

Action–Sound Coincidences Suppress Evoked Responses of the Human Auditory Cortex in EEG and MEG

János Horváth¹, Burkhard Maess², Pamela Baess³,
and Annamária Tóth⁴

Abstract

■ The N1 auditory ERP and its magnetic counterpart (N1[m]) are suppressed when elicited by self-induced sounds. Because the N1(m) is a correlate of auditory event detection, this N1 suppression effect is generally interpreted as a reflection of the workings of an internal forward model: The forward model captures the contingency (causal relationship) between the action and the sound, and this is used to cancel the predictable sensory refference when the action is initiated. In this study, we demonstrated in three experiments using a novel coincidence paradigm that

actual contingency between actions and sounds is not a necessary condition for N1 suppression. Participants performed time interval production tasks: They pressed a key to set the boundaries of time intervals. Concurrently, but independently of keypresses, a sequence of pure tones with random onset-to-onset intervals was presented. Tones coinciding with keypresses elicited suppressed N1(m) and P2(m), suggesting that action–stimulus contiguity (temporal proximity) is sufficient to suppress sensory processing related to the detection of auditory events. ■

INTRODUCTION

Goal-directed behavior is impossible without knowing the consequences of our actions. Our actions have a number of simple and immediate as well as complex and far-reaching consequences, which can be represented in different ways. The most fundamental of these representations connect actions to their immediate sensory consequences. Many studies suggest that causal relationships between self-produced movements and their sensory effects are represented by internal forward models (Miall & Wolpert, 1996). When engaging in an action, a copy of the outgoing motor commands (the efference copy; von Holst & Mittelstaedt, 1950) is produced, which is translated by an internal forward model into a special sensory signal representing the associated consequences (a corollary discharge; Sperry, 1950). The corollary discharge is special because it does not only allow one to compare the predicted sensory consequences of the action with the actual sensory input after the action took place (feedback), but it can also be used in parallel with the action to adjust sensory processing so that it can accommodate (some of the) predictably occurring sensory events because of the action itself. Forward modeling supports many functions of the neural system, and it has been recently suggested that it may play an important role in auditory perception as well. The goal of this study was to test whether characteristic auditory processing

changes related to combined action–sound events necessitate the assumption that a causal action–sound relationship is represented by an internal forward model.

There is strong evidence that forward modeling of the sensory consequences of self-produced movements plays an important role in sensorimotor integration (Wolpert, Ghahramani, & Jordan, 1995). Besides the sensory input, forward modeling adds a source of information that can be used to improve movement performance (Shadmehr, Smith, & Krakauer, 2010; Vaziri, Diedrichsen, & Shadmehr, 2006). Forward models also make it possible to cancel refference, that is, the stimulation inherently resulting from the action due to mechanics of the actor's own body.

Whereas it seems plausible that proprioceptive refference originating from the moving body parts is represented by forward models, predictions provided by internal forward models are also used by various cognitive subsystems beyond those directly involved in the control of the given effector (Davidson & Wolpert, 2005). For example, ticklishness on the palm (as well as the concurrent activation in somatosensory cortex) is reduced when the stimulation is self-produced (e.g., Blakemore, Wolpert, & Frith, 1998). Active, voluntary head movements lead to the cancellation of refference in the vestibular system (Cullen, 2004; Roy & Cullen, 2004). Forward modeling supports the stabilization of the visual field despite eye movements (Duhamel, Colby, & Goldberg, 1992); and arm movements influence the saccadic eye-movement system (Thura, Hadj-Bouziane, Meunier, & Boussaoud, 2011; Ariff, Donchin, Nanayakkara, & Shadmehr, 2002) as well as motor imagery (Gentili, Cahouet, Ballay, & Papaxanthis, 2004).

¹Hungarian Academy of Sciences, ²Max-Planck Institute for Cognitive and Brain Sciences, Leipzig, Germany, ³Aalto University, School of Science and Technology, Espoo, Finland, ⁴Eötvös Loránd University, Budapest, Hungary

Because forward models capture causal (contingent) action effect mappings connecting different cognitive sub-systems, it seems plausible that forward modeling may also play a role in most functions where our own controlled actions produce consistent patterns of stimulation. Research on internal forward modeling in the auditory modality has recently gained substantial momentum based on the finding that relatively high-level neural correlates of auditory event detection, most prominently the N1 auditory ERP (see Näätänen & Picton, 1987) and its event-related magnetic field (ERF) counterpart (N1[m]), are suppressed when the eliciting sounds are self-produced or self-initiated. Because N1(m) reflects the detection of auditory events (sound onsets or first-order changes in sound parameters), it is generally assumed that N1 suppression reflects the cancellation of auditory refference.

A number of studies suggest that the N1(m) response is suppressed for own vocalizations of speech sounds in comparison with when replayed vocalizations are listened to (e.g., Ventura, Nagarajan, & Houde, 2009; Heinks-Maldonado, Nagarajan, & Houde, 2006; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Curio, Neuloh, Numminen, Jousmäki, & Hari, 2000; Numminen & Curio, 1999). Because we have extensive experience with the control of and sensory stimulation produced by our own speech production system, the notion that speech-related N1 suppression reflects the workings of an internal forward model seems plausible. Some studies, however, suggest that N1 suppression may not be constrained to speech production: Non-speech-related actions may also lead to the suppression of the N1 response elicited by concurrent (speech or nonspeech) sounds (Baess, Horváth, Jacobsen, & Schröger, 2011; Aliu, Houde, & Nagarajan, 2009; Baess, Jacobsen, & Schröger, 2008; Ford, Gray, Faustman, Roach, & Mathalon, 2007; Martikainen, Kaneko, & Hari, 2005; McCarthy & Donchin, 1976; Schäfer & Marcus, 1973).

