



# Auditory event-related potentials reflect dedicated change detection activity for higher-order acoustic transitions

Annekathrin Weise<sup>a,\*</sup>, Erich Schröger<sup>a</sup>, Balázs Fehér<sup>b</sup>, Tímea Folyi<sup>b</sup>, János Horváth<sup>c,1</sup>

<sup>a</sup> Kognitive einschl. Biologische Psychologie, Institut für Psychologie, Universität Leipzig, Seeburgstr. 14–20, 04103 Leipzig, Germany

<sup>b</sup> Faculty of Education and Psychology, Eötvös Loránd University, Budapest, Hungary

<sup>c</sup> Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

## ARTICLE INFO

### Article history:

Received 13 December 2011

Accepted 3 June 2012

Available online 12 June 2012

### Keywords:

Asymmetry

Change detectors

Event-related potentials

N1

P2

Transition

## ABSTRACT

The processing of auditory changes at cortical level relies partly on dedicated change-detectors whose activity is reflected in the elicitation of the N1 and P2 event-related potentials (ERPs). In previous studies, N1 and P2 have been found only for first-order frequency transitions (i.e. constant-to-glide) but not for higher-order transitions (i.e. glide-to-constant). We tested whether this asymmetry is due to the complete lack, or the smaller number of dedicated higher-order change detectors compared to first-order change detectors by recording ERPs to constant-to-glide and glide-to-constant frequency transitions within pure and complex tones. For constant-to-glide transitions ERP amplitudes increased with the rate of frequency change and spectral complexity. Importantly, for glide-to-constant transitions, N1 was elicited, even though only for spectrally rich tones when the frequency-change rate was fastest. Thus, the asymmetry in auditory change-related N1 elicitation is attributable not to the lack of higher-order change detectors, but to their relatively low number.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Change detection is a fundamental function of the sensory system. Changes in sensory input convey information about potentially relevant changes in the environment, which may require exploration or indicate the necessity to change the ongoing behavior. Whereas sensory changes can be detected by comparing memory traces of previous and current stimulation (Chait et al., 2007; Näätänen et al., 1978, 2010), change-detection may also be implemented by dedicated low-level sensory processes (for a review see e.g. Fishbach et al., 2001; Phillips et al., 2002), which may allow faster detection times and more rapid responses than those based on memory comparisons. Dedicated change detectors may also provide accurate timing information on auditory transitions when there are no stationary or regular periods in the auditory signal, or these are too brief to engage higher-order, memory-based change-detection functions. Because of these advantages rapid change-detection may subserve a number of functions: for example, it may allow rapid parsing of the acoustic environment into separate events by identifying sound-onsets (Phillips et al.,

2002; Schwartz et al., 2012; Winkler et al., 2009), as well as analyzing speech or musical signals through the detection of spectral or intensity changes (Moore, 2003). The goal of the present study was to investigate the nature of low-level sensory change-detection as reflected by event related potentials (ERPs).

Auditory changes elicit a number of ERP waveforms. Transient waveforms follow acoustic changes in a short time and are present for relatively brief time periods only, whereas sustained ERP activity accompanies long lasting sounds (exceeding ~500 ms) for their whole duration (Picton et al., 1978a,b). One of the most prominent transient auditory ERP component is the N1. This is a negative ERP peaking frontocentrally between 100 and 150 ms following a change. When the electroencephalogram is recorded with a nose reference, N1 also shows a polarity inversion at mastoid sites, suggesting a supratemporal generator structure (for a review on N1 see Cone-Wesson and Wunderlich, 2003; Hyde, 1997; Martin et al., 2008; Näätänen and Picton, 1987). Often, N1 is preceded and followed by positive deflections, termed P1 and P2, that occur about 50–100 ms earlier and later, respectively, than N1 (Näätänen and Picton, 1987). The individual components of the P1–N1–P2 complex are mainly exogenous, that is, their parameters reflect the characteristics of the eliciting sound. For example, their latency and amplitude vary as a function of tone frequency (Dimitrijevic et al., 2008; Pratt et al., 2009), rate of frequency change (Arlinger et al., 1976, 1982; Noda et al., 1999; Ruhm, 1970, 1971; Yingling and Nethercut, 1983), and spectral complexity (Maiste and Picton, 1989; Seither-Preisler et al., 2003; Shahin et al., 2005, 2007).

\* Corresponding author. Tel.: +49 341 9735983; fax: +49 341 9735969.

E-mail addresses: [akweise@uni-leipzig.de](mailto:akweise@uni-leipzig.de) (A. Weise), [schroger@uni-leipzig.de](mailto:schroger@uni-leipzig.de) (E. Schröger), [feher.balazs.feher@gmail.com](mailto:feher.balazs.feher@gmail.com) (B. Fehér), [folyi.timea@gmail.com](mailto:folyi.timea@gmail.com) (T. Folyi), [horvath@cogpsyphy.hu](mailto:horvath@cogpsyphy.hu) (J. Horváth).

<sup>1</sup> Tel.: +36 1 354 2290; fax: +36 1 354 2416.

However, even though P1, N1, and P2 are elicited together, it is assumed that they reflect functionally different processes (e.g. Chait et al., 2008; Ross and Tremblay, 2009; for a review see Crowley and Colrain, 2004) because they are affected differentially by stimulus parameters (e.g. Shahin et al., 2005, 2007), stimulus expectancy (Viswanathan and Jansen, 2010), and participant variables (e.g., age, arousal state, etc., for a review see Crowley and Colrain, 2004). Furthermore, these waveforms originate from spatially separate generators (Chait et al., 2008; Meyer et al., 2006; Pantev et al., 1996).

Transient ERP responses can be elicited by a change in various sound parameters. For example, changes in spectrum (Dimitrijevic et al., 2008; Harris et al., 2008; Jones and Perez, 2001, 2002; Pratt et al., 2009), intensity (Dimitrijevic et al., 2009) or changes from consonant to vowel (Kaukoranta et al., 1987; Ostroff et al., 1998) are known to elicit N1 and P2.

Interestingly, however, not all types of changes within continuous sounds have been found to elicit transient auditory ERPs. Whereas *first-order transitions*, that is, changes in a constant (temporally invariant) sound parameter lead to the elicitation of transient ERP responses, *higher-order transitions*, that is, transitions from non-constant parameters (i.e. parameters which continually change in time) do not elicit these ERPs. For example, in a continuous pure tone a transition from a constant-frequency segment to a gradually rising- or falling-frequency segment (termed as *glide segment* in the following) elicits a transient ERP response, however, transitions from the same glide segments to constant-frequency segments do not (Clynes, 1969; Kohn et al., 1978). This asymmetry in the ERP elicitation pattern is paralleled by an asymmetry in the reaction times, for which glide-to-constant transitions lead to larger variance than constant-to-glide transitions (Kohn et al., 1978). This asymmetry was not only found to changes in frequency but also to changes in other stimulus parameters like intensity (Clynes, 1969; Kohn et al., 1980).

