Research Report

THE "TICKTOCK" OF OUR INTERNAL CLOCK: Direct Brain Evidence of Subjective Accents in Isochronous Sequences

Renaud Brochard,[1,2] Donna Abecasis,[2,3] Doug Potter,[4] Richard Ragot,[5] and Carolyn Drake[3]

[1]*Laboratoire d'Etude des Apprentissages et du Développement, CNRS UMR 5022, Université de Bourgogne, Dijon, France;*
[2]*Unit for the Study of Musical Skill and Development, Psychology Department, Keele University, Keele, England, United Kingdom;*
[3]*Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Institut de Psychologie, Université René Descartes (Paris 5), Boulogne-Billancourt, France;* [4]*Department of Psychology, Dundee University, Dundee, Scotland, United Kingdom; and*
[5]*LENA, CNRS UPR 640, Hôpital de la Pitié-Salpétrière, Paris, France*

Abstract—*The phenomenon commonly known as subjective accenting refers to the fact that identical sound events within purely isochronous sequences are perceived as unequal. Although subjective accenting has been extensively explored using behavioral methods, no physiological evidence has ever been provided for it. In the present study, we tested the notion that these perceived irregularities are related to the dynamic deployment of attention. We disrupted listeners' expectancies in different positions of auditory equitone sequences and measured their responses through brain event-related potentials (ERPs). Significant differences in a late parietal (P3-like) ERP component were found between the responses elicited on odd-numbered versus even-numbered positions, suggesting that a default binary metric structure was perceived. Our findings indicate that this phenomenon has a rather cognitive, attention-dependent origin, partly affected by musical expertise.*

When one is listening to a succession of identical tones occurring at a regular pace (an equitone isochronous sequence), some tones are heard as more salient (louder, longer, or both) than others. This phenomenon, used to explain why people perceive a "ticktock" rather than a "ticktick" when they hear even a quartz watch, was explored as early as the 19th century (Bolton, 1894). It was originally referred to as *subjective accenting* because no physical characteristic of the sounds accounted for the differences perceived. Several researchers have attempted since to examine this issue in a great variety of ways (Parncutt, 1994; Povel & Okkerman, 1981; Temperley, 1963). For instance, Povel and Okkerman (1981) assessed the strength of perceived accents in equitone sequences of two alternating time intervals, showing that in order to make the tones perceptually identical the physical intensity of one tone of each pair had to be increased. Using isochronous sequences, Parncutt (1994) asked listeners to rate whether a specific target tone within a sequence occurred on or off the beat and obtained a variety of responses even when the sounds were physically identical. Nevertheless, this perceptual phenomenon has been explored mainly through listeners' own conscious responses and has never been investigated using physiological measurement.

The fact that some events in equitone isochronous sequences appear to be more accented than others may be related to the way listeners organize sounds perceptually. Models of auditory perceptual organization have explained it in various ways, as, for example, in terms of an "internal clock" that tends to synchronize with the auditory sequences (e.g., Essens, 1995; Povel & Essens, 1985). According to time-based *dynamic attending* theory, postulated by Jones (e.g., Jones & Boltz, 1989; Jones, Boltz, & Kidd, 1982; Large & Jones, 1999; Large & Palmer, 2002), when presented with an auditory sequence listeners extract regularities from the first events within the sequence and thus anticipate the properties of future events, giving rise to rhythmic expectancies. These expectancies direct attention toward particular events, which are thus perceived as relatively accented.

This dynamic process would lead to the perception of a metrical structure consisting of alternating "strong" (accented) and "weak" (unaccented) beats. Based on the assumption that such accents correspond to peaks of attention, several studies in experimental psychology (e.g., Jones et al., 1982; Monahan, Kendall, & Carterette, 1987; Smith & Cuddy, 1989) have introduced violations of expectancies in auditory sequences and shown that they are, in fact, better processed if they occur on accented events than on unaccented ones. For instance, in the study by Jones et al. (1982), listeners had to judge whether the second of a pair of nine-tone melodic sequences was identical to the first sequence or was slightly changed. The results showed that correct detection of the melodic alterations was greater in accented than unaccented positions. Most of these studies, however, required overt responses from the listeners and have been criticized for relying too heavily on subjects' memory (Mondor & Terrio, 1998).

