general but different phrasing schemata, whereas quantitatively differing MEG signals appear to reflect gradual differences in the bottom-up processing of acoustic boundary markers. ■

INTRODUCTION

Successive tones make up a melody when they unfold in time with a specific rhythm and when they follow the rules of a musical system. Since the middle of the 18th century, melody has been an issue in science (Mattheson, 1737/1976). In this epoch, the term *musical phrase* was used for the first time, characterizing the caesura (cut) itself as well as the resulting segments within a melody.

For the perception of a musical phrase, the correct detection of phrase boundaries (caesuras) is necessary. Such boundaries are indicated by specific acoustic cues. For the marking of phrase boundaries, Riemann (1900) considered two musical cues as the most relevant: small pauses after a strong beat dividing a melodic entity into subunits, and the lengthening of the last tone preceding the boundary, which normally coincides with a strong beat. Other cues mentioned by Riemann include the implicit harmony of the phrase-limiting tone as well as leaps (intervals larger than a major second) immediately after a regular sequence of tone steps.

Modern psychology has also taken up the issue of perceptual segmentation and phrasing in music. Using behavioral methods, a considerable number of studies have dealt with this matter (e.g., Trainor & Adams, 2000; Wilson et al., 1999; Chiappe & Schmuckler, 1997; Krumhansl & Jusczyk, 1990; Palmer & Krumhansl, 1987; Stoffer, 1985; Tan, Aiello, & Bever, 1981; Sloboda & Gregory, 1980; Dowling, 1973). Palmer and Krumhansl (1987) investigated the roles of time and pitch structure in melodic phrase perception. They demonstrated that in musically trained adults, mental percepts of musical phrases could be explained by perceptual independence rather than by perceptual interaction of both factors. In a subsequent study of Krumhansl and Jusczyk (1990), the perceptiveness of infants to phrase boundaries was investigated. It was found that certain acoustic cues facilitate the process of phrase detection, including drops in pitch and contour as well as the lengthening of the tone preceding the phrase boundary. The authors suggested that phrase perception merely depends on the detection of certain acoustic markers and that musical expertise is not an essential precondition. Similar findings were reported by Trainor and Adams (2000). A powerful means to investigate phrase processing in both speech and music is the *click migration paradigm*

[1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, [2]Friedrich Schiller University, Jena

(Stoffer, 1985; Sloboda & Gregory, 1980; Ladefoged and Broadbent, 1960). Stoffer (1985), for instance, used a dichotic listening task with artificial songs presented to the left and clicks presented to the right ear. Participants had to mark the perceived click position. When comparing the subjective judgement with the actual click location, a migration tendency, preferentially towards higher-level phrase boundaries, could be observed. Musically trained subjects solved the click detection task faster than musically untrained ones. Differences between musicians and nonmusicians using musical phrase structure were also found by Chiappe and Schmuckler (1997). They demonstrated that in musicians, memory for tone information preceding a phrase boundary was disrupted, whereas no such effect could be found for nonmusicians. The authors concluded that musical phrases guide the parsing of musical sequences during perception and, while offering only a limited amount of melodic information, also serve as a natural means to avoid memory overload. Effects of musical expertise on musical phrase perception were also reported by Tan et al. (1981) and Dowling (1973).

Using electroencephalography (EEG) and magnetoencephalography (MEG), which provide a more direct access to the underlying neural processes, Knösche et al. (2005) identified a neural correlate for phrase boundary perception in music. While performing an out-of-key note detection task to keep attention constant, trained musicians were asked to listen to phrased piano melodies and their unphrased (but otherwise identical) counterparts. In event-related potentials (ERPs), a positive-going wave was found. It was characterized by a peak latency of approximately 550 msec after the offset of the phrase-boundary-defining pause and by a bilateral centroparietal scalp distribution. In neuromagnetic event-related fields (ERFs), a doublepeaked component stretching over a longer time window was observed, reflecting a different aspect of the underlying neuronal network. Source localization revealed that brain structures of the limbic system, in particular the anterior and posterior cingulate cortex as well as the posterior hippocampus, seem to be the most likely generators for these brain responses. Because these structures are known to be involved in memory and attention processes in general, the authors interpreted the identified components as reflecting chunking and attention redirection processes. Moreover, the observed ERP wave happened to be similar to a positive component, which had been found in response to prosodic phrase boundaries in speech (Steinhauer, 2001; Steinhauer & Friederici, 2001; Steinhauer, Alter, & Friederici, 1999). Because that component had been interpreted as a correlate of prosodic phrase closure and had been termed closure positive shift (CPS), the corresponding component for phrase closure in music was called *music CPS*.[1]

The report of Knösche et al. (2005) gives rise to two relevant questions. First, the music CPS has so far been demonstrated in trained musicians only. Although it has been observed that musicians are superior to nonmusicians in many aspects of music perception, including the processing and structuring of temporal patterns (e.g., Jongsma, Desain, & Honing, 2004; van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004; Drake et al., 2000; Chiappe & Schmuckler, 1997), there is evidence that also nonmusicians are able to process structural aspects of music quite well (Koelsch, Gunter, Schroeger, & Friederici, 2003; Koelsch, Gunter, Friederici, & Schroeger, 2000). It is therefore of great interest, whether a music CPS can also be observed in nonmusicians and whether there are any other ERP or ERF effects related to the perception of musical phrase boundaries. Moreover, differences in the early components with respect to musical expertise can be expected. Pantev et al. (1998) found an increased N1m for piano tones in musicians as compared to nonmusician controls. In contrast Shahin, Bosnyak, Trainor, and Roberts (2003) observed an enhancement of the N1c (which they observed at extreme temporal electrode sites) and of the P2, but not of the N1.

The second open question concerns the acoustic cues, which subjects might use to detect phrase boundaries in music. The two most important cues listed by Riemann (1900) are the pause and the prolonged boundary tone preceding the pause. They coincide with the two most important prosodic phrase boundary markers in spoken language, namely, pauses for breath and the lengthening of preboundary words and syllables (see, e.g., Pannekamp, Toepel, Alter, Hahne, & Friederici, 2005). Nonetheless, it is plausible that also the implicit harmonic function of the phrase-limiting tone (boundary tone) has an effect on phrase closure, in particular the tonic tone, which confirms the tonal key of a melody. Whereas we mainly expect effects on the CPS, the early components, that is, N1 and P2, might be influenced as well. In particular, the P2, which is presumably generated in the anterior planum temporale (Knösche, Neuhaus, Haueisen, & Alter, 2003), is thought to reflect a great deal of higher-order cognitive processing (for a review, see Crowley & Colrain, 2004). However, these components are also known to be subject to refractory influences (e.g., Budd, Barry, Gordon, Rennie, & Michie, 1998), which will inevitably complicate the interpretation of any results of the current study involving the factor pause length.

For the present experiments, we partially used the tone material from the previous study with musicians (Knösche et al., 2005). This material consisted of 101 melodies with a binary form—a principle of construction, which is in line with Riemann's opinion of the "bisection" as the prototype of phrased melodies. The boundary between the phrases is indicated by a pause,

but also by some properties of the last tone preceding the pause (boundary tone), namely, its length and its implicit harmonic function (tonic or dominant). In order to determine the influence these markers have onto the perception of phrase boundaries, melodies were sorted according to these cues. Because we were using natural stimuli, confounds between some cues were inevitable: Although the *implicit harmonic function* was almost completely independent of the other criteria (correlations < .1), *boundary tone length* was strongly correlated to *pause length* (correlation .4, long pauses coincided with long boundary tones and short pauses coincided with short boundary tones). For this reason, we subsumed the two factors boundary tone length and pause length under a meta-parameter, called *phrase boundary strength* (strong phrase boundaries have long pauses and long boundary tones). Thus, we examined the phrase closure effect for two (almost) independent melodic aspects: for the meta-parameter phrase boundary strength (strong vs. weak), and for the implicit harmonic function (tonic vs. dominant). In order to further specify the respective roles of the variables pause and boundary tone, another experimental manipulation was needed to disentangle the influences of these factors. For this purpose, two stimulus variants were additionally generated, so that each melody existed in altogether four versions:

(O) The original biphrasal version.
(C) The continuous version for reference and comparison, where the pause was replaced by one or several notes in such a way that the result sounded natural. All other parameters, in particular the properties of the boundary tone, were left unchanged.
(L) The long-pause condition, where the pause was extended and (inevitably, to hold the meter) the boundary tone shortened, without changing any other parameters.
(S) The short-pause condition, where the pause was shortened and the boundary tone was prolonged while keeping all other parameters constant.

Note that in the variation represented by conditions L and S, pause length and boundary tone length are correlated in a way (long pauses coincide with short boundary tones and vice versa) that is different from the sorting of the natural music stimuli, as has been described before (long pauses coincide with a long boundary tone and vice versa). This gives us two comparisons—one within-stimulus with a negative correlation between pause length and boundary tone length, and another one between-stimuli with a positive correlation between both factors—that should help to disentangle the effects on the CPS.

Stimuli were presented to two groups of subjects— musicians and nonmusicians. Because we were interested in the most natural and common situation of listening to music, we chose a task that was unrelated to the issue of phrasing. Subjects were required to detect an out-of-key tone, which was present at a random position in 10% of the melodies. These task stimuli were later excluded from analysis.

If we had chosen a task that is more directly related to the purpose of the study, for example, detecting phrase boundaries, we would possibly have introduced a bias. The observed brain processes might have been different from those taking place in a normal listening situation. Subjects would have focused their attention mainly on the temporal structure of the melodies and might have used more analytic or strategic ways to detect the boundaries and solve the task. Here, musicians might have had a stronger advantage over the nonmusicians than in a normal listening situation.

Data were collected with EEG and MEG. In an additional behavioral test, the ability of subjects to recognize the phrase boundaries in the given examples was examined.

METHODS

Subjects

Two groups of subjects, differing in their musical expertise, participated in the experiment. The first group consisted of 12 professional musicians, 6 men and 6 women (age 23 to 31 years, mean age 25.8). Subjects of this group had had their first instrumental lessons at the average age of 9 years. According to the Edinburgh Handedness Inventory (Oldfield, 1975), 11 of them were right-handed and one was left-handed. The second group consisted of 12 students with little experience in classical music (6 men and 6 women, age from 18 to 30, mean age 23.92). None of them had played an instrument or sung in a choir for at least 2 years prior to the experiment. Eleven students were right-handed, and one was left-handed. No subject had known hearing deficits or neurological disorders. Before the experiment, each volunteer gave a written consent to participate.

Stimuli

We used 101 binary phrased melodies based on various forms of key and meter. Ninety percent of them had the length of a musical period (four plus four measures) and were chosen from the baroque, classical, and romantic eras. The remaining 10% were invented in a modern style. All melodies included one principal phrase boundary approximately in the middle. Of each melody, four different versions were created, which solely differed in parameters of the phrase boundary (length of pause and length of boundary tone). Version O (phrased) denotes the original melody, whereas L (long pause) and S (short pause) are versions with prolonged and shortened pauses, respectively. In

version C (unphrased), the pause was eliminated altogether by filling it up with notes. In phrased melody versions, the average pause length was 498 msec for condition O, 942 msec for condition L, and 112 msec for condition S (Figure 1 gives note examples for all conditions).