The goal of the present study was to investigate the nature of the asymmetry in auditory change detection as reflected by ERPs. We contrasted two hypotheses explaining this asymmetry:

- (1) According to the *qualitative-difference* hypothesis, the presence of a transient ERP response to first-order auditory transitions and the absence of such a response to higher-order transitions may indicate that whereas there are sensory structures dedicated to the detection of first-order changes, no comparable dedicated structures exist for the detection of higher-order transitions. Because in natural environments sounds are rarely constant over time, and relevant sounds like speech show considerable temporal variation, a complete lack of dedicated higher-order change detectors is not an ecologically valid assumption.
- (2) A more persuasive explanation for the asymmetry in transient ERPs is provided by the *quantitative-difference* hypothesis. According to this assumption, the presence of a transient ERP response to first-order auditory transitions may indicate the activation of a large number of first order-change detectors, which lead to the elicitation of observable change-related transient ERPs. The absence of a transient ERP response to higher-order auditory transitions, however, may indicate that the stimulus material used in previous studies (Clynes, 1969; Kohn et al., 1978) only allowed the activation of a small number of higher order-change detectors, which did not lead to the elicitation of observable change-related transient ERPs. Since the amplitude of transient ERPs elicited by first-order changes generally grows with the magnitude of change (that is, the rate of frequency change; e.g. Arlinger et al., 1976; Noda et al., 1999; Ruhm, 1970), we expect that increasing the magnitude of change increases the amplitude of higher-order change-related ERPs as well. Furthermore, we hypothesize that spectrally rich

tones might also activate a larger proportion of the hypothetical higher-order change detectors (Maiste and Picton, 1989; Shahin et al., 2005, 2007), and therefore contribute to the elicitation of observable higher-order change-related brain potentials.

To test these hypotheses, we recorded ERPs elicited by glide-to-constant transitions within continuous tones. Whereas previous studies used relatively slow glides ( $\sim 7$  semitone/s [ST/s]: Kohn et al., 1978;  $\sim 9$  ST/s: Clynes, 1969) the rate of frequency change in the present study was in the 10–40 ST/s range. Moreover, besides pure tones, we also presented spectrally complex tones. According to the quantitative-difference hypothesis, the extended parameter range might result in the elicitation of observable transient ERPs to glide-to-constant transitions, reflecting the existence of higher-order change detectors. The absence of such ERPs would, on the other hand, support the qualitative-difference hypothesis. For control purposes, we also presented constant-to-glide transitions with the same set of parameters. In accord with the literature (Arlinger et al., 1976; Noda et al., 1999; Ruhm, 1970), constant-to-glide transitions should elicit transient ERPs with amplitudes increasing with the magnitude of change.

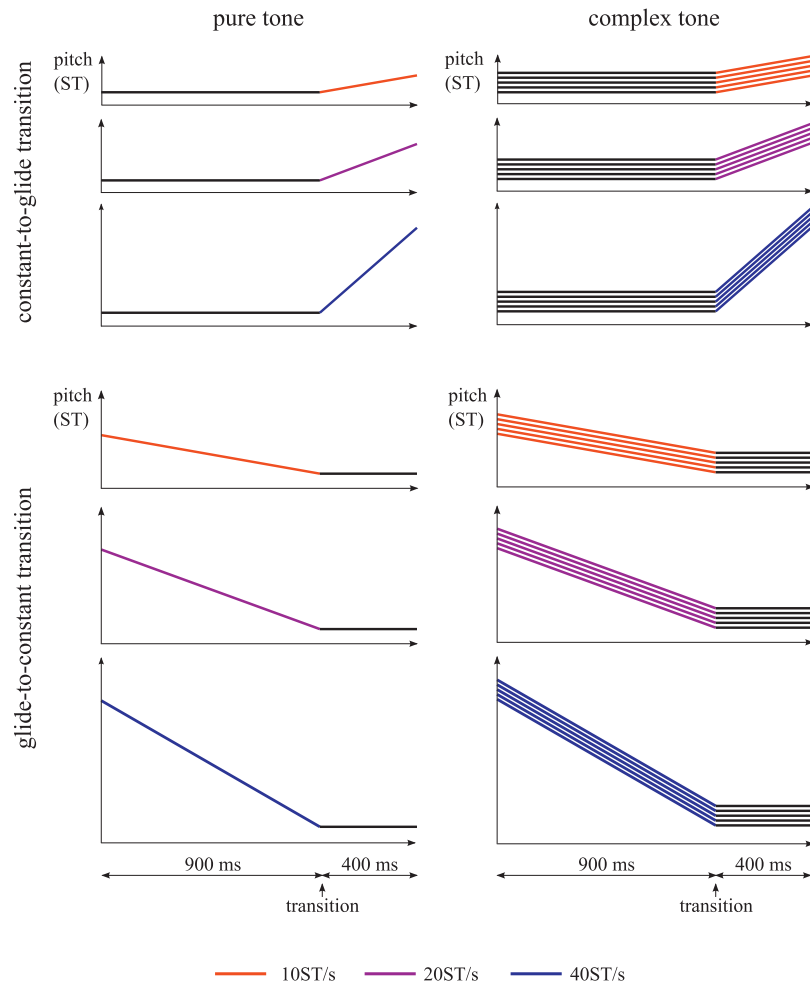
## 2. Materials and methods

Fourteen young adult volunteers (8 women, 1 left-handed) with self-reported normal hearing (21–32 years, mean age: 25 years) participated in the experiment either for course credit or payment. Data from one additional participant were excluded from analysis due to too many EEG-artifacts. All participants gave written informed consent after the experimental procedures were explained to them prior to the beginning of the measurements. The experimental protocol conformed to the Declaration of Helsinki and the ethics guidelines of the German Association of Psychology (ethics board of the Deutsche Gesellschaft für Psychologie, DGPs: <http://www.dgps.de/dgps/aufgaben/ethikr12004.pdf>). During EEG recording the participants were sitting in a chair in an acoustically attenuated and electrically shielded booth, and watched a self-selected, silent, subtitled movie, and were instructed to ignore the sounds.

