This is the accepted version of
Horváth, J. (2013). Attenuation of auditory ERPs to action-sound coincidences is not explained by voluntary allocation of attention. *Psychophysiology, 50*(3), 266–273. doi:10.1111/psyp.12009
as originally submitted for publication to Psychophysiology. Wiley Periodicals, Inc.

The definite version is available at http://onlinelibrary.wiley.com/doi/10.1111/psyp.12009/abstract

Attenuation of auditory ERPs to action-sound coincidences is not explained by voluntary allocation of attention

János Horváth

Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences,

Hungarian Academy of Sciences, Budapest, Hungary

RUNNING HEAD: Action-sound coincidence-effect is not attentional

Address correspondence to:

János Horváth

Institute of Cognitive Neuroscience and Psychology
RCNS, Hungarian Academy of Sciences
P.O.B. 398, Szondi u 83/85
H-1394 Budapest
HUNGARY

Phone: +36 1 354 2290
Fax: +36 1 354 2416
E-mail: horvath.janos@ttk.mta.hu

29 pages (including 1 table and 5 figures)

Abstract

The event-related potential (ERP) correlates of sound detection are attenuated when eliciting sounds coincide with our own actions. The role of attention in this effect was investigated in two experiments by presenting tones separated by random intervals. In the Homogeneous Condition of Experiment 1 and in Experiment 2 the same tone was repeated, whereas in the Mixed Condition of Experiment 1 tones with five different frequencies were presented. Participants performed a time interval production task by marking intervals with key-presses in Experiment 1, and tried to produce key-press-tone coincidences in Experiment 2. Although the auditory ERPs were attenuated for coincidences, no modulation by the multiplicity of tone-frequencies in Experiment 1, or by the task-relevancy of tones and coincidences in Experiment 2 was found. This suggests that coincidence-related ERP attenuation cannot be fully explained by voluntary attentional mechanisms.

Introduction

Sensory processing is affected in various ways during goal-directed behavior. We direct our attention to goal-relevant stimuli while disregarding others, and we perform actions, thereby bringing new sensory events about as we interact with the environment. Research on the sensory processing of auditory events related to one's actions gained substantial momentum in recent years. Numerous studies have demonstrated that the event-related potential (ERP) and -field (ERF) correlates of sound detection (especially the N1, see Näätänen & Picton, 1987; and its magnetic counterpart - N1m) are attenuated when the eliciting sounds are generated (Numminen & Curio, 1999; Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000; Ford, Mathalon, Kalba, Whitfield, Faustman, Roth, 2001; Houde, Nagarajan, Sekihara & Merzenich, 2002; Ford & Mathalon, 2004; Heinks-Maldonado, Mathalon, Gray & Ford, 2005; Heinks-Maldonado, Nagarajan & Houde, 2006; Ventura, Nagarajan, Houde, 2009) or initiated by our own actions (Schäfer & Marcus, 1973; McCarthy & Donchin, 1976; Martikainen, Kaneko, & Hari, 2005; Ford, Gray, Faustman, Roach, & Mathalon, 2007; Baess, Jacobsen, & Schröger, 2008; Aliu, Houde, & Nagarajan, 2009; Baess, Horváth, Jacobsen, & Schröger, 2011; Knolle, Schröger, Baess & Kotz, 2012; Sowman, Kuusik, Johnson, 2012). The attenuated responses were assumed to reflect the cancellation of auditory re-afference, that is, a preparatory influence on the auditory system related to performing the action. The bases of such influences are so-called internal forward models, which represent action-stimulus contingencies, and allow the translation of outgoing motor commands into preparatory sensory signals (Holst & Mittelstaedt, 1950; Sperry, 1950; for a summary, see Miall & Wolpert, 1996). A set of studies, on the other hand, demonstrated that a contingent action-sound relationship is not necessary for the attenuation of auditory ERPs or ERFs; the coincidence of a tone and an action is sufficient (Hazemann, Audin, Lille, 1975; Horváth, Maess, Baess, Tóth, 2012). Because the temporal proximity of sounds and actions is

a common feature in stimulation protocols focusing on action-sound contingency and coincidence, these results open up the possibility that auditory ERP-attenuation for action-sound coincidences or self-initiated sounds is caused by other, not forward-model-based mechanisms (for a summary of factors influencing attenuation-effects for self-initiated sounds, see also Hughes, Desantis & Waszak, in press). The goal of the present study was to investigate the mechanisms behind the coincidence effect: Two alternative hypotheses based on well-known voluntary attentional effects on auditory processing were tested in variations of the coincidence paradigm introduced by Horváth et al. (2012).

Because the auditory N1 is enhanced when the eliciting sounds are attended (Hillyard, Hink, Schwent, Picton, 1973, Schwent, Hillyard, Galambos, 1976, Hansen, Hillyard, 1980, Rif, Hari, Hämäläinen, Sams, 1991; Okamoto et al., 2007), and attenuated when a sound-focused attention set is disrupted (Horváth & Winkler, 2010), Horváth et al. (2012) suggested that the coincidence effect may be caused by a dynamic change in the distribution of attentional resources. That is, the N1 and P2 difference between tones coinciding with an action and tones separated by longer intervals from preceding actions may be caused by a difference in the amount of attention allocated to these tones. In the experiments demonstrating the coincidence effect, tones were presented relatively infrequently (with 1.5-6 s interstimulus interval) and they were task-irrelevant. Participants made key-presses in a time-interval reproduction task, every 4 seconds on average. Due to the scarcity of stimulation, attending the tones may allow one to maintain vigilance in this situation, and performing the action may lead to a short-term redirection of attention from the tones to the key-press. This change in the distribution of attention is made visible by the impacted processing of tones presented in close temporal proximity to key-presses, which is manifested in attenuated N1 in comparison to tones separated by longer time intervals from the actions. That is, pressing the key “removes” the attentional N1 enhancement from tones in close

temporal proximity to the key-press, whereas this enhancement is unaffected by key-presses occurring temporally far away.

Evidence for the attentional enhancement of auditory N1 comes from paradigms in which participants directed attention towards the sounds because of their task-relevancy, that is, attending the sounds was voluntary. In such paradigms, it was found that this type of attentional effort influenced auditory processing and evoked auditory responses in two ways (Okamoto, Stracke, Wolters, Schmael, Pantev, 2007): by a "gain" and a "sharpening" mechanism. The "gain" mechanism generally improves processing when attention is directed towards an auditory channel, whereas the "sharpening" mechanism (demonstrated by Okamoto et al., 2007 for frequency) allows one to selectively enhance the processing of sounds with a given feature-level while suppressing others. In the present context, the hypothetical key-press-related attenuation of the auditory ERPs may reflect the workings of the "gain", as well as the "sharpening" functions: One may generally increase auditory "gain" to maintain vigilance in a situation which is impoverished in sensory input. Because in the studies of Hazemann, et al. (1975) and Horváth et al. (2012) the same tone was presented over and over again, one may also assume that through the tone repetition a selective attention set to the given tone (and tone frequency) might have been established ("sharpening").

