

# Effects of Context on Electrophysiological Response to Musical Accents

Caroline Palmer,<sup>a</sup> Lisa R. Jewett,<sup>a</sup> and Karsten Steinhauer<sup>b</sup>

<sup>a</sup>Department of Psychology and <sup>b</sup>School of Communication Sciences and Disorders,  
McGill University, Montreal, Quebec, Canada

Listeners' aesthetic and emotional responses to music typically occur in the context of long musical passages that contain structures defined in terms of the events that precede them. We describe an electrophysiological study of listeners' brain responses to musical accents that coincided in longer musical sequences. Musically trained listeners performed a timbre-change detection task in which a single-tone timbre change was positioned within 4-bar melodies composed of 350-ms tones to coincide or not with melodic contour accents and temporal accents (induced with temporal gaps). Event-related potential responses to (task-relevant) attended timbre changes elicited an early negativity (MMN/N2b) around 200 ms and a late positive component around 350 ms (P300), reflecting updating of the timbre change in working memory. The amplitudes of both components changed systematically across the sequence, consistent with expectancy-based context effects. Furthermore, melodic contour changes modulated the MMN/N2b response (but not the P300) to timbre changes in later sequence positions. In contrast, task-irrelevant temporal gaps elicited an MMN that was not modulated by position within the context; absence of a P300 indicated that temporal-gap accents were not updated in working memory. Listeners' neural responses to musical structure changed systematically as sequential predictability and listeners' expectations changed across the melodic context.

**Key words:** musical accents; context effects; perceptual independence; ERP; mismatch negativity; P300; melodic contour; rhythmic grouping; timbre

## Introduction

Music comprises a rich variety of structural relationships that make listening enjoyable. The perceptual salience of different musical structures, such as melodic contour, rhythm, meter, and harmony, is complex because of the fact that the structures co-occur and change quickly, often at different rates. How listeners integrate the different musical structures in real time is an important perceptual question. The current study builds on behavioral and electrophysiological experiments that have probed

the perceptual salience of musical accents.<sup>1-3</sup> The goal is to investigate neural responses to timbral, temporal, and melodic accents as they coexist in a longer melodic context. We utilize a real-time timbre-change detection paradigm that measures the perceptual salience of musical accents as they unfold in a melodic context. Similar to click-localization tasks used first in language, the timbre-change detection task probes listeners' sensitivity to timbre changes at different melodic locations that contain combinations of musical accents.

## Behavioral Response to Musical Accents

A musical event is perceived as accented when it is marked for consciousness in some way relative to other surrounding events.<sup>4,5</sup> Accents are important because they direct

Addresses for correspondence: Dr. Caroline Palmer, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal QC H3A 1B1, Canada. Voice: 514-398-6128; fax: 514-398-4896. caroline.palmer@mcgill.ca [or] Karsten Steinhauer, School of Communication Sciences and Disorders, Centre for Research on Language, Mind and Brain, 1266 Pine Avenue, Montreal, QC H3G 1A8, Canada. karsten.steinhauer@mcgill.ca

attention to particular points in a melody and segment a continuous acoustic signal into smaller units, such as musical bars and phrases. Some musical accents arise from changes that unfold across successive events. *Contour accents* arise from directional changes in the rise and fall of successive pitches. Some studies suggest that the accent is perceived on the extreme pitch of a contour change (the point of change in pitch direction), whereas others posit that the accent occurs on the first event after the change in direction.<sup>4,5</sup> For simplicity we refer here to the melodic contour accent as occurring on the extreme pitch event. *Rhythmic grouping accents* arise when shortened or lengthened durations occur in a sequence of successive (equal duration) tones; changes in lengthening of tones are usually interpreted as accentuated.<sup>5</sup> *Timbral accents* arise when listeners perceive a change in timbre or tone quality arising from spectral features that distinguish tones otherwise similar in pitch, loudness, and duration. Evidence for the perceptual interaction of contour, rhythmic grouping, and timbre accents is mixed, and often depends on the particular task. For example, behavioral studies have addressed the perceptual independence of timbre and pitch, with mixed results.<sup>6,7</sup> The number of sequence events over which accents are defined and perceived may contribute to the mixed results.

Click-localization methods were first used to examine perceptual grouping and segmentation in language. When a click was superimposed on a spoken sentence, subjects often reported hearing the click one or two syllables away from the location at which it was actually positioned, called perceptual “migration.”<sup>8</sup> Fodor and Bever<sup>9</sup> found that superimposed clicks tended to migrate toward major syntactic and prosodic boundaries in language. Click migration was interpreted as a perceptual tendency to maintain the integrity of discrete perceptual units. Variants of the click-localization task have been applied to music perception. Gregory<sup>10</sup> found a significant tendency for clicks to be perceived as migrating

toward phrase boundaries. When clicks were placed near major and minor musical phrase boundaries, Stoffer<sup>11</sup> found a stronger tendency for clicks to migrate to major than to minor musical boundaries. However, Gregory and Stoffer used offline tasks; listeners responded in written form on a musical score, making it difficult to separate potential memory effects from those of perceptual salience. Furthermore, all click-localization tasks employ a task-relevant stimulus (the click) that is external to the stimulus structure under investigation (sentence or melody), and the findings may be influenced by divided attention.

