

Perceiving rhythm where none exists: Event-Related Potential (ERP) correlates of subjective accenting

D. Potter, M. Fenwick, D. Abecasis, Renaud Brochard

► **To cite this version:**

D. Potter, M. Fenwick, D. Abecasis, Renaud Brochard. Perceiving rhythm where none exists: Event-Related Potential (ERP) correlates of subjective accenting. *Cortex*, Elsevier, 2009, 1 (45), pp.103-109. 10.1016/j.cortex.2008.01.004 . hal-00570452

HAL Id: hal-00570452

<https://hal-univ-bourgogne.archives-ouvertes.fr/hal-00570452>

Submitted on 5 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting

Douglas D. Potter, Maggi Fenwick, Donna Abecasis and Renaud Brochard

Abstract

Previous research suggests that our past experience of rhythmic structure in music results in a tendency for Western listeners to subjectively accent equitonal isochronous sequences. We have shown in an earlier study that the occurrence of a slightly softer tone in the 8th to 11th position of such a sequence evokes a P300 event-related potential (ERP) response of different amplitudes depending on whether the tone occurs in putatively subjectively accented or unaccented sequence positions (Brochard et al., 2003). One current theory of rhythm processing postulates that subjective accenting is the result of predictive modulations of perceptual processes by the attention system. If this is the case then ERP modulations should be observed at an earlier latency than the P300 and these should be observed in ERPs to both standard and softer tones. Such effects were not observed in our previous study. This was possibly due to the use of a linked-mastoid reference which may have obscured lateralized differences. The aim of the present study was to replicate the previous auditory P300 subjective accenting findings and to investigate the possibility that these effects are preceded by ERP changes that are indicative of rhythmic modulation of perceptual processing. Previous auditory P300 findings were replicated. In addition and consistent with current theories of rhythm processing, early brain ERP differences were observed both in standard and deviant tones from the onset of the stimulus. These left lateralized differences are consistent with a rhythmic, endogenously driven, modulation of perception that influences the conscious experience of equitonal isochronous sequences.

Keywords: Music; Rhythm; Perception; Asymmetry; Attention

1. Introduction

While rhythm surrounds us throughout our whole life and is inherent in many mental activities, the neural mechanisms underlying rhythm perception remain largely unclear. The perception of rhythm is a dynamic process which involves the synchronisation of external musical stimuli with internal rhythmic processes (Jones and Boltz, 1989). Rhythm often refers to the organization of events in time, such that they are organized perceptually into groups. For instance, the perception of meter, i.e. the tendency to periodically group sound events, perceiving an alternation of accented (“strong”) and unaccented (“weak”) beats, takes place even in perfectly regular sequences of identical tones. This type of subjective accent imposed by listeners has long been described in behavioural studies (Bolton, 1894; Woodrow, 1909; Fraisse, 1982; Drake, 1993; Parncutt, 1994) where spontaneous grouping and accenting of tones, most frequently by twos or fours, have been reported. While the underlying cause of these simple forms of subjective accenting is not clear, it is evident that cultural differences in experience of musical rhythms influence the accuracy of perception of more complex rhythms (Hannon and Trehub, 2005). These phenomena are consistent with most theoretical conceptions of meter as a hierarchical structure. In the present study we are interested in determining whether there is evidence of the most basic level of metrical structure, which corresponds to an alternation of strong and weak beats. An important assumption is that the first stimulus in the sequence is more salient and receives more attention than following items (Thomassen, 1982) and, as a consequence, establishes the pattern of accenting within an isochronous sequence. Such accenting effects are consistent with a generative model of representation structure in the brain (Friston, 2002). In this model, the brain is continually predicting current spatio-temporal patterns of input on the basis of past patterns of input and new stimuli are accommodated within pre-existing representational structures.