To test the hypothetical involvement of the "sharpening" mechanism, tones with five different frequencies were presented in randomly mixed and homogenous sequences in a coincidence paradigm (Experiment 1). If attentional "sharpening" allowed by the constant repetition of a single frequency played a role in the coincidence-related attenuation of auditory ERPs, then attenuation should be absent or less in the mixed than in the homogeneous sequences.

In Experiment 2, the tones were made task-relevant by instructing participants to press the key so that key-presses coincided with tones as often as possible (a "coincidence game"), motivating them by a small bonus payment for each "hit". In this arrangement, attention should not be diverted from the tones by key-presses, since tones occurring at the time of the key-press have high task- (and motivational) relevance. If attention (either through a "gain" or a "sharpening" mechanism) played a role in coincidence-related ERP attenuation, then its magnitude should be smaller when the "coincidence game" was played in contrast to when a time-interval reproduction task was performed in which tones were task-irrelevant.

Methods

Participants

Experiment 1.

15 young adult volunteers (20-26 years, 23 years average age; five women, one left-handed) participated in Experiment 1 for monetary compensation. To motivate participants to attend both sessions, a bonus was paid for completing both sessions. Data from an additional participant was discarded because of the completion of only the first session of the experiment. All of them reported normal hearing status. Before the experimental sessions, participants gave written informed consent after the experimental procedures were explained to them.

Experiment 2.

14 volunteers from the group in Experiment 1 participated in Experiment 2 for monetary compensation. Data from one participant was discarded because of extensive eye-movement artifacts in the EEG. Before the experiment, participants gave written informed consent after the experimental procedures were explained to them.

Stimuli and procedures

Experiment 1.

Experiment 1 comprised two sessions, with 15 experimental blocks presented in each session. Blocks were separated by short breaks as needed, with a longer (~10 min) break around the middle of each session. The two sessions were administered within 5 weeks. During the experimental sessions, participants were sitting in a comfortable armchair in a well-lit, sound-proofed room. To minimize eye-movements, participants were instructed to rest their gaze at a fixation puppet placed in front of them during the experimental blocks.

Participants performed a time interval production task: they were required to press a rod-mounted key held in their dominant hand to mark time intervals, so that the histogram of the between-key-press-intervals would show a uniform distribution in the 2-6 s range, with no intervals outside this range. The experimental sessions started with a training phase in which participants learned the task with on-line visual feedback: the histogram of the produced intervals was updated on a screen in front of them after each key-press. During the experiment, this feedback was provided only at the end of each experimental block.

During task-performance, a sequence of tones was presented through headphones (HD-600, Sennheiser, Wedemark, Germany). The duration of the tones was 50 ms, including 10 ms rise and 10 ms linear fall times. Tone intensity was individually adjusted to 50 dB sensation level (above hearing threshold level). In each experimental block 75 tones were presented with onset-to-onset inter-stimulus intervals randomly sampled from a 2-6 s uniform distribution (the average block duration was 5 min). There were five tone frequencies: 988, 1109, 1245, 1397 and 1568 Hz. The choice of frequencies was based on the fact that Okamoto et al. (2007) found that the attentional “sharpening” effect on N1m significantly increased when the separation of frequencies interfering with a 1000 Hz tone decreased from ± 80 Hz to

± 40 Hz and lower separations. Also, the equivalent rectangular bandwidths of auditory filters is between 131 and 198 Hz for the given frequency range (Glasberg & Moore, 1990), so this range of frequencies should provide sufficient frequency-separation to make the hypothesized attentional "sharpening" effect visible. In the Mixed condition each tone was delivered 15 times in random order in each block. In the Homogenous condition, only one of the frequencies was presented in each block (75 times). There were 15 Mixed and 15 Homogenous (3 with each of the five frequencies) blocks presented in the two sessions. The blocks were presented in an interwoven order: "MHHMMHHM…" or "HMMHHMMH…" where "M"s denote Mixed, and "H"s denote Homogeneous blocks. The order of the frequencies between Homogenous blocks was random with the constraint that no frequency-repetitions between consecutive blocks were allowed.

The pre-generated, random tone presentation schedule was manipulated on-line during the experiment to produce key-press-tone coincidences (see Horváth et al., 2012). When a key-press occurred, the tone schedule was revised: tone presentation times following the key-press were shifted uniformly towards the moment of key-press so that the next tone to be delivered was delivered right away (if it was scheduled within 250 ms of the key-press) or after the integer multiple of 250 ms preceding its originally scheduled presentation time (e.g. if a tone was scheduled to be delivered 983 ms after the key-press, it was re-scheduled to 750 ms following the key-press). This adjustment was made only for the last key-press preceding a tone.

Due to the constraints of the stimulation equipment, for coincidences, there was 8 ± 2 ms (mean ± standard error) delay between key-presses and tone-presentations; also due to a programming error, for a number of coincidences (36 ± 6% of the coincidence trials in Experiment 1, and 39 ± 5% in Experiment 2) this delay was 18 ± 2 ms. These delays were taken into account in ERP processing (see below).

Experiment 2.

Experiment 2 was conducted within 4 months after Experiment 1 was completed. In Experiment 2, the stimulation arrangement was the same as in the Homogenous Condition of Experiment 1. The only difference was the task: In Experiment 2, participants performed a "coincidence game": they were instructed to press the key so that key-presses would coincide with ("hit") a tone as often as possible. For each coincidence, a small bonus was paid (which resulted in an accumulated premium of about +50-80% of the participation fee). To make the key-pressing activity similar to that in Experiment 1, we also introduced the following constraints in the instructions: a maximum of 75 key-presses could be made in a block, and key-presses must not have occurred closer than 2 s, or with a delay exceeding 10 s; also failure to comply with these constraints would result in the loss of the bonuses accumulated in the previous and the current block. Whereas there are some strategies which could help one to achieve a higher-than-random coincidence-rate (for example, not pressing the key if a tone was presented in the last two seconds), due the random nature of tone presentation, this is a game of chance.

