BRAIN RESEARCH 1585 (2014) 99–107

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Research Report

Musical phrase boundaries, wrap-up and the closure positive shift

Susana Silva[a,b], Paulo Branco[a], Fernando Barbosa[a], João Marques-Teixeira[a], Karl Magnus Petersson[b], São Luís Castro[a,*]

[a]Faculty of Psychology and Educational Sciences, University of Porto, Rua Alfredo Allen, 4200-135 Porto, Portugal
[b]Department of Psychology, IBB-CBME, University of Algarve, 8005-139, Faro, Portugal

ARTICLE INFO

Article history:
Accepted 11 August 2014
Available online 17 August 2014

Keywords:
Musical phrase
Boundary perception
Wrap-up
Music and language
Closure Positive Shift

ABSTRACT

We investigated global integration (wrap-up) processes at the boundaries of musical phrases by comparing the effects of well and non-well formed phrases on event-related potentials time-locked to two boundary points: the onset and the offset of the boundary pause. The Closure Positive Shift, which is elicited at the boundary offset, was not modulated by the quality of phrase structure (well vs. non-well formed). In contrast, the boundary onset potentials showed different patterns for well and non-well formed phrases. Our results contribute to specify the functional meaning of the Closure Positive Shift in music, shed light on the large-scale structural integration of musical input, and raise new hypotheses concerning shared resources between music and language.

1. Introduction

Electrophysiological responses to phrase boundaries in music have been reported in several studies (Knoesche et al., 2005; Nan et al., 2006, 2009; Neuhaus et al., 2006). The paradigm used in these studies contrasted musical phrases ending with a boundary pause (phrased condition) with the same musical phrases having the boundary pause filled with tones (unphrased condition). In both conditions, a second phrase continued the melody, and event-related potentials (ERPs) were measured at the onset of this second phrase. A positive centro-parietal positivity peaked roughly 550 ms after boundary pauses and was absent after the filling tones. The positive peak, the Closure Positive Shift (CPS), was interpreted as a marker of phrase boundary perception. However, experimental manipulations of musical phrase boundaries based on a different paradigm failed to elicit a CPS component (Istók et al., 2013). A possible reason for this is that the functional meaning of the CPS is not yet fully understood (Koelsch, 2011a), leading to failure in targeting the critical phrase boundary processes. Source localization of the CPS identified brain regions related to both memory and attention (Knoesche et al., 2005). A fMRI study of the CPS paradigm (Nan et al., 2008) highlighted brain regions related to the maintenance of past events in on-line processing memory, as well as regions related to selective attention. The component's sensitivity to the pre-boundary tone has been associated with boundary strength (Nan et al., 2006; Neuhaus et al., 2006), but it might also relate to the representation of past events (retrospective processing) in a more general sense. Consider Istók et al. (2013) experiment: they presented atonal scale-like isochronous melodies to

*Corresponding author. Fax: +351 226 079 725.
E-mail address: slcastro@fpce.up.pt (S.L. Castro).

http://dx.doi.org/10.1016/j.brainres.2014.08.025
0006-8993/

participants, and inserted a pitch leap within the sequence to define a phrase boundary. The failure to elicit a CPS component was attributed to the impairment of future-oriented processing: since there was no boundary pause, the listener would be unable to shift attention to the upcoming phrase. However, it is also possible that the failure to elicit CPS was due to phrases lacking internal structure, in the sense that there was not enough coherence between tones to induce a unified perception of the phrase. According to this latter hypothesis, the level of past-oriented processing may have been too low to elicit the CPS. While it seems likely that future-oriented processing (shift of attention) and past-oriented processing (memory) are both reflected in the CPS, several issues remain unspecified. Namely, we know little about the extent to which past events are involved.

A yet untested hypothesis on the functional meaning of CPS is that it reflects a unified representation of the musical phrase preceding the boundary. It is possible that the transition to a new phrase involves some memory of the previous one. This hypothesis is particularly relevant since it may clarify how listeners make sense of music. One view is that musical segments, including phrases, are first represented and then related to each other as parts of the whole musical piece (McAdams, 1989; Nan et al., 2008; Neuhaus, 2013; Pearce et al., 2010; Peretz, 1989). Consistent with this, probe tone paradigms have shown that within-phrase chord sequences have processing advantages over between-phrase chord sequences (Tan et al., 1981). A different view is that listening to music requires nothing more than a focus on the present (e.g., Levinson, 1997), and indeed behavioral research has shown that the actual comprehension of large-scale musical structures, including phrases, may well fall short of composers' expectations (see Lalitte et al., 2004, for a review). While according to the first view listeners are expected to represent musical phrases as holistic entities, according to the second view they are not.

Forming a unified representation of a musical phrase parallels *wrap-up processes* in language. The concept of wrap-up is well-established in language; it captures the retrospective processing of a sentence as a whole by the time the last word is perceived. Wrap-up processes have been inferred from longer reading times at sentence-final words (Aaronson and Scarborough, 1976, 1977; Just and Carpenter, 1980; Mitchell and Green, 1978). ERPs locked to sentence-final words show increased positivity between 300 and 700 ms compared to sentence-intermediate words (Friedman et al., 1975; Van Petten and Kutas, 1991), and violations in sentence-final position deviate from violations in intermediate positions (Hagoort, 2003; Hagoort and Brown, 1999; Hagoort et al., 2003; Kotz et al., 2003; Weber and Lavric, 2008). Analogous wrap-up processes in the music domain might also exist, namely if music listening is conceived as an active process going beyond the present moment. Even though, unlike reading, there is no obvious form of comprehension that must underlie communication by music, expectations and phrasing are involved in at least some forms of music listening. In this sense, it is a good option to induce active listening when aiming to investigate wrap-up processes in music.

In the current study, we tested whether the CPS indexes wrap-up processes in music during active listening. We manipulated the well-formedness of the phrase in order to induce different levels of wrap-up demands. Since well-formedness depends on universal as well as cultural expectations (Purwins et al., 2008; Narmour, 1992), we added two types of non-well formed phrases to the paradigm. Violations of universal expectations were represented by phrases with an intruder segment. In these phrases, the segments or *motifs* (Riemann, 1929; Schoenberg, 1967) are unrelated and violate the expectation of figural similarity (e.g., Bigand, 1990; McAdams and Matzkin, 2001, 2003; Minati et al., 2010, see Fig. 1 for examples). Violations of cultural expectations were represented by three bar phrases. A four bar phrase is typical in Western music (Berg, 1965; Cooper and Meyer, 1960; Davie, 1966; Feldstein, 1995; Riemann, 1929; Schoenberg, 1967; Stein, 1962), and familiarized listeners expect a boundary cue every fourth bar (Lerdahl and Jackendoff, 1983; Narmour, 1992; Sloboda, 1985). We also modified the experimental task. The CPS paradigm has been investigated in either a single out-of-key tone detection task (Knoesche et al., 2005; Neuhaus et al., 2006), or a style classification task (Chinese vs. Western: Nan et al., 2006, 2009). In the present study, we adopted a task that tuned attention to phrases with an intruder. Like tone detection, our task implies the detection of a violating event (a whole segment and not just a tone). Similar to style classification, it implies a broad time window for listening. Since the CPS is more robust in musicians than in non-musicians (Nan et al., 2009), we collected data from musicians.