The twelve types of tones presented during the experiment (see Fig. 1) can be described by three attributes with different levels: (1) transition type (glide-to-constant or constant-to-glide, described below), (2) rate of frequency change (10, 20 or 40 ST/s), and (3) spectral complexity (pure or complex).

- (1) *Transition type*: Each tone consisted of two segments. The duration of the first segment was 900 ms, the second one 400 ms (for a total of 1300 ms, including 5–5 ms attack and decay ramps). One of the segments was a *glide segment*, the other a *constant-pitch segment*. Tones with *glide-to-constant* transitions started with a glide segment with exponentially decreasing frequency, which was followed by a constant-pitch segment. Tones with *constant-to-glide* transitions started with a constant-pitch segment, which was followed by a glide segment with exponentially increasing frequency. The segment preceding the transition was chosen to be of 900 ms duration, because for constant-to-glide transitions the transient ERP amplitude is known to increase with the duration of the constant segment (Clynes, 1969; Maiste and Picton, 1989).
- (2) *The rate of frequency change* was 10, 20, or 40 ST/s for the glide segments. Note that the rate of change was exponential in frequency (as measured in Hz units), and it was linear in pitch (as measured in semitones). This is in contrast to previous studies which presented linear (as measured in Hz) frequency glides (Clynes, 1969; Kohn et al., 1978). The range of frequency change rates was similar and transcended those used in previous studies ( $\sim 7$  ST/s: Kohn et al., 1978;  $\sim 9$  ST/s: Clynes, 1969).
- (3) *Spectral complexity*: Complex tones consisted of five equally weighted harmonics ( $1 \times f_0$ ,  $2 \times f_0$ ,  $3 \times f_0$ ,  $4 \times f_0$ ,  $5 \times f_0$ ) with the base frequency ( $f_0$ ) equaling that of the corresponding pure tone. For pure tones, the frequency of the constant segment was 440 Hz.

The tones were offline generated via CSOUND (<http://www.csounds.com/>). Stimuli were presented binaurally via headphones with an intensity of 65 dB(A) (for the 440 Hz pure and complex constant-pitch segments) with an inter-stimulus interval of 700 ms. The stimulation was split into 21 experimental blocks. Each block was split into 4 microblocks, in which one of the 12 stimulus types was repeated for one minute. Microblocks were presented with random permutations. Overall, each of the 12 stimulus types was presented 210 times. The stimulation was run via MATLAB (<http://www.mathworks.com>) using the Cogent2000 toolbox ([http://www.vislab.ucl.ac.uk/cogent\\_2000.php](http://www.vislab.ucl.ac.uk/cogent_2000.php)). Stimulation time was 84 min. The experiment lasted about 3–4 h, including the time for electrode application and removal and individual breaks as needed.



**Fig. 1.** Schematic illustration of the twelve types of tones presented in the experiment. The tones consisted of a constant frequency, and a frequency glide segment. The duration of the first segment was 900 ms, that of the second 400 ms. The tones varied in three attributes: transition type: constant-to-glide (top), glide-to-constant (bottom), rate of frequency change (indicated by different linetypes), and spectral complexity: pure (left) or complex (right). Complex sounds consisted of five equally weighted harmonics with the base frequency equaling that of the corresponding pure tone. The base-frequency of the constant segment was the same for all tones.

Using a BIOSEMI Active-Two amplifier, electroencephalographic (EEG) activity was recorded with Ag/AgCl electrodes from 64 standard channel locations according to the 10% system (Nuwer et al., 1998) and from the left and right mastoids. Additionally, eye movements were monitored by recording horizontal and vertical electrooculogram (EOG). EEG and EOG signals were sampled at 512 Hz. An electrode placed on the tip of the nose served as offline-reference.

The EEG was filtered with a 20 Hz lowpass filter. We refrained from high-pass filtering the data in order to show that the transition-related ERP responses were superimposed on the sustained negativity.<sup>2</sup> For the ERP analysis, a two-pass procedure was used. In the first pass, for each trial a 1400 ms long epoch was extracted, including a 100 ms long pre-stimulus baseline interval. Epochs with signal range exceeding 150  $\mu$ V on any channel were discarded from further processing, as well as epochs corresponding to the first two trials of each microblock. The remaining epochs were averaged separately for each stimulus type. This first pass (Fig. 2, left) allowed the assessment of the sustained responses and their development preceding the transition, justifying the conclusion that the post-transition ERP responses are caused by the physical transition, and not by the preceding glide.

The second pass (Fig. 2, right), however, allowed focusing on the constant-to-glide and glide-to-constant transient ERPs. Here, for each trial an epoch of 500 ms duration including a 100 ms pre-transition baseline interval (i.e. the 800–1300 ms post-tone-onset interval) was extracted. Due to the pre-transition baseline, potential amplitude level differences in sustained activities are compensated, allowing the comparison of transition-related ERPs across different conditions. Epochs with signal range exceeding 100  $\mu$ V on any channel were discarded from further

processing, as well as epochs corresponding to the first two trials of each microblock. Epochs were averaged separately for each stimulus type. The transition-related response, when present under visual inspection in the group-average waveform, was measured at the Fz lead and the mastoids (the average of the left and right mastoid signals) from the individual ERPs as the mean amplitude in the 30-ms window (Table 1) centered at the latency of the group-average N1 and P2 peak, respectively. Whenever applicable, N1–P2 amplitude differences were calculated. N1 and P2 peak latencies (Table 3) for constant-to-glide transitions in complex sounds were identified in the 50–200 ms intervals for N1 and 150–300 ms post-transition intervals for P2.