EEG-recording and analysis

Experiment 1.

The electroencephalogram (EEG) was recorded by a Synamp2 amplifier (Compumedics Neuroscan, Victoria, Australia), with 1000 Hz sampling rate and on-line low-pass filtering of 200 Hz, from Ag/AgCl electrodes placed at the Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz. C4, T4, T5, P3, Pz, P4, T6, O1, O2 (10-20 system, Jasper, 1958) sites and the left and right mastoids (Lm, and Rm respectively). The reference electrode was placed at the tip of the nose. Horizontal electro-oculogram (EOG) was recorded between the outer canthi of

the two eyes by a bipolar setup, similarly to the vertical EOG, which was recorded from electrodes placed above and below the right eye.

The EEG was off-line low-pass filtered (30 Hz). Epochs corresponding to coincidences, and tones following key-presses by at least 1000 ms (1000 + ms post-key-press tones) were extracted. Because tones following key-presses were shifted towards the key-press in time, the inter-tone interval is shorter for these tones than for those preceded immediately by another tone (and not a key-press). Because inter-tone interval affects auditory ERPs (see Näätänen & Picton, 1987), only tone-events following key-presses were included in the analyses. To estimate the auditory activity for coincidences and the 1000+ ms post-key-press tones, the temporally corresponding key-press-related ERP was subtracted from them: For the coincidence ERP, epochs corresponding to key-presses without any other events within ±1 sec were extracted. For the 1000+ ms post-key-press tone ERP, epochs following key-presses by at least by 1000 ms with integer multiples of 250 ms with no actual events in them were extracted (1000+ ms post-key-press epochs). The variability in the key-press-tone delay for coincidences (as described above) was taken into account by shifting the key-press-related ERPs used for correction by 8 ms or 18 ms with random epoch selection in proportion to the occurrence of the two delays. That is, for example, if the key-press-tone delay for coincidences was 18 ms in 30% of the coincidences (and 8 ms for the rest), then 30% of the key-press-related epochs were shifted by 18 ms, the rest by 8 ms before averaging.

The duration of the extracted epochs was 700 ms, including a 200 ms pre-event interval. Amplitude calculations were referred to this pre-event interval. Epochs within the first 10 s of each block, as well as those with signal range exceeding 100 μV on any channel were discarded from the analyses.

Individual N1 and P2 amplitudes were measured as average signals in 20 ms long intervals centered at the group-average peak latencies of the action-corrected tone-related waveforms. The amplitudes at the Fz, Cz, Pz and Oz leads were submitted to repeated-measures analyses of variance (ANOVA) with Condition (Homogeneous vs. Mixed), Tone (corrected coincidence vs. 1000+ ms post-key-press tone) and Electrode (Fz, Cz, Pz, Oz) factors (Oz was calculated as the average signal of the O1 and O2 leads). For N1, a separate ANOVA for the average of the mastoid signals was also conducted with Condition and Tone factors. Partial eta-squared measures are reported. Greenhouse-Geisser corrections were calculated when appropriate; in such cases uncorrected degrees of freedom, ε-values and corrected p-values are reported. Interactions involving the two-level Tone factor were explored further through pair-wise Student's t-tests. All significant effects are reported.

Experiment 2.

In Experiment 2, EEG-recording and processing was identical to that in Experiment 1. N1 amplitude was analyzed in a repeated-measures ANOVA with Tone (corrected coincidence vs. corrected 1000 + ms post-key-press tone) and Electrode (Fz, Cz, Pz, Oz) factors. To compare N1-attenuation magnitudes between experiments, N1-amplitudes (elicited in the Homogeneous conditions) were normalized by dividing the individual amplitudes by the group mean N1-amplitude measured for the corrected 1000 + ms post-key-press tones in the same experiment. The normalized N1 amplitudes were submitted to a repeated-measures ANOVA with Experiment (1 or 2), Tone, and Electrode factors. In this analysis significant effects involving the Experiment and Tone factors were of interest. A further repeated-measures Experiment × Tone × Electrode ANOVA was calculated for the amplitudes measured in the time ranges of the P3b waveforms observable in Experiment 2.

Results

Behavioral measure

Participants complied with the instructions in both experiments (see Figure 1). The proportion of tones coinciding with key-presses was 5.0 ± 0.8 % in Experiment 1, and 5.5 ± 0.8 % in Experiment 2. This difference did not reach significance (paired t-test calculated for those participating in both experiments, t[12] = 2.09, p =.06).

ERPs

The group average ERPs recorded in the two experiments are presented in Figures 2, 3 and 4. Figure 2 shows that in both experiments a clear linear trend was present after more than 1000 ms following a key-press. The corrected 1000+ ms post-key-press tone waveform was obtained by subtracting this trend from the 1000+ ms post-key-press tone ERP. Figure 3 shows the derivation of the corrected coincidence waveform as the coincidence-minus-key-press-related waveform, and Figure 4 shows the corrected waveforms.

ERPs – Experiment 1.

In Experiment 1 the corrected 1000+ ms post-key-press tone-related N1 waveform peaked at 110 ms in the Mixed, and at 109 ms in the Homogeneous condition; the P2 peaked at 200 and 202 ms, respectively (Figure 4, left and center columns). The corrected coincidence waveform showed the same ERP pattern. The ANOVA of the N1 amplitudes at the midline showed a Tone main effect: $F(1,14) = 6.06$, $\eta^2_p = 0.30$, $p<.05$, indicating lower (less negative) amplitude for coincidences; and an Electrode main effect: $F(3,42) = 85.99$, $\varepsilon = 0.49$, $\eta^2_p = 0.86$, $p<.001$. The ANOVA of the N1 amplitudes at the mastoids showed no significant effects. Although no significant Tone × Electrode interaction was found, the topographical distributions (Figure 5, top and middle rows) show that the coincidence effect was numerically more posterior than the tone-related N1.

The ANOVA of the P2 amplitudes showed a Tone main effect: $F(1,14) = 35.99$, $\eta^2_p = 0.72$, $p<.001$, indicating lower (less positive) amplitudes for coincidences; an Electrode main effect: $F(3,42) = 12.32$, $\varepsilon = 0.46$, $\eta^2_p = 0.47$, $p<.01$; and a Tone × Electrode interaction: $F(3,42) = 31.21$, $\varepsilon = 0.50$, $\eta^2_p = 0.69$, $p<.001$. Pair-wise t-tests showed that the 1000+ ms post-key-press tone-minus-coincidence difference was larger at Cz, and smaller at Oz than at Fz or Pz (t-scores > 3.91, p-scores < .01).