Capota<sup>1</sup> assessed listeners’ ability to detect timbre changes that coincided with different musical accents in melodic contexts. A single-tone timbre-detection task kept participants’ attention focused on a melody. One-tone timbre changes were inserted at varying locations to coincide with different accents unfolding within the melody: temporal accents (induced with temporal gaps), and melodic accents (changes in pitch contour). Participants responded faster to timbre changes that occurred later in the melody. In addition, participants were slower at detecting timbre changes when they occurred near a melodic accent. Temporal accents resulted in faster responses to timbre changes at early melodic locations, but only when they preceded timbre changes. Overall, the findings suggested that early in the melody, listeners did not have sufficient information to form expectations and this resulted in greater processing time. Capota<sup>1</sup> further suggested that participants’ slower reaction times to timbre changes near melodic contour changes indicated that these accents in combination were less salient than when presented alone. In sum, the timbre-detection task demonstrated that sensitivity to musical accents was modulated by the sequential context.

### Neural Responses to Musical Accents

Event-related potentials (ERPs) have been used to explore the neural processing

implicated in the perception of musical accents. Two primary ERP components have been found for musical accents: the mismatch negativity (MMN) and P300 components.<sup>2,3</sup> The MMN is an early negative neural response that typically occurs 100–250 ms after the auditory presentation of a deviant stimulus and displays a fronto-central scalp distribution.<sup>12</sup> The MMN is classically obtained in an “oddball” paradigm in which a large number of identical acoustic stimuli (standards) are presented intermixed with occasional “deviants.” The MMN response is elicited irrespective of the participants’ attention during stimulus presentation, and does not require an overt behavioral response.<sup>13</sup>

In contrast, P300 components comprise positive-going ERP deflections that typically occur only with attention to the stimulus. The P3a is a fronto-central component around 250–400 ms following a stimulus recognized as rare, unusual, and/or relevant. The P3a is often preceded by N200 (or N2b), a negativity reflecting stimulus evaluation (resulting in an N2–P3 complex). Unlike the P3a, the late parietal P3b (300–600 ms) reflects conscious processing after identification of a task-relevant stimulus and is thus related to “context updating”—the revision of a representation of the environment in working memory.<sup>14</sup> The P3b/P300 amplitude for deviants increases with task relevance, low subjective target probability, and the participant’s level of certainty in identifying target stimuli.<sup>15</sup>

MMN responses have been noted to various types of musical accents, including durational contrasts, melodic contour, and timbre changes. Using a nonattended ERP paradigm, Tervaniemi *et al.*<sup>16</sup> tested deviant stimuli that differed in terms of several features (e.g., intensity, duration, contour) from a sequence of harmonically rich standard tones. MMN components were found for all deviants and showed longer latencies for duration manipulations (7-ms gaps of silence). Grimm and Schröger<sup>17</sup> conducted an experiment in which the length of 1000-ms sine waves was varied. The 600-ms

deviants elicited MMNs in both nonattended and attended conditions, whereas only the attended conditions yielded subsequent N2b and P300 waves. To explore MMN responses to contour changes, Tervaniemi *et al.*<sup>16</sup> employed a melody of five 50-ms sinusoidal tones in which a deviant melodic contour was created by replacing the penultimate tone with the first tone. Once again, an MMN was elicited in response to the deviant tones. Trainor *et al.*<sup>3</sup> also tested melodic contour, using an ascending four-tone sequence as a standard stimulus. Infrequent deviant contours with a descending interval on the final tone elicited an MMN to the final tone.

The MMN has also been found in response to timbre change. Timbre can be defined as the tone quality or the spectral features that distinguish instruments or voices; timbre distinguishes two tones from each other that are equal in pitch, subjective duration, and loudness.<sup>18</sup> Goydke *et al.*<sup>2</sup> manipulated timbral qualities in a comparison of attended and nonattended conditions. A series of 600-ms standard tones was presented in one timbre (piano), and participants actively detected a one-tone timbre deviant (flute) by pressing a button. Two neural responses were generated: an MMN in response to the deviant timbre changes, and additional P300 components in the attended condition only (see also Crummer *et al.*<sup>19</sup> for similar results). Thus, MMN responses have been noted for a variety of accents presented in isolation, including durational accents, melodic contour, and timbre change.

### The Current Study

The current experiment examined listeners’ neural responses to musical accents in longer melodic contexts containing multiple co-occurring accents. Following Capota,<sup>1</sup> we placed timbral, temporal, and melodic accents in different combinations at early, middle, and late contextual locations in melodies. Participants were instructed to indicate timbre changes at the end of each trial, to ensure

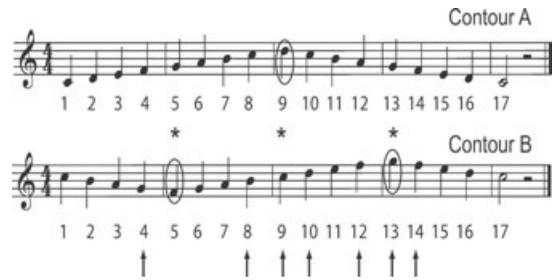
attentive processing. Detection accuracy and ERPs were measured for the accent manipulations. We hypothesized that both MMNs and P300 components would be elicited in response to the timbral accents, because the timbre-detection task required participants' directed attention. We also predicted an MMN in response to temporal accents induced with shortened tone durations. As these "deviant" accents were not relevant to the task, no P300 response was expected. Finally, we hypothesized that contour changes might interact with the processing of timbral accents, as the detection of timbre changes was slowed down near contour changes in Capota's<sup>1</sup> speeded-response task.