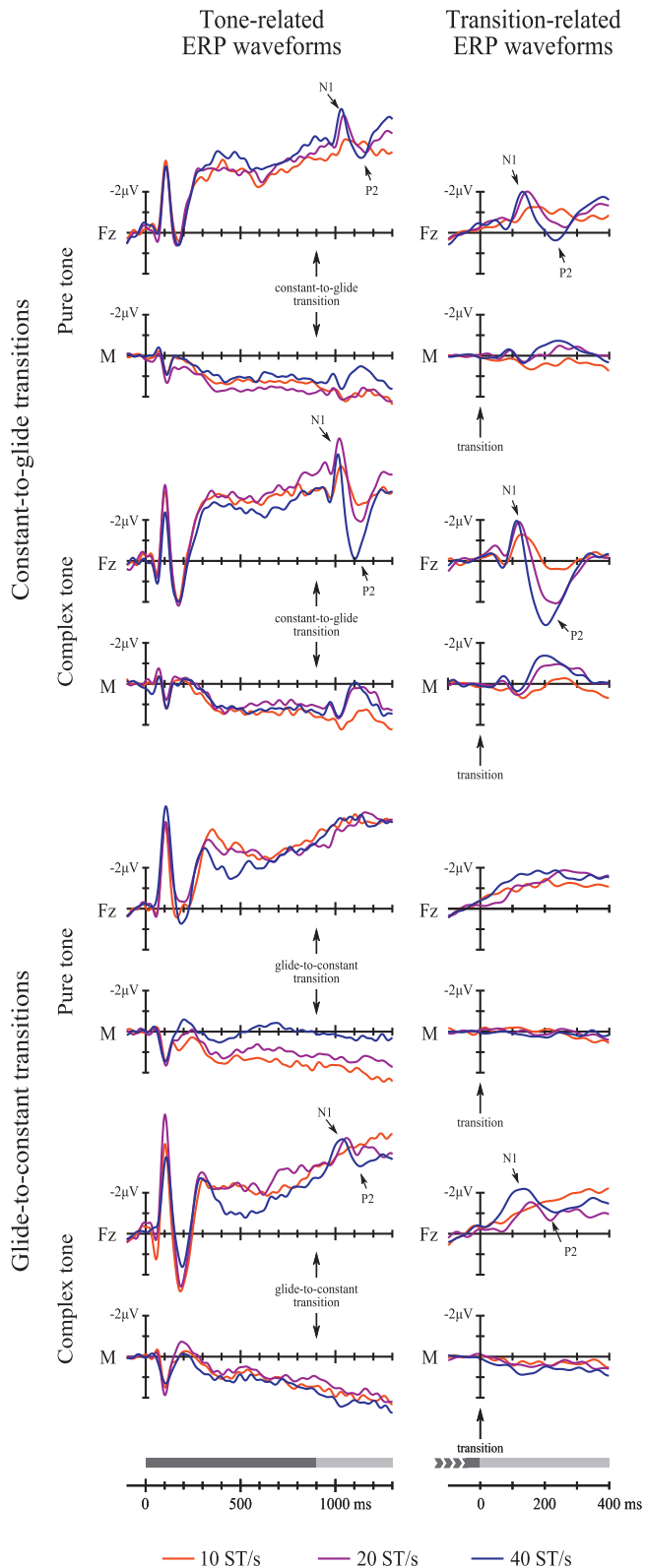
The primary analyses concerned whether the rate of frequency change affected the transient ERPs, therefore one-way repeated measures ANOVAs were calculated using the Rate factor (10, 20 or 40 ST/s). Significant rate-effects in these analyses signal the presence of a transient ERP response. Nonetheless, a response may also be present when the rate-effect is non-significant, due to a plateau-effect. That is, the transition-related response may have already reached its full amplitude at the 10 ST/s rate. In such cases, the presence of the ERP response was verified by two-tailed Student's one sample *t*-tests of the respective amplitudes against zero.

Furthermore, to assess potential differences in sensitivity to pure and spectral complex tones, ERP amplitudes were analyzed in two-way repeated measures ANOVAs with Complexity (Pure, Complex) and Rate (10, 20, or 40 ST/s) factors for tones with constant-to-glide transitions. Greenhouse–Geisser adjustment was applied throughout. In the result section the Greenhouse–Geisser epsilon ( $\epsilon$ ) values and the uncorrected degrees of freedom are reported.

### 3. Results

Fig. 2 (left column) shows that the onset of each stimulus elicited a clear P1–N1–P2 waveform. This was followed by a frontally

<sup>2</sup> The N1 effects, which are in the focus of this study, were similarly present when applying a 1–20 Hz band-pass filter whereas the sustained activity was not. In order to not “hide” the existence of the sustained activity we preferred the present way of analyzing the data.



**Fig. 2.** Group-average ERP waveforms measured at Fz and the average of left and right mastoid (M) signal, elicited by the twelve types of tones, arranged in rows according to tone-complexity and transition type. The tone-related waveforms shown in the left column and the transition-related waveforms in the right column result from different ERP processing settings. For the tone-related waveforms (left column), tone-onset (0 ms) is at the crossing of the axes; for the transition-related waveforms (right column), the crossing of the axes (0 ms) indicates the onset of the transition. The 100-ms baseline interval is either before sound onset (left column) or before the transition (right column). Within each graph the three ERPs corresponding to the rates of frequency change (indicated by different linetypes) are shown.

maximal sustained negativity. After the transition at 900 ms, distinct, more rapid ERP waveforms are observable for the constant-to-glide transitions, and for the complex tones with glide-to-constant transitions. These transient ERPs are clearly superimposed on the sustained negativity. They were analyzed from the ERPs resulting from the second-pass processing (Fig. 2, right), which removes level differences due to the sustained activity by using a 100 ms pre-transition baseline.

In the following the results obtained to the transition are reported, separately for transition type and spectral complexity, in dependence of the rate of frequency change.

### 3.1. Pure tones with glide-to-constant transitions

No transition-related ERP response was observable for pure sounds with glide-to-constant transitions (Fig. 2). Rather, the sustained negativity that established during the continuous sound remained uninterrupted irrespective of the occurrence of the transition within the stimulus until the offset of the stimulus.

### 3.2. Complex tones with glide-to-constant transitions

The glide-to-constant transition within spectral complex sounds elicited a clearly identifiable N1 (Fig. 2) with fronto-central maximum (Fig. 3) only for the 40 ST/s rate. A one-way repeated measures ANOVA of the ERP amplitudes calculated in the 30-ms interval centered at the group-average peak latency of the N1 that was observable in the 40 ST/s rate condition showed a significant rate effect ( $F(2,26) = 7.92$ ,  $p = 0.005$ ,  $\epsilon = 0.79$ ). Follow-up comparisons showed that the ERP amplitude was significantly higher (more negative) for the 40 ST/s than for the other two rates ( $F(1,13)$ -scores  $> 12.02$ ,  $p$ -scores  $< 0.01$ ), but no other significant difference was found. No significant rate effect was found at the mastoids (Fig. 2).

### 3.3. Pure tones with constant-to-glide transitions

The constant-to-glide transition elicited N1–P2 waveforms (Fig. 2), with inverted polarity at the mastoids (even though less clearly for the 10 ST/s rate). One-way repeated measures ANOVAs of the N1 amplitudes showed no significant differences at Fz or the mastoids. The one-way repeated measures ANOVA of the P2 amplitudes measured at Fz showed no significant effect, but the ANOVA of the mastoid signals showed a significant rate effect ( $F(2,26) = 12.13$ ,  $p < 0.001$ ,  $\epsilon = 0.93$ ). Follow-up analyses showed that in comparison to the 10 ST/s rate, amplitudes were larger for the 40 and 20 ST/s rates ( $F(1,13)$ -scores  $> 16.40$ ,  $p$ -scores  $< 0.01$ ), but no other significant difference was found.