Each melody was played and recorded in a piano timbre on a programmable keyboard. For each melody, a target version was created by replacing an arbitrarily selected tone with a false (dissonant) one. For presentation, all MIDI files were transformed to audio format and the amplitude of the sound envelopes was equalized. The stimuli were mixed in a pseudo-randomized manner, where ordering was balanced to counter sequence effects. Ten percent of the stimuli were replaced by their respective target version.

In order to investigate if the CPS was modulated by acoustic cues in the close vicinity of the boundary, we sorted the stimuli into classes according to the following criteria: (1) length of boundary tone and (2) harmonic ending of the first phrase. The length of the boundary tone correlated strongly with the length of the pause (see introduction), leading to the definition of the meta-parameter *phrase boundary strength*. We created two equal-sized classes of stimuli: with *strong phrase boundaries* (long pauses and long boundary tones) and with *weak phrase boundaries* (short pauses and short boundary tones). With respect to the harmonic ending of the first phrase, we formed two approximately equal-sized classes of stimuli: with *tonic endings* and with *dominant endings*.

To test the listening expertise of each subject group, an additional rating test was performed with a subset of the participants, using all stimuli of one EEG or MEG session. Six musicians and eight nonmusicians answered the following questions: ''Have you ever heard this melody somewhere else before?'' ''How many melodic sections (phrases) have you noticed?'' ''If there was more than one phrase, was the impression of the pause clear or unclear for you?'' For the evaluation of pause clarity, subjects made use of a 10-point scale.

Paradigm and Procedure

Brain activity was measured in two EEG and two MEG sessions. In each recording session, 186 different melody versions split up into six blocks of 31 tunes each were presented. Subjects did not get any particular information about the stimulus structure and the original purpose of the study.

For EEG recording, subjects were comfortably seated in a soundproof chamber. Melodies were played via loudspeaker placed at a distance of 1.5 m. Before the first block was run, stimulus intensity and the centering of the sound source were adjusted. Subjects were asked to relax their facial muscles, to keep their hands, arms, and fingers as motionless as possible, and to fixate a cross at the center of the screen.

Figure 1. Example score of a melody with manipulations of pause length. Daniel Gottlieb Türk ''Crossing with the hand'' (from *Easy Pieces in C and G Major*). ''O'': original version, two phrases. ''C'': continuous version, pause filled with notes. ''L'': original pause lengthened, boundary tone shortened. ''S: original pause shortened (in this example, only caesura), boundary tone lengthened.

In the MEG experiment, participants were requested to lie on a stretcher in a magnetically shielded room. They had to focus their eyes on a monitor positioned 50 cm above the head. Melodies were binaurally presented via ear tubes.

After recording, participants had to fill out a short questionnaire about their current mental state and their impression about the experimental run. For stimulus presentation, we used the software package ERTS (Experimental Run Time System, Version 3.11; BeriSoft 1995). Each trial started with a fixation cross. After 2 sec, a melody was presented. During the response period, the question ''False tone?'' was displayed. For a positive answer, subjects had to press the left key button of the response device and vice versa (maximum response time, 6 sec).

Recording

Brain electrical activity was recorded with Ag/AgCl electrodes inserted into a special electrode cap attached to the scalp (Electro Cap International Inc., Eaton, OH). According to the 10-10 system, 25 electrode sites were chosen (see, e.g., Oostenveld & Praamstra, 2001). Position C2 in the right hemisphere was used for the ground electrode, and the nose was the electrically inactive reference point. For registering blink artifacts, a vertical and a horizontal electrooculogram were recorded from above and below the right eye and from the outer canthus of both eyes, respectively. Electrode impedance was kept below 5 kΩ, and the time constant was infinite as direct current amplifiers were used. For MEG recordings, a 148-channel whole-head magnetometer system was used (MAGNES WHS-2500, 4D Neuroimaging, San Diego, CA). The head position was computed using five coils attached to specific anatomical sites, which were then localized by both the MEG device and a 3-D digitizer (Polhemus 3Space Fastrak 3SF0002, Colchester, VT).

Data Analysis

Preprocessing of EEG Data

EEG data were digitized on-line with a sampling rate of 250 Hz. Trigger points marking the onset of the melodies were automatically recorded. Additional trigger points at the onsets of the last tone of the first phrase and the first tone of the second phrase were computed off-line (for the unphrased version, analogous time points were chosen). The last tone of the first phrase was the last fixed point in time, where both phrased and unphrased melody versions were still absolutely identical. The onset of the first tone of the second phrase marks the end of the period, where both versions are different. Due to the large variability in pause length, activity time-locked to one of these events would be more or less blurred when averaged with respect to the

other one. Therefore, both trigger points are necessary to capture all possible effects of the processing of the phrase boundary.

EEG raw data were filtered off-line with a 0.50-Hz high-pass filter. For each subject, only those trials were averaged that were not contaminated by artifacts (e.g., by eye blinks larger than 50 µV or muscle activity). Data sets with more than 50% rejected trials were excluded from analysis altogether, resulting in 15 data sets for 10 musicians and 15 data sets for 11 nonmusicians being available for analysis. For averaging, time windows of −0.2 to 1.2 sec around the respective trigger points were chosen.

Baseline correction is a technique used to remove brain responses that have nothing to do with the event of interest. In most cases, the chosen baseline interval immediately precedes the trigger point, which usually is the onset of the stimulus. This technique is based on the assumption that in the time interval immediately preceding the stimulus, there are no transient brain responses, but only long-time potential shifts, which would last almost unchanged over the entire time interval of interest. Also, and above all, there should be no possibility of differential activity between the experimental conditions in the baseline window. Normally, a flat and close to zero baseline trace for all conditions is a good indication for a successful baseline correction. The first assumption (no transient activity in baseline interval) cannot be held here, because the trigger points are situated somewhere in the middle of an ongoing piece of music. The second assumption (no differential activity in baseline window) can be easily fulfilled for the first trigger (onset of last tone in first phrase) by using a window of, for example, 200 msec prior to the trigger point. If we, however, use the same technique for the second trigger (onset of first tone in the second phrase), there is the possibility of differential activity in the baseline window, because during this time period, there was silence in the phrased condition, but tones being played in the unphrased condition. Alternatively, one could use a baseline interval prior to the offset of the first phrase. This ''detached baseline'' should guarantee that there is no differential activity in the baseline window, but could allow some longer lasting activity from the pause interval to contaminate the response to the first note of the second phrase. Moreover, it can be shown that for the detached baseline, averages and statistics are more contaminated with noise. Hence, there is no perfect solution. Fortunately, however, it turned out that both choices of baseline yield almost identical results in most cases. We therefore present here the results using a 200-msec baseline interval immediately preceding the trigger point, because this technique is more common and yields traces that are better aligned prior to and at the trigger point. In cases, where the detached baseline yielded different results, this is reported explicitly.

Finally, individual averages were combined condition-wise to grand averages for display and interpretation.

Preprocessing of MEG Data

For MEG analysis, raw data were digitized on-line with a sampling rate of 508 Hz. Using the same high-pass filtering and artifact rejection techniques as for EEG (eye blink rejection threshold, 30 fT), we obtained 18 data sets for nine musicians and 14 data sets for nine nonmusicians. As it is not possible to restore a certain position of the magnetometer array with precision after the subject has moved, subject's data of different sessions and blocks were first averaged within blocks (per condition) and were then interpolated to a set of average sensor positions using a method based on linear inverse techniques (Knösche, 2002). Because of this procedure, values from different blocks represent the magnetic field at the same positions with respect to the head; they were then averaged over blocks for each of the two sessions. Subsequent processing, including choices of trigger points and baseline windows, was identical to the processing of EEG data.

Statistical Analysis

Grand average curves were first analyzed by visual inspection in order to identify time windows of interest. For each of the identified time windows, several analyses of variance (ANOVAs) were performed. First, a three-way ANOVA was computed to test the general phrase boundary effect. Within-subject factors were COND (condition: phrased vs. continuous), VERT (anterior–posterior topography: anterior—F3, FZ, F4; central—C3, CZ, C4; posterior—P3, PZ, P4), and LAT (left–right topography: left—F3, C3, P3, midline—FZ, CZ, PZ, right—F4, C4, P4). In addition, a 4-way ANOVA was performed to examine group differences between musicians and nonmusicians regarding the phrase boundary effect. In this design, "musical expertise" was added as a between-subjects factor (musicians vs. nonmusicians). A second four-way ANOVA was computed to study the effect of acoustic cues. Here, "stimulus class" was added as a fourth within-subject factor (e.g., "tonic" vs. "dominant" for the implicit harmonic function of the boundary tone). Tests were done separately for musicians and nonmusicians.

MEG data were analyzed in a similar way, with two exceptions. First, both topographic factors consisted of two instead of three levels. Thus, all 148 MEG channels, which were equally distributed over the helmet surface, were subsumed to four regions of interest (left anterior, right anterior, left posterior, and right posterior). Second, for those mean amplitude values that belonged to the right half of the brain, the sign was reversed. This way, pseudo-interactions with LAT caused by certain peculiarities of MEG topographies could be avoided.[2] For

each chosen time window, the mean amplitude value was taken as the dependent variable. Degrees of freedom were corrected with Huynh and Feldt's epsilon.

RESULTS

Behavioral Data for Musicians and Nonmusicians

Subjects were asked to press a button to indicate whether the previously presented musical piece included a false tone or not (see Methods section). In most cases, both musicians and nonmusicians solved the task correctly. Small differences between groups were not significant. (For details, refer to Table 1.)

Fourteen participants (6 musicians and 8 nonmusicians) also took part in an additional rating test to reexamine phrase perception after recording. For evaluation of phrased and unphrased melody versions, subjects were asked questions on stimulus familiarity, on the number of perceived melodic sections (phrases), and on the clarity of the perceived phrase boundary. For stimulus familiarity, rating differences between groups were significant. (For detailed results, see Table 2.)

Musicians vs. Nonmusicians and the Effect of Phrasing

Averages with respect to the onsets of the boundary tone and the first tone in the second phrase were computed (see Methods section). For the latter trigger point, we used two different baselines: the detached baseline, −1000 to −500 msec prior to the onset of the pause and the "direct baseline" −200 to 0 msec prior to the trigger point (offset of the pause). The reported and depicted results refer to the latter baseline. (Fundamentally different results for the other baseline are mentioned in the text.)

Averages with respect to the onset of the boundary tone did not yield any significant effect, except for a slight negative variation (maximum ca. −1 μV, broad scalp distribution) for the unphrased compared to the

Table 1. Button Press Responses of Musicians and Nonmusicians when Detecting Out-of-Key Notes

	Musicians	Nonmusicians	Significance (p)
Hits	244	223	0.07
Correct rejections	30	27	0.22
False alarms	23	45	0.07
Misses	7	11	0.07
Omissions	2	0	0.39

Altogether, 306 stimuli were analyzed. The significance of differences between the groups was evaluated by a two-tailed unpaired *t* test.