The core assumption of these studies is that capturing a contingent action–stimulus relationship occurs rapidly, at least within the order of minutes (Aliu et al., 2009), and the resulting forward model is then used to derive predictive sensory information, which is manifested in the suppression of the N1 response. The goal of this study is to investigate whether N1 suppression can be explained by less complex, more economic assumptions without relying on a hypothetical internal forward model representing an action–sound contingency.

The studies interpreting auditory N1 suppression in the framework of internal forward modeling exclusively used contingent stimulation: Actions always brought about a sound event. In these studies, action-contingent stimulation also involved a consistent temporal relationship between action and stimulus (i.e., stimuli were delivered at least within a couple hundred of milliseconds after the action). Therefore, it seems possible that the necessary condition for auditory N1 suppression is not contingency but temporal contiguity, that is, the temporal proximity

of an action and a sound. That auditory processing may be affected by concurrent but not causally related motor activity is not without support: Makeig, Müller, and Rockstroh (1996) found that the amplitude and phase of the auditory steady-state response in the EEG was perturbed by concurrent, voluntary finger movements. Hazemann, Audin, and Lille (1975) presented a sound sequence with random ISIs and instructed participants to produce an even-paced keypress sequence. They found that the amplitude of the N1 and P2 ERP waveforms elicited by sounds close to keypresses was smaller than for sounds far from keypresses. Whereas Hazemann and colleagues did not directly remove keypress-related ERPs from the sound-locked waveform (see Methods below), the contributions of these ERPs to the N1 and P2 effects were probably low because of the randomness of keypress–stimulus separation (in a range of 0–220 msec).

The goal of this study was to investigate whether keypress–tone contiguity without a keypress–tone contingency was sufficient to produce an N1 suppression effect. We utilized a coincidence paradigm: Participants pressed a button to set boundaries in a time interval production task, while a concurrent but temporally independent sound sequence was presented with random intersound intervals. The coincidence paradigm has a number of methodological advantages in comparison with previously used paradigms to measure N1 suppression, which are presented in detail in the Methods section (below). The main question was whether keypress–sound coincidences resulted in suppressed auditory processing as reflected by the N1(m) event-related response. Because there was no contingent action–tone relationship in this setting, a potential N1 suppression effect could not be attributed to the cognitive system capturing a causal action–tone relationship in the form of a forward model. We recorded EEG in two experiments (Experiments 1 and 2) with different interval production instructions and recorded magnetoencephalogram (MEG) in Experiment 3 (with the same experimental setting as in Experiment 1).

The goal of recording MEG in Experiment 3 was to assess whether the observed N1 ERP attenuation was the result of attenuated auditory processing activity. This was necessary because the scalp-recorded, fronto-centrally negative N1 ERP waveform elicited by sounds is the sum of at least two subcomponents. One of these is stimulation-specific and is generated by tangentially oriented dipoles in each superior temporal lobe. The other is a nonspecific component with unknown origin (Näätänen & Picton, 1987). Whereas the ERP sums these components, the ERF reflects the activity of the supratemporal generator (N1m; Näätänen, 1988) because of the “insensitivity” of MEG to nontangentially oriented sources (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). Therefore, ERF is highly useful in assessing the contribution of the supratemporal (auditory) component to N1 effects. In some cases, ERPs may also allow conclusions regarding the involvement of the supratemporal generators, because the N1 elicited by sinusoid

tones often exhibits a polarity reversal at the mastoids when the EEG is recorded with a nose reference (Vaughan & Ritter, 1970). An N1 effect showing such a polarity reversal signals that the effect (at least in part) originates from the supratemporal generator. The lack of a polarity reversal, on the other hand, does not mean that the supratemporal component is not affected because it may simply be overlapped by the nonspecific component or other ERPs.

Because N1 suppression was followed by the suppression of the P2 ERP in the study by Hazemann et al. (1975), it was also investigated whether a P2 suppression occurred as well. Note that studies using contingent stimulation often did not explicitly investigate ERP amplitudes in the P2 time interval, but P2 suppression effects could be seen on the ERP figures (see, e.g., Baess et al., 2011; Ford & Mathalon, 2004; Schäfer & Marcus, 1973). The present setup also allowed the assessment of whether N1 and P2 were affected similarly by the experimental manipulation or not.