In the study presented here, listeners' brain responses were measured directly in real time through auditory event-related potentials (ERPs). Several ERP components have already been reported as being sensitive to violations of expectancies in auditory stimuli (e.g., Besson, 1997; Näätänen, 1990). For instance, the preattentive mismatch negativity component (MMN), peaking at about 150 to 250 ms after stimulus onset, is elicited by infrequent deviations in some physical feature of a repeated sound event or pattern (Näätänen, 1990; Schröger, Tervaniemi, & Näätänen, 1995; Schröger & Winkler, 1995). Later positive components have also been observed in response to violations of expectancies in musical contexts (Besson & Faïta, 1995; Hantz, Crummer, Wayman, Walton, & Frisina, 1992; Janata, 1995; Trainor, Desjardins, & Rockel, 1999). Unlike the MMN, these late components are dependent on listeners' attention and are generally elicited in tasks involving detection of rare events (e.g., Schröger, 1996). Their latencies vary from around 300 to 600 ms after stimulus onset, and they are commonly referred to as the "P300 family" of components. The easier the detection of deviance, the larger the amplitudes and the shorter the latencies of these ERP components (e.g., Sussman, Ritter, & Vaughan, 1998). In the case of the P300, these characteristics have been related to the amount of attention deployed

Address correspondence to Renaud Brochard, Laboratoire d'Etude des Apprentissages et du Développement, CNRS UMR 5022, Université de Bourgogne, Pôle AAFE, BP 26513, 21065 Dijon, France; e-mail: renaud.brochard@u-bourgogne.fr.

 Copyright © 2003 American Psychological Society

and the difficulty of categorizing the event (e.g., Donchin & Coles, 1988; Janata, 1995).

In the experiment we report here, violations of expectancies were introduced in different positions in isochronous sequences so we could assess the presence of subjective accents. These violations consisted of infrequent decreases in loudness (deviant tones) that had to be detected by the listeners. We hypothesized that the ERP response to these violations would differ according to their location within the sequence. As Western listeners have been reported to more frequently produce and perceive binary metrical patterns than other patterns (Drake, 1993; Fraisse, 1982), and as the first tone of an auditory sequence has been shown to be perceptually salient (Thomassen, 1982, 1983), we expected that odd-numbered events would correspond to accented positions ("strong" beats) and even-numbered events to unaccented ones ("weak" beats). We therefore hypothesized that the ERP response to the infrequent deviant tones would be larger in odd than in even positions, because more attention would be allocated to these presumably strong beats, despite the fact that all beats had the same physical properties. We tested musically trained and untrained listeners to explore a possible effect of musical expertise, because musicians have been described as having stronger metrical expectancies than nonmusicians. For instance, Drake, Penel, and Bigand (2000) showed that musicians synchronized with musical excerpts more accurately than nonmusicians, extracting metrically important events over longer time spans.

METHOD

Participants

Twelve paid university students (5 males, 7 females) participated in the experiment. Half of them had received more than 5 years of formal musical training; the other half had not received any musical training at all.

Stimuli

The stimuli (see Fig. 1) consisted of 96 isochronous sequences of 13 to 16 identical tones (equitone sequences) containing occasional deviant tones (10.5% deviant tones and 89.5% standard tones per sequence, on average). The sequences had a variable number of events so as not to favor any particular (binary or ternary) metrical pattern. The interonset interval (IOI) within each sequence was 600 ms. This presentation rate has been described by several authors as a preferred "spontaneous tempo" and an optimal IOI in temporal discrimination tasks (Drake, Jones, & Baruch, 2000; Fraisse, 1982), although subjective accenting has been observed to be quite independent of the sequence presentation rate (Parncutt, 1994). Standard tones had a duration of 50 ms, a frequency of 440 Hz, and a loudness level of 70 dB SPL. The occasional deviant tones were 4 dB softer than the standards. This small decrease was equivalent to the size of a subjective accent measured within pairs of identical tones (Povel & Okkerman, 1981) and was intended to make discrimination possible, though difficult enough for the subtle subjective phenomenon to be evidenced. In each sequence, the deviant tone was introduced in one of four possible positions, corresponding to hypothetically strong (Positions 9 and 11) or weak (Positions 8 and 10) beats. A second deviant tone (which was not included in the analyses) was introduced in a later position in half of the sequences, to make the number and position of the changes unpredictable. We asked participants to count the number of loudness changes they heard within each sequence to ensure that they were attending to the stimuli.