For several reasons it is difficult to extrapolate specific neural processes associated with multiple musical accents based on previous findings. First, researchers have typically employed short, simple stimuli in order to dissociate outcomes from co-occurring variables. The current study presents accents in longer contexts and directly compares early and late influences of context and expectancy on neural responses. Second, in contrast to the majority of studies, which present one accent type at a time, the present study manipulated three accent types systematically and measured ERP responses to their combination. To our knowledge, no ERP study has addressed neural response to timbral, temporal, and melodic contour accents in combination. By studying multiple accents in context, we can test whether ERP components associated with a given accent type are modulated by the presence of other types of accents.

## Method

### Participants

Sixteen musically trained listeners with a minimum of 5 years of private lessons on any instrument (mean = 9.8 years) were recruited from the McGill University community. Participants were healthy right-handed young



**Figure 1.** Example stimuli for melodic Contours A (top) and B (bottom). Arrows indicate possible locations of timbre changes; asterisks indicate temporal-gap locations and circles indicate contour-change locations.

adults with no history of hearing or neurologic impairment.

### Stimulus Materials

Six melodies consisting of 17 isochronous tones were composed, three for each of two contours (A and B): Contour A contained an ascending pattern up to tone 9, followed by a descending pattern (an example is shown in the top panel of Fig. 1). Contour B contained a descending pattern to tone 5, followed by an ascending pattern to tone 13, and then terminated with a descending pattern (example in bottom panel of Fig. 1). The melodic contour accents thus occurred on tone 9 for Contour A and on tones 5 and 13 for Contour B. All melodies were presented with 350-ms inter-onset intervals (IOIs), with tone durations equal to the IOI.

Three temporal accent conditions were created: (a) unmodified melodies with no temporal accents; (b) melodies with one temporal accent at position 9; and (c) melodies with two temporal accents at positions 5 and 13. Thus, temporal accents coincided either with contour accents in Contour A (position 9) or with the alternate contour accent locations in Contour B (positions 5 and 13). Temporal grouping accents were induced by shortening a single tone from 350 to 125 ms and inserting 225 ms of silence (a temporal gap), thus keeping the 350-ms IOIs constant. The temporal gap was placed on

the extreme tone of the melodic contour accent because listeners cannot perceive the shortened duration as a temporal gap until after it occurs (consistent with the method of Capota<sup>1</sup>).

Timbral accents were induced with a single-tone timbre change on 87.5% of trials. Possible timbre-change locations included tones 4 (a control condition, before any other accents occurred), 8, 9, 10 (around the location of Contour A change), and 12, 13, 14 (around the location of Contour B change). These locations were chosen to occur before, on, or after the melodic contour accents and temporal gaps. Foils formed 12.5% of the trials, on which no timbre change occurred, to maintain listeners' attention. We focus here on representative results for timbre changes at positions 9 and 13 (manipulations of melodic contour changes and temporal-gap accents) and position 4 (timbre change only). The melodic context tones were presented with a piano timbre and the timbre deviants were presented with a steel guitar timbre, chosen because of their similar amplitude envelopes. Each melody was 5950 ms in total duration. Stimuli were presented over insert earphones.

### Design and Procedure

The independent variables (temporal gaps, contour change, timbre change) were manipulated across trials within subjects. Dependent measures were response accuracy of timbre-change detection (recorded by keypress) and brain responses to the three accent structures, analyzed as ERPs. This resulted in 2 (contour)  $\times$  3 (temporal gaps)  $\times$  8 (timbre-change positions, among which only positions 4, 9, and 13 were analyzed) within-subjects design, yielding 144 unique stimuli. These were presented in three blocks of 48 trials each, and followed by a fourth block that was a repetition of the first (trials presented in reversed order), thus resulting in a total of 192 pseudo-randomly ordered trials. A block of 12 practice trials containing all melody subtypes and examples of all accent types was presented to familiarize listeners with

all melodies (including the location of contour accents) prior to the experimental trials.

Participants first completed musical background and neurologic questionnaires. They were told that they would hear a short melody on each trial, and that some of the trials would contain a one-note timbre change. Each melody was followed by a computer prompt that asked whether or not a timbre change had occurred. Participants completed a block of 12 practice trials, and then a block of 48 experimental trials. Participants were given a break at the end of each experimental block. The total duration of the experiment was approximately 2 h.