Thus the perception of stimuli may be subtly altered by prior expectations. If, however, a stimulus deviates significantly from these expectations, error signals will be generated to allow accommodation of this new information within pre-existing structures. These signals can be detected using event-related potential (ERP) measures. For example, such error signals

may be detected as a mismatch negativity (MMN) or, if the signal deviates considerably, this may result in the activation of attention mechanisms marked by an N2/P3 ERP complex. In previous research (Brochard et al., 2003) it was found that a 4dB reduction in tone amplitude, introduced in the latter part of an isochronous equitonal sequence, is processed differently depending on whether it occurs in odd (putatively subjectively accented) rather than even (putatively subjectively unaccented) positions. Softer tones in odd numbered positions evoked a larger P300 brain ERP response, reflecting an apparent binary pattern of metrical accentuation (Abecasis et al., 2005). This component, peaking at about 300–600 msec post-stimulus onset, is elicited by violations of listeners' expectancies and both its amplitude and latency depend upon listeners' attention and the degree of difficulty in the decision-making process of the task, in this case counting the number of infrequent lower amplitude tones (Donchin and Coles, 1988; Janata, 1995; Besson and Faïta, 1995; Polich and Kok, 1995; Granot and Donchin, 2002). The differences in P300 amplitude provide clear evidence of a subjective difference in the processing of softer (deviant) tones in odd and even sequence positions but did not provide any basis for determining how early this subjective accenting effect influences stimulus processing. Jones (Jones, 1976; Jones and Boltz, 1989; Drake et al., 2000) postulates that attention is synchronized to regular auditory sequences, through rhythmic expectancies for the occurrence of the next salient beat. On this basis one would predict that ERP modulations might distinguish subjectively accented and unaccented tones, possibly from stimulus onset or before stimulus onset. This was not, however, observed in our previous study (Brochard et al., 2003). The lack of earlier differences between the ERP responses to putatively accented and unaccented tone stimuli could have been due to the use of a linked-mastoid reference.

Although often chosen as neutral reference for ERP recordings, these sites are sensitive to activity in primary stages of auditory processing in the cortex. If activation at the two linked-mastoid electrode sites is different then a current will flow between the electrodes and cause local distortion of the recording of field potentials from the surface of the head. This may have resulted in the masking of low level accenting effects emanating from the temporal region and post-hoc rereferencing would not resolve this problem. In this study a midline reference was used to remove this confound. It is also likely

that dynamic modulations of perceptual processes should be lateralized to the left hemisphere (Platel et al., 1997; Potter et al., 2000; Vuust et al., 2005). In the positron emission tomography study of Platel et al. participants selectively attended to familiarity, pitch, rhythm and timbre of randomly arranged sequences of notes. Attending to familiarity, pitch and rhythm preferentially activated left hemisphere sites and attending to timbre activated frontal regions of the right hemisphere. In the Potter et al. ERP study participants were instructed simply to listen to modern polyrhythmic African music for a brief period of time. A single trial across-subject averaging technique was used to visualize common ERP deflections. ERP deflections that were synchronized to the music and located predominantly over the left hemisphere were observed. In this latter study regions of the brain associated with auditory processing appear to be driven by complex structure of the rhythmical sequences in the music. Vuust et al. (2005) used the MMN as a measure of sensitivity to rhythmic structure. They found that both musicians and non-musicians produce an MMN to temporal violations of rhythmic structure. However, musicians produced a larger response over the left than right hemispheres suggesting an effect of training on lateralization of rhythm processing. The MMN also had a shorter latency in musicians than non-musicians. In the present study the assumption is that our extensive experience with music will result in individuals imposing a simple implicit rhythmic structure on the isochronous equitonal stimuli that they listen to and that this will be more strongly lateralized to the left hemisphere in trained musicians. In the present study a nose reference was used and mastoid electrodes adjacent to the temporal lobe were included as active electroencephalogram (EEG) recording sites to maximize the likelihood of detecting evidence of subjective accenting effects occurring in cortical regions involved in auditory perception.