METHODS

General Considerations

The measurement of auditory processing activity in the presence of concurrent action-related activity is not trivial. Whereas neural responses to action–sound and sound-alone events may be compared directly, nonauditory contributions from the response to the action–sound event may contaminate the results. One approach to eliminate such confounds is to estimate the action-related contribution to the action–sound processing response, subtract it, and compare the result to that elicited by sound-alone events. Studies taking this approach exclusively use the responses to action-alone events recorded in separate experimental blocks to estimate the action-related contribution. Whereas this approach is probably less prone to confounds than the direct comparison approach, it is also based on the assumption that action-related neural responses do not differ when the action occurs on its own and when it occurs with a contingent stimulus event. Whereas this assumption might be correct, it is also obvious that the participant's cognitive state differs in the two conditions (e.g., because the overall level of stimulation or expectations regarding the sensory consequences of the keypress differs), which may lead to differences in the actions themselves, and the action-related response contributions as well, which, in turn, may lead to biased estimates. In contrast to previous studies, the coincidence paradigm allowed the derivation of an unbiased estimate for the sensory processing activity because all relevant events (keypress–sound coincidence and sound-alone and keypress-alone events) occurred within the same experimental condition.

Participants

Fourteen paid volunteers (six women, age = 19–24 years, two left-handed) participated in Experiment 1; 13 (eight

women, age = 19–24 years, three left-handed), in Experiment 2; and 20 (10 women, age = 23–31 years, all right handed), in Experiment 3. In all three experiments, participants gave written informed consent after the experimental procedures were explained to them. All participants reported normal hearing status and had no history of neurological disorders.

Stimuli and Procedure

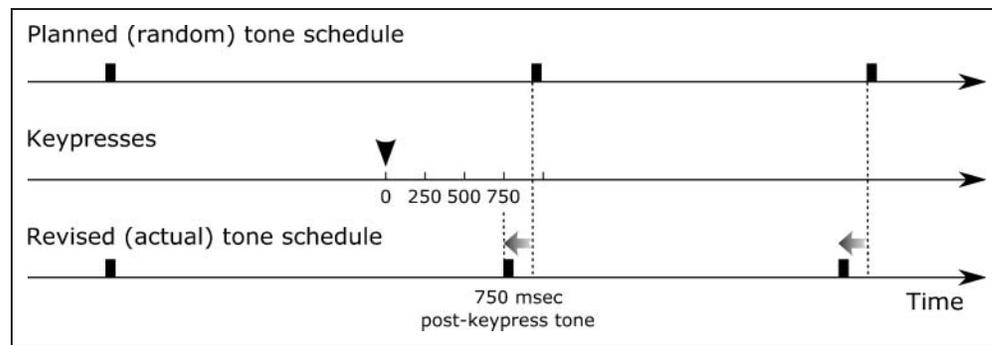
In all experiments participants performed time interval production tasks. In Experiments 1 and 3, participants were instructed to produce a sequence of keypresses in which between-keypress intervals showed a uniform distribution between 2 and 6 sec within each 5-min-long experimental block. In Experiment 2, on the other hand, a regular, even-paced sequence with a keypress every 4 sec was required. In all three experiments, the experimental session started with a training phase, during which participants performed the task with on-line visual feedback: A computer screen showed a histogram of their between-keypress intervals, which was updated after each keypress. During the experimental phase, feedback was provided only at the end of each experimental block.

In Experiments 1 and 2, each participant held a rod-mounted key in their dominant hand and pressed the key with the thumb; in Experiment 3, the key was mounted on a response box, which was positioned under the dominant hand, and participants used their index finger to press the key.

In all three experiments, a series of 50-msec-long (including 10-msec linear rise and 10-msec linear fall times), 1000-Hz sinusoid tones was presented. Tone intensity was individually adjusted to 60 dB SL (sensation level; above hearing level) in Experiments 1 and 3 and to 50 dB SL in Experiment 2. The tones were delivered through headphones (HD-600, Sennheiser, Wedemark, Germany) in Experiments 1 and 2 and through tubal insert phones (TIP-300, Nicolet Biomedical, Madison, WI) in Experiment 3.

In all three experiments, the schedule of tone presentation was pregenerated for each participant so that the onset-to-onset intervals were random in the range of 2–6 sec (with uniform distribution). The experiments were divided into 14 experimental blocks with 72 tones presented in each of them (1008 tones in total). Between blocks, short breaks were taken as needed, with a longer break around the middle of the session (after the seventh block). Keypress–tone coincidences were created through the following manipulation (see Figure 1): At every keypress, the preplanned tone presentation schedule was revised. The schedule was shifted so that the next tone was presented either right after the keypress (0 msec) or with a delay of integer multiples of 250 msec. That is, if the next tone was to be presented between 0 and 249 msec following the keypress, it was presented right away; if it was to be presented in 250–499 msec, it was presented 250 msec after the keypress, and so forth. If there were more keypresses before

Figure 1. Example for the manipulation of the tone presentation in the experiments. When a keypress occurred, the scheduled sounds were shifted toward the keypress so that the next tone was delivered with a delay of integer multiples of 250 msec. In this example, the next tone would be delivered between 750 and 1000 msec after the keypress according to the planned schedule. This is revised by shifting the tone schedule following the keypress so that the next tone is delivered exactly 750 msec after the keypress (750-msec post-keypress tone).



a tone, the manipulation was carried out in reference to the last keypress only. The result of this manipulation was that all tones preceded immediately by a keypress were shifted similarly, ensuring that the distribution of the intervals separating these tones from the previously presented tones was the same. In contrast, sounds not preceded by a keypress were not shifted at all; therefore, the tone-to-tone intervals preceding these tones were longer. Because N1 amplitude is known to increase with increasing tone-to-tone interval (Näätänen & Picton, 1987), a comparison between shifted and unshifted tones would be confounded by the systematic tone-to-tone interval differences. Therefore, only (shifted) sounds preceded by at least a keypress were included in the analyses.