The presence of the N1–P2 complex at Fz was supported only for the 20 and 40 ST/s rates by the results of one sample two-tailed  $t$ -tests of the N1–P2 amplitude differences against zero (see Table 2).

### 3.4. Complex tones with constant-to-glide transitions

The constant-to-glide transition elicited a clearly identifiable P1–N1–P2 complex for all rates of frequency change (Fig. 2). N1 and P2 showed a frontocentral voltage distribution across the scalp with polarity inversion at the mastoids (Fig. 3).

The one-way repeated measures ANOVAs of the N1 amplitudes (at Fz or at the mastoids) showed no significant effects, but the ANOVAs of the P2 amplitudes showed significant rate

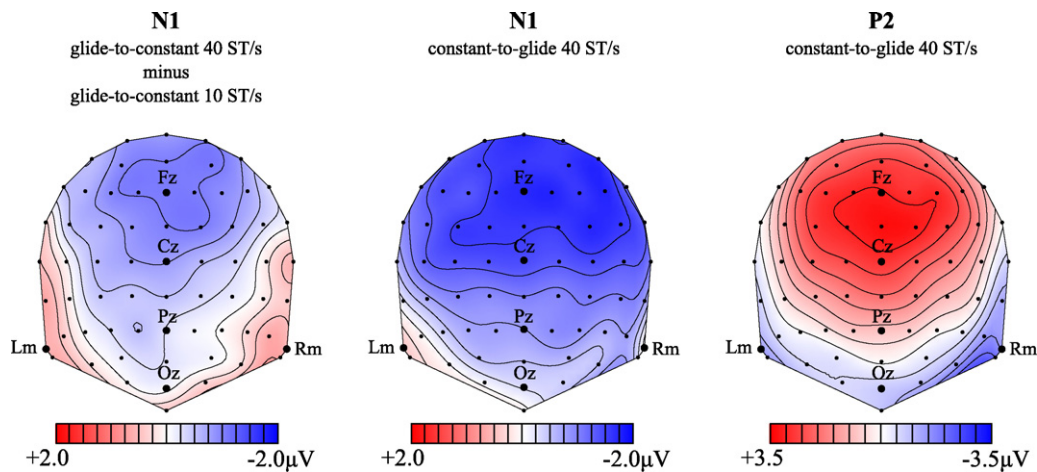
The temporal structure of the tone-segments is represented by gray bars above the time scale at the bottom of the figure. Arrows indicate the transition-related N1 and P2 waveforms.

**Table 1**  
Mean N1 and P2 amplitudes at Fz and mastoids in the given time window elicited to glide-to-constant and constant-to-glide transitions, respectively, in pure and complex sounds as a function of rate of frequency change.

Sound type	Rate (ST/s)	N1			P2		
		Window <sup>a</sup> (ms)	Amplitude in $\mu\text{V}$ (SE) at Fz	Amplitude in $\mu\text{V}$ (SE) at mastoids	Window <sup>a</sup> (ms)	Amplitude in $\mu\text{V}$ (SE) at Fz	Amplitude in $\mu\text{V}$ (SE) at mastoids
Complex glide-to-constant	10	120–150	–1.14 (0.25)				
	20	120–150	–1.27 (0.19)				
	40	120–150	–2.17 (0.19)				
Complex constant-to-glide	10	112–142	–1.23 (0.28)	0.62 (0.17)	208–238	0.38 (0.60)	–0.15 (0.19)
	20	104–134	–1.72 (0.27)	0.49 (0.17)	219–249	2.02 (0.39)	–0.95 (0.20)
	40	98–128	–1.67 (0.39)	0.29 (0.23)	188–218	3.04 (0.36)	–1.36 (0.26)
Pure constant-to-glide	10	167–197	–1.24 (0.32)	0.52 (0.16)	297–327	–0.83 (0.33)	0.44 (0.19)
	20	129–159	–1.92 (0.27)	0.25 (0.18)	247–277	–0.31 (0.35)	–0.39 (0.18)
	40	116–146	–1.86 (0.39)	0.29 (0.19)	217–247	0.33 (0.47)	–0.70 (0.22)

As glide-to-constant transitions in pure tones did not elicit a transition response they are not listed here. SE, standard error of means; ST, semitones.

<sup>a</sup> The window in which mean amplitudes were measured are related to the onset of the transition.



**Fig. 3.** Topographical distribution of the group-averaged transition-related N1 and P2 waveforms for the glide-to-constant (left) and constant-to-glide (middle and right) transitions within the complex tones for the 40 semitone/s (ST/s) rate. To estimate the topography of the transient ERP effect without the contribution of the slow wave (see Fig. 2), the ERP to the corresponding 10 ST/s rate transition was subtracted from the ERP elicited by the 40 ST/s glide-to-constant transition (left).

effects (at Fz:  $F(2,26) = 17.26$ ,  $p < 0.001$ ,  $\epsilon = 0.73$ ; at the mastoids:  $F(2,26) = 9.32$ ,  $p < 0.01$ ,  $\epsilon = 0.90$ ). Follow-up analyses revealed that at Fz, all amplitudes significantly differed from each other being largest for the highest rate ( $F(1,13)$ -scores  $> 8.20$ ;  $p$ -scores  $< 0.05$ ), while at the mastoids the amplitude for the 10 ST/s rate was lower (less negative) than for the other two rates ( $F(1,13)$ -scores  $> 9.17$ ;  $p$ -scores  $< 0.01$ ).

Due to the lack of a significant rate-effect on the N1, the presence of the N1–P2 complex was verified for each stimulus type by

**Table 2**  
Mean N1–P2 amplitude differences for ERPs elicited to constant-to-glide transitions in pure and complex sounds, respectively, as a function of rate of frequency change.

Sound type	Rate (ST/s)	N1–P2 amplitude in $\mu\text{V}$ (SE) at Fz	$T(df = 13)^a$
Complex constant-to-glide	10	–1.61 (0.44)	–3.66*
	20	–3.75 (0.50)	–7.45**
	40	–4.72 (0.61)	–7.79**
Pure constant-to-glide	10	–0.41 (0.25)	–1.67n.s.
	20	–1.61 (0.41)	–3.94*
	40	–2.19 (0.57)	–3.86*

ST, semitones; SE, standard error of means.