ERPs – Experiment 2.

In Experiment 2, the corrected 1000+ ms post-key-press tone-related N1 peaked at 112 ms and was followed by a P2 at 218 ms (Figure 2, right column). In the corrected coincidence-waveform the N1 was followed by a partially overlapping, centrally peaking negativity (N2) at around 190 ms. Because P2 was overlapped by the N2, only N1 amplitudes were analyzed. The ANOVA of the N1 amplitudes at the midline showed a Tone main effect: $F(1,12) = 5.11$, $\eta^2_p = 0.30$, $p<.05$, indicating lower (less negative) amplitudes for coincidences; and an Electrode main effect: $F(3,36) = 61.75$, $\varepsilon = 0.56$, $\eta^2_p = 0.84$, $p<.001$. Although no significant Tone × Electrode interaction was found, the topographical distributions (Figure 5, bottom row) show that the coincidence effect was numerically more posterior than the tone-related N1. The ANOVA of the N1 amplitudes at the mastoids showed that the reversed N1 amplitude was higher (more positive) for coincidences: $F(1,12) = 4.94$, $\eta^2_p = 0.29$, $p<.05$.

ERPs – Between-experiment comparisons.

Since significant N1-attenuation was found in both experiments, the question in the following analysis of the normalized amplitudes was whether interactions between the Experiment and Tone factors indicating a difference in the magnitude (ratio) of N1-attenuation would be found. The ANOVA of the normalized N1 amplitudes showed a Tone

main effect: $F(1,12) = 6.96$, $\eta^2_p = 0.37$, $p<.05$, an Electrode main effect: $F(3,36) = 71.48$, $\varepsilon = 0.53$, $\eta^2_p = 0.86$, $p<.001$, and a Tone × Electrode interaction: $F(3,36) = 3.73$, $\varepsilon = 0.58$, $\eta^2_p = 0.24$, $p<.05$. Pair-wise t-tests showed that the normalized 1000+ ms post-key-press tone-minus-coincidence difference was larger at Pz and Oz than at Fz (t-scores > 2.21, p-scores < .05). The mean N1-attenuation-ratio was 33 ± 12 % and 35 ± 12 % in Experiment 1 and 2, respectively. This analysis confirms that the coincidence effect (as presented in Figure 5) is more posterior than the tone-related N1.

In contrast with Experiment 1, in Experiment 2, a parietally distributed positivity, identified as a P3b waveform (see e.g. Polich, 2007) was observable, peaking at 296 ms and 319 ms for the corrected 1000+ ms post-key-press tone and corrected coincidence waveform, respectively (Figure 4). The ANOVA of the amplitudes in the P3b intervals (see Table 1) showed a significant three-way interaction, which was followed up by separate ANOVAs in the two Experiments. In Experiment 1, only an Electrode main effect: $F(3,36) = 5.95$, $\varepsilon = 0.51$, $\eta^2_p = 0.33$, $p<.05$; in Experiment 2, however, a Tone main effect: $F(1,12) = 28.43$, $\eta^2_p = 0.70$, $p<.001$, showing higher (more positive) amplitudes for the coincidence; and an Electrode main effect: $F(3,36) = 8.69$, $\varepsilon = 0.46$, $\eta^2_p = 0.42$, $p<.01$ was found.

Discussion

The results showed attenuated auditory ERP responses to tone-action coincidences in comparison with tones separated by at least 1 s from preceding key-presses in both experiments, which confirms previous studies (Hazemann et al., 1975; Horváth et al., 2012). This attenuation, however, was not significantly modulated by the multiplicity of tone-frequencies in Experiment 1, or by the task-relevancy of tones and coincidences in Experiment 2. Although the present experimental design cannot prove the absence of

voluntary attentional effects, the contribution of these to the attenuation of the auditory N1 (and P2 in Exp. 1) is unlikely to be substantial (as suggested by Ford et al., 2001).

When tones were task-irrelevant (Experiment 1) only the N1 and P2 waveforms were elicited. In Experiment 2, tones elicited an additional P3b, attributable to the task-relevancy of the tones. For coincidences, the P3b was enhanced, and it was preceded by an N2 waveform, which probably reflects the significance of this event in task- and motivational terms (see e.g. Ritter, Simson, Vaughan, Friedman, 1979; Polich, 2007). The emergence of these ERPs, however, made it impossible to assess P2-modulation in Experiment 2.

Horváth et al. (2012) suggested a number of post-hoc hypotheses which may explain the coincidence-related ERP-attenuation. Because the present results suggest that the contribution of voluntary attention is not substantial, the attenuation-effect should be explained by other mechanisms. One hypothesis is that N1 (and possibly P2) attenuation is caused by processes initiating the formation of action-sound contingency-representations which are triggered when action-sound coincidences occur. Some of these hypothetical processes may be reflected in the present results (as well as in those of Horváth et al., 2012) by the topography of the coincidence-effect in the N1 time range, which showed that the effect was larger at posterior than at frontal sites (note that this difference was statistically significant only in the analysis of the normalized amplitudes involving data from both experiments). The coincidence-related enhancement of the positive aspect of N1 at the mastoids in Experiment 2, may also suggest that N1-attenuation does not reflect a "pure", genuine auditory N1-effect, but, in part, may result from the emergence of a different (positive) ERP component in the same time range as the N1. It has to be noted, however, that the enhancement at the mastoids may also signal a stronger attentional focus on the auditory channel at the moment of the keypress, which is offset by the coincidence-related activity at the midline sites.

A further hypothesis suggests that these attenuation-effects are rooted in a general "expectation" that our actions should cause a sensory event, and this "expectation" may be encoded in an internal forward model, which, however, does not represent the specific features of the contingent sensory event. Since the ERP attenuation was present even if the tones with different frequencies were presented in a random mixed order, the results of Experiment 1 are compatible with this hypothesis.

One outstanding issue in this line of research is the yet unknown relationship of ERP attenuation measured in paradigms utilizing action-contingent stimulation protocols and that measured in the coincidence-paradigm. Although the present study did not address this issue, the results suggest that the coincidence-paradigm may in part reflect a different set of effects from those obtained in studies focusing on self-induced sounds. Whereas an attenuation of the auditory N1 subcomponent (as evidenced by the attenuation of its magnetic reflection) takes place in both types of paradigms (see e.g. Martikainen et al., 2005 and Horváth et al., 2012), the posterior ERP topography in the N1 time range seems to be a distinct feature of the coincidence-related effect. Due to the nature of the paradigm, one may speculate that this may be an ERP signature of processes initiating the formation of action-sound contingency-representations or binding actions and sensory events together (see Hughes et al., in press).