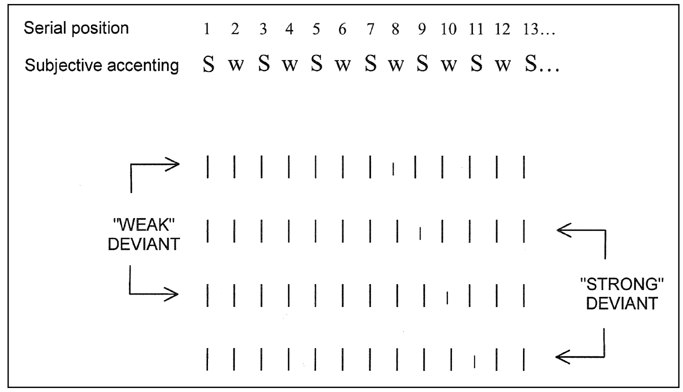

Fig. 1. The four types of isochronous tone sequences used in the experiment. Half the sequences contained a deviant tone (4 dB softer) in a hypothetically strong (S) position (9 or 11); the other half contained the deviant tone in a weak (w) position (8 or 10).

Electroencephalogram (EEG) Measurement

Subjects were seated in a soundproof room, and stimuli were delivered via headphones. EEG was recorded continuously from all 19 standard locations of the 10-20 system of electrode placement plus 7 additional sites (referred to linked mastoid electrodes). Vertical and horizontal electro-oculograms (EOGs) were recorded from bipolar electrode pairs placed above and below the right eye (vertical) and on the outer canthus of each eye (horizontal). EEG and EOG were recorded with bandwidths of 0.03 to 35 Hz and digitized on-line at a rate of 2 ms/point. Averages from 100 ms before stimulus onset to 600 ms after stimulus onset were produced off-line. An EOG eye-blink artifact-correction procedure was used on the data.

RESULTS

In the comparison of the ERPs corresponding to standard and deviant tones, a clear response to the disruption of listeners' expectancies could be observed through the components described earlier, namely, a negative frontal component peaking at about 150 to 250 ms after stimulus onset and a later positive parietal component at about 400 to 600 ms after stimulus onset.

As stated earlier, deviant sounds could occur in four different positions within the sequences (8, 9, 10, and 11). In order to test our hypothesis that events were processed differentially following a binary pattern, we averaged together ERPs corresponding to deviants that occurred in odd-numbered positions (i.e., either Position 9 or Position 11, both hypothetically strong), and then compared this average with the average of ERPs to deviants occurring in even-numbered positions (i.e., either Position 8 or Position 10, both hypothetically weak; see Fig. 2, top panel). In a control analysis, weak and strong positions were mixed together: ERPs for Positions 8 and 9 were averaged, and this average was compared with the average for Positions 10 and 11 (see Fig. 2, bottom panel). We expected the first analysis to yield larger differences than the second.