<sup>a</sup> To verify the presence of the transition response, one sample two-tailed Student's  $t$ -tests were applied to test against zero level (\*\* $p < 0.001$ , \* $p < 0.01$ , n.s., not significant). The windows, in which corresponding mean amplitudes were measured, are specified in Table 1.

one-sample, two-tailed  $t$ -tests comparing N1–P2 amplitude differences at Fz against zero (see Table 2).

For the ERPs elicited by these tones, the signal-to-noise ratio also allowed the unequivocal individual assessment of N1 and P2 peak latencies (see Table 3) at Fz. Therefore, one-way repeated measures ANOVAs of N1 and P2 peak latencies were calculated at Fz, which showed significant rate effects (N1:  $F(2,26) = 10.838$ ,  $p < 0.01$ ,  $\epsilon = 0.63$ ; P2:  $F(2,26) = 18.98$ ,  $p < 0.001$ ,  $\epsilon = 0.87$ ). Follow-up analyses showed that N1 latency decreased with increasing rate, with significant differences between all rates ( $F(1,13)$ -scores  $> 6.29$ ;  $p$ -scores  $< 0.05$ ). P2 latency was significantly shorter for the 40 ST/s rate than for the other two rates ( $F(1,13)$ -scores  $> 20.53$ ;  $p$ -scores  $< 0.001$ ).

**Table 3**  
N1 and P2 peak latencies to constant-to-glide transitions in spectral complex sounds as a function of rate of frequency change.

Sound type	Rate (ST/s)	N1 latency (SE) at Fz in ms	P2 latency (SE) at Fz in ms
Complex constant-to-glide	10	128 (0.005)	246 (0.008)
	20	121 (0.003)	233 (0.004)
	40	113 (0.003)	199 (0.005)

Note, that peak latencies were identified in 950–1100 ms intervals relative to sound onset for N1 and 1050–1200 ms intervals relative to sound onset for P2. N1 and P2 peak latencies were measured relative to the onset of the transition. ST, semitones; SE standard error of means.

### 3.5. Sensitivity to spectral complexity

Constant-to-glide transitions elicited larger transition responses for complex sounds than for pure sounds. Moreover, the rate of frequency change affected the transition response. The two-way repeated measures ANOVA of P2 amplitudes at Fz showed significant effects for spectral complexity ( $F(1,13)=82.52$ ,  $p<0.001$ ) and rate ( $F(2,26)=15.3472$ ,  $p<0.001$ ,  $\epsilon=0.77$ ), but their interaction was not significant ( $F(2,26)=2.7616$ ,  $p>0.05$ ,  $\epsilon=0.98$ ). An analogous ANOVA of the N1 amplitudes did not reveal any significant result. Follow-up analyses for the rate main effect revealed that all P2 amplitudes (aggregated over ERPs to pure and complex tones) significantly differed from each other being largest (more positive) for the 40 ST/s rate and smallest (less positive) for the 10 ST/s rate ( $F(1,13)$ -scores  $>6.55$ ;  $p$ -scores  $<0.05$ ).

## 4. Discussion

The goal of the present study was to investigate the nature of asymmetry in the detection of a first-order vs. a higher-order auditory transition within continuous sounds, as reflected by transient auditory ERPs.

The main finding is that, in contrast with the previous notions (Clynes, 1969; Kohn et al., 1978), even glide-to-constant transitions elicited N1-like activity over frontocentral brain regions. This shows that there are cortical structures dedicated to higher-order change detection besides first-order auditory change detectors. This finding supports the notion that the detection of auditory transitions in sounds with considerable temporal variation – for example in natural sounds and especially speech – relies on dedicated higher-order change-detection structures. We hypothesized that these structures provide accurate change detection function for auditory signals for which higher-order, memory-based processes cannot provide a timely or accurate transition-information.

The present findings fit and extend the results of previous experiments (Clynes, 1969; Kohn et al., 1978) in that N1-like activity was not present for pure tones with relatively slow rates of frequency change, whereas N1-like activity was elicited by spectrally rich tones with glide-to-constant transitions that were characterized by a relatively high rate of frequency change (40 ST/s). This is in accord with the assumption that the number of such higher-order change detectors is relatively low compared to those involved in first-order change detection whose activity is already observable for pure tones with lower rates ( $\sim 7$  to 9 ST/s; Clynes, 1969; Kohn et al., 1978). Because of the low number of dedicated higher-order change-detectors, the ERP correlates of their activity can be observed only when a larger number of these higher-order change detectors is activated, for example by increasing the rate of frequency change (as known for first-order transitions: e.g. Arlinger et al., 1976; Noda et al., 1999; Ruhm, 1970), or spectral complexity (Maiste and Picton, 1989; Shahin et al., 2005, 2007).

In line with previous studies (Arlinger et al., 1976, 1982; Clynes, 1969; Kohn et al., 1978; Noda et al., 1999; Ruhm, 1970), the constant-to-glide transitions elicited distinct transient ERPs for pure and spectrally rich sounds (even though less clearly for the 10 ST/s rate in pure sounds) that varied as a function of rate of frequency change and spectral complexity. The frontocentral ERP-response was accompanied by a polarity inversion at the mastoids.

Whereas the rate of frequency change had no significant effect on the constant-to-glide transition-related N1 amplitude, P2 amplitude significantly increased with increasing rate of frequency change (though for pure tones only at mastoid sites), and was higher for spectrally rich sounds. Moreover, the latencies of N1 and P2 elicited by constant-to-glide transitions within complex sounds decreased with increased rate of frequency change. These

results are largely compatible with those by Noda et al. (1999), who found that the amplitudes of the magnetoencephalographically (MEG) measured N1 and P2 event-related fields (ERFs) increased with the rate of frequency change. Also, the current findings resemble those obtained via simultaneous EEG/MEG recordings by Shahin et al. (2005), who found the P2 amplitude (but not that of N1) to increase with spectral complexity. The apparent ceiling effect on the N1 amplitude in the present study might be due to the overlap of the N1 and P2 ERPs, which might not be reflected by the ERFs in response to the rate-of-frequency change (Noda et al., 1999).

The direction-dependent asymmetry in the detection of auditory changes in the current approach between glide and constant frequency segments (see also Clynes, 1969; Kohn et al., 1978) is interpreted as the reflection of the activity of different numbers of different types of dedicated change detectors. We suggest that the larger transient ERPs to constant-to-glide transitions do reflect a larger number of (synchronously) responsive cortical first-order change detectors, whereas the absent or smaller ERPs to glide-to-constant transitions reflect a smaller number of (synchronously) responsive higher-order change detectors. The larger transient ERPs to constant-to-glide transitions might reflect the contribution of cortical neurons that are tuned to changes in frequency (Phillips et al., 1985; Whitfield and Evans, 1965), which are not involved in the processing of a constant-frequency segment following a glide segment. Similar asymmetries due to different proportions of responsive cortical neurons in dependence of the stimulus type are suggested, for example, for increasing vs. decreasing glides (Maiste and Picton, 1989) as well as for ramped vs. damped sinusoids (Lu et al., 2001).