### Data Recording and Analysis

The EEG was continuously recorded from 19 Ag/AgCl electrodes referenced against the left mastoid (10/20 system, 500 Hz/24 bit sampling rate; Neuroscan Synamps2 amplifiers, Compumedics USA, Charlotte, NC); FCz served as the ground electrode. Electro-ocular artifacts were monitored using electrodes at the outer canthi of each eye (horizontal) and above and below the left eye (vertical). Data were analyzed using EEProbe 3.2 software (ANT Inc., Enschede, the Netherlands) and were filtered offline with a bandpass of 0.16–30 Hz; epochs containing eye blinks or movement artifacts were rejected and excluded from the ERP averages. Averages were computed both across the entire melody and for the critical accent locations (timbre-change positions, temporal-gap positions, and contour-change positions), time-locked to tone onset using a 200-ms prestimulus base-line. ERP components were quantified as mean amplitudes in representative time intervals and were analyzed with repeated-measures analysis of variance (ANOVAs) performed separately for three midline electrodes (Fz, Cz, Pz) and for 12 lateral electrodes (F3/4, C3/4, P3/4, F7/8, T3/4, T5/6) with topographical factors Anterior/Posterior (3), and Hemisphere (2). Where appropriate, we report Huyn/Feldt-corrected

*P*-values (but original degrees of freedom) to protect against Type-I errors.

## Results

### Behavioral Responses

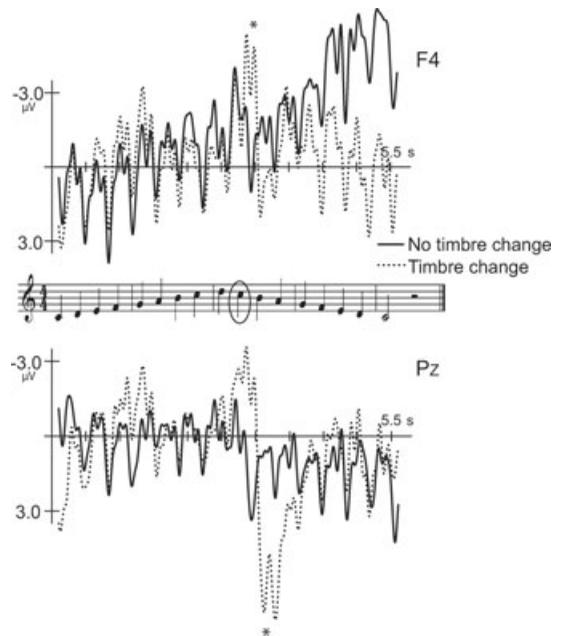
Listeners' accuracy in the offline timbre-detection task was assessed in an ANOVA on melodic contour-, temporal gap-, and timbre-change locations. A main effect of timbre-change location,  $F(6, 90) = 3.151$ ,  $P < 0.01$ , indicated that listeners responded more accurately to the timbre changes when they occurred at later context positions, consistent with Capota's speeded-response measures.<sup>1</sup> There was also a significant difference in response accuracy to trials with timbre changes (99.1%) compared with trials without timbre changes (95.66%), reflecting a 4.34% false-alarm rate and only 0.9% missed timbre changes. Thus, listeners attended well over all conditions.

### ERP Responses

ERP results are presented as grand average group data. Analyses revealed robust timbre-change effects across the accent positions and local interactions between accent types. Because of space limitations, we focus here on representative results for timbre changes on positions 9 and 13 (those trials in which timbre change, melodic contour change, and temporal gaps were crossed) as well as the control position 4 (timbre changes in the absence of other accents). To maximize data comparability across accent types and positions, the effects of timbre change and temporal gaps occurring at the same position (i.e., on the 9th and 13th melodic tones) were examined separately.

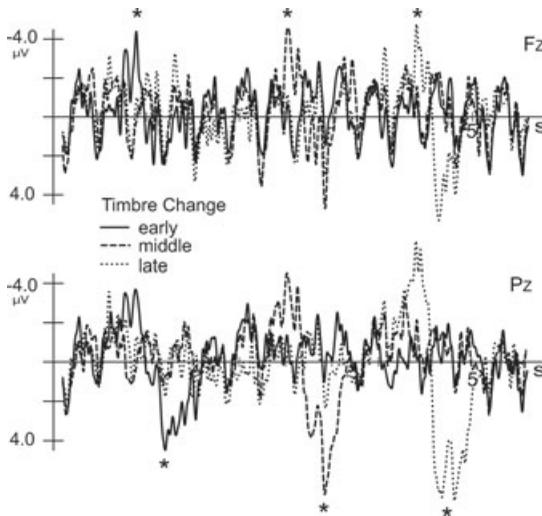
### Responses to Timbre Changes

Figure 2 shows ERP responses to timbre change across the entire melodic context, time-aligned to a sample melody. The dotted line indicates the ERP response to a timbre change



**Figure 2.** Grand average ERP across the entire melodic context at frontal F4 (top) and parietal Pz electrodes (bottom), time-aligned with one of the melodies used in the experiment. Dotted lines indicate response to timbre change on tone 10 (circled in musical notation), showing MMN/P300 pattern (asterisks); solid lines indicate response to no timbre change, with increased negative slow wave across the melodic context.