Previous research would suggest that dynamic modulations of attention predicted by Jones and collaborators (Jones, 1976; Jones and Boltz, 1989; Drake et al., 2000) might take the form of a "processing negativity" (PN) in the event-related brain potentials generated by the presentation of tone sequences (Näätänen, 1982, 1992). These ERP modulations were first described in dichotic listening tasks as the negative shift found in attended as compared to unattended channels (Hillyard et al., 1973; Näätänen et al., 1978) and these differences may start as early as 50 msec post-stimulus onset. A potentially

confounding effect resulting from the short interstimulus interval (ISI) used in this and previous studies is that any early effects that were observed could result from overlapping ERP deflections from the previous stimulus. For instance, Starr et al. (1997) found a negative slow wave in ERP to frequent standard tones around 380–680 msec post-stimulus onset that increased in amplitude throughout a stimulus sequence, being larger before, than after, a deviant tone occurred. However, in the present study we predict a dynamic 'subjective accenting' modulation that will affect tones based on their position in the sequence and not their actual amplitude. In addition the accented beat naturally precedes the unaccented beat in the simplest rhythm structures suggesting that an opposite pattern of relative negativity would be observed in the present study. In summary, previous research suggests that individuals subconsciously impose rhythmic structure to isochronous equitonal sequences and this can be demonstrated as modulations of the P300 ERP using a target detection paradigm (Brochard et al., 2003). The main aim of the present study was to test the hypothesis that the P300 effects, associated with the end of stimulus evaluation, are preceded by differences in ERP deflections that mark dynamic modulations of perceptual processes by attention mechanisms (Jones, 1976; Jones and Boltz, 1989; Drake et al., 2000) or the activation of temporally bound rhythmic representation structures as suggested by generative models of perception (Friston, 2002). These ERP deflections should occur from stimulus onset as they represent an imposition of structure on the input rather than the detection of deviation from expected input. As such the modulation should be present in response to both infrequent deviant and frequent standard tones that occur in putatively subjectively accented positions in the latter part of tone sequences. The same oddball paradigm as employed by Brochard et al. (2003) was used. Only musically trained participants were recruited for this experiment since, in our previous study, musicians produced more robust effects of subjective accenting on the P300.

2. Methods

Ten volunteer participants (seven male, three female) with normal hearing took part in this study. The age range was 22–55 years old (mean age $\frac{1}{4}$ 43.3 yrs). All participants had a minimum of 8 years formal music training

(mean \pm 9.4 yrs). Each participant gave their written consent after the nature of the experiment was fully explained to them. Stimuli consisting of isochronous sequences of 13–16 70dB SPL standard tones were created (to avoid inducing a 4/4 meter). One or two of the tones in each sequence were replaced by 66dB SPL deviant tones. The first deviant tone could occur in one of four different positions in each sequence. These positions corresponded to either a subjectively accented beat (positions 9 or 11) or a subjectively unaccented beat (positions 8 or 10). A 4dB decrease in volume is considered a slight change for an individual to detect and is equivalent to the size of a subjective accent (Povel and Okkerman, 1981; see Brochard et al., 2003). Each tone had a frequency of 440 Hz and duration of 50 msec and rise and fall time of 10 msec. Half of the sequences contained one 66dB deviant tone, and half contained two 66dB deviant tones to reduce predictability and maintain attentiveness. Where there were two 66dB deviant tones, only the first of these were used in the analysis. The ISI in a sequence was 600 msec. During the EEG recordings, the stimuli were presented binaurally via headphones. Participants were instructed to visually fixate on a small red circle placed at a distance of 2m and to minimize both body and eye movements. Participants were then instructed to count the number of infrequent, deviant, soft tones they heard in each sequence and report this at the end of the sequence. When participants reported no soft tones; one soft tone when there were actually two; two soft tones when there was actually one; or more than two soft tones, these were counted as error trials. It should be noted that these deviant soft tones are effectively target stimuli for the participant and are sometimes labeled as such in P300 experiments that involve active detection of infrequent deviant stimuli. No feedback on accuracy was provided during the task. In order to minimize guessing and predictability, within each block of trials the sequences were presented randomly. A block of trials consisted of 16 isochronous sequences and each participant was presented with six blocks of trials. The blocks were separated by short rest periods of 30 sec. The test duration was approximately 25 min. Participants were briefly interviewed post-testing to obtain feedback regarding the degree of difficulty of the task. Each participant commented that the occurrence and frequency of the soft tones were unpredictable which suggests that no simpler response bias was operating. Continuous EEG was recorded using Contact Precision amplifiers and