For coincidences (i.e., when a tone was presented right after a keypress), there was a short delay between the keypresses and the tone because of the necessary processing time after keypresses: This was 4.3 ± 0.1 msec (mean \pm *SD*) in Experiments 1 and 2 and 9.3 ± 0.1 msec in Experiment 3. These delays were taken into account during the analyses, but for convenience, in the following, we will not include these short delays into the references to keypress-tone intervals and refer to the corresponding events only as coincidences, 250-msec post-keypress tones, 500-msec post-keypress tones, and so on. Also, because of a programming error, for some of the coincidences, a further, additional delay occurred: In Experiment 1, this additional delay was 9.1 ± 0.1 msec and affected $33\% \pm 8\%$ of the cases; in Experiment 2, it was 9.0 ± 0.1 msec and affected $44\% \pm 7\%$ of the cases; in Experiment 3 this was 8.4 ± 0.1 msec and affected $31\% \pm 7\%$ of the cases. Coincidences with such unwanted, additional tone delays were discarded from the event-related response analyses.

EEG Recording and Processing—Experiments 1 and 2

In Experiments 1 and 2, participants sat in a comfortable chair in a noise-attenuated chamber. The EEG was re-

corded with a Synamp2 amplifier (Compumedics Neuroscan, Victoria, Australia) 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, and O2 (10–20 system; Jasper, 1958) sites and the left and right mastoids. Because the auditory N1 elicited by pure tones often shows a polarity inversion between electrodes placed at the two sides of the Sylvian fissure when a nose reference is used (Vaughan & Ritter, 1970), the reference electrode was placed on the tip of the nose. The horizontal EOG was recorded by a bipolar electrode setup placed near the outer canthi of the two eyes; the vertical EOG was recorded by electrodes placed above and below the right eye. Sampling rate was 1000 Hz, and on-line low-pass filtering of 200 Hz was used. The continuous recording was band-pass filtered off-line (0.1–20 Hz). Epochs of 600-msec duration, including a 200-msec pre-event interval, were extracted for various events described below in the ERP and ERF Analyses section. Epochs with a signal range exceeding 100 μ V on any channel were rejected from further processing.

MEG Recording and Processing—Experiment 3

In Experiment 3, the MEG was recorded with a 306-channel (204 orthogonal planar gradiometers and 102 magnetometers at 102 locations) Neuromag Vectorview MEG system (Elekta Oy, Helsinki, Finland) in supine position in an electromagnetically shielded room (Vacuumschmelze, Hanau, Germany). Horizontal and vertical EOG was recorded with a bipolar electrode setup from the outer canthi of the two eyes and from above and below the left eye, respectively. Sampling rate was 1000 Hz, and an on-line low-pass filtering at 330 Hz was used. Five head-position indicator coils (three positioned at forehead and two behind the ears) were used to continuously monitor head movements. The analysis of the head positions was based on 500-msec windows of data shifted by 250-msec intervals. Head movements were always less than or equal to 4 mm for all participants; therefore, we did not apply head

movement correction. The signal space separation method (Taulu, Kajola, & Simola, 2004) was used for external interference suppression, for the interpolation of bad channels, and to recompute the MEG data for an identical head position across all blocks. The continuous recording was off-line band-pass filtered (0.8–16 Hz). Epochs of 350-msec duration, including a 100-msec pre-event interval, were extracted, corresponding to events described below in the ERP and ERF Analyses section. Epochs with a signal range exceeding 200 pT/m (gradiometer), 4 pT (magnetometer), or 80 μ V (horizontal and vertical EOG) were excluded from the analyses.

ERP and ERF Analyses

In all three experiments, epochs corresponding to the following events were extracted from the EEG or MEG recordings: coincidences (i.e., tones immediately presented after a keypress); 250-msec post-keypress tones (i.e., tones following keypresses by 250 msec), 500-msec post-keypresses tones (i.e., tones following keypresses by 500 msec), 750-msec post-keypress tones (i.e., tones following keypresses by 750 msec), and 1000+ msec post-keypress tones (i.e., tones following keypresses by at least 1000 msec). For all these events, no other event occurred between the initial keypress and the last sampling point of the epoch.

To subtract the contribution of the action (motor)-related activity from the tone-locked event-related responses (see below), a number of keypress-locked responses were also extracted: epochs corresponding to keypresses at least 1 sec away from any other event and epochs following such keypresses at integer multiples of 250 msec with no actual events in them (250, 500, 750, and 1000+ msec post-keypress epochs). The zero time points of these post-keypress epochs were at 250, 500, 750, and at least 1000 msec following the keypress.

To estimate the auditory activity for coincidences, the event-related response elicited by keypresses was subtracted from that elicited by coincidences (in the following, this is referred to as corrected coincidence tone response). To estimate the auditory activity in responses to 250, 500, 750, and 1000+ msec post-keypress tones, the corresponding post-keypress epochs were subtracted, respectively. These are termed corrected 250, 500, 750, and 1000+ msec post-keypress tone responses, respectively.