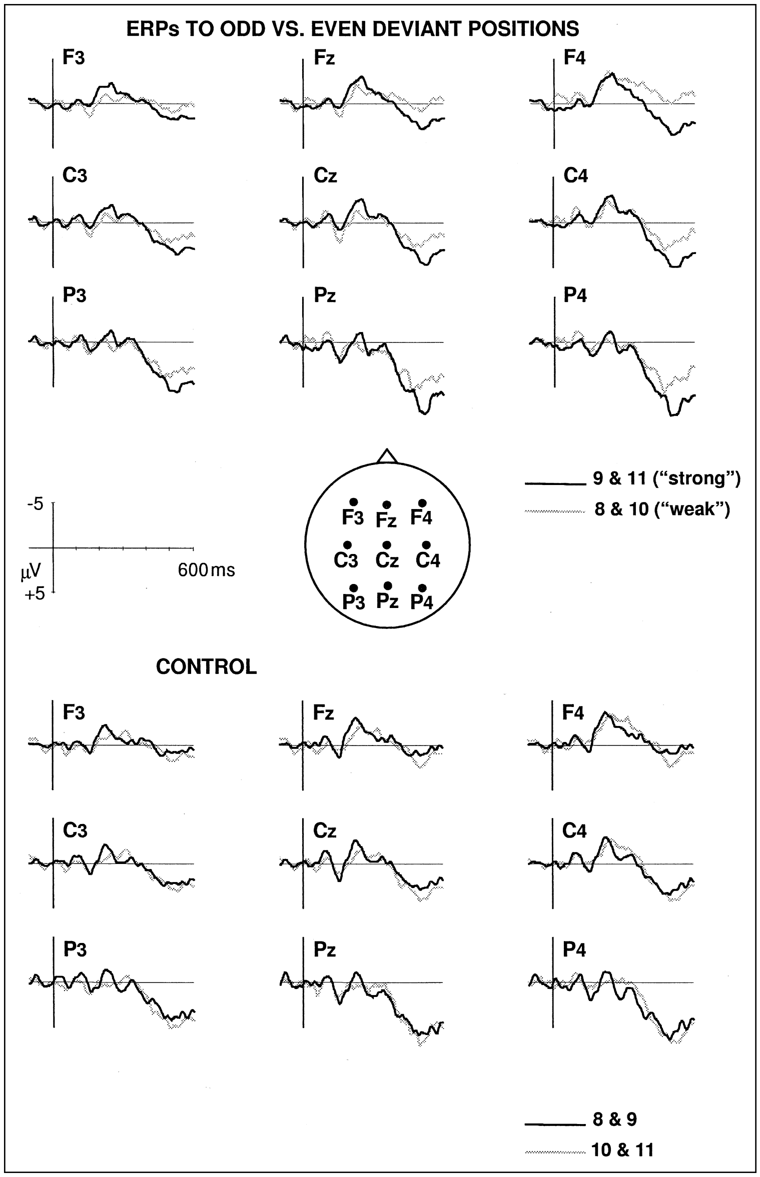

Fig. 2. Event-related potential (ERP) responses to deviant tones in isochronous sequences. The top panel shows grand-average ERP responses to the deviant tones in odd-numbered positions (9 and 11, in black) versus even-numbered positions (8 and 10, in gray). The bottom panel (control analysis) shows grand-average ERP responses to these same deviant tones in consecutive positions, so that responses to strong and weak positions are averaged (8 and 9, shown in black, vs. 10 and 11, shown in gray). Results are shown for nine electrodes, whose placement is indicated in the illustration in the middle of the figure.

The results of the first analysis indeed showed significant differences between the responses elicited by deviant odd-numbered and even-numbered positions. Odd-numbered, hypothetically strong, positions showed significantly larger amplitudes in the late positive parietal component of the ERPs. At 500 to 600 ms after stimulus onset, for example, significant differences were found at electrode Pz, $t(11) = 3.36$, $p < .01$, and electrode P4, $t(11) = 3.91$, $p < .01$. In the second (control) analysis, performed on averaged consecutive positions, no such significant differences were found, suggesting that the differences found in the first analysis were due to a metrical perception and were not simply related to the time lag between the events or any effect of stimulus repetition.

The effect of long-term musical training was tested through repeated measures analyses of variance with musical expertise as a between-subjects factor. A significant interaction was observed at electrode P4 400 to 450 ms after stimulus onset, $F(1, 10) = 5.35$, $p <$

.05. This result suggests that the ERP differences between deviant strong and weak positions arose earlier in subjects who had musical training than in those who did not, perhaps because musicians have more efficient processing or stronger temporal expectancies (see Fig. 3).