In summary, mechanisms of voluntary attention cannot explain the action-sound coincidence-related attenuation of the auditory ERPs. Though the present results rule out some of the potential explanations, further research is required to understand the cause of the attenuation-effect.

Acknowledgements

This research was supported by the European Community's Seventh Framework Programme (under grant agreement PERG04-GA-2008-239393), and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. I thank Judit Roschéné Farkas and Annamária Tóth for collecting the data, and two anonymous reviewers for constructive comments on the manuscript.

References

Aliu, S. O., Houde, J.F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience, 21*, 791-802. doi: 10.1162/jocn.2009.21055

Baess, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology, 70*, 137-143. doi: 10.1016/j.ijpsycho.2008.06.005

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology, 48*, 1276-1283. doi: 10.1111/j.1469-8986.2011.01196.x

Curio, G., Neuloh, G., Numminen, J., Jousmäki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping, 9*, 183-91. doi: 10.1002/(SICI)1097-0193(200004)9:4<183::AID-HBM1>3.0.CO;2-Z

Ford, J.M., Gray, M., Faustman, W.O., Roach, B.J., & Mathalon, D.H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology, 44*, 522-529. doi: 10.1111/j.1469-8986.2007.00533.x

Ford, J.M., & Mathalon, D.H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. *Journal of Psychiatric Research, 38*, 37-46. doi: 10.1016/S0022-3956(03)00095-5

Ford, J. M., Mathalon, D. H., Kalba, S., Whitfield, S., Faustman, W.O., Roth, W.T. (2001). Cortical responsiveness during talking and listening in schizophrenia: an event-related

brain potential study. *Biological Psychiatry, 50*, 540-549. doi: 10.1016/S0006-3223(01)01166-0

Glasberg, B. R. & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*, 103-138. doi: 10.1016/0378-5955(90)90170-T

Heinks-Maldonado, T.H., Mathalon, D.H., Gray, M., & Ford, J.M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology, 42*, 180-90. doi: 10.1111/j.1469-8986.2005.00272.x

Heinks-Maldonado, T.H., Nagarajan, S.S., & Houde, J.F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport, 17*, 1375-9. doi: 10.1097/01.wnr.0000233102.43526.e9

Holst E. von, & Mittelstaedt H. (1950). Das Reafferenzprinzip. *Naturwissenschaften, 37*, 464-476. doi: 10.1007/BF00622503

Horváth, J., Maess, B., Baess, P., Tóth, A. (2012) Action-sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience 24*, 1919-1931. doi:10.1162/jocn_a_00215

Horváth, J., & Winkler, I. (2010). Distraction in a continuous-stimulation detection task. *Biological Psychology, 83*, 229-238. doi: 10.1016/j.biopsycho.2010.01.004

Houde, J.F., Nagarajan, S.S., Sekihara, K., & Merzenich, M.M. (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience, 14,* 1125-1138. doi:10.1162/089892902760807140

Hughes, G., Desantis, A., & Waszak, F. (in press). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological Bulletin*. doi:10.1037/a0028566

Hansen JC, Hillyard SA (1980): Endogenous brain potentials associated with selective auditory attention. *Electroencephalography and Clinical Neurophysiology, 49,* 277–290. doi:10.1016/0013-4694(80)90222-9.

Hazemann, P., Audin, G., & Lille, F. (1975). Effect of voluntary self-paced movements upon auditory and somatosensory evoked potentials in man. *Electroencephalography and Clinical Neurophysiology, 39*, 247-254. doi: 10.1016/0013-4694(75)90146-7

Hillyard, S.A., Hink, R.F., Schwent, V.L., & Picton, T.W. (1973). Electrical signs of selective attention in the human brain. *Science, 182*, 177-80. doi:10.1126/science.182.4108.177.

Jasper, H.H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10*, 371-375

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *Journal of Cognitive Neuroscience, 24*, 698-706. doi:10.1162/jocn_a_00167

Martikainen, M.H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex, 15*, 299-302. doi: 10.1093/cercor/bhh131

McCarthy, G., & Donchin, E. (1976). The effects of temporal and event uncertainty in determining the waveforms of the auditory event related potential (ERP). *Psychophysiology, 13*, 581-590. doi: 10.1111/j.1469-8986.1976.tb00885.x

Miall, R.C., & Wolpert, D.M. (1996). Forward models for physiological motor control. *Neural Networks, 9*, 1265-1279. doi: 10.1016/S0893-6080(96)00035-4

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*, 375-425. doi:10.1111/j.1469-8986.1987.tb00311.x.

Numminen, J., & Curio, G. (1999). Differential effects of overt, covert and replayed speech on vowel-evoked responses of the human auditory cortex. *Neuroscience Letters, 272*, 29-32. doi: 10.1016/S0304-3940(99)00573-X

Okamoto, H., Stracke, H., Wolters, C.H., Schmael, F., Pantev, C. (2007). Attention Improves Population-Level Frequency Tuning in Human Auditory Cortex. *The Journal of Neuroscience 27*(39), 10383–10390. doi: 10.1523/JNEUROSCI.2963-07.2007

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72(2)*, 184–187. doi:10.1016/0013-4694(89)90180-6

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1990). Corrigendum. *Electroencephalography and Clinical Neurophysiology, 76*, 565.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*, 2128‑2148. doi: 10.1016/j.clinph.2007.04.019

Rif, J., Hari, R., Hämäläinen, M.S., Sams, M. (1991). Auditory attention affects two different areas in the human supratemporal cortex. *Electroencephalography and Clinical Neurophysiology 79 (6)*, 464–472. doi:10.1016/0013-4694(91)90166-2.

Ritter, W., Simson, R., Vaughan, H. G. Jr., Friedman, D. (1979) A brain event related to the making of a sensory discrimination. *Science, 203*, 1358-1361. doi: 10.1126/science.424760

Schäfer, E.W., & Marcus, M.M. (1973). Self-stimulation alters human sensory brain responses. *Science, 181*, 175-177. doi: 10.1126/science.181.4095.175

Schwent, V.L., Hillyard, S.A., Galambos, R. (1976). Selective attention and the auditory vertex potential. II. Effects of signal intensity andmasking noise. *Electroencephalography and Clinical Neurophysiology 40 (6)*, 615–622. doi:10.1016/0013-4694(76)90136-X.

Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43*, 482-489. doi: 10.1037/h0055479

Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental Brain Research, 222(1-2)*, 149–157. doi:10.1007/s00221-012-3204-7

Ventura, M.I., Nagarajan, S.S., & Houde, J.F. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neuroscience, 10*, 58. doi:10.1186/1471-2202-10-58

Table

Effect	**Result**
Experiment	$F(1,12) = 47.80$, $\eta^2_p = 0.80$, $p<.001$
Tone	$F(1,12) = 5.52$, $\eta^2_p = 0.31$, $p<.05$
Electrode	$F(3,36) = 10.84$, $\varepsilon = 0.58$, $\eta^2_p = 0.47$, $p<.001$
Experiment × Tone	$F(1,12) = 36.79$, $\eta^2_p = 0.75$, $p<.001$
Experiment × Electrode	$F(3,36) = 5.51$, $\varepsilon = 0.44$, $\eta^2_p = 0.31$, $p<.05$
Experiment × Tone × Electrode	$F(3,36) = 4.52$, $\varepsilon = 0.44$, $\eta^2_p = 0.27$, $p<.05$

Table 1. The significant effects in the Experiment × Tone × Electrode ANOVA of the amplitudes in the P3b time intervals (20 ms long windows centered at the P3b peak amplitudes for the corrected 1000+ ms post-key-press tone-, and coincidence-related group-averaged ERP waveforms measured in Experiment 2).

Figure captions

Figure 1. Histograms of all inter-key-press intervals from all participants in the Homogeneous and Mixed Conditions of Experiment1, and in Experiment 2.

Figure 2. Group-average event related potentials elicited by tones separated by at least 1000 ms from preceding key-presses, and the corresponding average epochs with no actual events, and their differences (corrected 1000+ ms post-key-press tones), measured in the two experiments at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

Figure 3. Group-average event related potentials elicited by tone-key-press coincidences and the corresponding key-presses, and their differences (corrected coincidence), measured in the two experiments, at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

Figure 4. Group-average corrected coincidence and corrected 1000+ ms post-key-press tone ERP waveforms measured in the two experiments, at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

Figure 5. Group-average topographical distributions of the corrected 1000+ ms post-key-press tone ERPs (left column), and the corresponding coincidence-minus-tone differences (coincidence-effect, right column) in the N1 time interval in the Mixed (top row) and Homogeneous (middle row) conditions of Experiment1, and Experiment 2 (bottom row). The topographies were interpolated as described by Perrin, Pernier, Bertrand, & Echallier (1989, 1990), using a spline order of 4, a Legendre-polynomial order of 50, and no smoothing.

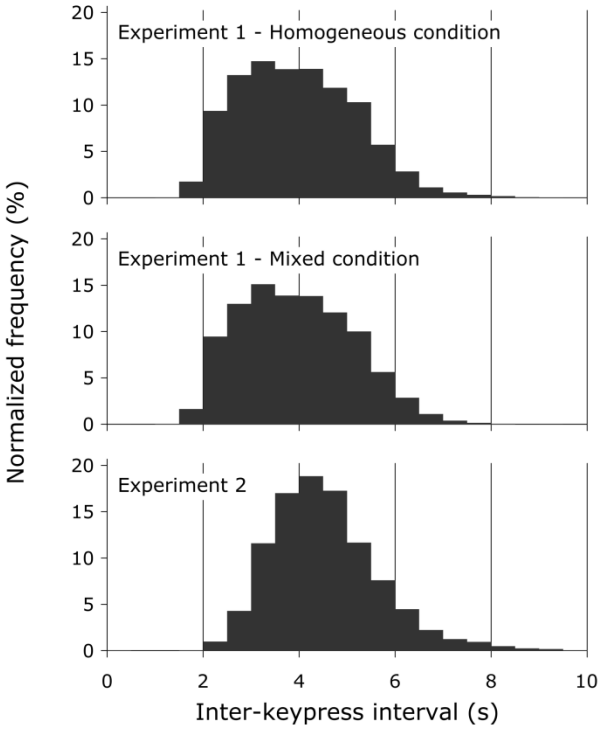

Figure 1. Histograms of all inter-key-press intervals from all participants in the Homogeneous and Mixed Conditions of Experiment1, and in Experiment 2.