Two lines of analyses were conducted: In the first analysis, the corrected coincidence tone response was compared with the corrected 1000+ msec post-keypress tone response. Second, a trend analysis was calculated for the corrected coincidence, 250, 500, 750, and 1000+ msec post-keypress tone responses.

For the ERPs, individual N1 and P2 amplitudes were measured as the average signal in 20-msec-long intervals centered at the group-average peak latency of the waveform. The amplitudes at the F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 leads were submitted to repeated-measures ANOVAs with Sound, Laterality (3, z, and 4), and Anterior–

Posterior (AntPos; F, C, and P) factors. 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 Stimulus factor were explored further through pairwise Student's t tests. For the N1 amplitudes, a further Sound \times Laterality (left, right) repeated-measures ANOVA was calculated for the mastoid signals as well. To assess whether the N1 and P2 amplitudes were different functions of keypress–tone separations, a further Component (N1, P2) \times Keypress–tone separation (0, 250, 500, 750, 1000+ msec) ANOVA was conducted over the amplitudes measured at Cz. Because, in this analysis, the shapes of the functions were investigated, the amplitudes were z-transformed (producing distributions with a mean of 0 and an SD of 1), separately for each component, each pooling the amplitudes for different keypress–tone separations. Whereas this transformation eliminated the Component main effect, the keypress–tone separation effect was preserved for both components. In this analysis, the difference between the shapes of the amplitude functions is signaled by a significant Component \times Keypress–tone separation interaction.

For the N1m ERFs, the measured variable was the source strength of single dipoles individually fitted in each hemisphere (as described below). Source strengths were analyzed through pairwise Student's t tests and Sound \times Hemisphere repeated measures ANOVAs. Because the dipole fitting approach was not successful for the P2m, root-mean-squared (RMS) ERF amplitudes (calculated over all magnetometers in a 20-msec interval centered at the group-average N1m and P2m latencies) were also analyzed in pairwise Student's t tests and trend analyses. The similarity of the N1m and P2m RMS amplitudes as functions of keypress–tone separations were assessed as in the analysis of the ERPs described above.

RESULTS

Participants complied with the instructions in all three experiments (Figure 2). Mean coincidence rates (i.e., the proportion of tones coinciding with a keypress [out of the 1008 tones delivered in total]; with standard deviations) were $5\% \pm 1\%$ in Experiment 1, $6\% \pm 1\%$ in Experiment 2, and $5\% \pm 1\%$ in Experiment 3.

ERPs

In Experiments 1 and 2, tones elicited a clear succession of N1 and P2 waveforms (Figure 3).

Experiment 1

N1

The N1 peaked at 98 msec in the corrected coincidence ERP and at 101 msec in the corrected 1000+ msec post-

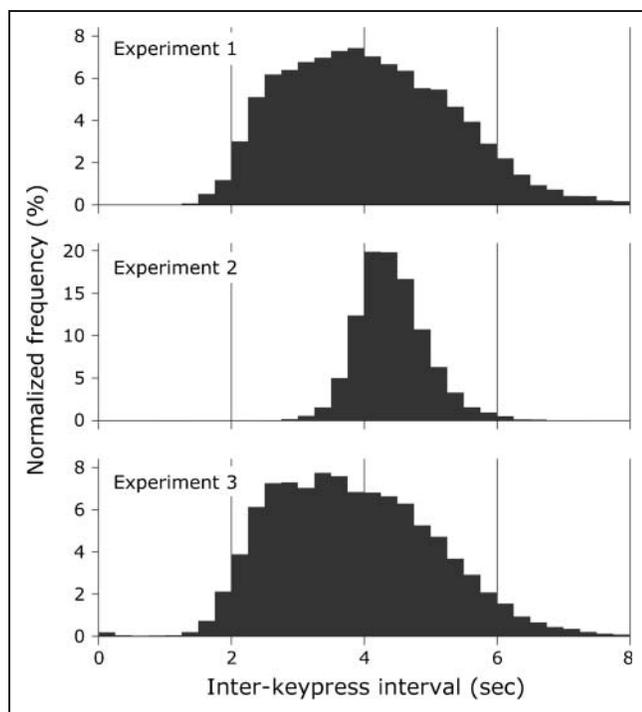


Figure 2. Histograms of all inter-keypress intervals from all participants in the three experiments.

keypress tone ERP at Cz in the group-average waveform. The ANOVA of the N1 amplitudes measured in the range of 89–109 msec showed a Stimulus main effect, $F(1, 13) = 4.73, p < .05$, indicating a lower (less negative) N1 amplitude for the coincidence; an AntPos main effect, $F(2, 26) = 32.24, \epsilon = 0.80, p < .001$; a Laterality main effect, $F(2, 26) = 24.56, \epsilon = 0.88, p < .001$; Stimulus \times AntPos interaction, $F(2, 26) = 8.42, \epsilon = 0.65, p < .01$; and a Stimulus \times Laterality interaction, $F(2, 26) = 6.20, \epsilon = 0.97, p < .01$. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes at the F, C, and P leads (averaged over the Laterality levels) indicated that the N1 suppression effect was stronger at central and parietal than at frontal sites [$t(13) > 2.97, p < .05$]. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes on the left, middle, and right leads (averaged over the levels of the AntPos factor) indicated that the N1 suppression effect was stronger at central than at lateral sites [$t(13) > 2.62, p < .05$].