Neuroscan software with silver/silver chloride electrodes mounted in an EasyCap headcap. The EEG was recorded (.03–100 Hz band pass; 400 Hz digitization rate; offline low-pass filter: 45 Hz, 48 dB/oct) with 11 electrodes attached to the scalp along the midline (Fz, Cz, Pz), temporal region (LT, between T7 and FT7, RT, between T8 and FT8), parietal region (P3, P4) and the left and right mastoids (LM, RM). The electro-oculogram (EOG) was monitored from electrodes placed on the infraorbital and supraorbital ridges of the right eye (vertical eye movements, VEOG) and at the outer canthus of both eyes (horizontal eye movements, HEOG). The reference electrode was placed on the nose. Impedances for all participants were 4–7 kohms. EEG epochs (–100 to 900 msec with respect to the stimulus onset) were averaged separately for 66dB and 70dB stimuli for both the putatively subjectively accented and unaccented stimuli. The pre-stimulus interval was used for baseline correction. All samples containing EEG artifacts greater than ± 60 mV were rejected. Repeated measures analysis of variance (ANOVA) was carried out on mean amplitude measurements from selected time windows of the ERPs. The factors used in the analysis were tone amplitude (standard/deviant), subjective accenting (subjectively accented/subjectively unaccented) and electrode site. Separate analyses were carried out on midline and lateral electrodes depending on the specific feature of the ERP deflection that was being analyzed. The Greenhouse–Geisser correction was applied in cases where there were more than two levels in a factor. Original degrees of freedom and corrected significance levels are given.

3. Results

Participants made an average of 15% errors in identifying the correct number of 66dB tones in the sequences. The grand averaged waveforms for the subjectively accented and unaccented standard and deviant tones are illustrated in Fig. 1. Deviant tones evoke an N2/P300 complex and the P300 is larger when the deviant tone occurs in a putatively subjectively accented position in the tone sequence. Mean amplitude measures at the midline sites Fz, Cz and Pz in the latency range 500–600 msec are illustrated in Fig. 2. ANOVA of these data, with factors of tone amplitude, subjective accenting and electrode site (Fz, Pz, Cz) produced the following. As predicted, deviant tones evoked significantly larger ERPs than standards ($F(1,9) = 11.65$, $p = .008$, $ES = .4$

.564). There was also a significant interaction between tone amplitude and subjective accenting ($F(1,9) = 7.36, p = .02, ES = .450$) due to P300 responses to deviants being larger in the subjectively accented positions than the unaccented ones. There was no three-way interaction between accenting, tone amplitude and site. It is evident in Fig. 1 that accented standard tones are relatively more negative than unaccented standard tones in the latency range 200–500 msec. Mean amplitude measures in this latency range were used to characterise this difference. ANOVA with factors of accenting, hemisphere, and site (temporal, mastoid, parietal) revealed the following. ERPs to putatively subjectively accented tones were significantly more negative than those to putatively unaccented tones ($F(1,9) = 6.11, p = .035, ES = .404$). A significant interaction between accenting and hemisphere was also observed ($F(1,9) = 6.01, p = .037, ES = .400$) and this was due to a larger accenting effect over the left hemisphere than the right hemisphere. There was no significant interaction between site (anterior–posterior) and hemisphere or accent. Both the attention synchronisation theory of Jones and Boltz (1989) and predictive coding theory suggest that early subjective accenting effects may be observable in ERPs at sites close to auditory cortex.