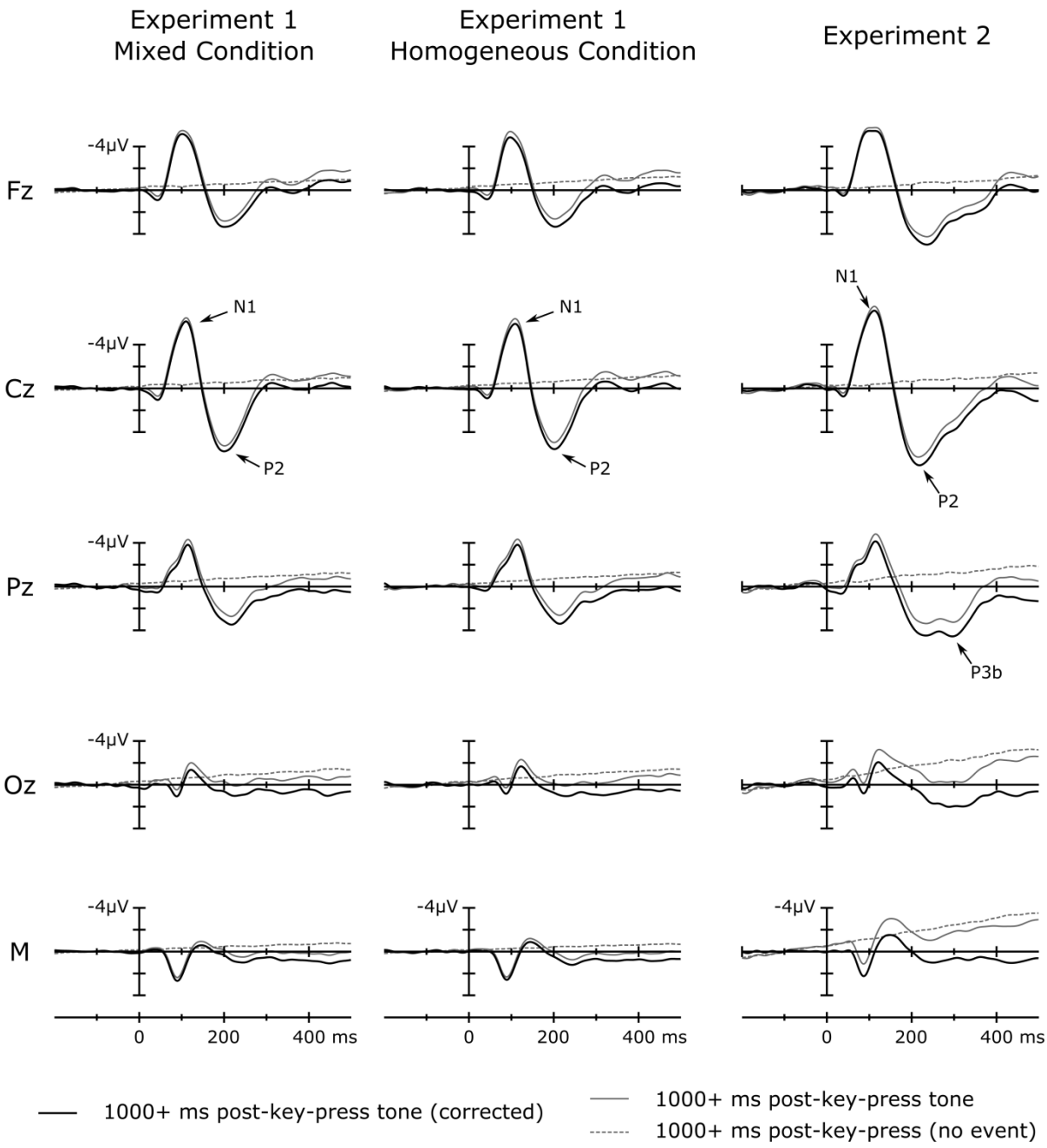

Figure 2. Group-average event related potentials elicited by tones separated by at least 1000 ms from preceding key-presses, and the corresponding average epochs with no actual events, and their differences (corrected 1000+ ms post-key-press tones), measured in the two experiments at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

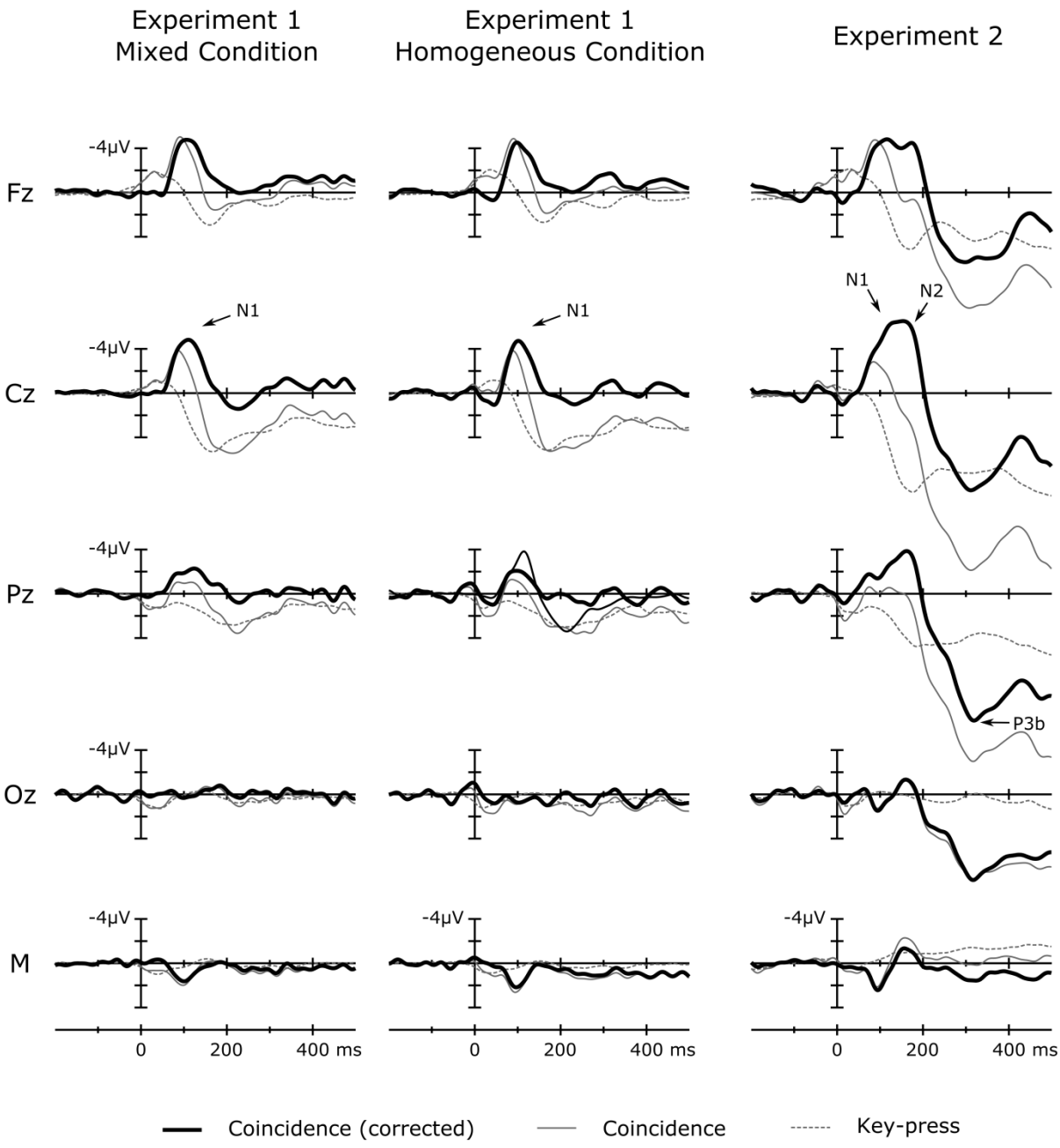

Figure 3. Group-average event related potentials elicited by tone-key-press coincidences and the corresponding key-presses, and their differences (corrected coincidence), measured in the two experiments, at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