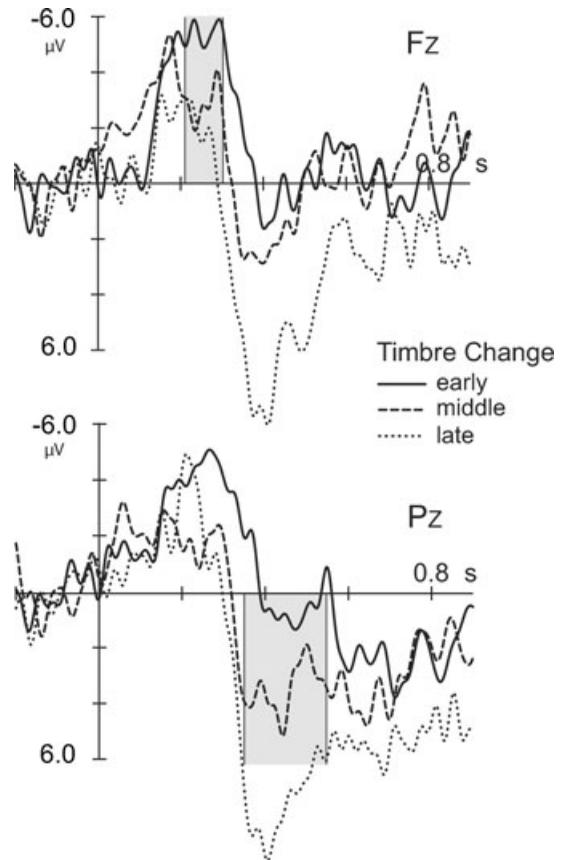
occurring on the 10th tone (circled) and the solid line represents the response to matched control trials with no timbre change. The timbre change elicited an early negative component (visible at both the frontal F4 and the parietal Pz electrodes) approximately 120–300 ms after tone onset, which is compatible with the topography and time course of the MMN/N2b. This negative deflection was followed by a parietal P300 peaking at about 400 ms. Figure 2 also shows that the negativity was superimposed with a P200 response that peaked at about 230 ms after each tone onset; the MMN/N2b peak associated with timbre accents thus seems to be split into two apparent peaks separated by the P200. In addition, the top panel of Figure 2 (frontal electrode) displays a negative slow wave, which seems to reflect the listeners' expectancy for an upcoming timbre



**Figure 3.** Grand average ERP response to timbre changes at early position 4 (solid line), middle position 9 (dashed line), and late position 13 (dotted line) in the melodic context. Asterisks in top figure indicate MMN at frontal electrode Fz; asterisks in bottom figure indicate P300 at parietal electrode Pz.

change. After encountering a timbre change (dotted line) at position 10, this negative trend discontinues, whereas it is still present in the control condition (solid line).

Figure 3 demonstrates the effect of melodic context on the MMN/N2b and P300 in frontal and parietal electrodes, respectively, at early (tone 4), middle (tone 9) and late (tone 13) timbre-change positions (a 0.5-Hz highpass filter was applied to eliminate slow waves, such as those shown in Fig. 2). Figure 3 demonstrates a consistent biphasic pattern across context positions. However, the peak amplitude of the MMN/N2b and P300 components appears to change across context positions. To further investigate position effects, statistical analyses were performed on smaller average windows (−200 to 900 ms) time-locked to the onset of the target tones at each position. The difference waves over these average windows (timbre change minus matched control) for trials with timbre change at early (4), middle (9), and late (13) context positions are superimposed in Figure 4. Trials containing gap accents (shortened tones) at these positions were analyzed



**Figure 4.** Difference wave (timbre change minus no timbre change) time-locked to target tones at early position 4 (solid line), middle position 9 (dashed line), and late position 13 (dotted line) in the melodic context. Shaded significance regions indicate analysis windows for MMN/N2b (210–300 ms) at Fz (top) and for P300 (350–550 ms) at Pz (bottom).

separately and are reported below (the presence of gaps did not alter the findings reported for timbre changes).

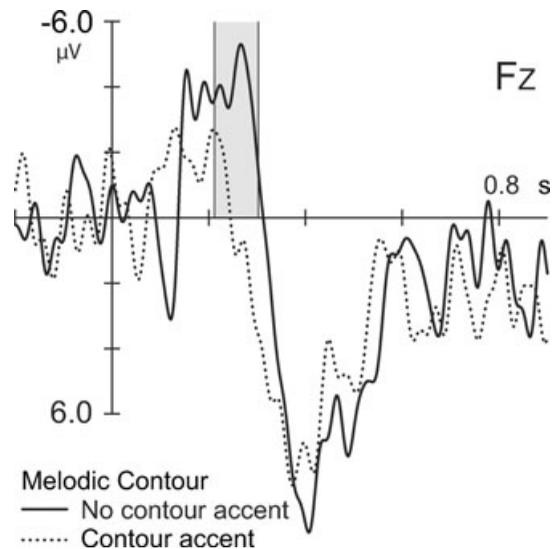
The MMN was quantified in two consecutive time windows of 120–210 ms and 210–300 ms (regions representing the timbre-deviant MMN and MMN/N2b components, respectively) and the P300 was quantified at 350–550 ms. A global ANOVA on the early MMN window (120–210 ms) revealed only main effects of timbre change [midline:  $F(1,15) = 42.96$ ;  $P < 0.0001$ ; lateral:  $F(1,15) = 75.67$ ,  $P < 0.0001$ ], as well as Timbre  $\times$  AntPost interactions reflecting the frontal distribution [midline:  $F(2,30) = 2.93$ ;