Table 2. Rating Results of Musicians and Nonmusicians when Listening to 153 Melodies, Judging "Stimulus Familiarity," "Number of Phrases" and "Clarity of Phrase Boundary"

	Musicians (6 Subjects)	Nonmusicians (8 Subjects)	Significance (p)
Familiarity of melodies			
Familiar	13	76	0.018*
Unfamiliar	140	77	0.014*
Reported number of phrases for continuous version (C)			
Single phrase	10	13	0.452
Two phrases	19	17	0.689
More than two phrases	2	1	0.287
Reported number of phrases for original version (O)			
Single phrase	8	9	0.824
Two phrases	25	24	0.741
More than two phrases	1	0	0.686
Clarity of phrase boundary for O version (on a scale from 1 to 10)			
≤5	9	11	0.603
>5	23	18	0.370
Omissions	2	4	0.572
Reported number of phrases for long pause version (L)			
Single phrase	6	7	0.650
Two phrases	29	26	0.314
More than two phrases	1	3	0.193
Clarity of phrase boundary for L version (on a scale from 1 to 10)			
≤5	9	13	0.434
>5	26	19	0.302
Omissions	1	4	0.337
Reported number of phrases for short pause version (S)			
Single phrase	7	9	0.519
Two phrases	25	23	0.570
More than two phrases	2	2	0.909

Table 2. (continued)

	Musicians (6 Subjects)	Nonmusicians (8 Subjects)	Significance (p)
Clarity of phrase boundary for S version (on a scale from 1 to 10)			
≤5	9	11	0.490
>5	24	19	0.394
Omissions	1	4	0.429

A two-tailed unpaired *t* test was used to evaluate the significance of differences between groups. Results are listed separately for conditions C, O, L, and S. Task items (versions with out-of-key notes) are only included in the familiarity judgement.

*p < .05.

phrased version during the first 100 msec post trigger. This result was very surprising because the earliest differences between both melody versions started around 130 msec after the onset of the phrase final tone (boundary tone). Post hoc analysis revealed that slight imbalances in the stimulus sequences due to the pseudorandom ordering of stimulus conditions were most probably responsible for this finding. Unphrased compared to phrased items were slightly more likely to be preceded by a succession of phrased ones. For a more thorough investigation of this effect, another study would be necessary because the present material does not contain enough stimuli for reliable statistics.[3] Thus, in the following, only ERP and MEG results related to the onset of the second phrase are reported.

Brain responses elicited by phrased (condition O) and unphrased (condition C) melody versions differed between musicians and nonmusicians (Figure 2; see Table 3a and b for within-group statistics and Table 4 for comparisons between groups). In response to phrased melody versions, musicians showed a centroparietally distributed positive ERP component between 400 and 600 msec after pause offset (CPS; also significant interaction between phrasing condition and subject group, 400–600 msec) and a double peak ERF wave in frontocentral areas of the sensor array extending from 400 to 700 msec (CPSm; also significant interaction between phrasing condition and group, 500–700 msec). The magnetic effect is apparently right lateralized, although significant only for the later time window between 600 and 700 msec. ERP curves for nonmusicians revealed a negative component roughly between 300 and 500 msec with a broadly distributed, left lateralized frontocentral topography, but no later positive CPS-like effect (also significant interaction between phrasing condition and subject group, 400–500 msec). In neuromagnetic recordings, nonmusicians showed a response similar to the one in musicians (CPSm), although amplitude values, in particular at right frontal channel positions, were significantly smaller.
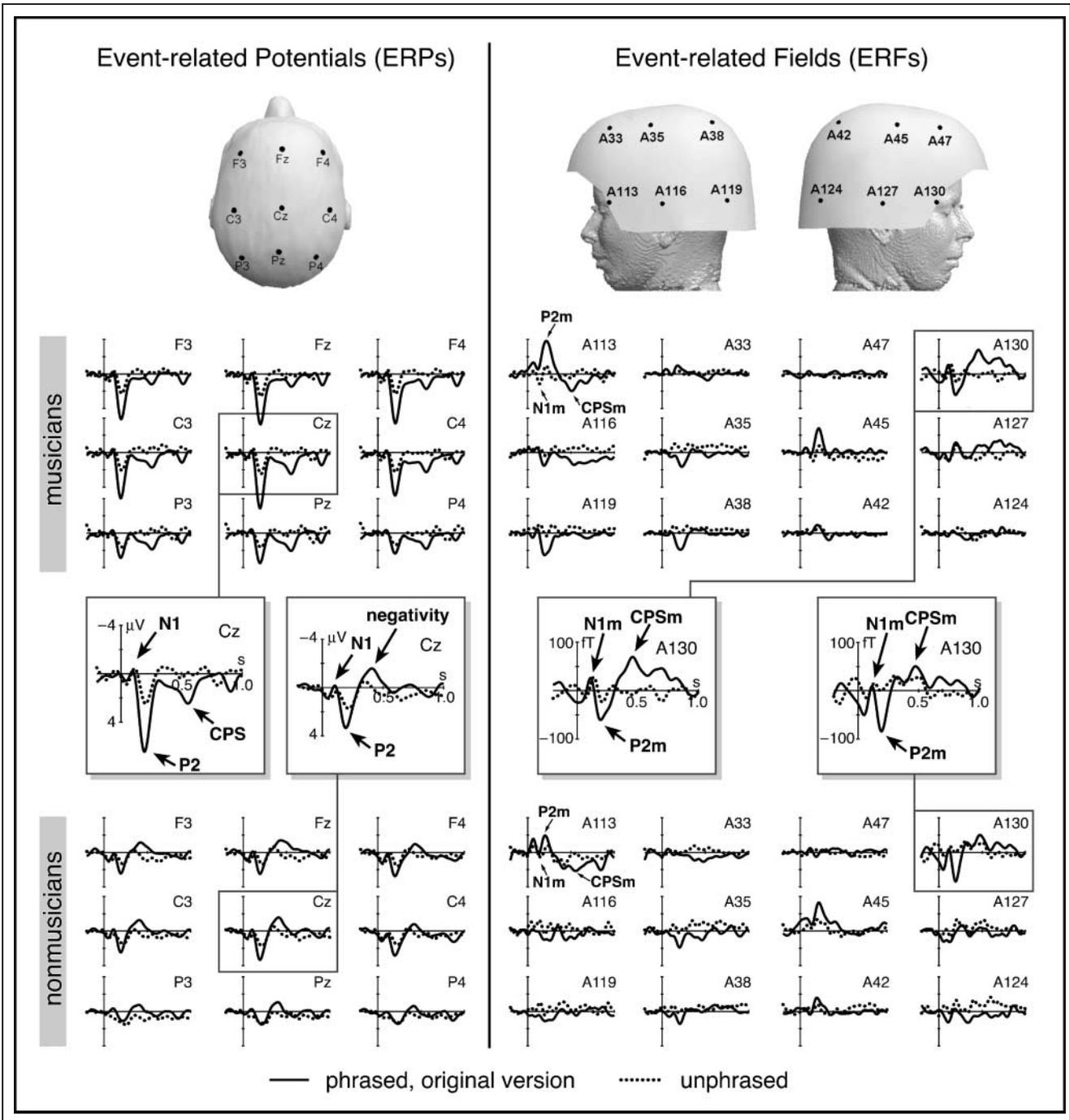
Figure 2. Grand average ERPs and ERFs at selected channels for phrased and unphrased melodies.

The P2 component was generally stronger for the phrased compared to the unphrased items. Furthermore, it was larger for musicians than for nonmusicians. The P2 featured highest amplitude values in frontal areas of the brain, thus differing in topographic distribution from the more parietally peaking CPS. For the P2, there was no significant interaction between phrasing condition and subject group. The magnetic counterpart of the P2, the P2m, was also diminished for nonmusicians as compared to musicians, but selectively only for the phrased condition and mainly in the left hemisphere (see Figure 2; interaction COND × GROUP × VERT × LAT).

Using the detached baseline, we additionally observed main effects for both subject group and phrasing condition in the N1 time window of the EEG data. Because these effects depended upon the choice of baseline, it seems likely that they were due to some sustained effect originating from the pause interval.

Table 3. Results of Repeated-Measures ANOVAs: Effect of Phrasing [Phrased vs. Unphrased (O–C)]

	80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
a. Musicians						
COND						
ERP (1/14)[a]	$F = 1.12$	$F = 27.44$, $p < .0001$	$F = 3.22$, $p = .094$	$F = 4.69$, $p = .048$	$F = 9.41$, $p = .0083$	$F = 1.19$
ERF (1/16)	$F < 1$	$F = 2.04$	$F < 1$	$F = 9.23$, $p = .0078$	$F = 10.97$, $p = .0044$	$F = 26.30$, $p = .0001$
COND × VERT						
ERP (2/28)	$F < 1$	$F = 8.42$, $p = .008$	$F = 1.20$	$F < 1$	$F = 1.11$	$F < 1$
ERF (1/16)	$F = 10.20$, $p = .0056$	$F = 14.64$, $p = .0015$	$F = 2.41$	$F = 11.70$, $p = .0035$	$F = 3.59$, $p = .076$	$F = 10.80$, $p < .0047$
COND × LAT						
ERP (2/28)	$F < 1$	$F = 1.83$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
ERF (1/16)	$F < 1$	$F = 2.60$	$F < 1$	$F = 1.15$	$F = 1.06$	$F < 1$
COND × VERT × LAT						
ERP (4/56)	$F < 1$	$F = 4.72$, $p = .0034$	$F = 2.04$	$F = 1.46$	$F = 1.05$	$F = 1.23$
ERF (1/16)	$F = 2.38$	$F = 19.18$, $p = .0005$	$F = 3.24$, $p = .091$	$F = 1.26$	$F < 1$	$F = 4.66$, $p = .046$
b. Nonmusicians						
COND						
ERP (1/14)	$F = 1.37$	$F = 1.82$	$F < 1$	$F = 3.05$	$F < 1$	$F < 1$
ERF (1/12)	$F = 1.15$	$F < 1$	$F = 1.55$	$F = 1.29$	$F < 1$	$F < 1$
COND × VERT						
ERP (2/28)	$F < 1$	$F = 5.74$, $p = .013$	$F = 1.84$	$F = 1.26$	$F = 3.39$, $p = .067$	$F < 1$
ERF (1/12)	$F = 2.58$	$F = 5.42$, $p = .037$	$F = 6.97$, $p = .020$	$F = 10.41$, $p = .0066$	$F = 19.17$, $p = .0007$	$F = 22.75$, $p = .0004$
COND × LAT						
ERP (2/28)	$F = 1.78$	$F < 1$	$F = 5.62$, $p = .009$	$F = 3.11$, $p = .074$	$F = 3.70$, $p = .05$	$F = 2.88$, $p = .073$
ERF (1/12)	$F = 2.13$	$F = 1.50$	$F < 1$	$F = 5.59$, $p = .034$	$F = 3.69$, $p = .077$	$F = 3.11$
COND × VERT × LAT						
ERP (4/56)	$F = 1.33$	$F < 1$	$F < 1$	$F < 1$	$F = 1.16$	$F = 1.48$
ERF (1/12)	$F = 3.58$, $p = .081$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 3.30$, $p = .093$

[a]Degrees of freedom.