If the CPS reflects wrap-up processes, both types of non-well formed phrases should differ from well-formed phrases at the boundary offset point, and the differences should be observed in the CPS time window (500–600 ms). In addition, we expected differences in the P2 time window (150–250 ms) (Knoesche et al., 2005; Nan et al., 2006, 2009; Neuhaus et al., 2006). In order to validate the perception of non-well formedness we also analyzed ERPs at the intruder onset, where we expected responses to the upcoming violation of figural similarity. To make sure that the retrospective processing of melodies could be captured by the EEG, we further looked into ERPs at the intruder offset (brain responses to impaired integration already expected for the intruder version) and at the end of the first phrase (responses

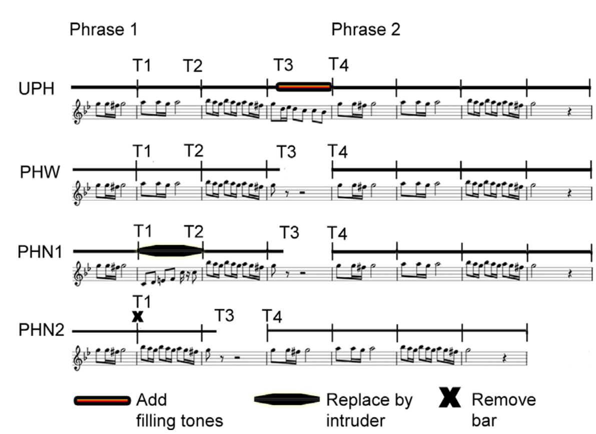

Fig. 1 – An example taken from the set of melodies to illustrate the four experimental conditions. Vertical lines represent the end of a bar. The second bar in PHN1 represents a violation of figural similarity, since neither rhythm nor pitch is similar to the first bar. T1–T4 refer to the four trigger points. UPH = unphrased; PHW = phrased well-formed; PHN1 = phrased (non-well formed) with intruder; and PHN2 = phrased (non-well formed) with three-bars.

to impaired integration expected for both intruder and three-bar phrases).

2. Results

2.1. Behavioral results

Overall performance accuracy was clearly above chance levels (mean% correct±SD=84±8%; $t(23)=21.0$, $p<0.001$). The mean hit rate (correct Yes responses to melodies with an intruder) was 0.84, and the mean false alarm rate (incorrect Yes responses to melodies without an intruder) was 0.15, yielding a mean discrimination index d-prime of 2.21. Response accuracy did not differ significantly across conditions ($F(3,69)=2.42$, $p=0.074$).

2.2. EEG results

2.2.1. Intruder segment (T1 and T2)

At the intruder onset (T1, Fig. 2A), the omnibus ANOVA for the 0–500 ms time window showed no significant effect of melodic condition, but there was an interaction with region ($F(3.66,84.1)=5.4$, $p=0.001$) suggesting topographically narrow effects. Comparisons between phrases with intruder (PHN1 hereafter) and all other phrase types revealed significant interactions (PHN1 vs. unphrased, UPH: $F(1.31,30.1)=11.7$, $p<0.001$; PHN1 vs. phrased well-formed, PHW: $F(1.41,32.5)=8.4$, $p=0.003$; PHN1 vs. phrased with three bars, PHN2: $F(1.33,30.5)=8.3$, $p=0.004$), while comparisons between UPH, PHW and PHN2 did not. At posterior sites, PHN1 showed increased negativity compared to PHN2 ($p<0.001$), UPH ($p=0.003$) and PHW ($p=0.009$). Differences extended to the central region in the comparison PHN1/PHN2 ($p=0.007$).

Between 500 and 800 ms, there was a main effect of condition ($F(2.33,53.6)=4.7$, $p=0.010$) with no significant interactions between melodic condition and region. PHN1 phrases differed from all other phrase types (PHN1 vs. UPH: $F(1,23)=10.0$, $p=0.004$; PHN1 vs. PHW: $F(1,23)=7.2$, $p=0.013$; PHN1 vs. PHN2: $F(1,23)=7.0$, $p=0.015$). None of these comparisons displayed interactions between condition and region, indicating a widespread positive component for the intruder condition. Comparisons between UPH, PHW and PHN2 revealed no significant main effects or interactions.

At the intruder offset (T2, Fig. 2B), there was a near-to-significant effect of melodic condition between 300 and 500 ms ($F(2,46)=2.52$, $p=0.091$). Phrases with an intruder showed a trend for lower negativity compared to PHW ($F(1,23)=3.87$, $p=0.061$), and they were significantly less negative than UPH ($F(1,23)=4.81$, $p=0.039$). UPH and PHW did not differ from each other. There were no interactions with region.

2.2.2. Boundary onset (T3)

Between 0 and 500 ms there were no effects of melodic condition. These started between 500 and 700 ms ($F(2,46)=3.9$, $p=0.028$, see Fig. 3). PHW melodies showed increased positivity compared to PHN1 ($F(1,23)=5.3$, $p=0.031$) and PHN2 ($F(1,23)=4.3$, $p=0.050$). No interactions with region were observed in any of the two comparisons. Conditions PHN1 and PHN2 did not differ from each other. Between 700 and 1000 ms, there was a marginal effect of