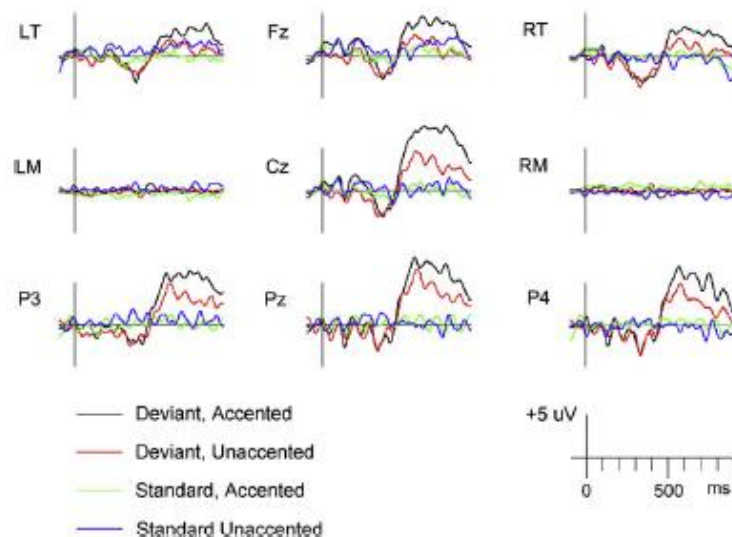


Fig. 3 – Grand averaged ERPs evoked by frequent standard 70dB and infrequent deviant 66dB tones in putative subjectively accented and unaccented sequence positions, re-referenced to linked-mastoids. Although the effect of subjective accenting on P300 is preserved, the early onset accenting effect over the left hemisphere is no longer present.

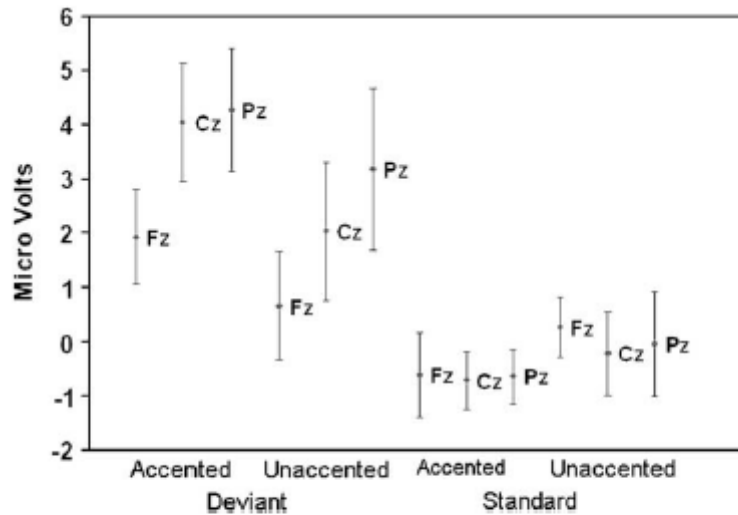
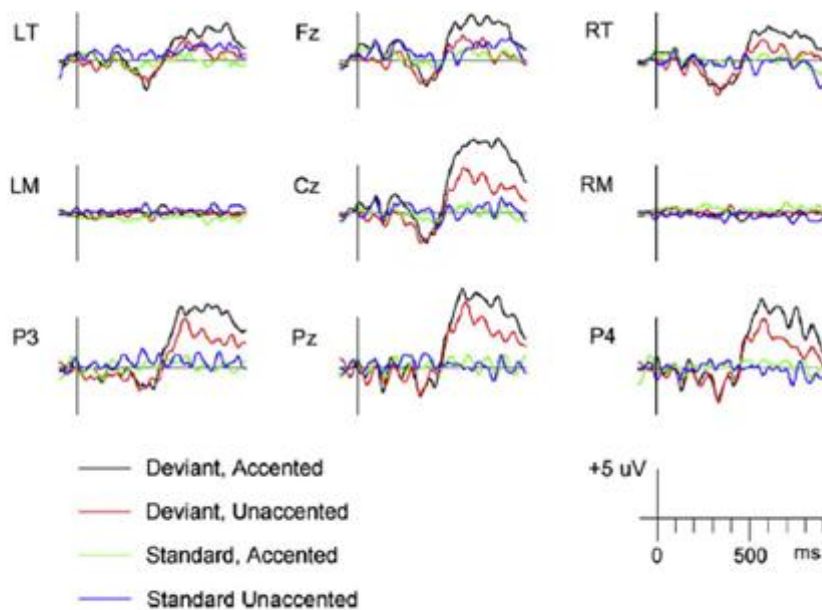


Fig. 2 – Distribution of mean amplitude of P300 at midline electrode sites in the latency range 500–600 msec evoked by putative subjectively accented and unaccented deviant and standard tones. Error bars indicate ± 1 standard error of mean.