$P < 0.07$ ; lateral:  $F(2,30) = 5.86$ ,  $P < 0.03$ ], but no interactions with other factors. The global ANOVA for the MMN/N2b window (210–300 ms) revealed significant main effects of timbre change at both midline [ $F(1,15) = 35.50$ ;  $P < 0.0001$ ] and lateral electrodes [ $F(1,15) = 55.57$ ;  $P < 0.0001$ ]. An additional Timbre  $\times$  Position  $\times$  Ant/Post interaction [ $F(2,30) = 3.19$ ;  $P < 0.05$ ] was found at lateral electrodes as well as a marginally significant Timbre  $\times$  Position effect at the midline [ $F(2,30) = 2.65$ ;  $P < 0.09$ ], reflecting the larger negativity for early positions and its frontal distribution. Pairwise comparisons between positions revealed that the MMN/N2b amplitude at position 4 was larger than both positions 9 [midline: Timbre  $\times$  Position,  $F(1,15) = 6.51$ ;  $P < 0.03$ ] and 13 [lateral: Timbre  $\times$  Position  $\times$  Ant/Post,  $F(1,15) = 6.76$ ;  $P < 0.02$ ], whereas positions 9 and 13 did not differ from each other.

Analyses of the P300 window (350–550 ms) likewise confirmed main effects of Timbre [lateral:  $F(1,15) = 10.35$ ;  $P < 0.006$ ; midline:  $F(1,15) = 9.61$ ;  $P < 0.01$ ] as well as an interaction with position [Timbre  $\times$  Pos; lateral: ( $F(1,15) = 6.75$ ;  $P < 0.004$ ; midline:  $F(1,15) = 6.76$ ;  $P < 0.004$ ]. Pairwise comparisons revealed significant P300 differences between positions 13 and 9 [Timbre  $\times$  Position, lateral:  $F(1,15) = 5.29$ ;  $P < 0.04$ ; midline:  $F(1,15) = 3.96$ ;  $P < 0.07$ ] and between positions 13 and 4 [Timbre  $\times$  Position, lateral:  $F(1,15) = 11.76$ ;  $P < 0.004$ ; midline:  $F(1,15) = 10.83$ ;  $P < 0.005$ ], whereas positions 4 and 9 differed only marginally ( $P > 0.05$ ). Thus, both MMN/N2b and P300 responses to timbre change were modulated by contextual position in opposite directions; N2b amplitude decreased and P300 response increased at later contextual positions.

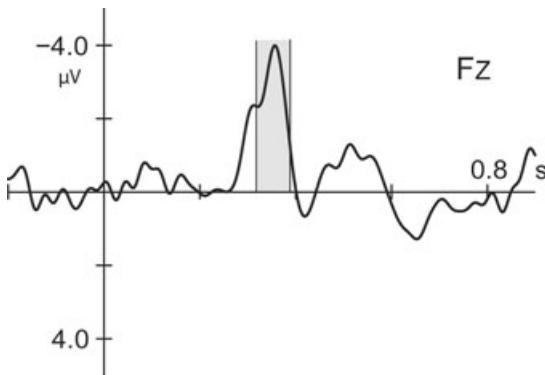
### Melodic Contour Interactions with Timbre Change

Intriguingly, position effects on the MMN/N2b component (210–300 ms) for timbre



**Figure 5.** Difference wave (timbre change minus no change) time-locked to target tones at position 13 by melodic contour condition, illustrated at Fz. Solid line indicates no coinciding contour accent; dotted line indicates coinciding contour accent. Shaded significance region indicates analysis window for MMN/N2b (210–300 ms).

change (but not the P300) were additionally influenced by the respective melodic contour accents. This effect is illustrated in Figure 5 for a comparison of contour effects on timbre-change detection at electrode Fz: the presence of a contour change decreased the MMN/N2b amplitude to timbre change at position 13. No corresponding effects of contour change were found to timbre changes at position 9. An ANOVA contrasting positions 9 and 13 reflected this pattern in a significant Position  $\times$  Timbre  $\times$  Contour  $\times$  AntPost interaction at both the midline [ $F(1,15) = 4.79$ ;  $P < 0.03$ ] and at lateral electrodes [ $F(2,30) = 4.36$ ;  $P < 0.04$ ]. Follow-up analyses for each position revealed a significant Timbre  $\times$  Contour  $\times$  AntPost interaction at position 13 for both midline [ $F(2,30) = 6.91$ ;  $P < 0.007$ ] and lateral electrodes [ $F(2,30) = 5.92$ ;  $P < 0.02$ ], while no such effect was present at position 9 (both  $F$ -values  $< 1$ ). This interaction is of particular interest because speeded-response-time measures with the same stimuli<sup>1</sup> yielded a slowed response for combined timbre and contour accents.