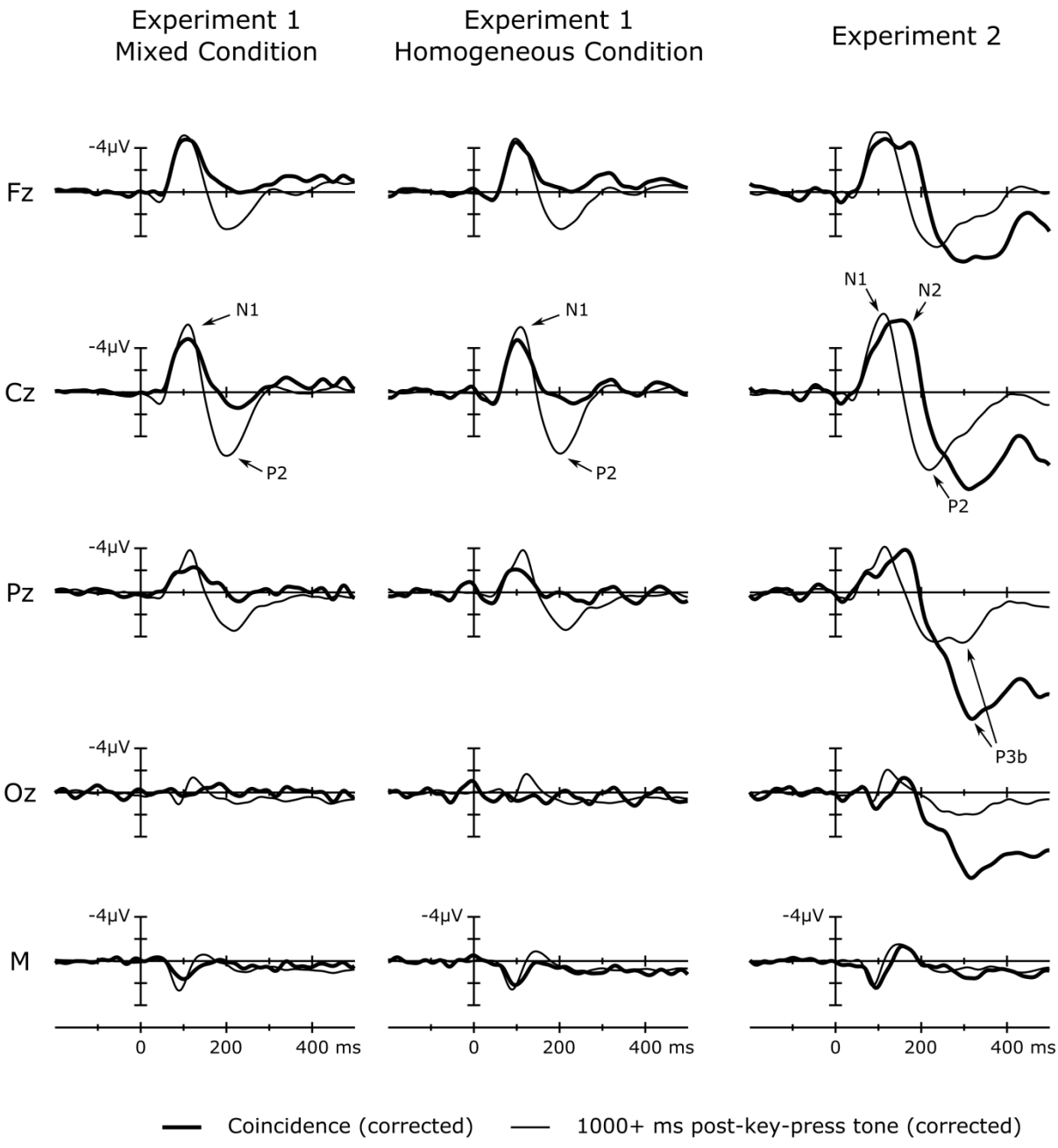

Figure 4. Group-average corrected coincidence and corrected 1000+ ms post-key-press tone ERP waveforms measured in the two experiments, at the Fz, Cz, Pz, the average of the O1 and O2 (Oz), and the average of the mastoid (M) signals. Tone onset is at the crossing of the axes.

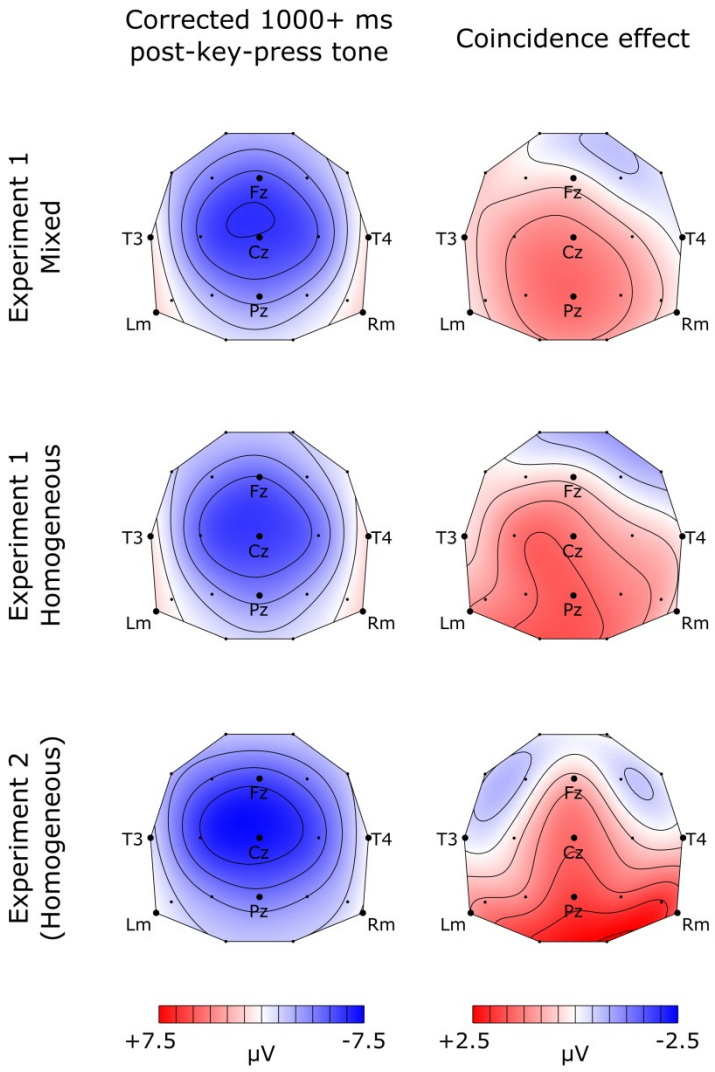

Figure 5. Group-average topographical distributions of the corrected 1000+ ms post-key-press tone ERPs (left column), and the corresponding coincidence-minus-tone differences (coincidence-effect, right column) in the N1 time interval in the Mixed (top row) and Homogeneous (middle row) conditions of Experiment1, and Experiment 2 (bottom row). The topographies were interpolated as described by Perrin, Pernier, Bertrand, & Echallier (1989, 1990), using a spline order of 4, a Legendre-polynomial order of 50, and no smoothing.