Table 4. Results of Repeated Measures ANOVAs: Effects of Musical Expertise, Musicians vs. Nonmusicians, Factor MUSEXP [Phrased vs. Unphrased (O–C)]

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
MUSEXP	ERP (1/28)	$F < 1$	$F = 11.68$, $p = .002$	$F < 1$	$F = 1.38$	$F = 1.36$	$F < 1$
	MEG (1/28)	$F = 2.60$	$F = 1.02$	$F < 1$	$F = 1.12$	$F < 1$	$F < 1$
COND	ERP (2/28)	$F = 2.49$	$F = 15.97$, $p = .0004$	$F < 1$	$F < 1$	$F = 1.77$	$F < 1$
	MEG (2/28)	$F = 1.42$	$F < 1$	$F = 2.30$	$F = 6.78$, $p = .014$	$F = 5.77$, $p = .023$	$F = 5.56$, $p = .025$
COND × MUSEXP	ERP (1/28)	$F < 1$	$F = 3.02$, $p = .093$	$F = 2.50$	$F = 7.03$, $p = .013$	$F = 3.93$, $p = .057$	$F = 1.68$
	MEG (1/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 3.86$, $p = .065$	$F = 10.29$, $p = .0033$
COND × MUSEXP × VERT	ERP (2/56)	$F < 1$	$F = 1.16$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/28)	$F < 1$	$F < 1$	$F = 2.45$	$F < 1$	$F = 1.23$	$F < 1$
COND × MUSEXP × LAT	ERP (2/56)	$F < 1$	$F < 1$	$F = 1.45$	$F = 1.62$	$F = 1.92$	$F < 1$
	MEG (1/28)	$F < 1$	$F < 1$	$F < 1$	$F = 6.58$, $p = .016$	$F = 2.07$	$F = 2.99$, $p = .095$
COND × MUSEXP × LAT × VERT	ERP (4/112)	$F = 1.38$	$F = 1.38$	$F = 1.04$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/28)	$F < 1$	$F = 5.23$, $p = .027$	$F < 1$	$F < 1$	$F < 1$	$F < 1$

Influence of Phrase Boundary Strength

ERP and ERF effects for the strength of the phrase boundary (characterized by pause length and the length of the prefinal tone of the first phrase—boundary tone) are depicted in Figure 3 (for statistics, see Table 5a and b). In musicians, the CPS appears earlier (about 400–500 msec) and weaker for weak phrase boundaries (shorter boundary tones and shorter pauses) than for strong phrase boundaries (longer boundary tones and longer pauses). In the corresponding neuromagnetic patterns, such a latency shift between conditions could not be observed. Instead, for strong phrase boundaries, we observed a CPSm also for the unphrased condition (most pronounced in the right hemisphere). However, because this effect depends on the choice of baseline (with detached baseline, it is not present), it is probably influenced by some pretrigger activity. Conclusions have therefore to be drawn with caution.

For nonmusicians, the negative shift in the grand average ERPs was smaller for weak than for strong phrase boundaries. ERF responses of nonmusicians, however, showed less clear amplitude differences between strong and weak phrase boundaries at all channels. Moreover, results depended strongly upon the choice of baseline, indicating a substantial interaction between the time windows of interest and the baseline time window prior to the trigger point.

For both groups of subjects, there seemed to be stronger N1 components for strong as compared to weak phrases boundaries, probably due to different refractory effects for different pause length in both classes of stimuli. This effect could not be demonstrated for the N1m.

Disentangling Phrase Boundary Strength: Dissociation between the Influences of Pause Length and Boundary Tone Length

We analyzed brain responses for stimulus variants manipulated in pause length, that is, for melodies with a prolonged pause (condition L), with original pause length (condition O) and with pause reduction (condition S). In order to keep the meter, the length of the boundary tone had to be changed into the opposite direction. This way, we obtained a negative correlation between pause length and boundary tone length (long pauses correspond with short boundary tones, condition L and vice versa, condition S) in contrast to the
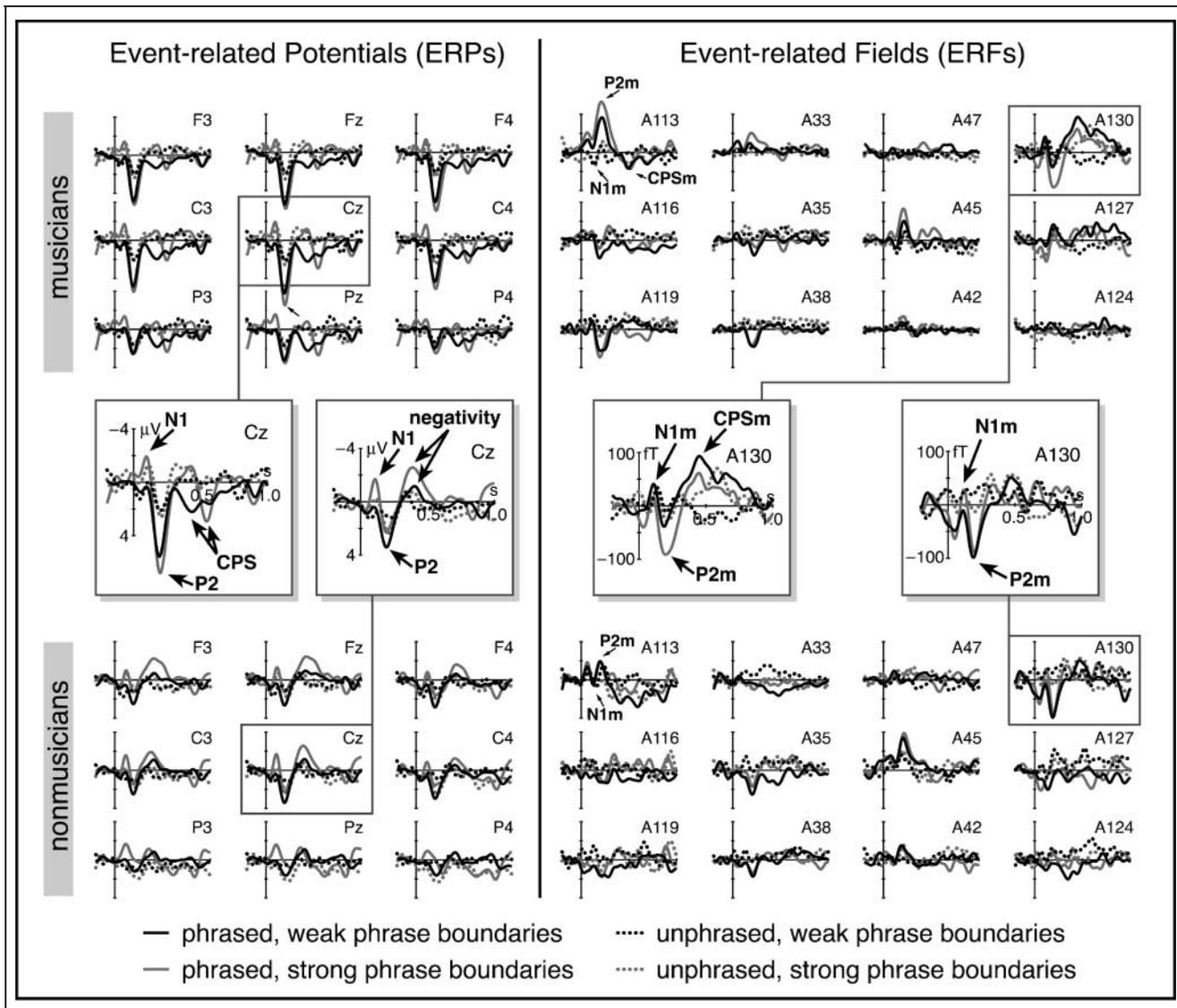
Figure 3. Grand average ERPs and ERFs at selected channels for examples with weak (red) and strong (green) phrase boundaries. For the locations of the displayed channels, see Figure 2, top row.

positive correlation by a natural sorting of melodies as used in the previous subsection (where long pauses tended to correspond to long boundary tones and vice versa). This enabled us to dissociate the influences of the two factors. ERP and ERF traces for pause length variation are depicted in Figure 4; for statistics refer to Table 6a and b.

For musicians, we obtained the following results. Whereas the amplitudes of the N1 turned out to be sorted according to the pause length (short < original < prolonged pauses), the apparently reduced P2 for long pauses (Figure 4) was not significant. Moreover, we found a reduced CPS for both shortened and prolonged pauses. This effect was most pronounced at central sites and over the right hemisphere (interaction COND × VERT × LAT; see Figure 4 and Table 6a). Differences between shortened and lengthened pauses could not

be found for the direct baseline. However, a larger reduction for short pauses was observed if the baseline interval before the offset of the first phrase was used (see detached baseline). Neuromagnetic patterns of musicians revealed a double peak CPSm, which was maximal in (mainly right) frontal regions of the sensor array and had different topographies for different pause lengths. The amplitudes for shortened pauses were lowest.

For nonmusicians, we found a larger N1 for prolonged pauses and no differences for the P2. Moreover, nonmusicians showed a distinct early negativity only for the original pause. Prolonged and shortened pause lengths did not evoke any significant component. Similar to musicians, ERF responses of nonmusicians revealed different amplitude values for different pause conditions, which were more distinct at frontal channel

Table 5. Results of Repeated Measures ANOVAs: Effects of Phrase Boundary Strength [Phrased vs. Unphrased (O–C)]

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
a. Musicians							
COND × STRENGTH	ERP (1/14)	$F = 3.86$, $p = .070$	$F < 1$	$F < 1$	$F = 4.87$, $p = .044$	$F < 1$	$F = 1.76$
	MEG (1/16)	$F < 1$	$F = 2.14$	$F < 1$	$F < 1$	$F < 1$	$F = 1.92$
COND × STRENGTH × VERT	ERP (2/28)	$F < 1$	$F = 1.12$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/16)	$F < 1$	$F = 1.59$	$F < 1$	$F = 1.36$	$F = 1.88$	$F = 8.43$, $p = .010$
COND × STRENGTH × LAT	ERP (2/28)	$F = 1.00$	$F < 1$	$F = 3.60$, $p = .053$	$F = 3.24$, $p = .080$	$F < 1$	$F < 1$
	MEG (1/16)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 4.72$, $p = .045$	$F = 9.36$, $p = .0075$
COND × STRENGTH × LAT × VERT	ERP (4/56)	$F < 1$	$F = 1.05$	$F < 1$	$F < 1$	$F = 2.80$, $p = .036$	$F < 1$
	MEG (1/16)	$F < 1$	$F = 1.73$	$F < 1$	$F = 1.67$	$F < 1$	$F < 1$
b. Nonmusicians							
COND × STRENGTH	ERP (1/14)	$F = 1.15$	$F = 1.93$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/13)	$F = 1.14$	$F < 1$	$F = 2.28$	$F = 2.30$	$F = 1.59$	$F < 1$
COND × STRENGTH × VERT	ERP (2/28)	$F = 2.73$, $p = .096$	$F < 1$	$F < 1$	$F = 4.05$, $p = .053$	$F < 1$	$F = 1.48$
	MEG (1/13)	$F = 3.93$, $p = .069$	$F < 1$	$F = 4.12$, $p = .063$	$F = 2.53$	$F = 5.53$, $p = .036$	$F < 1$
COND × STRENGTH × LAT	ERP (2/28)	$F = 1.80$	$F = 4.46$, $p = .036$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/13)	$F < 1$	$F < 1$	$F = 7.82$, $p = .015$	$F = 4.84$, $p = .046$	$F = 3.05$, $p = .10$	$F = 2.74$
COND × STRENGTH × LAT × VERT	ERP (4/56)	$F < 1$	$F < 1$	$F < 1$	$F = 1.12$	$F = 1.07$	$F = 1.03$
	MEG (1/13)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 2.08$	$F = 4.57$, $p = .052$

positions of the left than for those of the right hemisphere. Frontal channels of the left hemisphere showed a CPSm with higher amplitude values for the original pause and slightly lower ones for the prolonged and short-pause conditions.