DISCUSSION

The main purpose of the present study was to provide physiological evidence for the phenomenon of subjective accenting. The results obtained indeed demonstrate that equal sound events are not processed equally in isochronous sequences. Significant differences were found between the ERP responses to violations of expectancies in different positions.

Furthermore, the ERP responses showed significantly larger amplitudes when elicited by deviant sounds in odd-numbered (i.e., subjectively strong) positions as compared with even-numbered (i.e., subjectively weak) positions. These differences were not solely due to the time lag between the events, because they were not found in the control analysis performed on averaged consecutive positions. Therefore, this result supports Jones's theory (Jones & Boltz, 1989; Large & Jones, 1999) in that attention seems to be deployed periodically. As predicted, this dynamic oscillation of attention seems to follow, at least predominantly, a default binary pattern. Exploratory data (not presented here) suggest that when a metrical structure is induced by a physical cue at the beginning of each sequence, binary and ternary induction yield different patterns of ERP components. This issue is currently under investigation, but the results obtained in the present study seem to resemble the binary-induced pattern much more than the ternary one.

It is important to note that the significant differences between the responses to odd-numbered and even-numbered deviant positions were observed in the late positive parietal component of the ERPs, possibly corresponding to the so-called P3b component described in relation to the attentive processing of deviance (Janata, 1995). The fact that the differences were found in such a late component suggests that they might not be due to preattentive perceptual mechanisms, but perhaps instead reflect more cognitive processing. Thus, subjective accenting may not correspond to an automatic, early auditory process but to a higher-level one. A potential influence of early processes cannot be ruled out, as increased statistical power might have yielded significant differences in earlier components. However, top-down modulation effects have recently been shown to go back to very early stages in processing, though temporally such effects may be delayed in the ERP signature (Olson, Chun, & Allison, 2001). A study involving a more classical MMN paradigm, with an analysis focusing on a 100- to 300-ms time window, might further clarify the stage of processing at which subjective accenting originates.

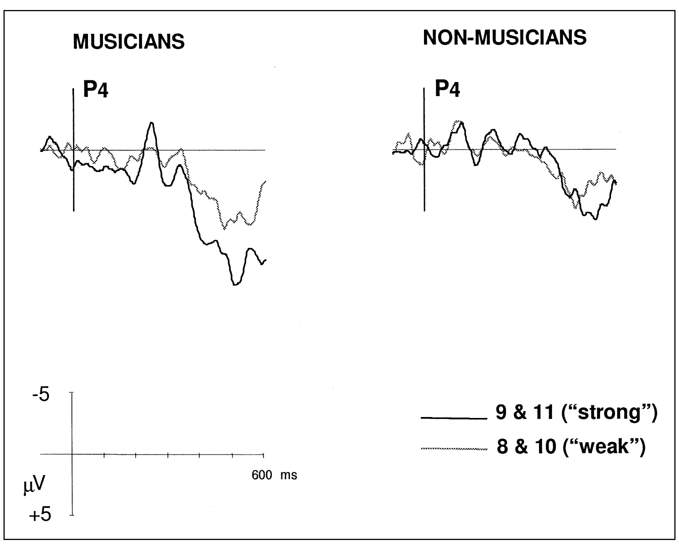

Fig. 3. Event-related potential (ERP) differences between odd-numbered (black) and even-numbered (gray) deviant positions in musicians versus nonmusicians. The results shown are from the right parietal site P4.

The late significant differences found in our exploratory analysis of musician and nonmusician listeners seem to be consistent with a rather cognitive interpretation of subjective accenting. Previous reports have already shown that musical training influences mainly high- rather than low-level auditory organization (e.g., Drake, Penel, & Bigand, 2000). These studies have suggested that musically trained listeners organize sound events over longer time spans than non-musically trained listeners do, and the differences found here seem to reflect this more efficient temporal processing. Finally, our results are also consistent with studies that have used subjective accenting as an index of temporal integration (Szelag, Kowalska, Rymarczyk, & Pöppel, 1998; Szelag, von Steinbüchel, Reiser, Gilles de Langen, & Pöppel, 1996) and demonstrated an effect of cognitive development. It would be interesting to extend this developmental approach by using the methodology presented here, which allows for the exploration of processes less controlled than the active perceptual integration of tones, as well as by examining a possible effect of acculturation on the tendency to perceive binary-like patterns. Studying effects of acculturation, in addition to those of formal training, might contribute to the understanding of how and at what level rhythmic expectancies are generated and how they relate to attentional processes. Overall, the subjective-accenting phenomenon seems to lend itself to electrophysiological exploration, and its implications certainly deserve further investigation.