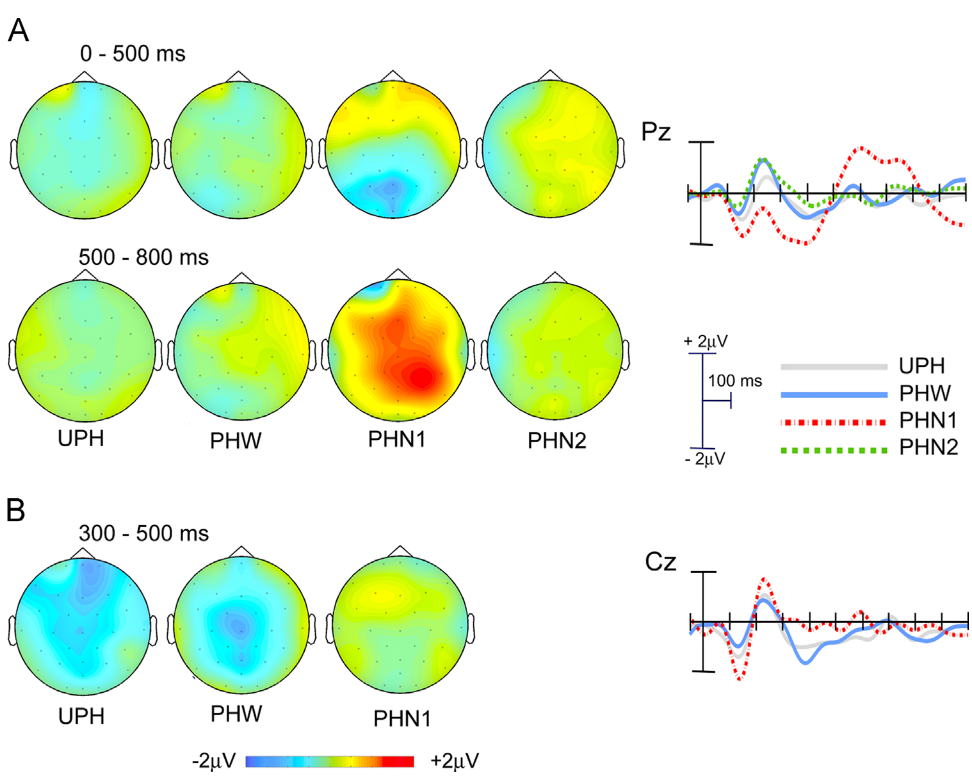

Fig. 2 – Topographic maps and representative waveform of ERPs locked to (A) the onset and (B) the offset of the intruder segment, where responses to PHN1 were expected to differ from the other conditions. UPH=unphrased; PHW=phrased well-formed; PHN1=phrased (non-well formed) with intruder; PHN2=phrased (non-well formed) with three-bars. Waveforms were low-passed filtered (<8 Hz) for visualization purposes only.

condition ($F(2,46)=2.6$, $p=0.086$). Negativity was marginally higher in PHW compared to PHN1 melodies ($F(1,23)=3.9$, $p=0.061$), and in PHW compared to PHN2 ($F(1,23)=3.1$, $p=0.091$). There were no interactions between melodic condition and region. Conditions PHN1 and PHN2 did not differ from each other.

2.2.3. *Boundary offset (T4)*

The P2 time window (150–250 ms) showed a main effect of melodic condition ($F(3,69)=22.2$, $p<0.001$), with lower positivity for the UPH condition compared to the other three (UPH vs. PHW: $F(1,23)=43.6$, $p<0.001$; UPH vs. PHN1: $F(1,23)=31.6$, $p<0.001$; UPH vs. PHN2: $F(1,23)=33.1$, $p<0.001$, see Fig. 4). In the three comparisons, there were significant interactions with region (UPH vs. PHW: $F(1.16,26.6)=24.3$, $p<0.001$; UPH vs. PHN1: $F(1.15,26.6)=37.5$, $p<0.001$; UPH vs. PHN2: $F(1.15,23.4)=33.8$, $p<0.001$). Significant differences were observed at anterior and central regions ($p<0.001$), matching the fronto-central distribution of the P2. Comparisons between PHW, PHN1 and PHN2 revealed no significant main effects or interactions.

The CPS time window did not display effects of melodic condition but, rather, an interaction between condition and region ($F(2.96,68)=7.19$, $p<0.001$). This is likely due to a frontal negativity in conditions PHW, PHN1 and PHN2 (Fig. 4, topographic maps) that counteracted the posterior positivity (CPS) in the whole-scalp average, and therefore reduced differences between those three conditions and UPH melodies. The interaction replicated in the comparisons between UPH and the other three conditions (UPH vs. PHW: $F(1.21,27.9)=14.27$, $p<0.001$; UPH vs. PHN1: $F(1.31,30.2)=11.15$, $p=0.001$; UPH vs. PHN2: $F(1.40,32.2)=14.81$, $p<0.001$), and it was absent in the remaining comparisons. At central and posterior regions, UPH versions elicited lower positivity than PHW, PHN1 and PHN2 ($p<0.001$ in all) matching the music CPS effect.

3. Discussion

Our goal was to investigate whether the CPS component reflects the retrospective processing of the musical phrase as a whole (phrase wrap-up). We found an EEG marker of musical phrase wrap-up, but this did not correspond to the CPS. The P2/CPS responses at the phrase boundary offset were not modulated by the wrap-up demands of the preceding phrase: well- and non-well formed phrases elicited components that were statistically indistinguishable. The amount of information in the first phrase had no impact either: our three-bar phrases, which had an irregular structural quality, were shorter than phrases with intruder, and well-formed phrases. Yet they did not elicit smaller CPS amplitudes. Therefore, it seems it is not the case that the unified representation of the phrase is reflected in the CPS component.

Our findings agree with previous results on the role of local boundary cues (pause, length and harmonic function of the pre-boundary tone) in eliciting the CPS (Knoesche et al., 2005; Nan et al., 2006, 2009; Neuhaus et al., 2006). In our study, the pause was crucial to generate the component, which was not elicited in the unphrased condition. Moreover, our findings add to the evidence that the CPS indexes attention and memory processes as suggested by Knoesche et al. (2005) from source analysis. However, the fact that the CPS did not respond to the well-formedness of the phrase lends more support to an attention-based approach to the component rather than to a memory-based one. According to our results, the unified representation (memory) of the whole phrase does not interfere with brain activity at the boundary offset, even if some memory of, at least, the ending segment is likely to be involved. The fact that properties of the pre-boundary tone such as length and harmonic function modulate the CPS (Neuhaus et al., 2006) argues in this favor: tone *lengthening* can only be perceived in context, and the same goes for harmonic function, which implies the build-up of a context.

One remark on unphrased melodies deserves mention. As expected, replacing the boundary pause with filling tones suppressed the CPS response, but it is interesting to note that a particularly pronounced P2 was observed by the time the filling tones started (Fig. 3, compare with boundary offset P2, Fig. 4). High P2 amplitudes have been related to the effects of musical expertise on low-level acoustic processing (see Shahin et al., 2003, 2005, for complex tones). This is consistent with the fact that our subjects were musicians, but it does not explain