One may argue that the relatively early negative ERP wave-form elicited by the 40 ST/s glide-to-constant transition in the complex tone reflects a higher-order auditory function, which detects a mismatch by comparing the sensory memory representation of previous stimulation (here: the glide segment) with the incoming auditory change (here: segment of constant frequency). Indeed, the memory-based mismatch negativity (MMN) ERP component (Näätänen et al., 1978, 2010) reflects violations of regularities extracted by the auditory system from previous regular stimulation. However, there are several arguments against MMN elicitation: (1) an infrequent but predictable, change in auditory stimulation only elicits MMN if the repeating pattern (here: e.g. sounds comprising both glide and constant-frequency segment) cannot be represented as a single unit (that is, if the elements of the pattern cannot be perceptually grouped; see e.g. Scherg et al., 1989; Sussman et al., 1998). When the stimulation parameters allow grouping in such a stimulation (Sussman et al., 1998) or when information on the repetitive pattern is available (Sussman et al., 2002) as it might be the case in the current approach, the potential regularity violations do not elicit an MMN). Thus, it is unlikely that the present outcome reflects the memory-based MMN. Also, (2) even the slowest glides in the current study are quite obviously recognizable as glides, and thus easily distinguishable from constant frequency segments (even though no behavioral data provided in the current article). Given the fact that the genuine MMN is not affected by the magnitude of the change if the change is highly conspicuous (Horváth et al., 2008) the putative MMN in the current approach should not only be manifested for the 40 ST/s rate in spectrally complex sounds but also for pure tones and slower rates. (3) Moreover, studies focusing on the processing of changes within long lasting sounds at different temporal positions by utilizing the MMN (e.g. Grimm and Schröger, 2005; Timm et al., 2011; Weise et al., 2010) reported that MMN decreased (before 350 ms) or even vanished (beyond 350 ms) with increased temporal distance of the change relative to sound onset. The fact that we obtained a response in the ERP to this late transition (at 900 ms) and that the temporal distance effect on MMN amplitude

is not known for N1 suggests that we measured rather N1 than MMN.

In summary, the N1-like activity to glide-to-constant transitions provides evidence for the existence of higher-order change detectors. As the transient ERPs, whose occurrence and extent are assumed to reflect the activity of dedicated change detectors, varied as a function of transition type and the magnitude of change, the present data speak in favor for the quantity-difference hypothesis. That is, the number of higher-order change detectors is markedly lower than that of the first-order change detectors. To engage a sufficient number of these detectors to produce an N1-like ERP activity to glide-to-constant transitions, the presentation of spectrally complex tones with high rates of frequency change was necessary. In contrast, constant-to-glide transitions in pure tones and more modest frequency change rates engaged a sufficient number of first-order change detectors to produce similar N1 potentials.

### Acknowledgements

This research was supported by the German Research Foundation (no. Schr 375/19-2), the European Community's Seventh Framework Programme (PERG04-GA-2008-239393), the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD, Project 50345549), the Hungarian Scholarship Board (Magyar Ösztöndíj Bizottság, MÖB – P/853), and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. It was realized by using Cogent 2000 developed by the Cogent 2000 team at the FIL and the ICN (<http://www.vislab.ucl.ac.uk>). The authors wish to thank Martin Reiche and Anne Steiger for assistance in data collection.