Acknowledgments—We thank S. Nte and C. Woods for technical support.

REFERENCES

Besson, M. (1997). Electrophysiological studies of music processing. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 217–250). East Sussex, England: Psychology Press.

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

Bolton, T.L. (1894). Rhythm. *American Journal of Psychology*, *6*, 145–238.

Donchin, E., & Coles, M.G.H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374.

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

Drake, C., Jones, M.R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*, 251–288.

Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception*, *18*, 1–25.

Essens, P. (1995). Structuring temporal sequences: Comparison of models and factors of complexity. *Perception & Psychophysics*, *57*, 519–532.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.

Hantz, E.C., Crummer, G.C., Wayman, J.W., Walton, J.P., & Frisina, R.D. (1992). Effects

of musical training and absolute pitch on the neural processing of melodic intervals: A P3 event-related potential study. *Music Perception*, *10*, 25–42.
Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*(2), 153–164.
Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.
Jones, M.R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics*, *32*, 211–218.
Large, E.W., & Jones, M.R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.
Large, E.W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.
Monahan, C.B., Kendall, R.A., & Carterette, E.C. (1987). The effect of melodic and temporal contour on recognition memory for pitch change. *Perception & Psychophysics*, *41*, 576–600.
Mondor, T.A., & Terrio, N.A. (1998). Mechanisms of perceptual organisation and auditory selective attention: The role of pattern structure. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1628–1641.
Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*, 201–288.
Olson, I.R., Chun, M.M., & Allison, T. (2001). Contextual guidance of attention: Human intracranial event-related potential evidence for feedback modulation in anatomically early, temporally late stages of visual processing. *Brain*, *124*, 1417–1425.
Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.
Povel, D.J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2*, 411–440.
Povel, D.J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception & Psychophysics*, *30*, 565–572.
Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology, Evoked Potentials Section*, *100*, 517–526.
Schröger, E., Tervaniemi, M., & Näätänen, R. (1995). Time course of loudness in tone patterns is automatically represented by the human brain. *Neuroscience Letters*, *202*, 117–120.
Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience Letters*, *193*, 185–188.
Smith, K.C., & Cuddy, L.L. (1989). Effects of metric and harmonic rhythm on the detection of pitch alterations in melodic sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 457–471.
Sussman, E., Ritter, W., & Vaughan, H.G., Jr. (1998). Attention affects the organisation of auditory input associated with the MMN system. *Brain Research*, *789*, 130–138.
Szelag, E., Kowalska, J., Rymarczyk, K., & Pöppel, E. (1998). Temporal integration in a subjective accentuation task as a function of child cognitive development. *Neuroscience Letters*, *257*, 69–72.
Szelag, E., von Steinbüchel, N., Reiser, M., Gilles de Langen, E., & Pöppel, E. (1996). Temporal constraints in processing of nonverbal rhythmic patterns. *Acta Neurobiologiæ Experimentalis*, *56*, 215–225.
Temperley, N.M. (1963). Personal tempo and subjective accentuation. *Journal of General Psychology*, *68*, 267–287.
Thomassen, M.T. (1982). Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America*, *71*, 1596–1605.
Thomassen, M.T. (1983). Erratum. *Journal of the Acoustical Society of America*, *73*, 373.
Trainor, L.J., Desjardins, R.N., & Rockel, C. (1999). A comparison of contour and interval processing in musicians using event-related potentials. *Australian Journal of Psychology*, *51*, 147–153.

(RECEIVED 5/15/02; REVISION ACCEPTED 9/16/02)