Furthermore, to differentiate more precisely between the modulating effects of pause length and boundary tone duration on phrase perception, we compared pause length and ERP amplitude values for condition L (longer pauses and shorter boundary tones) with those for the strong phrase boundary class (natural sorted melodies with longer pauses and longer phrase boundary tones). Melodies with strong phrase boundaries had an average pause length of 589 msec and an aver-

age boundary tone length of 850 msec. They evoked a CPS of 2.51 μV^4 in musicians and an early negativity of approximately -1.74 μV^4 in nonmusicians. Stimulus versions of the L condition had an average pause length of 942 msec and a boundary tone length of only 110 msec, which caused a CPS of 1.39 μV in musicians and an early negativity of $+0.08$ μV (i.e., no negativity at all) in nonmusicians. Thus, the L condition, featuring longer pause but shorter boundary tones, produces smaller effects than the melody subset with strong phrase boundaries. This indicates that besides the pause, the length of the phrase boundary tone also influences the amplitudes of the CPS and early negativity.
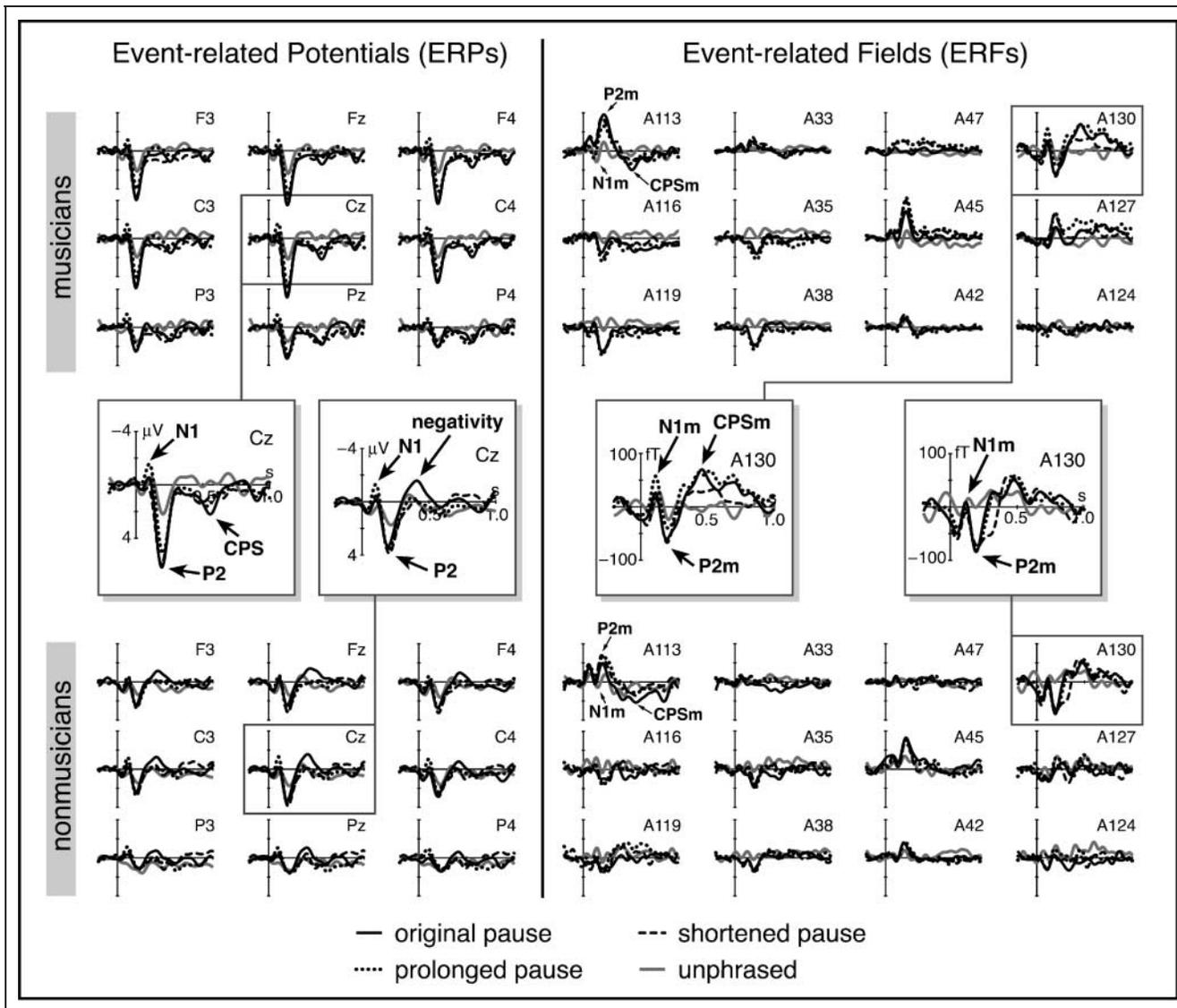
Figure 4. Grand average ERPs and ERFs at selected channels for examples with original (black), prolonged (red), and shortened (green) pause lengths. Boundary tone lengths are varied the opposite direction. For the locations of the displayed channels, see Figure 2, top row.

Influence of Harmonic Closure

Waveforms for phrase boundaries characterized by harmonic closure on the tonic or on the dominant are depicted in Figure 5; for statistics see Table 7a and b. Although the curves indicate differences for the N1 time window, these effects did not prove reliable in the statistics (Table 7). ERPs for musicians revealed distinct differences in the distributions and morphology of the P2 and CPS components. Apparently, strong closure on the tonic produces significantly more posterior distributions than closure on the dominant, which exhibits a quite frontal CPS. Moreover, the CPS on tonic endings seems to feature an additional earlier peak around 400 msec. ERF results for musicians show no clear differences between the two kinds of harmonic endings.

For nonmusicians, there were very clear effects in the time window of the early negativity, especially concerning the unphrased condition. For the dominant ending, there is an early negativity similar to the one in the phrased condition, whereas for tonic endings, a strong positivity around 400 msec was observed. Magnetic recordings did not show a similar effect.

When using the detached baseline (-1000 to -500 msec preboundary), the grand average curves remain qualitatively the same, although they did not reach statistical significance.

DISCUSSION

We analyzed ERP and ERF responses of musicians and nonmusicians to phrased and unphrased melody versions. In addition, we examined if variation of specific acoustic cues, that is, pause length, length, and harmonic function

Table 6. Results of Repeated Measures ANOVAs: Effect of Pause Length

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
a. Musicians							
Original vs. shortened vs. lengthened pause (O–S–L)							
COND	ERP (2/28)	$F = 4.50$, $p = .022$	$F = 1.55$	$F < 1$	$F = 1.12$	$F < 1$	$F = 1.66$
	ERF (2/32)	$F < 1$	$F = 4.76$, $p = .024$	$F = 3.67$, $p = .043$	$F = 3.01$, $p = .080$	$F = 1.19$	$F < 1$
COND × VERT	ERP (4/56)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	ERF (2/32)	$F = 5.18$, $p = .016$	$F = 3.07$, $p = .062$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × LAT	ERP (4/56)	$F < 1$	$F < 1$	$F = 1.32$	$F = 1.04$	$F = 1.58$	$F = 1.08$
	ERF (2/32)	$F < 1$	$F = 1.26$	$F < 1$	$F = 1.24$	$F < 1$	$F < 1$
COND × VERT × LAT	ERP (8/112)	$F = 3.67$, $p = .0008$	$F < 1$	$F = 1.35$	$F = 1.15$	$F = 1.94$, $p = .085$	$F = 2.59$, $p = .030$
	ERF (2/32)	$F = 1.05$	$F < 1$	$F = 2.59$, $p = .093$	$F = 2.13$	$F = 6.11$, $p = .0073$	$F = 1.75$
Original vs. shortened pause (O–S)							
COND	ERP (1/14)	$F = 1.30$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 2.98$
	ERF (1/16)	$F < 1$	$F = 6.77$, $p = .019$	$F < 1$	$F = 3.55$, $p = .078$	$F < 1$	$F < 1$
COND × VERT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	ERF (1/16)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F = 1.73$	$F < 1$	$F = 2.69$, $p = .098$	$F = 2.57$, $p = .095$
	ERF (1/16)	$F < 1$	$F = 2.73$	$F < 1$	$F = 1.71$	$F < 1$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F = 3.14$, $p = .031$	$F < 1$	$F < 1$	$F < 1$	$F = 1.39$	$F = 3.92$, $p = .011$
	ERF (1/16)	$F = 1.48$	$F < 1$	$F = 1.13$	$F = 2.39$	$F = 4.89$, $p = .042$	$F < 1$
Shortened vs. lengthened pause (S–L)							
COND	ERP (1/14)	$F = 6.14$, $p = .027$	$F < = 1.12$	$F = 1.22$	$F = 1.80$	$F < 1$	$F = 2.19$
	ERF (1/16)	$F < 1$	$F = 1.28$	$F = 4.43$, $p = .051$	$F = 1.20$	$F < 1$	$F < 1$
COND × VERT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	ERF (1/16)	$F = 6.33$, $p = .023$	$F = 4.33$, $p = .054$	$F < 1$	$F < 1$	$F = 1.0$	$F < 1$
COND × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F = 2.49$	$F = 2.44$	$F = 2.37$	$F < 1$
	ERF (1/16)	$F = 3.26$, $p = .090$	$F < 1$	$F < 1$	$F = 1.66$	$F < 1$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F = 4.53$, $p = .0031$	$F < 1$	$F < 1$	$F = 1.10$	$F = 1.92$	$F = 1.70$
	ERF (1/16)	$F < 1$	$F < 1$	$F = 1.53$	$F < 1$	$F = 1.63$	$F = 3.99$, $p = .063$

Table 6. (*continued*)

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
Original vs. lengthened pause (O–L)							
COND	ERP (1/14)	$F = 4.74$, $p = .047$	$F = 3.46$, $p = .084$	$F < 1$	$F < 1$	$F = 2.09$	$F < 1$
	ERF (1/16)	$F < 1$	$F = 7.17$, $p = .016$	$F = 5.53$, $p = .032$	$F = 4.41$, $p = .052$	$F < 1.98$	$F = 1.08$
COND × VERT	ERP (2/28)	$F < 1$	$F = 1.04$	$F = 1.27$	$F < 1$	$F < 1$	$F < 1$
	ERF (1/16)	$F = 8.0$, $p = .012$	$F = 4.30$, $p = .055$	$F < 1$	$F = 1.27$	$F < 1$	$F = 2.00$
COND × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F = <1$	$F < 1$	$F = 1.28$
	ERF (1/16)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F = 3.29$, $p = .028$	$F < 1$	$F = 1.30$	$F = 1.94$	$F = 2.53$, $p = .060$	$F = 2.50$, $p = .065$
	ERF (1/16)	$F = 1.19$	$F = 1.20$	$F = 514$, $p = .037$	$F = 4.29$, $p = 0.055$	$F = 14.14$, $p = .0017$	$F = 1.34$

b. Nonmusicians

Original vs. shortened vs. lengthened pause (O–S–L)