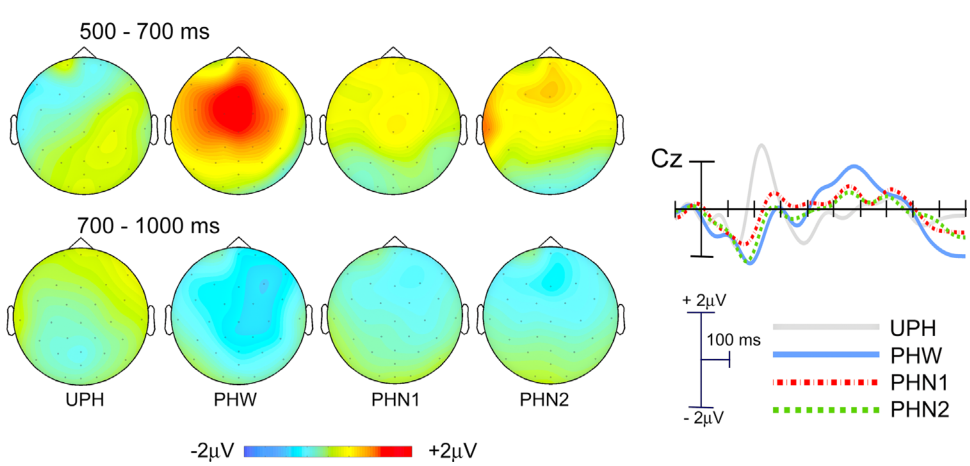

Fig. 3 – Topographic maps and representative waveform of ERPs locked to the boundary onset, where wrap-up demands should be higher for PHN1 and PHN2 compared to PHW. UPH = unphrased; PHW = phrased well-formed; PHN1 = phrased (non-well formed) with intruder; PHN2 = phrased (non-well formed) with three-bars. Waveforms were low-passed filtered (<8 Hz) for visualization purposes only.

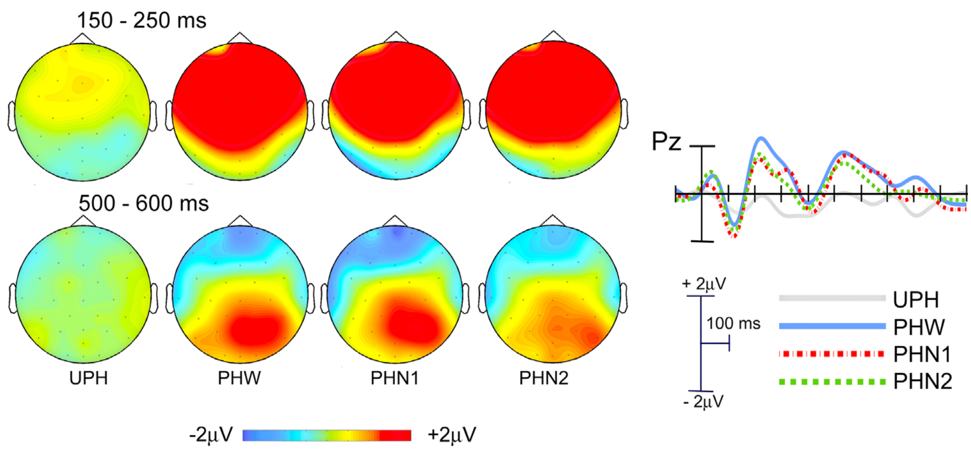

Fig. 4 – Topographic maps and representative waveform of ERPs locked to the boundary offset, where wrap-up demands should be higher for PHN1 and PHN2 compared to PHW. UPH = unphrased; PHW = phrased well-formed; PHN1 = phrased (non-well formed) with intruder; PHN2 = phrased (non-well formed) with three-bars. Waveforms were low-passed filtered (<8 Hz) for visualization purposes only.

why the P2 at the boundary onset was more pronounced than at the offset. Silva et al. (2014) found this very same pattern with the typical CPS paradigm (e.g., Knoesche et al., 2005), and proposed that the increased P2 at the onset of filling tones is an index of the listener's violated expectation of a boundary pause. The approach was based on the notion that the P2 component is related to high-level processing (e.g., De Diego Balaguer et al., 2007; Minati et al., 2010; Snyder et al., 2006), an idea shared by Neuhaus et al. (2006) when pointing to a functional link between P2 and CPS. Based on fMRI results, Nan et al. (2008) also suggested that the filling tones impair the integration of slowly evolving, predictive information (phrase contour, metrics, etc.) with the local information provided by the boundary pause. Interestingly, this difficulty was reflected in the activation of the planum temporale, which has been proposed as the generator of P2 (Knoesche et al., 2003).

3.1. *A signature of phrase wrap-up at boundary onset*

The wrap-up demands of the phrase modulated a widespread positive peak ranging from 500 to 700 ms after the boundary onset, as well as a negative peak from 700 to 1000 ms. Well-formed phrases elicited larger mean amplitudes in these two time windows, whereas the two types of non-well-formed phrases did not differ. The positive peak showed increased amplitudes for phrases with a predictable structure and lower wrap-up demands (well-formed). This ERP pattern is consistent with reports of increased late positivity (300–700 ms) at the offset of correct sentences compared to internal positions (e.g., Van Petten and Kutas, 1991) and, hence, with a wrap-up effect.

ERPs at the intruder offset suggest that there may have been retrospective processing even before the phrase ends. At this point, phrases without an intruder (UPH, PHW) showed a trend for a late widespread negativity, and phrases where an intruder had just been heard did not. Intermediate retrospective processing may be linked with the chunking of motives into sections shorter than the whole phrase (note that an intruder segment replaces one motif). However it is also possible that the differences are due to the fact that there is a new discontinuity in PHN1 at this point, i.e., return to the original melodic context, so we should be cautious with interpretation here.

Our findings are consistent with the view that expert listeners do retrospective processing of musical materials and that they are sensitive to the global structure of phrase-length units, at least under active listening conditions. Musicians seem to be able to retrieve a phrase concept (four-bar, figural similarity) from long-term memory, and check whether the incoming information matches this template. This reliance on top-down strategies in phrase processing has already been stressed by Neuhaus et al. (2006), who suggested that non-musicians are more dependent than musicians on bottom-up processes with regard to detecting local discontinuities. Further research is needed to investigate wrap-up effects in non-musicians and to compare these with the responses we found in musicians.

3.2. *Implications for shared resources hypotheses*

The designation "Closure Positive Shift" was first applied to a positive peak elicited at the onset of the boundary pause following speech clauses (Kerkhofs et al., 2007; Li and Yang, 2009, 2010; Steinhauer et al., 1999), purely prosodic materials (Pannekamp et al., 2005), or sub-vocalizations of text (Steinhauer and Friederici, 2001). The language CPS has been elicited at the boundary onset (Kerkhofs et al., 2007; Pauker et al., 2011), and the music component at the boundary offset. These latency differences have been accounted for by the fact that speech provides more and earlier boundary cues than music (Knoesche et al., 2005; Neuhaus et al., 2006). Despite differences in the experimental paradigm and in the latency of the component, the idea of a common correlate of boundary processing (Koelsch, 2011a; Nan et al., 2009) was proposed as evidence of shared neural resources for music and language (Koelsch, 2011a, 2011b; Nan et al., 2009; Patel, 2003).