At the mastoids, a significant Stimulus main effect was found, $F(1, 13) = 8.74, p < .05$, showing that the corrected coincidence ERP was more positive than the corrected 1000+ msec post-keypress tone ERP (i.e., for the coincidence, the polarity-reversed N1 amplitude was higher).

The trend analysis (see Figure 6, top left) of the N1 amplitudes at Cz showed a significant linear trend, $F(1, 13) = 14.89, p < .001$, indicating an amplitude increase with growing keypress–tone separation.

P2

At the Cz lead, in the corrected coincidence waveform, the P2 peaked at 178 msec, whereas in the corrected 1000+ msec post-keypress tone ERP, it peaked at 181 msec. The ANOVA of the P2 amplitudes measured in the range of 169–189 msec showed a Stimulus main effect, $F(1, 13) = 32.39, p < .001$, indicating a lower (less positive) P2 amplitude for the coincidence; an AntPos main effect, $F(2, 26) = 30.76, \epsilon = 0.70, p < .001$; a Laterality main effect, $F(2, 26) = 17.13, \epsilon = 0.73, p < .001$; a Stimulus \times AntPos interaction, $F(2, 26) = 20.66, \epsilon = 0.93, p < .001$; a Stimulus \times Laterality interaction, $F(2, 26) = 13.88, \epsilon = 0.98, p < .001$; and an AntPos \times Laterality interaction, $F(4, 52) = 5.30, \epsilon = 0.66, p < .01$. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes at the F, C, and P leads (averaged over the Laterality levels) indicated that the P2 suppression effect was strongest at the central and weakest at the parietal sites, with significant differences between each pair of sites [$t(13) > 2.55, p < .05$]. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes on the left, middle, and right leads (averaged over the levels of the AntPos factor) indicated that the P2 suppression effect was stronger at central than at lateral sites [$t(13) > 3.68, p < .01$].

For the P2 amplitudes measured at Cz (see Figure 6, top right), significant linear and quadratic trends were found: $F(1, 13) = 58.11, p < .001$ and $F(1, 13) = 15.53, p < .001$, respectively. These indicate an amplitude increase with growing keypress–tone separation.

N1 and P2 Amplitudes as the Function of Keypress–Tone Separation

The Component \times Keypress–tone separation ANOVA of the z-transformed amplitudes showed a significant interaction, $F(4, 52) = 3.61, \epsilon = 0.72, p < .05$, indicating that the two component amplitudes behave differently as the function of keypress–tone separation.

Experiment 2

N1

The N1 peaked at 101 msec in the corrected coincidence ERP and at 104 msec in the corrected 1000+ msec post-keypress tone ERP at Cz in the group-average waveform. The ANOVA of the N1 amplitudes measured in the range of 92–112 msec showed a Stimulus main effect, $F(1, 12) = 4.79, p < .05$, indicating a lower (less negative) N1 amplitude for the coincidence; an AntPos main effect, $F(2, 24) = 15.90, \epsilon = 0.69, p < .001$; a Laterality main effect, $F(2, 24) = 23.70, \epsilon = 1.00, p < .001$; and a Stimulus \times AntPos interaction, $F(2, 24) = 6.13, \epsilon = 0.65, p < .05$. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes at the F, C, and P leads (averaged over the Laterality levels) indicated that the N1 suppression effect was stronger at central and parietal

than at frontal sites [$t(12) > 2.53, p < .05$]. No significant effects were found at the mastoids.

The trend analysis of the N1 amplitudes at Cz (see Figure 6, top left) showed a significant linear trend, $F(1, 12) = 4.80, p < .05$, indicating an amplitude increase with growing keypress–tone separation.

P2

At the Cz lead, in the corrected coincidence waveform, the P2 peaked at 186 msec, whereas in the corrected 1000+ msec post-keypress tone ERP, it peaked at 187 msec. The ANOVA of the P2 amplitudes measured in the range of 177–197 msec showed a Stimulus main effect, $F(1, 12) = 11.71, p < .01$, indicating a lower (less positive) P2 amplitude for the coincidence; an AntPos main effect, $F(2, 24) = 14.62, \epsilon = 0.89, p < .001$; a Laterality main effect, $F(2, 24) = 5.96, \epsilon = 0.98, p < .01$; a Stimulus \times AntPos interaction,

$F(2, 24) = 25.01, \epsilon = 0.82, p < .001$; an AntPos \times Laterality interaction, $F(4, 48) = 4.70, \epsilon = 0.55, p < .05$; and a Stimulus \times AntPos \times Laterality interaction, $F(4, 48) = 3.27, \epsilon = 0.69, p < .05$.