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
COND	ERP (2/28)	$F = 3.15$, $p = .060$	$F < 1$	$F = 1.65$	$F = 2.36$	$F < 1$	$F < 1$
	ERF (2/24)	$F < 1$	$F < 1$	$F = 2.44$	$F < 1$	$F < 1$	$F < 1$
COND × VERT	ERP (4/56)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 1.76$
	ERF (2/24)	$F = 1.05$	$F < 1$	$F = 2.90$, $p = 0.091$	$F = 198$	$F = 1.20$	$F = 3.32$, $p = .060$
COND × LAT	ERP (4/56)	$F < 1$	$F = 1.95$	$F = 2.21$, $p = .095$	$F = 1.09$	$F = 1.84$	$F = 1.45$
	ERF (2/24)	$F < 1$	$F < 1$	$F = 1.14$	$F = 3.04$, $p = .069$	$F = 2.74$	$F < 1$
COND × VERT × LAT	ERP (8/112)	$F < 1$	$F = 1.18$	$F = 1.39$	$F = 1.29$	$F = 1.34$	$F < 1$
	ERF (2/24)	$F < 1$	$F < 1$	$F = 1.34$	$F < 1$	$F < 1$	$F < 1$

Original vs. shortened pause (O–S)

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
COND	ERP (1/14)	$F < 1$	$F < 1$	$F = 2.46$	$F = 3.69$, $p = .075$	$F < 1$	$F < 1$
	ERF (1/12)	$F < 1$	$F < 1$	$F = 6.88$, $p = .022$	$F < 1$	$F < 1$	$F < 1$
COND × VERT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	ERF (1/12)	$F = 1.59$	$F < 1$	$F = 6.07$, $p = .030$	$F = 3.59$, $p = .082$	$F < 1$	$F = 6.40$, $p = .026$
COND × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F = 2.71$, $p = .084$	$F < 1$	$F = 1.82$	$F < 1$
	ERF (1/12)	$F < 1$	$F < 1$	$F < 1$	$F = 4.30$, $p = .060$	$F < 1$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F = 1.54$	$F = 1.43$	$F = 1.44$	$F = 1.38$	$F < 1$	$F < 1$
	ERF (1/12)	$F < 1$	$F = 1.62$	$F = 1.54$	$F < 1$	$F < 1$	$F = 1.01$

Table 6. (continued)

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
Shortened vs. lengthened pause (S–L)							
COND	ERP (1/14)	$F = 3.07$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 2.62$
	ERF (1/12)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × VERT	ERP (2/28)	$F = 1.44$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 3.94$, $p = .040$
	ERF (1/12)	$F = 1.17$	$F < 1$	$F = 1.38$	$F = 1.69$	$F = 2.69$	$F = 3.28$, $p = 0.095$
COND × LAT	ERP (2/28)	$F < 1$	$F = 1.51$	$F < 1$	$F < 1$	$F < 1$	$F = 2.02$
	ERF (1/12)	$F < 1$	$F < 1$	$F = 1.64$	$F < 1$	$F = 6.08$, $p = .023$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F < 1$	$F = 1.01$	$F = 2.63$, $p = .068$	$F = 2.02$	$F = 2.90$, $p = .036$	$F < 1$
	ERF (1/12)	$F < 1$	$F < 1$	$F = 1.09$	$F < 1$	$F < 1$	$F < 1$
Original vs. lengthened pause (O–L)							
COND	ERP (1/14)	$F = 5.29$, $p = .037$	$F < 1$	$F = 3.98$, $p = .066$	$F = 4.09$, $p = .063$	$F < 1$	$F < 1$
	ERF (1/12)	$F = 1.40$	$F < 1$	$F = 1.99$	$F < 1$	$F < 1$	$F < 1$
COND × VERT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 1.80$
	ERF (1/12)	$F < 1$	$F = 1.40$	$F = 1.62$	$F < 1$	$F = 2.$	$F < 1$
COND × LAT	ERP (2/28)	$F < 1$	$F = 3.62$, $p = .044$	$F = 4.50$, $p = .020$	$F = 2.63$, $p = .095$	$F = 2.49$	$F = 2.49$
	ERF (1/12)	$F < 1$	$F < 1$	$F = 1.45$	$F = 5.69$, $p = .035$	$F = 6.05$, $p = .03$	$F < 1$
COND × VERT × LAT	ERP (4/56)	$F < 1$	$F = 1.09$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	ERF (1/12)	$F < 1$	$F < 1$	$F = 1.01$	$F < 1$	$F = 1.03$	$F < 1$

of the boundary tone, had a modifying effect on phrase boundary perception.

First, it seems appropriate to discuss briefly some methodological issues, that is, the implications of the choices that have been made concerning the points of alignment (trigger points) and the baseline correction window. As pointed out in the Methods section, the choice of the trigger point emphasizes or suppresses brain responses depending on their temporal relationship to the various constituents of the phrase boundary. Roughly speaking, we can distinguish between activities that are correlated to (but not necessarily caused by) the closure of the first phrase and the onset of the second phrase. We did not find any significant ERP or ERF effects that were aligned to the end of the first phrase, except for a small variation before the point of divergence between phrased and unphrased melodic versions, which was most probably due to slight ordering imbalances of the stimuli (see Results section). In con-

trast, we found major effects related to the end of the pause. Here, the question of the correct baseline interval is relevant. We used two choices of baseline: (1) the detached baseline prior to the onset of the pause (-1000 to -500 msec), and the direct baseline directly preceding the trigger point at the pause offset (-200 to 0 msec). With respect to the averages, both baselines can be contaminated by pretrigger activity: in case of the direct baseline by transient, for the detached baseline by sustained differences between the conditions. Thus, any major differences between the two choices of baselines could indicate that the activity in question at least partially stems from the pause interval. Therefore, although only results for the direct baseline are reported in detail in the Results section (for space reasons), it is mentioned explicitly if there are any major discrepancies between the two baseline choices. Additionally, it is plausible that effects that are visible in the grand averages with both baselines prove significant only for the
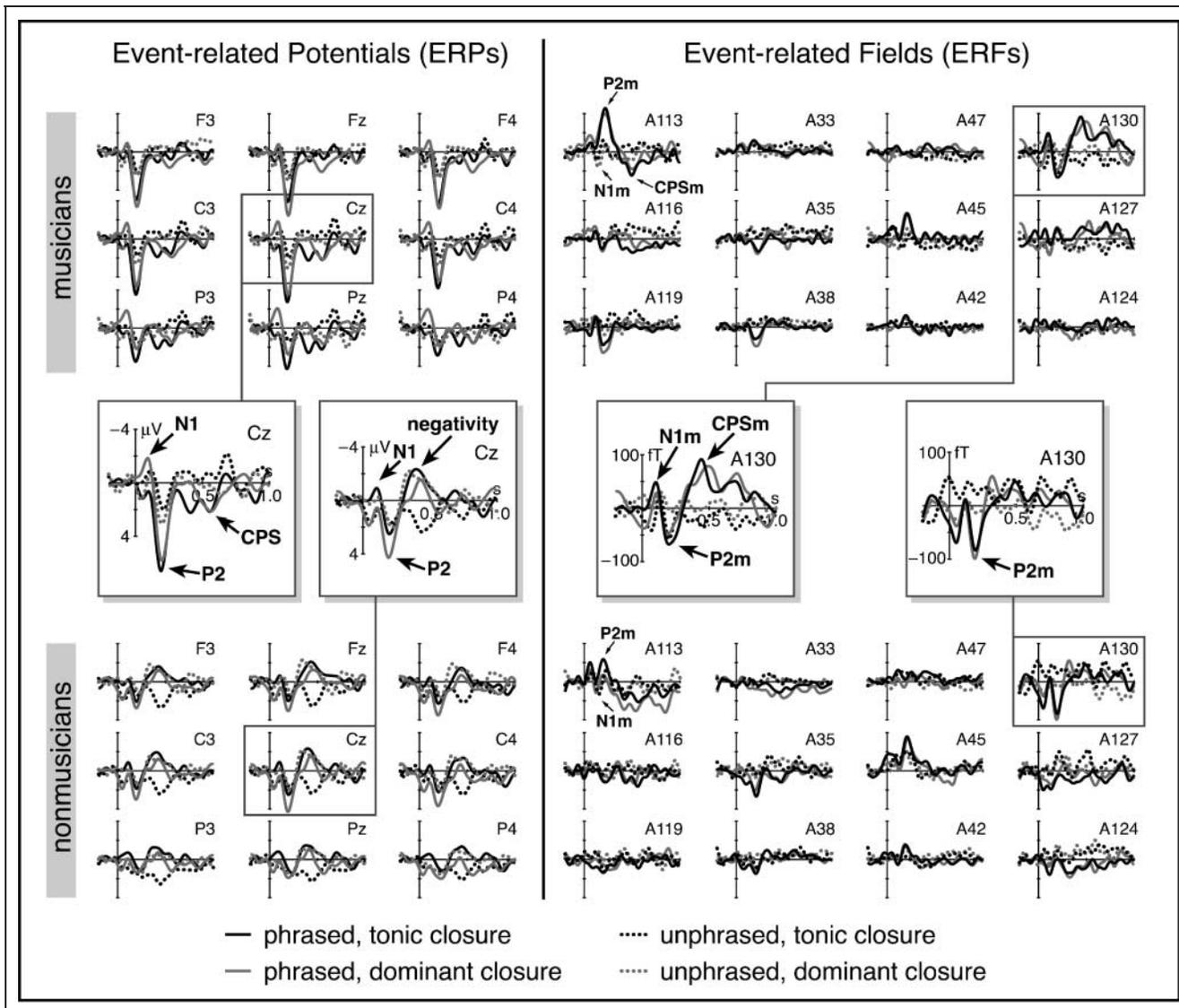
Figure 5. Grand average ERPs and ERFs at selected channels for examples with tonic (red) and dominant (green) closure of the first phrase. For the locations of the displayed channels, see Figure 2, top row.

direct baseline, because here the noise in the active and baseline windows is higher correlated. In other words, low-frequency fractions of the noise will be the same in both the baseline and the active time window, and therefore will be cancelled out in the variance of the difference between the two. This explains, why, for example, effects of harmonic endings of the first phrase become significant only when using the direct baseline.

Early Components—Refractory Effects

The N1 (and N1m) is not different between phrased and unphrased musical pieces (see Figure 2). However, for very long pauses (melody subset with strong phrase boundaries [Figure 3] and prolonged pauses, condition L [Figure 4]), the component is increased. There is no clear difference between musicians and nonmusicians.

This effect is probably due to the different duration of the silent period prior to the first note of the second phrase. Such refractory effects are well known in the literature (see, e.g., Budd et al., 1998).