We found that the retrospective processing of the whole phrase had no effect on the music CPS, and we gave further support to the idea that a boundary pause is crucial for eliciting the CPS (see Neuhaus et al., 2006, for evidence that the boundary tone is also determinant). In contrast, the language CPS seems to be sensitive to retrospective processing, since the component is modulated by the hierarchical level of the preceding prosodic unit (Li and Yang, 2010). Also in contrast to music, the language CPS does not depend on the presence of a boundary pause (Steinhauer, 1999). Overall, this suggests that the functional meaning of the music and the language CPS may not overlap completely, and the CPS may not stand as a common cross-domain correlate when specific boundary-related processes (wrap-up, attention shifts) are considered. One possibility is that the CPS is a correlate of attention shifts in music and language (shared function), while only the language CPS reflects retrospective processes (non-shared function). According to this, music and language share resources for making the listener "move on" along the musical/linguistic stream, but not for integrating the information from each phrase/sentence (see Nan et al., 2008). Another possibility is that the language CPS reflects phrase wrap-up, the music CPS reflects attention shifts, and neither process is common to the two CPS components. This possibility is raised by confronting our findings with the lack of sensitivity to pauses (attention shifters) in the language CPS. However, we should keep in mind that phrase boundaries in language often manifest themselves as shifts of intonation and less often as distinctive pauses/rests. Therefore, the lack of sensitivity to pauses may simply reflect the irrelevance of pauses in typical speech structure.

3.3. *Notes on the response to melodic violations*

The onset of the intruder elicited central-posterior early negativities, as well as a late positive deflection peaking around 600 ms. The topographies of our early negativity seem more compatible with the widespread distribution of the N1/N2 components in studies on melodic violations (Besson and Macar, 1987; Nittono et al., 2000; Paller et al., 1992), than with the right anterior negativities (ERAN/RATN) found in response to harmonic violations (Koelsch et al., 2005, 2002; Patel et al., 1998). This finding does not look surprising, although the reason why harmonic and melodic violations should lead to different topographies is not well-understood. Concerning latency, our negative peak spread over a broader time window (0–500 ms) than we should expect based on results from studies on melodic violations (50–200 ms). One explanation might be that studies on melodic violations compared deviant with non-deviant terminal notes, while we compared deviant with non-deviant segments. In our study, the first intruder tone was followed by other incongruent tones (3 to 6 tones, average 4.6 within a time window up to 1000 ms). Thus, successive negativities, elicited by the successive individual notes of the segment, added into the broad negativity we observed. The enhanced positivity after the first note of the intruder segment is consistent with studies on harmonic (Koelsch et al., 2005, 2002; Maess et al., 2002; Patel et al., 1998) and melodic violations (Besson and Macar, 1987; Besson and Faita, 1995; Nittono et al., 2000; Paller et al., 1992), in which late positivities have been reported in response to the violating event.

4. Conclusion

In this EEG study, we narrowed the scope of hypotheses on the functional meaning of the music CPS by showing that it is unrelated to phrase wrap-up. We also found evidence for on-line wrap-up effects related to musical phrases, thus contributing to a better understanding of the neural correlates of music perception. Our findings raise new hypotheses concerning neural resources shared by language and music in the processing of phrase boundaries, namely that retrospective processes, attention processes, or both may not recruit the same resources in the two domains. Future within-subjects studies manipulating phrase well-formedness and boundary discontinuities in both music and language may shed light on this matter.

5. Experimental procedure

5.1. *Subjects*

Thirty subjects participated in the experiment, and 6 were excluded from analysis because of the amount of artefacts in their EEG data. Thus, 24 musicians (14 female, mean age=22.8, SD=4.1) took part in the experiment. On average, they had 10 years of formal music training and started to learn their instrument at 9 years of age. They reported a mean of 8 h per week of musical practice and 14 public performances per year. None reported hearing problems, psychiatric and/or neurological disorders, and none was taking any medication. They were all right handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971). All subjects signed informed consent and received a voucher as reward for their participation.

5.2. *Stimulus material and paradigm*

An initial set of 50 melodies was selected from not well-known baroque, classical and romantic pieces, as well as from examples provided in CPS studies (Knoesche et al., 2005; Nan et al., 2006, 2009; Neuhaus et al., 2006). We looked for sequences of two phrases with four bars each, in which the last (fourth) bar of the first phrase comprised a pause (Fig. 1). Most of them were highly parallel, in the sense that the material in the first phrase was repeated in the second one. From each of these 50 original melodies (condition *phrased well-formed*, *PHW*), we derived three additional versions (cf. Fig. 1).

In the *unphrased* (*UPH*) version, pauses were filled with tones intended to convey the impression of connecting both phrases. The filling tones followed the harmony, register and rhythmic features of the first phrase, and their melodic contour was congruent with the second phrase. We established that the time intervals between filling tones in the unphrased melodies should not be larger than any time interval within the first phrase. To generate melodies with non-well-formed phrases, we looked for motifs with the length of one bar. In the *phrased with intruder* (*PHN1*) version, we replaced one bar/motif of the first phrase with a bar/motif from another melody. In the *phrased with three bars* (*PHN2*) version, we removed the segment that was replaced in the PHN1 version. Since motifs provide segmentation points within the phrase, removing a motif should not affect the continuity of the phrase. So, while PHN1 versions had a typical length of four bars and a segment that did not relate to the rest of the phrase, PHN2 versions had an atypical length, but its constituent motifs were related. In order to avoid predictability, the intruder segment was placed in variable points of the phrase. In 29 out of the final set of 34 melodies (see below), the intruder was inserted at the second or third bar; it was placed at the beginning of the phrase in four melodies, and at the end in one. Melodies were recorded as audio files, using a high-quality piano-like synthesized timbre (Propellerhead Reason, v. 4.0 soundbank). In order to keep the focus on structure, human interpretation was eliminated, and intensity was kept constant. The mean length of the first phrase was 6266±2267 ms, excluding the pause. In the PHN2 condition, phrases were shorter (4520±1643 ms). The average pause length was 952±404 ms. We ran a pre-test to select an optimal stimulus subset. It was our goal to choose those melodies in which the perception of an intruder segment was limited to the PHN1 version. We presented all melodies in all versions to a group of subjects ($n=4$), and asked them to state if each of the 50×4 melodies contained an intruder segment. We chose a final set of 34 melodies.

Four different triggers (cf. Fig. 1) were placed in each melody. The first and second triggers (T1, T2) delimited the intruder segment. Triggers T1 and T2 marked the *onset* and *offset of the intruder segment* in the PHN1 version. In PHW and UPH versions, they were placed at the corresponding time points (onset/offset of the congruent segment). In PHN2, T1 marked the removal of the one-bar motif. We excluded T2 from this version because we had no equivalent time point. Unlike PHN1, the retrospective processing of a violation was not expected immediately after the segment removal in PHN2. Unlike PHW and UPH, PHN2 had no events between T1 and T2 (T2=T1), and comparisons of retrospective processing would be made between conditions with different lengths (PHW/UPH vs. PHN2). The third trigger (T3) was placed at the offset of the first phrase (the *boundary onset*). The last (T4) is the relevant trigger in CPS studies, and it corresponds to the *boundary offset*. The names of the third and fourth triggers refer to the phrased versions. In the unphrased version, they correspond to the onset and offset of filling tones.