To resolve the three-way interaction, Student's t tests were conducted between the post-keypress-minus-coincidence amplitudes measured at the left, midline, and right electrodes at each levels of the AntPos factor. The suppression effect was maximal at the middle for all three AntPos levels. For the frontal leads, the suppression effect was stronger on the left and midline than on the right [$t(12) > 2.37, p < .05$]; at the central leads, a significant difference in suppression was only found between the right and midline electrodes [$t(12) = 2.66, p < .05$]. No significant suppression difference was found at the parietal sites. Student's t tests conducted between the post-keypress-minus-coincidence amplitudes at the F, C, and P leads (averaged over the Laterality levels) indicated that the P2 suppression

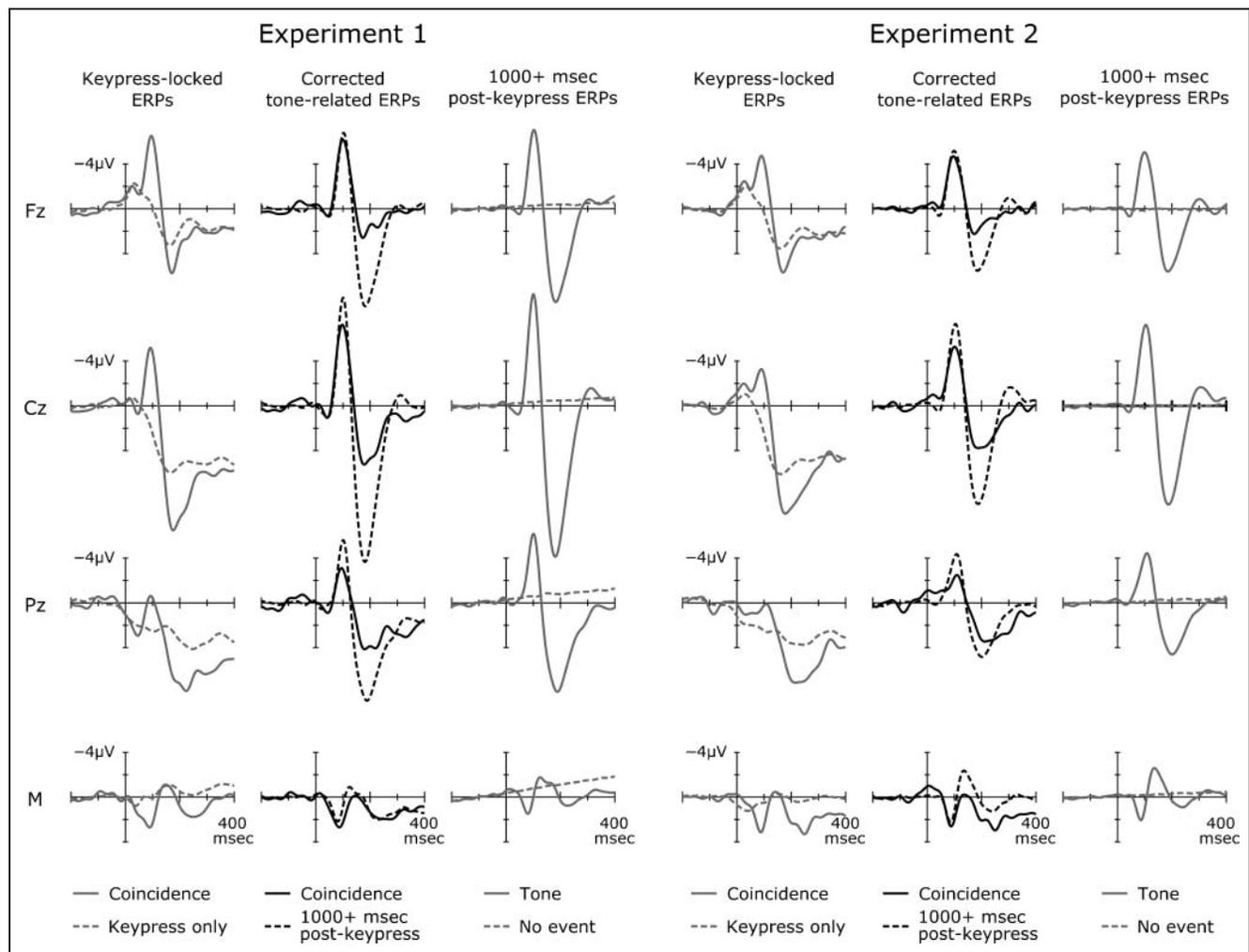


Figure 3. Group-mean ERPs elicited in Experiments 1 (left) and 2 (right) at the Fz, Cz, Pz, and the average of the mastoid leads. For each experiment, the keypress-locked ERPs are presented in the left column (for tone–keypress coincidences and for keypresses separated from any other event by at least 1 sec); in the right column, ERPs to tones following keypresses by at least 1000 msec and the average of epochs following keypresses by at least 1000 msec with no actual events in them are presented. In the middle column, the corrected tone-related ERPs are shown: The corrected coincidence tone ERP is the coincidence-minus-keypress difference waveform of the ERPs presented in the left column, whereas the corrected 1000+ msec post-keypress tone ERP is the difference of the waveforms presented in the right column.

effect was strongest at the central and weakest at the parietal sites, with significant differences between each pair of sites [$t(12) > 2.96, p < .05$].

For the P2 amplitudes measured at Cz (see Figure 6, top right), significant linear and quadratic trends were found, $F(1, 12) = 17.99, p < .001$ and $F(1, 12) = 11.15, p < .01$, respectively, indicating an amplitude increase with growing keypress–tone separation.

N1 and P2 Amplitudes as the Function of Keypress–Tone Separation

The Component \times Keypress–tone separation ANOVA of the z-transformed amplitudes showed no significant interaction.