Grand averaged ERPs evoked by frequent standard 70dB and infrequent deviant 66dB tones in putative subjectively accented and unaccented sequence positions, re-referenced to linked-mastoids. Although the effect of subjective accenting is preserved, the early onset accenting effect over the left hemisphere is no longer present.

It can be seen in Fig. 1 that the ERP deflections are more negative in the subjectively accented conditions at the mastoid electrodes in the first 100 msec after stimulus onset. ANOVA of mean amplitudes of the ERPs at the mastoid sites in the latency range 0–100 msec with

factors of subjective accenting, tone amplitude and hemisphere was carried out. Mean amplitudes in this latency range were significantly more negative in the accented conditions than the unaccented conditions ($F(1,9) = 9.95$, $p = .012$, $ES = .525$). The predicted interaction between subjective accenting and hemisphere was not significant ($F(1,9) = 2.07$, $p = .184$, $ES = .187$). However, separate exploratory analyses of the subjective accenting effects at left and right mastoid electrodes indicated that the effect was significant at the left electrode ($F(1,9) = 7.31$, $p = .024$, $ES = .448$) but not the right electrode ($F(1,9) = 3.34$, $p = .101$, $ES = .271$). Averages were digitally re-referenced to a linked-mastoid to allow comparison with the previous study of Brochard et al. (2003) and are illustrated in Fig. 3. ANOVA of mean amplitude measures in the latency range 500–900 msec at Cz, with factors of subjective accent and tone amplitude revealed that the P300 evoked by 66dB deviant tones was significantly different from ERPs to standard 70dB tones ($F(1,9) = 11.53$, $p = .008$, $ES = .562$). There was no main effect of subjective accent but there was a significant interaction between subjective accent and tone amplitude ($F(1,9) = 5.77$, $p = .040$, $ES = .391$). This was due to a significant difference in the amplitude of the P300 evoked by putatively subjectively accented and unaccented 66dB tones ($F(1,9) = 6.13$, $p = .035$, $ES = .405$). It can be seen in Fig. 3 that using this reference masks the early accenting effect that is present in the 0–100 msec latency range, though subsequently such an effect can be seen in the latency range 200–400 msec at temporal and parietal sites. These latter differences are not, however, significant.

4. Discussion

The main aim of this study was to test the hypothesis that extensive exposure to rhythmic structures in music will lead to a tendency to perceive isochronous equitonal sequences as having, by default, a

binary accented structure that is probably synchronized to the first beat in the sequence. Such an observation would be consistent with the finding that we are, in general (and musicians in particular), very sensitive to discrepancies of timing in music and that these effects can be detected in brain responses associated with pre-attentive processing (Vuust et al., 2005). Our previous research (Brochard et al., 2003) provided indirect evidence of accenting effects in the form of modulations of P300 amplitude but not earlier ERP modulations. The present finding of an ERP modulation in the latency range 0–100 msec that may be associated with subjective accenting is consistent with the dynamic attending theory of rhythm processing (Jones, 1976; Jones and Boltz, 1989; Drake et al., 2000). However, it is also possible that dynamic modulations of perceptual processing occur, as a result of predictive processes that are an inherent part of pre-attentive processing (Friston, 2002). The present study does not distinguish between these two theoretical accounts of the basis of subjective accenting effects. Putative subjectively accented standard tones were relatively more negative than subjectively unaccented standard tones in the 200–500 msec latency range. In contrast the ERP responses to deviant tones no longer differed at the latency of the N2. The active detection of the deviant tones may lead to the reorienting of attention (Astafiev et al., 2006) and the resetting of subjective accenting. However, this remains to be determined. A concern regarding the early negative deflections that were observed in this study is the possibility that they were a nonspecific effect associated with anticipation of infrequent deviant stimuli. Negative shifts have been hypothesized to reflect anticipatory activity (Kotchoubey, 2006) that may well overlap the early part of the ERP to the following stimulus (Starret et al., 1995; Kotchoubey, 2006). For example, a late slow wave observed by Starr et al. (1997) in ERPs to standard tones showed negative polarity and a frontal distribution