Participants were asked to listen to the melodies and decide whether they contained an intruder, that is, a group of tones that was not congruent with the rest of the melody. They were told that we had replaced one segment of the melody with a segment from another melody. Thus, their attention was focused on the detection of PHN1 versions. In the pre-stimulus period we presented a fixation cross for 500 ms, which remained on the monitor during stimulus presentation. Participants responded in a self-paced manner, by pressing Yes or No in response to the question “Intruder?”. Half the participants used the left hand for the Yes response; the other half used the right hand. There was a 1500 ms inter-trial interval. Participants were asked to relax their facial muscles and to blink between the question and the next fixation cross. The 34×4 melodies were pseudo-randomized and organized into three blocks. Each block had approximately the same number of PHW, UPH, PHN1 and PHN2 versions. The block duration was approximately 14 min and the block order was counterbalanced. A sequence of four practice trials started the experimental session. Stimuli were presented with Presentation software (http://nbs.neuro-bs.com/presentation). Participants listened to the melodies through high-quality headphones (Technics, RP-F290).

5.3. *Recording procedure*

We recorded EEG activity on 29 electrode positions distributed over the entire scalp (Fpz, Fp2, F7, F3, Fz, F4, F8, M1, C3, Cz, C4, M2, P7, P3, Pz, P4, P8, O1, Oz, O2 from the 10–20 system; FC5, FC1, FC2, FC6, CP5, CP1, CP2, CP6, POz from the 10–10 extension). The ground electrode was placed at electrode position AFz. We used an elastic cap (Waveguard 32, ANT) with Ag–AgCl electrodes inserted. For the recording of electro-oculographic (EOG) activity, two flat-type electrodes were placed at the outer canthi of the eyes, and another under the right eye. The latter was off-line re-referenced to Fp2 in order to visualize vertical EOG activity. The EEG was digitized on-line with a sampling rate of 512 Hz, with software ASA 4.1 (ANT Software B. V., Enschede, Netherlands). The 32-channel amplifier (ANT, refa32) had a resolution of 71.5 nV per bit. Recordings of all channels were referenced to the left mastoid, and active channels were off-line re-referenced to the average of the two mastoids. Impedances were kept below 10 KΩ throughout the experiment.

5.4. *Signal analysis*

ERPs were analyzed at four different trigger points: the onset/offset of the intruder (T1–T2), the boundary onset (T3) and the boundary offset (T4). Preprocessing and grandaveraging was performed with the fieldtrip toolbox for Matlab (Oostenveld et al., 2011). Data were segmented into epochs of 1200 ms, from −200 ms to 1000 ms around the trigger points. ERPs were baseline-corrected with respect to the 100 ms pre-stimulus interval. In T1 and T2 analyses, we used a reference period of 50 ms pre-trigger in order to minimize cross-condition imbalances in the pre-stimulus period. EOG artefacts were first identified by visual means, and subsequent threshold-based analyses marked other deviant trials. All contaminated trials were rejected (7.9% for T1, 9.4% for T2, 7.1% for T3, and 12.3% for T4). Preprocessed epochs were bandpass filtered from 0.01 Hz to 25 Hz.

Time windows were defined by visual inspection in T1 (0–500, 500–800 ms), T2 (0–300, 300–700, 700–1000 ms) and T2 (0–300, 300–500, 500–700, 700–1000 ms) analysis. In T4 analysis, we focused on time windows pointed out as relevant in the music CPS literature (Knoesche et al., 2005; Nan et al., 2006, 2009; Neuhaus et al., 2006): the P2 time window (150–250 ms), and the CPS time window (500–600 ms). We discarded melodies with the intruder at phrase-initial position ($n=4$) from T1 analysis, and melodies with an intruder at phrase-final position ($n=1$) in T2 and T3. The three Regions Of Interest (ROIs) in T1 and T2 analysis comprised electrodes located near the midline, where effects were apparent: F3, Fz, F4, FC1 and FC2 (anterior); C3, Cz, C4, CP1 and CP2 (central); P3, Pz, P4, Poz and Oz (posterior). In T3 and T4, the following electrodes were considered: frontal (F7, F3, Fz, F4, and F8) and fronto-central sites (FC5, FC1, FC2, and FC6) in the anterior region, central (C3, Cz, and C4) and centro-parietal sites (CP5, CP1, CP2, and CP6) in the central region, and parietal (P7, P3, Pz, P4, and P8), one parietal-occipital (POz) and occipital (O1, Oz, and O2) in the posterior region.

We tested the effects of melodic condition and region (3 levels: anterior, central, posterior) by means of repeated measure ANOVAs. In T1 and T4 analyses, melodic condition had 4 levels (UPH, PHW, PHN1, and PHN2). In the intruder offset (T2) analysis, PHN2 was excluded for reasons stated above (see 5.2). In the boundary onset (T3) analysis, UPH was excluded because we did not expect a comparable wrap-up process (there was no boundary pause). Therefore, only 3 levels of melodic condition were considered in T2 (UPH, PHW, PHN1) and T3 (PHW, PHN1, PHN2). Separate ANOVAs were carried out for comparing melodic conditions two at a time, following the hypothesized contrasts: local violation (PHN1) vs. no local violation (PHW, UPH, PHN2) for the onset of the intruder (T1); impaired (PHN1) vs. non-impaired (PHW, UPH) retrospective processing before the end of the phrase for the offset of the intruder (T2); high (PHN1, PHN2) vs. low phrase wrap-up demands (PHW) for the boundary onset (T3); boundary pause (PHW, PHN1, PHN2) vs. no boundary pause (UPH) for the P2/CPS effect at the boundary offset (T4). Interactions with region were explored by means of Tukey HSD post hoc comparisons. A critical p-value of 0.05 was adopted. Greenhouse–Geisser corrections were applied to violations of sphericity. Corrected degrees of freedom are specified in the text.

Concerning behavioral results, we tested accuracy against chance levels by means of one-sample t-tests, and we compared accuracy levels across melodic conditions with a repeated measures ANOVA.

Acknowledgments

This work was supported by Fundação para a Ciência e a Tecnologia (PTDC/PSI/66641/2006, PEst-C/PSI/IU0050/2011, PEst-OE/EQB/LA0023/2011, PEst-OE/EQB/LA0023/2013, PTDC/PSI-PCO/110734/2009, IBB/CBME, LA, FEDER/POCI 2010), Vetenskapsrådet.