The P2, in contrast, is clearly larger for musicians than for nonmusicians. This phenomenon has been reported repeatedly in the literature (Shahin et al., 2003; Trainor, Shahin, & Roberts, 2003). Moreover, it is evident (see, e.g., Figure 2) that the P2 is greatly diminished in the unphrased condition. Its amplitudes and topography, however, are influenced in a complex way by phrase boundary strength, pause length, harmonic closure, and musical expertise (see Figures 3, 4, and 5). Hence, P2 variations after the phrase boundary might not only reflect simple recovery of neuronal populations after the pause, but also some higher bottom-up order feature extraction processes necessary for the recognition of

Table 7. Results of Repeated Measures ANOVAs: Effects of Harmonic Closure [Phrased vs. Unphrased (O–C)]

		80–120 msec (N1, N1m)	180–220 msec (P2, P2m)	300–400 msec	400–500 msec	500–600 msec	600–700 msec
a. Musicians							
COND × HARMONY	ERP (1/14)	$F = 2.07$	$F < 1$	$F = 1.18$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/16)	$F < 1$	$F < 1$	$F = 3.80$, $p = .069$	$F = 4.40$, $p = .052$	$F < 1$	$F < 1$
COND × HARMONY × VERT	ERP (2/28)	$F < 1$	$F = 3.44$, $p = .067$	$F = 6.27$, $p = .016$	$F = 2.06$		$F = 2.99$, $p = .099$
	MEG (1/16)	$F = 5.90$, $p = .027$	$F = 6.36$, $p = .023$	$F = 3.02$	$F = 2.87$	$F < 1$	$F < 1$
COND × HARMONY × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F = 1.34$	$F < 1$	$F = 1.18$	$F = 1.12$
	MEG (1/16)	$F < 1$	$F = 1.43$	$F < 1$	$F = 1.11$	$F = 2.38$	$F < 1$
COND × HARMONY × LAT × VERT	ERP (4/56)	$F < 1$	$F = 1.56$	$F = 4.46$, $p = .0094$	$F = 2.47$, $p = .055$	$F = 2.436$, $p = .074$	$F = 1.16$
	MEG (1/16)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
b. Nonmusicians							
COND × HARMONY	ERP (1/14)	$F = 2.21$	$F = <1$	$F = 8.18$, $p = .013$	$F = 18.47$, $p = .0007$	$F = 2.84$	$F < 1$
	MEG (1/13)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × HARMONY × VERT	ERP (2/28)	$F = 2.29$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F = 2.32$
	MEG (1/13)	$F < 1$	$F = 1.31$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
COND × HARMONY × LAT	ERP (2/28)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/13)	$F = 3.54$, $p = .082$	$F < 1$	$F = 1.61$	$F < 1$	$F < 1$	$F = 2.05$
COND × HARMONY × LAT × VERT	ERP (4/56)	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$	$F < 1$
	MEG (1/13)	$F = 1.75$	$F < 1$	$F = 1.01$	$F < 1$	$F < 1$	$F < 1$

a boundary between two phrases. This is particularly true when these variations involve not only simple amplitude effects but also topological modifications. Figure 5 shows that the distribution of the P2 for musicians becomes more frontal for dominant and more parietal for tonic closures—a behavior very similar to that of the CPS. In other cases (e.g., the P2 effects observed for different pause lengths, Figure 4), interpretation is difficult, due to the inevitable confound between this and the refractory and habituation effects. The specific relationships between the early components and the later CPS will be discussed below for the different experimental conditions.

An important question is whether the CPS might be influenced by these early components or if it even originates from superimposed P2 components elicited by the following tones. It was found, however, that the CPS reflects activity of a quite different neural substrate, as it is clearly centroparietally distributed (see Figures 2 and 3), whereas the P2 features much more frontal topologies. Similar topology differences between CPSm and P2m can be observed for the MEG data. Source localizations have suggested that the P2 might originate from the anterior planum temporale (Knösche et al., 2003), whereas the CPS is generated from various limbic structures, including the cingulate cortex and hippocampus (Knösche et al., 2005). This, however, does not exclude the possibility of a partial overlap in neuronal substrate and/or function between the two processes, as will be discussed in the later subsections.

It is of course legitimate to ask if and to what extent refractory effects might also modify the CPS similarly to

the early N1 and P2 components. Refractoriness of neuronal populations describes their state of fatigue or inexcitability after stimulus-evoked excitation. It is usually visible in a decrease of amplitude of the respective ERP response. The faster the activation rate of a neuronal population (presentation rate of stimuli) is, the more the components are reduced. Refractory effects are usually reported for the early, mainly exogenous, components (in the auditory domain up to 200 msec after stimulus onset) and not for the later cognitive components. One exception is the work of Dell'Acqua, Jolicoeur, Vespignani, and Toffanin (2005), who found a refractory effect on the P3, which, however, concerned a modulation of latency and not of amplitude. If the different amplitudes of the CPS in the phrased and unphrased conditions would be due to refractory effects (similar to the P2), the respective neuronal population should have been activated in short succession. That, in turn, would mean that the CPS is not related to the phrase boundary as such (which is a relatively rare event) but to a type of event occurring at a faster rate—most probably to all note onsets. In this case, it should be detected after any note following a sufficient recovery period. This is not the case. If one, for example, compares the ERPs with respect to the onset of the phrase boundary tone with those for the onset of the first tone of the second phrase after having made the mean recovery periods comparable by appropriately selecting a subset of stimuli,[5] one finds a P2 of approximately equal amplitude, but a CPS only for the latter case. Hence, the CPS seems to be restricted to the reonset after phrase boundaries and therefore a decisive influence of refractory effects is considered unlikely.

Effects of Musical Expertise

When processing phrased and unphrased melody versions (conditions O and C), musicians revealed a clear bilateral CPS in centroparietal brain regions and a right-lateralized CPSm at frontocentral channel positions. Nonmusicians showed a slightly left lateralized early negativity peaking at frontocentral electrode placements and a frontally distributed CPSm similar to musicians, although of smaller amplitude. Qualitatively different brain responses in different time intervals measured with ERP suggest different processing strategies in both subject groups. Presumably, trained musicians encode phrase units in music similar to phrase units in language. Thus, in the broader sense, musicians who could be considered experts in the music domain, just like native speakers are considered experts in the domain of language, seem to process segmental time structures in music basically in the same way as in language. The music CPS/CPSm could therefore represent a general form of information processing independent of modality. Thus, musicians' ERP/ERF data give support to a recent conclusion drawn by Besson and Schön (2001),

who suggested that special ERP components might reflect more general cognitive principles rather than being domain-specific for language or music. In contrast, nonmusicians in our study, who were not skilled in listening to musical phrase structures in the analytical way, did not seem to have established a similar relationship between the segmental structure of melodies and that of spoken sentences. Instead, they might have solved the pitch-detection task under the tacit assumption of uninterrupted sound information, as no specific information regarding phrase structure was given in advance. For nonmusicians in this sense, the pause could have indicated a violation in the continuity and buildup of a melodic context. Thus, the frontally distributed early negativity might reflect violation of sound expectancy. This interpretation is backed by recent results of Koelsch et al. (2000, 2003) for the processing of chord sequences. For violations of sound expectancy by out-of-key (i.e., nondiatonic) chords, they demonstrated that nonmusicians do respond with an early right-hemispheric negativity of frontal topography (ERAN). Nonetheless, because of different time windows[6] and a slightly left lateralized instead of a dominantly right-hemispheric occurrence of the early negativity in this study, it is not obvious whether these negativities are members of the same family. Furthermore, current ERP results for pause variation revealed that nonmusicians responded with a significant early negativity to the original pause but not to prolonged and short pauses. Thus, a mere exogenous reaction to pure discontinuation of sound does not seem to be likely (Figure 4). We therefore conclude that nonmusicians as well as musicians must have a precise idea of an adequate pause length, although ERP indicators (CPS and early negativity) are qualitatively different. Accordingly, as suggested by Stoffer (1985), a clear mental representation of general phrasing schemata might exist in long-term memory for both subject groups. In nonmusicians, however, the finding of an early negativity even for unphrased melodies with a dominant closure points out that these general schemata seem to be represented less distinctly than in musicians. Additionally, nonmusicians might rely on harmonic cues (as compared to pause and boundary tone length) to a higher degree than musicians.

Consequentially, when perceiving phrase units, musicians and nonmusicians apparently make use of two different strategies: top-down activation of a general schema for phrasing retrieved from long-term memory as well as the use of acoustic boundary markers for bottom-up processing of the melodic input. Present results are also in line with previous findings of Krumhansl and Jusczyk (1990) who stated in their infants' study, that musical expertise does not seem to be a necessary precondition for musical phrase perception per se. However, present ERP results (i.e., the CPS and the early negativity) make clear that different neural markers reflect different pro-

cessing strategies based on different degrees of musical expertise: When listening to phrased melodies, musicians seem to catch the structure in a speechlike sense, whereas nonmusicians seem to react primarily to expectancy violation regarding continuity. It is interesting to note that similar results have been reported for hummed speech in L1 and L2 speakers of German. Whereas for native speakers, there was a clear CPS (Pannekamp et al., 2005), for L2 learners an earlier and more frontally distributed negative wave was observed instead (Isel & Friederici, 2005; Isel, personal communication).

The right lateralization of the CPSm in musicians is in line with theories that fast temporal changes are mainly processed in the left hemisphere, whereas the processing of slower features is dominated by the right hemisphere (e.g., Pöppel, 2003; Zatorre, Belin, & Pehune, 2002).

Different forms of harmonic phrase closure (i.e., the ending of the first phrase on the tonic or on the dominant) significantly change the ERPs for both subject groups, whereas the magnetic recordings did not reveal any clear differences (Figure 5). For musicians, the CPS was modified in its topology and morphology (frontal for dominant and parietal for tonic closure; starting earlier for tonic closure). This suggests the active use of this cue for the processing of the phrase boundary and strengthens the role of the CPS as an indicator for top-down processes, rather than for mere pause perception. For nonmusicians, in contrast, there is a similar early negativity for both types of harmonic ending. Additionally, such a negative wave is also present for the unphrased condition with the previous phrase ending on the dominant. This finding suggests that due to a different degree of musical expertise, a general concept for phrasing might be stored less precisely in the long-term memory of nonmusicians than of musicians. Further studies should clarify the influence of harmonic cues on phrase boundary perception in detail.

In addition to these results, musicians also reacted with a more pronounced P2 wave to the onset of the second phrase (Figure 2). Such differences in P2 amplitude are also found when averaging with respect to the onset of other tones. Similar findings have been reported before for isolated musical tones (Shahin et al., 2003; Trainor et al., 2003). It is therefore likely that this effect reflects neuroplastic changes of the primary auditory regions due to musical training, as concluded by Shahin et al. (2003), and is therefore mainly related to bottom-up processing on the single tone level. Similarly to Shahin et al. and in contrast to, for example, Pantev et al. (1998), we found no significant differences of the N1 or N1m amplitude between musicians and nonmusicians.

Effects of Boundary Markers

In both subject groups, ERP and ERF results for phrase boundary strength (Figure 3) support the music-theoretical considerations of Riemann (1900) concerning the relevance of pause length and boundary tone length for the marking of phrase boundaries. The importance of pause length for the presence of the CPS and early negativity is proved by the differences between phrased and unphrased conditions (Figure 2).