**Figure 6.** Difference wave (temporal gap minus no temporal gap) time-locked to target tones, across positions. Shaded significance region at Fz indicates analysis window for MMN (320–390 ms) to temporal gaps.

### Responses to Temporal Gaps

ERP effects on temporal gaps at positions 9 and 13 elicited a single consistent MMN component that was not followed by a P300, as expected for the task-irrelevant accents. The difference wave for gap minus no-gap trials, shown in Figure 6, shows a frontal MMN peaking around 360 ms after tone onset (about 100 ms later than the timbre-related MMN/N2b peak), which was expected as the gap cannot be perceived as such until the tone offset at 125 ms after the onset. The frontal MMN for temporal gaps, quantified in a time window of 320–390 ms, was reflected in significant gap main effects [midline:  $F(1,15) = 15.96$ ;  $P < 0.002$ ; lateral:  $F(1,15) = 14.87$ ;  $P < 0.002$ ] and interactions with AntPost [midline:  $F(2,30) = 30.26$ ;  $P < 0.0001$ ; lateral:  $F(2,30) = 12.22$ ;  $P < 0.0001$ ]. Position and contour effects were absent from the gap-related MMN. Thus, MMN response to musical accents was not always modulated by position within the melodic context.

### Discussion

Electrophysiological responses to co-occurring musical accents indicated systematic context effects across short melodies. Responses

to attended timbre accents that occurred at early, middle, and late contextual positions changed in amplitude for both MMN/N2b and the P300 components. In general, the later the timbre accent occurred in the melodies, the smaller the MMN/N2b amplitude and the larger the P300 amplitude. These findings were obtained under high attending conditions, as evidenced by high performance accuracy in the offline timbre-change detection task (99%) for both early- and late-occurring accents. Thus, amplitude changes were more likely due to anticipation of upcoming events and less likely due to global fluctuations in attention across the melodic context. The contextual effects across early to late positions are consistent with Capota's<sup>1</sup> behavioral findings of reduced online response latency to tones presented late in the melodies. The different scalp distributions of MMN/N2b and P300 responses, combined with different interactions of timbral and contour accents on MMN/N2b but not on P300, suggest they are not simply the result of the same underlying processes.

The increasing P300 amplitude to timbre changes at later melodic context positions may reflect increasing expectancies associated with stimulus probabilities and are consistent with faster detection times for timbre changes presented at later context positions.<sup>1</sup> This interpretation is also consistent with the frontal slow negative wave observed prior to timbre changes (Fig. 2). Alternatively, this effect may be related to the larger number of standards preceding the deviant. The number of tones with no timbre change presented before the timbre accents carried information about both the contextual position within the melody and the likelihood of upcoming timbre changes. Future work may tease apart these possible influences.

A reduced MMN amplitude to timbre changes at later context positions would be surprising, given the common perspective that the MMN reflects an automatic (pre-attentive) response to deviations from expectancies. A larger number of standards preceding a

deviant stimulus usually tends to elicit a larger-amplitude MMN<sup>12</sup> and top-down expectancies have been shown not to affect the MMN (Rinne *et al.*<sup>20</sup>). Additional analyses of the *early* MMN time window (120–210 ms) response showed effects of timbre change only, unaffected by position or contour. Thus, these findings suggest a possible N2b component in the later (210–300 ms) time window, which in general occurs in response to attended task-relevant deviants, has a time window that tends to overlap with the MMN, and varies with expectancy and task demands.<sup>17</sup> Previous studies with timbre-change deviants indicate a relatively early MMN time range (150–200 ms) compared to the later N2b time range.<sup>12</sup> Of note is the fact that the MMN amplitude responses to temporal gaps did not decrease across sequence positions, even though the gaps were completely predictable from the preceding context. Thus, attention to the particular stimulus dimensions may be necessary to elicit the contextually modulated MMN/N2b amplitudes.

An intriguing contextual effect was the influence of contour accents on neural responses to timbre change. Timbre changes at late contextual positions only that co-occurred with contour changes yielded a reduced MMN/N2b amplitude compared with the absence of contour change. Consistent with other findings that multiple accents compete for attention,<sup>5,6</sup> sensitivity to the timbral accents was reduced in the presence of other salient accents (contour change). This modulation of timbre sensitivity by contour was apparent only at later sequence positions. These findings are also consistent with Capota's<sup>1</sup> reduction in speeded responses to timbre changes that co-occurred with contour changes. Given that later sequence positions contain high predictability for contour changes, listeners may have anticipated the contour change, which detracted from detection of the timbre change. Thus, the reduced MMN/N2b may arise from reduced attention at later sequence positions that contain multiple musical accents. The relatively early latency of the MMN/N2b (210–300 ms) suggests that

different types of musical accents may interact with each other at early stages of perception (see also Caclin *et al.*<sup>21</sup>).