REFERENCES

Aaronson, D., Scarborough, H.S., 1976. Performance theories for sentence coding: some qualitative evidence. J. Exp. Psychol.: Hum. Percept. Perform. 2, 56–70, http://dx.doi.org/10.1037/0096-1523.2.1.56.

Aaronson, D., Scarborough, H.S., 1977. Performance theories for sentence coding: some quantitative models. J. Verbal Learn. Verbal Behav. 16, 277–303, http://dx.doi.org/10.1016/S0022-5371(77)80052-2.

Berg, A., 1965. Why is Schoenberg's music so difficult to understand?. In: Reich, W. (Ed.), Alban Berg. Hartcout, New York, pp. 189–204.

Besson, M., Faita, F., 1995. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. J. Exp. Psychol.: Hum. Percept. Perform. 21, 1278–1296, http://dx.doi.org/10.1037/0096-1523.21.6.1278.

Besson, M., Macar, F., 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology 24, 14–25, http://dx.doi.org/10.1111/j.1469-8986.1987.tb01853.x.

Bigand, E., 1990. Abstraction of two forms of underlying structure in a tonal melody. Psychol. Music 18, 45–60, http://dx.doi.org/10.1177/0305735690181004.

Cooper, G., Meyer, L., 1960. The Rhythmic Structure of Music. University of Chicago Press, Chicago.

Davie, C.T., 1966. Musical Structure and Design. Dover, New York.

De Diego Balaguer, R., Toro, J.M., Rodriguez-Fornells, A., Bachoud-Lévi, A.-C., 2007. Different neurophysiological mechanisms underlying word and rule extraction from speech. PLoS One 2 (11), e1175, http://dx.doi.org/10.1371/journal.pone.0001175.

Feldstein, S., 1995. Alfred's pocket dictionary of music, third edition Alfred Publishing Co., Inc., Van Nuys.

Friedman, D., Simson, R., Ritten, W., Rapin, I., 1975. The late positive component and information processing in sentences. Electroencephalogr. Clin. Neurophysiol. 38 (3), 255–262, http://dx.doi.org/10.1016/0013-4694(75)90246-1.

Hagoort, P., 2003. Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. J. Cogn. Neuroscience 15 (6), 883–899, http://dx.doi.org/10.1162/089892903322370807.

Hagoort, P., Brown, C.M., 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. J. Psycholinguist. Res. 28 (6), 715–728, http://dx.doi.org/10.1023/A%3A1023277213129.

Hagoort, P., Wassenaar, M., Brown, C.M., 2003. Syntax-related ERP-effects in Dutch. Cogn. Brain Res. 16, 38–50, http://dx.doi.org/10.1016/S0926-6410(02)00208-2.

Istók, E., Friberg, A., Huotilainen, M., Tervaniemi, M., 2013. Expressive timing facilitates the neural processing of phrase boundaries in music: evidence from event-related potentials. PLoS One 8 (1), e55150, http://dx.doi.org/10.1371/journal.pone.0055150.

Just, M.A., Carpenter, P., 1980. A theory of reading: from eye fixations to comprehension. Psychol. Rev. 87, 329–354.

Kerkhofs, R., Vonk, W., Schriefers, H., Chwilla, D.J., 2007. Discourse, syntax and prosody: the brain reveals an immediate interaction. J. Cogn. Neurosci. 19 (9), 1421–1434, http://dx.doi.org/10.1162/jocn.2007.19.9.1421.

Knoesche, T.R., Neuhaus, C., Haueisen, J., Alter, K., 2003. The role of the planum temporale in the perception of musical phrases. In: Proceedings of the 4th International Conference on Noninvasive Functional Source Imaging (NFSI), Chieti, Italy, September 2013.

Knoesche, T.R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O.W., Friederici, A.D., 2005. Perception of phrase structure in music. Hum. Brain Map. 24 (4), 259–273, http://dx.doi.org/10.1002/hbm.20088.

Koelsch, S., 2011a. Toward a neural basis of music perception – a review and updated model. Front. Psychol. 2, 110, http://dx.doi.org/10.3389/fpsyg.2011.00110.

Koelsch, S., 2011b. Towards a neural basis of processing musical semantics. Phys. Life Rev. 8, 89–105, http://dx.doi.org/10.1016/j.plrev.2011.04.004.

Koelsch, S., Gunter, T.C., Wittfoth, M., Sammler, D., 2005. Interaction between syntax processing in language and in music: an ERP Study. J. Cogn. Neurosci. 17 (10), 1565–1577, http://dx.doi.org/10.1162/089892905774597290.

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 (5), 657–663, http://dx.doi.org/10.1111/1469-8986.3950657.

Kotz, S.A., Frisch, S., Von Cramon, D.Y., Friederici, A.D., 2003. Syntactic language processing: ERP lesion data on the role of the basal ganglia. J. Int. Neuropsychol. Soc. 9, 1053–1060, http://dx.doi.org/10.1017/S1355617703970093.

Lalitte, P., Bigand, E., Poulin-Charronat, B., McAdams, S., Delbé, C., D'Adamo, D., 2004. The perceptual structure of thematic materials in "The Angel of Death". Music Percept.: Interdiscip. J. 22 (2), 265–296, http://dx.doi.org/10.1525/mp.2004.22.2.265.

Lerdahl, F., Jackendoff, R., 1983. A Generative Theory of Tonal Music. MIT Press, Cambridge.

Levinson, J., 1997. Music in the Moment. Cornell University Press, Ithaca, NY.

Li, W., Yang, Y., 2009. Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. Neuroscience 158 (4), 1416–1425, http://dx.doi.org/10.1016/j.neuroscience.2008.10.065.

Li, W., Yang, Y., 2010. Perception of Chinese poem and its electrophysiological effects. Cogn. Behav. Syst. Neurosci. 168 (3), 757–768, http://dx.doi.org/10.1016/j.neuroscience.2010.03.069.

Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2002. Musical syntax is processed in Broca's area: an MEG study. Nat. Neurosci. 4 (5), 540–545, http://dx.doi.org/10.1038/87502.

McAdams, S., 1989. Psychological constraints on form-bearing dimensions in music. Contemporary Music Rev. 4 (1), 181–198, http://dx.doi.org/10.1080/07494468900640281.

McAdams, S., Matzkin, D., 2001. Similarity, invariance, and musical variation. Ann. N. Y. Acad. Sci. 930, 62–76, http://dx.doi.org/10.1111/j.1749-6632.2001.tb05725.x.

McAdams, S., Matzkin, D., 2003. The roots of musical variation in perceptual similarity and invariance. In: Peretz, I., Zatorre, E. (Eds.), The Cognitive Neuroscience of Music. Oxford University Press, Oxford, pp. 79–94.