However, also the length of the boundary tone seems to be a relevant cue for phrase boundary marking: For strong phrase boundaries (longer boundary tones and longer pauses), musicians revealed a stronger CPS and CPSm, whereas nonmusicians showed a greater early negativity. However, if the pause is prolonged at the expense of the boundary tone (Figure 4), these effects are less pronounced. Although pause lengths are, on average even greater than for the "strong phrase boundary" condition, the amplitude of the CPS is smaller. This proves that the boundary tone length also has a strong influence on the CPS. A long pause alone does not produce a large CPS—in fact, the CPS was largest for the original combination of pause and boundary tone length. This effect was even clearer for the early negativity observed for nonmusicians (Figures 3 and 4).

Hence, although the pause length is certainly the most relevant boundary marker, prolonged boundary tones modify both ERP components and should therefore be considered as a second relevant boundary marker within phrase perception.

The decisive role of pause and boundary tone lengths as the most relevant cues for phrase boundary detection is in line with previous results of Frankland and Cohen (2004). In their study, they reported on the relationship between empirical evidence and the local grouping rules as defined in the *Generative Theory of Tonal Music* by Lerdahl and Jackendoff (1983). According to Frankland and Cohen, only distances between the note attack points (corresponding to the combined boundary tone and pause lengths) and the rest periods between the tones (corresponding to pause length) were highly correlated with the empirical results for boundary location.

Furthermore, the CPS latency shift, which was observed between strong and weak phrase boundaries in the ERPs of musicians (Figure 3), seems to provide evidence for chunking activity and the retrospective processing of the first phrase as proposed in the previous study (Knösche et al., 2005). The suggestion is based on results of an additional regression analysis, revealing a quite strong correlation (.37) between the length of the first phrase and the length of the boundary tone. Short first phrases correlated strongly with short boundary tones and long first phrases correlated strongly with long boundary tones. It is therefore proposed that the earlier CPS with lower amplitude reflects a faster processing of the shorter first phrase as a lower amount of information was to be processed.

The length of pauses and boundary tones also had significant effects on the early components, that is, the N1 and P2. Enhancements of the N1 were observed

for strong as compared to weak phrase boundaries (Figure 3; statistically below significance level) and for prolonged as compared to original and shortened pauses (Figure 4), but not for phrased as compared to unphrased items in general (Figure 2). Moreover, also the type of harmonic closure seems to influence the component (significant only for MEG). Despite the expected confound with refractory effects, these results seem to demonstrate some top-down influences already on the N1, which are related to the investigated phrase boundary cues. The P2 is enhanced for phrased as compared to unphrased items, but not for prolonged as compared to shortened pauses. In the latter case, not the longest, but the original pause caused the larger amplitude, similar to the CPS. This might be some hint for a functional relationship (but not identity) of the underlying processes of P2 and CPS. With regard to harmonic phrase closure on the tonic or dominant, results for the CPS and P2 seem to be heterogeneous and allow only a preliminary and cautious discussion. For musicians in particular, P2 and CPS are similarly influenced (more frontal distribution for the dominant and more parietal distribution for the tonic closure, see Figure 5). This again might indicate some kind of functional relationship between the neuronal processes underlying the two components. On the other hand, the fact that only the CPS is modified in latency suggests separate mechanisms for the processing of harmonic cues (Figure 5).

Summarizing, there is evidence that both N1 and P2 reflect processes that are related to the detection and processing of phrase boundary cues. In particular, the P2 shows some traits that might indicate a kind of functional relatedness to the CPS. However, these effects are inevitably confounded by the known refractoriness of both components and should be interpreted with caution.

Conclusions

Present ERP and ERF results indicate that musicians and nonmusicians process phrase boundaries in melodies in a different way. When comparing present results with previous ones in the domain of language, correlates for phrase boundary perception (i.e., the music CPS and the language CPS) resemble each other in several parameters. Apparently, musicians but not nonmusicians process musical phrase structures in a speechlike manner. Nonmusicians, in contrast, might have perceived phrase boundaries in the sense of expectancy violation during the buildup of a musical context, thus detecting discontinuity in the auditory input. In addition, present findings for pause manipulation and phrase boundary strength provide evidence that both pauses and the lengthening of preboundary items (syllables or tones) appear to be relevant acoustical markers for phrase processing in music and language and, in a more general sense, for processing sequential structures in time. In particular, the pause insertion seems to play the decisive

role. Also, different harmonic closure before the phrase boundary seems to influence brain processes underlying the CPS and the early negativity in a significant way, thus adding to the notion that these are neurophysiological correlates of top-down processes related to the processing of phrase structure.

Furthermore, EEG and MEG methods seem to be sensitive to different aspects within phrase processing: In ERPs we observed two qualitatively different components reflecting perceptual differences between musicians and nonmusicians, whereas in ERFs, quantitative differences in the CPSm amplitudes of all participants could be depicted. Moreover, the harmonic function of the boundary tone did not have an influence on the CPSm. We therefore suggest that the CPS and early negativity both reflect a top-down process in the respective subject group, that is, the retrieval of general but probably different schemata stored in long-term memory, which seem to be less distinct in nonmusicians than in musicians. In contrast, ERF signals might primarily reflect the bottom-up processing of acoustical boundary markers similar for musicians and nonmusicians. In summary, we gained insight into the modifying effects of acoustical cues in the close vicinity of the phrase boundary. Further research should address the issue as to whether also general musical traits, for example, global musical aspects such as meter or the symmetrical structure of the first phrase modulate the CPS and CPSm in musicians and nonmusicians.

Notes

1. In order to simplify the handling of component names in the current study, we will use the term "CPS" for the positive shift measured with EEG and the term "CPSm" to describe the CPS-related equivalent shift measured with MEG.
2. Sign inversion was necessary because identical brain activity patterns located in different hemispheres cause MEG patterns of opposite polarity over these hemispheres.
3. The effect vanished if only those stimuli were averaged that were preceded by at least three phrased melodies in succession. Hence, it seems that the closure of the first phrase produced different expectations concerning the appearance of a pause, depending on the experience of the last few stimuli. However, sequence effects did not influence the presence or absence of the ERP effects related to the offset of the pause (CPS, see below).
4. Mean amplitude for Cz over time interval 500–600 msec for the CPS and 300–500 msec for the early negativity.
5. In this case, we compared the pause-offset response for those 18 stimuli with the shortest distance between the onset

of the boundary tone and the pause offset (average 517 msec) with the boundary tone response for those 48 stimuli with the longest onset–onset period before the boundary tone (average 508 msec). These stimulus selections were necessary to make the note distances comparable and rule out refractory effects.

6. Onset of the ERAN: approximately 150 msec, onset of the early negativity in this study: approximately 300 msec.

REFERENCES

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences, 930,* 232–258.

Budd, T. W., Barry, R. J., Gordon, E., Rennie, C., & Michie, P. T. (1998). Decrement of the N1 auditory event-related potential with stimulus repetition: Habituation vs. refractoriness. *International Journal of Psychophysiology, 31,* 51–68.

Chiappe, P., & Schmuckler, M. A. (1997). Phrasing influences the recognition of melodies. *Psychonomic Bulletin & Review, 4,* 254–259.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115,* 732–744.

Dell'Acqua, R., Jolicoeur, P., Vespignani, F., & Toffanin, P. (2005). Central processing overlap modulates P3 latency. *Experimental Brain Research,* [On-line], 1–26.

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics, 14,* 37–40.

Drake, C., Peuel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18,* 1–23.

Frankland, B. W., & Cohen, A. J. (2004). Parsing of melody: Quantification and testing of the local grouping rules of Lerdahl and Jackendoff's *A Generative Theory of Tonal Music*. *Music Perception, 21,* 499–543.

Isel, F., & Friederici, A. D. (2005). The processing of prosodic boundaries in a second language: ERP evidence from French late bilinguals in spoken German. Proceedings of the Annual Meeting of Cognitive Neuroscience Society, New York, 2005.

Jongsma, M. L. A., Desain, P., & Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological Psychology, 66,* 129–152.

Knösche, T. R. (2002). Transformation of whole-head MEG recordings between different sensor positions. *Biomedizinische Technik, 47,* 59–62.

Knösche, T. R., Neuhaus, C., Haueisen, J., & Alter, K. (2003). The role of the planum temporale in the perception of musical phrases. Proceedings of the 4th International Conference on Noninvasive Functional Source Imaging (NFSI), Chieti, Italy, September 2003.

Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Friederici, A. D., & Witte, O. (2005). The perception of phrase structure in music. *Human Brain Mapping, 24,* 259–273.

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

Koelsch, S., Gunter, T., Schroeger, E., & Friederici, A. D. (2003). Processing tonal modulations: An ERP study. *Journal of Cognitive Neuroscience, 15,* 1149–1159.

Krumhansl, C. L., & Jusczyk, P. W. (1990). Infants' perception of phrase structure in music. *Psychological Science, 1,* 70–73.

Ladefoged, P., & Broadbent, D. E. (1960). Perception of sequences in auditory events. *Quarterly Journal of Experimental Psychology, 12,* 162–170.

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

Mattheson, J. (1976). *Kern melodischer Wißenschafft* (Nachdruck der Ausgabe Hamburg). Hildesheim: Olms. (Original work published 1737).

Oldfield, R. C. (1975). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112,* 713–719.

Palmer, C., & Krumhansl, C. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 116–126.

Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven sentence processing: An event-related brain potential study. *Journal of Cognitive Neuroscience, 17,* 407–421.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392,* 811–814.

Pöppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time." *Speech Communication, 41,* 245–255.

Riemann, H. (1900). *Vademecum der Phrasierung*. Leipzig: Max-Hesse-Verlag.

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience, 23,* 5545–5552.

Sloboda, J. A., & Gregory, A. H. (1980). The psychological reality of musical segments. *Canadian Journal of Psychology, 34,* 274–280.

Steinhauer, K. (2001). *Hirnphysiologische Korrelate prosodischer Satzverarbeitung bei gesprochener und geschriebener Sprache* (MPI Series in Cognitive Neuroscience). PhD thesis, Leipzig.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience, 2,* 191–196.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research, 30,* 267–295.

Stoffer, T. H. (1985). Representation of phrase structure in the perception of music. *Music Perception, 3,* 191–220.

Tan, N., Aiello, R., & Bever, T. G. (1981). Harmonic structure as a determinant of melodic organization. *Memory & Cognition, 9,* 533–539.

Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception & Psychophysics, 62,* 333–340.

Trainor, L. J., Shahin, A., & Roberts, L. E. (2003). Effects of musical training on the auditory cortex in children. *Annals of the New York Academy of Sciences, 999,* 506–513.

van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of sequential sounds—An event-related potential study comparing musicians and nonmusicians. *Journal of Cognitive Neuroscience, 16,* 331–338.

Wilson, S. J., Dressing, J., Wales, R. J., & Pattinson, P. (1999). Cognitive models of musical function within the brain. *Australian Journal of Psychology, 51,* 125–139.

Zatorre, R. J., Belin, P., & Pehune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6,* 37–46.