before the occurrence of an infrequent deviant tone, and was assumed to be related to listeners' attention to and expectation of the deviant tone. This explanation seems unlikely in the present study as the early negativity observed at the mastoid site was observed to both deviant and standard tones. The sustained left-lateralised negative ERP deflection, evoked only by accented standard tones, does not fit with Kotchoubey's anticipation model either. It seems possible that the negative deflections observed in the present study might be properly classified as "processing negativities" that reflect the extent of attention allocation (Alho et al., 1987; Higashima et al., 2004). Whether these processing negativities are generated by the same mechanism that generates the Nd remains to be determined. The Nd consists of an early and a late component, originating in auditory and frontal areas, respectively (Giard et al., 2000). It is possible that the subjective accenting effect observed at the mastoid site is an example of the early Nd deflection that originates in the auditory cortex. However, the onset of the effect was earlier than is typically described in the case of the Nd. Finally, lateralization of the observed accenting effects to the left hemisphere is consistent with previous observations that attention to rhythm tends to activate regions of the left hemisphere more than the right hemisphere (Platel et al., 1997; Potter et al., 2000; Vuust et al. 2005).

The results of this study strongly support our previous findings (Brochard et al., 2003; Abecasis, et al. 2005) by providing replicable physiological evidence of subjective accenting. Early differences between "accented" and "unaccented" positions in the tone sequence, whatever the intensity of the tone, could reflect early segmentation of the tone sequence into groups of two events (Frasse, 1982; Handel, 1989). As stated earlier, the findings are also consistent with an attention based account of rhythm perception,

such as Jones' dynamic attending theory (Jones and Boltz, 1989; see Kotchoubey, 2006). Attention can alter neural activity in the auditory system at the level of the cochlea (e.g. Maison et al., 2001), brainstem and thalamic nuclei (e.g. Hirschhorn and Michie, 1990), as well as early positive and negative obligatory cortical components (see Woldorff et al., 1993) and the level at which the modulations observed in this study first occur remains to be determined. However, the present findings do not rule out the possibility that these effects result from the predictive nature of temporal representation structures (Friston, 2002) and this would not necessarily involve dynamic modulations of perceptual systems by attention. The findings of this study provide further evidence that our perceptions of rhythmic structure in auditory events are significantly influenced by our prior experience. Further lines of research might include a more detailed study of the stage at which auditory processing is affected by attention, a more detailed consideration of the effect of either implicit or explicit musical expertise, or indeed the effects of recent rhythmic pattern experience on the perception of subsequent isochronous stimulus sequences.

references

Abecasis D, Brochard R, Granot R, and Drake C. Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception*, 22: 549–562, 2005.

Alho K, Töttölä K, Reinikainen K, Sam M, and Näätänen R. Brain mechanism of selective listening reflected by event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 68: 458–470, 1987.

Astafiev SV, Shulman GL, and Corbetta M. Visuospatial reorienting signals in the human temporo-parietal junction are independent of

response selection. *European Journal of Neuroscience*, 23: 591–596, 2006.

Besson M and Faïta F. An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21: 1278–1296, 1995.

Bolton TL. Rhythm. *American Journal of Psychology*, 6: 145–238, 1894.

Brochard R, Abecasis D, Potter D, Ragot R, and Drake C. The ticktock of our internal clock: evidence of subjective accents in isochronous sequences. *Psychological Science*, 14: 362–366, 2003.

Donchin E and Coles MGH. Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11: 357–374, 1988.

Drake C, Jones MR, and Baruch C. The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77: 251–288, 2000.

Drake C. Reproduction of musical rhythms by children, adult musicians, and adult non-musicians. *Perception and Psychophysics*, 53: 25–33, 1993.