Minati, L., Salvatoni, L., Rosazza, C., Pietrocini, E., Visani, E., Panzica, F., Scaioli, V., Avanzini, G., Franceschetti, S., 2010. Event-related potential (ERP) markers of melodic processing: the N2 component is modulated by structural complexity, not melodic 'meaningfulness'. Brain Res. Bull. 83 (1/2), 23–28, http://dx.doi.org/10.1016/j.brainresbull.2010.06.011.

Mitchell, D.C., Green, D.W., 1978. The effects of context and content on immediate processing in reading. Q. J. Exp. Psychol. 30, 609–636, http://dx.doi.org/10.1080/14640747808400689.

Nan, Y., Knöesche, T.R., Friederici, A.D., 2006. The perception of musical phrase structure: a cross-cultural ERP study. Brain Res. 1094 (1), 179–191, http://dx.doi.org/10.1016/j.brainres.2006.03.115.

Nan, Y., Knöesche, T.R., Friederici, A.D., 2009. Non-musicians' perception of phrase boundaries in music: a cross-cultural ERP study. Biol. Psychol. 82 (1), 70–81, http://dx.doi.org/10.1016/j.biopsycho.2009.06.002.

Nan, Y., Knösche, T.R., Zysset, S., Friederici, A.D., 2008. Cross-cultural music phrase processing: an fMRI study. Hum. Brain Map. 29 (3), 312–328, http://dx.doi.org/10.1002/hbm.20390.

Narmour, E., 1992. The Analysis and Cognition of Melodic Complexity: The Implication-Realisation model. University of Chicago Press, Chicago.

Neuhaus, C., 2013. Processing musical form: behavioural and neurocognitive approaches. Musicae Sci. 17 (1), 109–127, http://dx.doi.org/10.1177/1029864912468998.

Neuhaus, C., Knoesche, T.R., Friederici, A.D., 2006. Effects of musical expertise and boundary markers on phrase perception in music. J. Cogn. Neurosci. 18 (3), 472–493, http://dx.doi.org/10.1162/jocn.2006.18.3.472.

Nittono, H., Bito, T., Hayashi, M., Sakata, S., Hori, T., 2000. Event-related potentials elicited by wrong terminal notes: effects of temporal disruption. Biol. Psychol. 52 (1), 1–16, http://dx.doi.org/10.1016/S0301-0511(99)00042-3.

Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh Inventory. Neuropsychologia 9 (1), 97–113, http://dx.doi.org/10.1016/0028-3932(71)90067-4.

Oostenveld, R., Fries, P., Maris, E., Schoffelen, J-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011http://dxdoi.org/10.1155/2011/156869 (Article ID 156869, 9 pp.).

Paller, K.A., McCarthy, G., Woods, C.C., 1992. Event-related potentials elicited by deviant endings to melodies. Psychophysiology 29 (2), 202–206, http://dx.doi.org/10.1016/0013-4694(87)90124-6.

Pannekamp, A., Toepel, U., Alter, K., Hahne, A., Friederici, A.D., 2005. Prosody driven sentence processing: an event-related brain potential study. J. Cogn. Neurosci. 17 (3), 407–421, http://dx.doi.org/10.1162/0898929053279450.

Patel, A.D., 2003. Language, music, syntax and the brain. Nat. Neurosci. 6 (7), 674–681, http://dx.doi.org/10.1038/nn1082.

Patel, A.D., Gibson, E., Ratner, J., Besson, M., Holcomb, P.J., 1998. Processing syntactic relations in language and music: an event-related potential study. J. Cogn. Neurosci. 10 (6), 717–733, http://dx.doi.org/10.1162/089892998563121.

Pauker, E., Itzhak, I., Baum, S.R., Steinhauer, K., 2011. Effects of cooperating and conflicting prosody in spoken english garden path sentences: ERP evidence for the boundary deletion hypothesis. J. Cogn. Neurosci. 23 (10), 2731–2751, http://dx.doi.org/10.1162/jocn.2011.21610.

Pearce, M.T., Müllensiefen, D., Wiggins, G.A., 2010. Melodic grouping in music information retrieval: new methods and applications. In: Ras, Z.W., Wieczorkowska, A. (Eds.), Advances in Music Information Retrieval. Springer, Berlin, pp. 364–388.

Peretz, I., 1989. Clustering in music: an appraisal of task factors. Int. J. Psychol. 24, 157–178, http://dx.doi.org/10.1080/00207594.1989.10600040.

Purwins, H., Herrera, P., Grachten, M., Hazan, A., Marxer, R., Serra, X., 2008. Computational models of music perception and cognition I: the perceptual and cognitive processing chain. Phys. Life Rev. 5 (3), 151–168, http://dx.doi.org/10.1016/j.plrev.2008.03.004.

Riemann, H., 1929. Composición Musical. Editorial Labor, Barcelona.

Schoenberg, A., 1967. Fundamentals of Musical Composition. Faber and Faber, London.

Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E., 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J. Neurosci. 23, 5545–5552.

Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B., 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. Neuroreport 16, 1781–1785, http://dx.doi.org/10.1097/01.wnr.0000185017.29316.63.

Silva, S., Barbosa, F., Marques-Teixeira, J., Petersson, K.M., Castro, S.L., 2014. You know when: event-related potentials and theta/beta power indicate boundary prediction in music. J. Integr. Neurosci. 13, 1–16, http://dx.doi.org/10.1142/S0219635214500022.

Sloboda, J., 1985. The Musical Mind. Clarendon Press, Oxford.

Snyder, J.S., Alain, C., Picton, T.W., 2006. Effects of attention on neuroelectric correlates of auditory stream segregation. J. Cogn. Neurosci. 18 (1), 1–13, http://dx.doi.org/10.1162/089892906775250021.

Stein, E., 1962. Form and Performance. Faber, London.

Steinhauer, K., Alter, K., Friederici, A.D., 1999. Brain potentials indicate immediate use of prosodic cues in natural speech processing. Nat. Neurosci. 2, 191–196, http://dx.doi.org/10.1038/5757.

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. J. Psycholinguist. Res. 30 (3), 267–295, http://dx.doi.org/10.1023/A:1010443001646.

Tan, N., Aiello, R., Bever, T., 1981. Harmonic structure as a determinant of melodic organization. Memory Cogn. 9, 533–539, http://dx.doi.org/10.3758/BF03202347.

Van Petten, C., Kutas, M., 1991. Influences of semantic and syntactic closure on open- and closed-class words. Memory Cogn. 19, 95–112, http://dx.doi.org/10.3758/BF03198500.

Weber, K., Lavric, A., 2008. Syntactic anomaly elicits a lexico-semantic (N400) ERP effect in the second language but not the first. Psychophysiology 45 (6), 920–925, http://dx.doi.org/10.1111/j.1469-8986.2008.00691.x.