In contrast to timbre and contour accents, MMN responses to temporal gaps were not modulated by context. Moreover, the gap response did not elicit a P300, which replicates former findings of ERPs to unattended deviants.<sup>3,16</sup> As expected, the MMN to temporal gaps peaked later than MMN response to timbre changes; this is most likely due to the time required to determine tone duration differences, which are triggered by tone offsets as opposed to tone onsets.

All accent types occurred on metrically strong beats (downbeats) in this experiment, for several reasons. We controlled for metrical accent by placing all accents on downbeats, so that the effects of individual accents could be evaluated above and beyond any metrical effects. We chose metrical downbeats because melodic contour accents and metrical accents commonly co-occur in Western tonal music, and thus this alignment is typical. Finally, the corresponding control conditions (without the melodic contour accents) maintained the metrical accent, and therefore meter did not contribute to the evaluation of timbral, melodic contour, or temporal accents. Interactions of these factors with metrical accent may be addressed in future research.

In sum, listeners' neural responses to timbre accents but not to temporal gaps were modulated by accents' position within the unfolding melodic context. The contextual modulations implicated task relevance and expectancy (stimulus probability) information; in some cases, the combination of accents interacted in neural responses (timbre accents and contour accents) and in other cases, the neural responses were independent of each other (temporal-gap accents). This exploratory study extended previous findings by demonstrating the complex time course of how listeners integrate different musical structures in longer musical contexts. Accent structures formed by melodic contour, temporal gaps, and timbre co-occur and unfold

at different rates; their combined impact can be fully measured only in longer, richer musical contexts that just a few electrophysiological studies have used to date.<sup>3,16</sup> In short, complex musical forms elicit complex brain responses, and ERPs offer a promising approach for studying their temporal relationship.

### Acknowledgments

This research was supported by the Canada Foundation for Innovation, the Canada Research Chairs program, and by CIHR Grant MOP-74 to the first and last authors. The authors thank Christine Capota, John Drury, and Frances Spidle for assistance and Pascale Lidji for comments on an earlier draft.

### Conflicts of Interest

The authors declare no conflicts of interest.

### References

1. Capota, C.A. 2005. The perception of phrase boundaries in music. Senior thesis, McGill University, Montreal, Canada.
2. Goydke, K.N., E. Altenmueller, J. Möller & T.F. Munte. 2004. Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cogn. Brain Res.* **21**: 351–359.
3. Trainor, L.J., K.L. Macdonald & C. Alain. 2002. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *J. Cogn. Neurosci.* **14**: 430–442.
4. Boltz, M. & M.R. Jones. 1986. Does rule recursion make melodies easier to reproduce? If not, what does? *Cogn. Psychol.* **18**: 389–431.
5. Drake, C., W.J. Dowling & C. Palmer. 1990. Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Percept.* **8**: 315–333.
6. Melara, R.D. & L.E. Marks. 1990. Interaction among auditory dimensions: timbre, pitch, and loudness. *Percept. Psychophys.* **48**: 169–178.
7. Semal, C. & L. Demany. 1991. Dissociation of pitch from timbre in auditory short-term memory. *J. Acoust. Soc. Am.* **89**: 2404–2410.
8. Ladefoged, P., & D.E. Broadbent. 1965. Perception of sequence in auditory events. *Q. J. Exp. Psych.* **13**: 162–170.
9. Fodor, J.A. & T.G. Bever. 1965. The psychological reality of linguistic segments. *J. Verbal Learning Verbal Behav.* **4**: 414–420.
10. Gregory, A.H. 1978. Perception of clicks in music. *Percept. Psychophys.* **24**: 171–174.
11. Stoffer, T.H. 1985. Representation of phrase structure in the perception of music. *Music Percept.* **3**: 191–220.
12. Näätänen, R. 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* **38**: 1–21.
13. Näätänen, R. 2007. The mismatch negativity: where is the big fish? *J. Psychophysiol.* **21**: 133–137.
14. Donchin, E. 1981. Surprise! . . . Surprise? *Psychophysiology* **18**: 493–513.
15. Donchin, E., M.G.H. Coles. 1988. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* **11**: 357–374.
16. Tervaniemi, M., A. Castaneda, M. Knoll & M. Uther. 2006. Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. *Neuroreport* **17**: 1225–1228.
17. Grimm, S. & E. Schröger. 2005. Pre-attentive and attentive processing of temporal and frequency characteristics within long sounds. *Cogn. Brain Res.* **25**: 711–721.
18. McAdams, S., S. Winsberg, S. Donnadieu, *et al.* 1995. Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychol. Res.* **58**: 177–192.
19. Crummer, G.C., J.P. Walton, J.W. Wayman, *et al.* 1994. Neural processing of musical timbre by musicians, nonmusicians, and musicians possessing absolute pitch. *J. Acoust. Soc. Am.* **95**: 2720–2727.
20. Rinne T., S. Antila & I. Winkler. 2001. MMN is unaffected by top-down predictive information. *Neuroreport* **12**: 2209–2213.
21. Caclin, A., M.H. Giard, B.K. Smith & S. McAdams. 2007. Interactive processing of timbre dimensions: a Garner interference study. *Brain Res.* **1138**: 159–170.