Fraisse P. Rhythm and tempo. In Deutsch D (Ed), *The psychology of music*. New York: Academic Press, 1982.

Friston K. Beyond phrenology: what can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience*, 25: 221–250, 2002.

Giard M-H, Fort A, Mouchetant-Rostaing Y, and Pernier J. Neurophysiological mechanisms of auditory selective attention in humans. *Frontiers in Bioscience*, 5: 84–94, 2000.

Granot R and Donchin E. Do Re MiFa Sol La Ti – constraints, congruity, and musical training: an event-related brain potentials study of musical expectancies. *Music Perception*, 19:487–528, 2002.

Handel S. *Listening: an introduction to the perception of auditory events*. Cambridge, MA: MIT Press, 1989.

Hannon EE and Trehub SE. Metrical categories in infancy and adulthood. *Psychological Science*, 16: 48–55, 2005.

Higashima M, Nagasawa T, Kawasaki Y, Oka T, Sakai N, Tsukada T, et al. Event-related potentials elicited by nontarget tones in an auditory oddball paradigm in schizophrenia. *International Journal of Psychophysiology*, 51: 189–200, 2004.

Hillyard SA, Hink RF, Schwent VL, and Picton TW. Electrical signs of selective attention in the human brain. *Science*, 182: 177–180, 1973. Hirschhorn TN and Michie PT. Brainstem auditory evoked potentials (BAEPS) and selective attention revisited. *Psychophysiology*, 27: 495–512, 1990.

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

Jones MR. Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review*, 83:323–355, 1976.

Jones MR and Boltz M. Dynamic attending and responses to time. *Psychological Review*, 96: 459–491, 1989.

Kotchoubey B. Event-related potentials, cognition, and behavior: a biological approach. *Neuroscience and Biobehavioral Reviews*, 30: 42–65, 2006.

Maison S, Micheyl C, and Collet L. Influence of focused auditory attention on cochlear activity in humans. *Psychophysiology*, 38:35–40, 2001.

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

Näätänen R, Ahissari E, Gaillard, and Mantysalo S. Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42: 313–329, 1978.

Näätänen R. Processing negativity: an evoked-potential reflection of selective attention. *Psychological Bulletin*, 92: 605–640, 1982.

Parncutt R. A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11: 409–464, 1994.

Platel H, Price C, Baron JC, Wise R, Lambert J, Frackowiak RSJ, et al. The structural components of musical perception – a functional anatomical study. *Brain*, 120: 229–243, 1997.

Polich J and Kok A. Cognitive and biological determinants of P300: an integrative review. *Biological Psychology*, 41: 103–146, 1995.

Potter DD, Sharpe H, Basbinar D, and Jory S. Insights into the functional organization of music processing revealed using continuous across-subject event-related potential averaging. In Woods C, Luck GB, Brochard R, O’neill SA, and Sloboda JA (Eds), *Proceedings of the sixth international conference on music perception and cognition*. Keele, Staffordshire, UK: Department of Psychology, 2000.

Povel DJ and Okkerman H. Accents in equitone sequences. *Perception and Psychophysics*, 30: 565–572, 1981.

Starr A, Aguinaldo T, Roe M, and Michalewski HJ. Sequential changes of auditory processing during target detection: motor responding versus mental counting. *Electroencephalography and Clinical Neurophysiology*, 105: 201–212, 1997.

Starr A, Sandroni P, and Michalewski HJ. Readiness to respond in a target detection task: pre- and post-stimulus event-related potentials in normal subjects. *Electroencephalography and Clinical Neurophysiology*, 96: 76–92, 1995.

Thomassen MT. Melodic accent: experiments and a tentative model. *Journal of the Acoustical Society of America*, 71: 1596–1605, 1982.

Vuust P, Pallesen KJ, Bailey C, Van Zuijen TL, Gjedde A, Roepstorff A, et al. To musicians, the message is in the meter: pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, 24: 560–564, 2005.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, et al. Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, 90: 8722–8726, 1993.

Woodrow H. A quantitative study of rhythm. *Archives of Psychology*, 14: 1–66, 1909.