### References

- Arlinger, S.D., Jerlwall, L.B., Åhrén, T., Holmgren, E.C., 1976. Slow evoked cortical responses to linear frequency ramps of a continuous pure tone. *Acta Physiologica Scandinavica* 98 (4), 412–424.
- Arlinger, S., Elberling, C., Bak, C., Kofoed, B., Lebech, J., Saermark, K., 1982. Cortical magnetic-fields evoked by frequency glides of a continuous tone. *Electroencephalography and Clinical Neurophysiology* 54 (6), 642–653.
- Chait, M., Poeppel, D., Simon, J.Z., 2008. Auditory temporal edge detection in human auditory cortex. *Brain Research* 1213, 78–90.
- Chait, M., Poeppel, D., de Cheveigné, A., Simon, J.Z., 2007. Processing asymmetry of transitions between order and disorder in human auditory cortex. *Journal of Neuroscience* 27 (19), 5207–5214.
- Clynes, M., 1969. Dynamics of vertex evoked potentials: the R–M function. In: Donchin, E., Lindsley, L.B. (Eds.), *Average Evoked Potentials: Methods, Results, Evaluation*. NASA, pp. 363–374.
- Cone-Wesson, B., Wunderlich, J., 2003. Auditory evoked potentials from the cortex: audiology applications. *Current Opinion Otolaryngology & Head and Neck Surgery* 11 (5), 372–377.
- Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology* 115 (4), 732–744.
- Dimitrijevic, A., Loll, B., Michalewski, H.J., Pratt, H., Zeng, F.G., Starr, A., 2009. Intensity changes in a continuous tone: auditory cortical potentials comparison with frequency changes. *Clinical Neurophysiology* 120 (2), 374–383.
- Dimitrijevic, A., Michalewski, H.J., Zeng, F.G., Pratt, H., Starr, A., 2008. Frequency changes in a continuous tone: auditory cortical potentials. *Clinical Neurophysiology* 119 (9), 2111–2124.
- Fishbach, A., Nelken, I., Yeshurun, Y., 2001. Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. *Journal of Neurophysiology* 85 (6), 2303–2323.
- Grimm, S., Schröger, E., 2005. Pre-attentive and attentive processing of temporal and frequency characteristics within long sounds. *Brain Research. Cognitive Brain Research* 25 (3), 711–721.
- Harris, K.C., Mills, J.H., He, N.-J., Dubno, J.R., 2008. Age-related differences in sensitivity to small changes in frequency assessed with cortical evoked potentials. *Hearing Research* 243 (1–2), 47–56.
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., Winkler, I., 2008. MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology* 45 (1), 60–69.
- Hyde, M., 1997. The N1 response and its applications. *Audiology & Neuro-Otology* 2 (5), 281–307.
- Jones, S.J., Perez, N., 2001. The auditory 'C-process': analyzing the spectral envelope of complex sounds. *Clinical Neurophysiology* 112 (6), 965–975.
- Jones, S.J., Perez, N., 2002. The auditory C-process of spectral profile analysis. *Clinical Neurophysiology* 113 (10), 1558–1565.
- Kaukoranta, E., Hari, R., Lounasmaa, O.V., 1987. Responses of the human auditory cortex to vowel onset after fricative consonants. *Experimental Brain Research* 69 (1), 19–23.
- Kohn, M., Lifshitz, K., Litchfield, D., 1978. Averaged evoked potentials and frequency modulation. *Electroencephalography and Clinical Neurophysiology* 45 (2), 236–243.
- Kohn, M., Lifshitz, K., Litchfield, D., 1980. Average evoked potentials and amplitude modulation. *Electroencephalography and Clinical Neurophysiology* 50 (1–2), 134–140.
- Lu, T., Liang, L., Wang, X., 2001. Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *Journal of Neurophysiology* 85 (6), 2364–2380.
- Maiste, A., Picton, T., 1989. Human auditory evoked potentials to frequency-modulated tones. *Ear and Hearing* 10 (3), 153–160.
- Martin, B.A., Tremblay, K.L., Korczak, P., 2008. Speech evoked potentials: from the laboratory to the clinic. *Ear and Hearing* 29 (3), 285–313.
- Meyer, M., Baumann, S., Jancke, L., 2006. Electrical brain imaging reveals spatio-temporal dynamics of timbre perception in humans. *NeuroImage* 32 (4), 1510–1523.
- Moore, B.C.J., 2003. *An Introduction to the Psychology of Hearing*, 5th ed. Academic Press, San Diego.
- Näätänen, R., Gaillard, A.W., Mäntysalo, S., 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica* 42 (4), 313–329.
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24 (4), 375–425.
- Näätänen, R., Astikainen, P., Ruusuvirta, T., Huotilainen, M., 2010. Automatic auditory intelligence: an expression of the sensory-cognitive core of cognitive processes. *Brain Research Reviews* 64 (1), 123–136.
- Noda, K., Tonoike, M., Doi, K., Koizuka, I., Yoshida, H., Yamaguchi, M., et al., 1999. A pitch glide activates an intermediate response between auditory N1 and mismatch negativity. *Neuroreport* 10 (9), 1909–1912.
- Nuwer, M.R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.M., Hinrichs, H., et al., 1998. IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology* 106 (3), 259–261.
- Ostroff, J.M., Martin, B.A., Boothroyd, A., 1998. Cortical evoked response to acoustic change within a syllable. *Ear and Hearing* 19 (4), 290–297.
- Pantev, C., Eulitz, C., Hampson, S., Ross, B., Roberts, L.E., 1996. The auditory evoked off response: sources and comparison with the on and the sustained responses. *Ear and Hearing* 17 (3), 255–265.
- Phillips, D.P., Hall, S.E., Boehnke, S.E., 2002. Central auditory onset responses, and temporal asymmetries in auditory perception. *Hearing Research* 167 (1–2), 192–205.
- Phillips, D.P., Mendelson, J.R., Cynader, M.S., Douglas, R.M., 1985. Responses of single neurones in cat auditory cortex to time-varying stimuli: frequency-modulated tones of narrow excursion. *Experimental Brain Research* 58 (3), 443–454.
- Picton, T.W., Woods, D.L., Proulx, G.B., 1978a. Human auditory sustained potentials. I. The nature of the response. *Electroencephalography and Clinical Neurophysiology* 45 (2), 186–197.
- Picton, T.W., Woods, D.L., Proulx, G.B., 1978b. Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology* 45 (2), 198–210.
- Pratt, H., Starr, A., Michalewski, H.J., Dimitrijevic, A., Bleich, N., Mittelman, N., 2009. Auditory-evoked potentials to frequency increase and decrease of high- and low-frequency tones. *Clinical Neurophysiology* 120 (2), 360–373.
- Ross, B., Tremblay, K., 2009. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research* 248 (1–2), 48–59.
- Ruhm, H.B., 1970. Rate of frequency change and the acoustically evoked response. *The Journal of Auditory Research* 10, 29–34.
- Ruhm, H.B., 1971. Directional sensitivity and laterality of electroencephalic responses evoked by acoustic sweep frequencies. *Journal of Auditory Research* 11 (1), 9–16.
- Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience* 1 (4), 336–355.
- Schwartz, M., Tavano, A., Schröger, E., Kotz, S.A., 2012. Temporal aspects of prediction in audition: cortical and subcortical neural mechanisms. *International Journal of Psychophysiology* 83 (2), 200–207.
- Seither-Preisler, A., Krumbholz, K., Lütkenhöner, B., 2003. Sensitivity of the neuromagnetic N100m deflection to spectral bandwidth: a function of the auditory periphery? *Audiology & Neuro-Otology* 8 (6), 322–337.
- Shahin, A.J., Roberts, L.E., Miller, L.M., McDonald, K.L., Alain, C., 2007. Sensitivity of EEG and MEG to the N1 and P2 auditory evoked responses modulated by spectral complexity of sounds. *Brain Topography* 20 (2), 55–61.
- Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B., 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport* 16 (16), 1781–1785.
- Sussman, E., Ritter, W., Vaughan, H.G., 1998. Predictability of stimulus deviance and the mismatch negativity. *Neuroreport* 9 (18), 4167–4170.
- Sussman, E., Winkler, I., Huotilainen, M., Ritter, W., Näätänen, R., 2002. Top-down effects can modify the initially stimulus-driven auditory organization. *Brain Research. Cognitive Brain Research* 13 (3), 393–405.

- Timm, J., Weise, A., Grimm, S., Schröger, E., 2011. An asymmetry in the automatic detection of the presence or absence of a frequency modulation within a tone: a mismatch negativity study. *Frontiers in Psychology* 2, 189.
- Viswanathan, D., Jansen, B.H., 2010. The effect of stimulus expectancy on dishabituation of auditory evoked potentials. *International Journal of Psychophysiology* 78 (3), 251–256.
- Weise, A., Grimm, S., Müller, D., Schröger, E., 2010. A temporal constraint for automatic deviance detection and object formation: a mismatch negativity study. *Brain Research* 1331, 88–95.
- Whitfield, I.C., Evans, E.F., 1965. Responses of auditory cortical neurons to stimuli of changing frequency. *Journal of Neurophysiology* 28, 655–672.
- Winkler, I., Denham, S.L., Nelken, I., 2009. Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences* 13 (12), 532–540.
- Yingling, C.D., Nethercut, G.E., 1983. Evoked responses to frequency shifted tones: tonotopic and contextual determinants. *The International Journal of Neuroscience* 22 (1–2), 107–118.