ERFs—Experiment 3

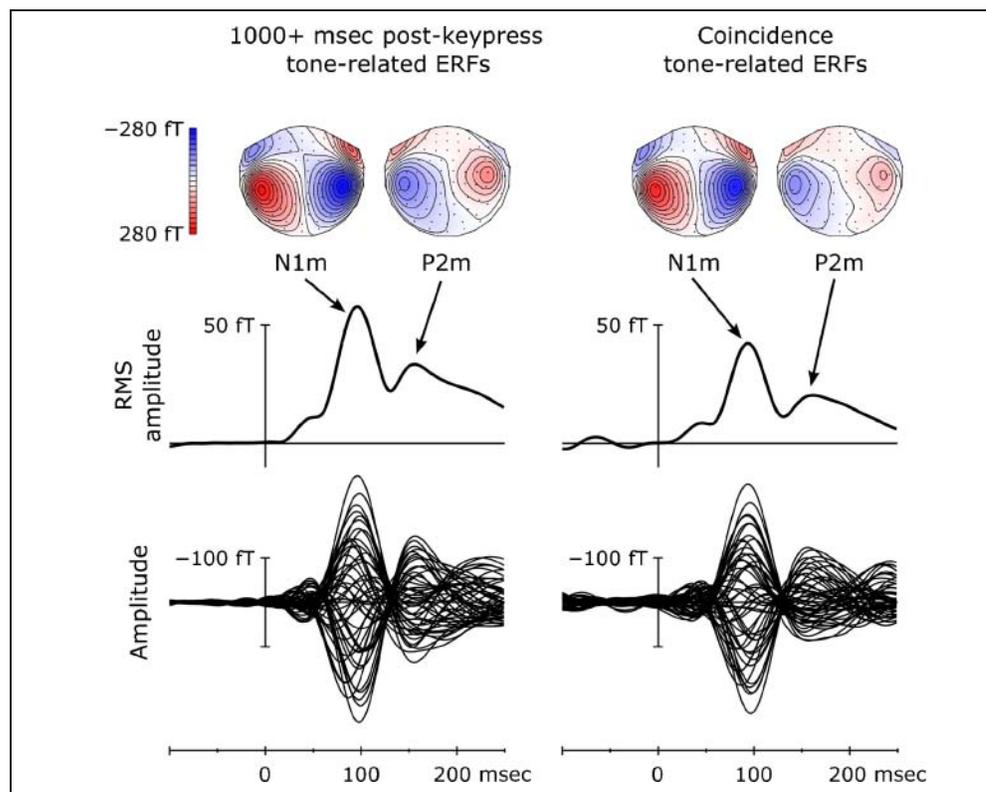
Group-mean magnetic N1 and P2 (N1m and P2m) distributions as well as event-related magnetometer and RMS amplitude signals are presented in Figure 4. N1m was observed for all participants at 97 ± 4 msec after stimulus onset in the corrected 1000+ msec post-keypress tone ERF response. For 18 participants, the ERFs were dipolar over both hemispheres: For these participants, two-dipole models were fitted using a spherical volume conductor model and data from all sensors; for two participants, a dipolar ERF was observable over the right hemisphere only. For these two participants, a single-dipole model was fitted using data from the right hemisphere sensors only.

The goodness of fit values were in the range of 84%–99% (median = 96%, first quartile = 93%, third quartile = 97%; mean = 94%, $SD = 5\%$). Each dipole was located in one of the supra-temporal cortices (as assessed on each participant’s MRI; see Figure 5). Using these dipole positions and orientations, dipole magnitudes were fitted at the same latency to the corrected coincidence tone ERF response and the corrected 250, 500, and 750 msec post-keypress tone ERF responses. Student’s t tests showed that source strength was smaller for the corrected coincidence tone response than for the corrected 1000+ msec post-keypress tone response on both sides [$t(17) = 3.54, p < .01$ on the left; $t(19) = 2.82, p < .05$ on the right].

The Stimulus (corrected coincidence vs. corrected 1000+ msec post-keypress) \times Hemisphere (left, right) ANOVA of the source strengths for the participants with a dipolar ERF on both sides showed a main effect of stimulus type, $F(1, 17) = 11.38, p < .01$, indicating a smaller source strength for the corrected coincidence tone response, and a Stimulus \times Hemisphere interaction, $F(1, 17) = 7.48, p < .05$, showing that the source strength differences between the two stimuli were larger on the left than on the right side.

The trend analyses of the source strengths (see Figure 6, bottom left) showed significant linear [$F(1, 17) = 15.72, p < .001$] and quadratic [$F(1, 17) = 6.54, p < .05$] trends in left as well as in the right hemisphere [linear: $F(1, 19) = 12.48, p < .001$; quadratic: $F(1, 19) = 4.85, p < .05$]. This indicates that N1m source strength increases with increasing keypress–tone separation.

Figure 4. Group-mean ERFs elicited in Experiment 3 for the corrected 1000+ msec post-keypress tone (left column) and for the corrected coincidence tone (right column). The bottom shows the amplitudes measured at the temporal magnetometers. The center shows the RMS amplitudes (calculated over all magnetometers). The average RMS amplitude in the 100-msec prestimulus interval (baseline) was subtracted from the RMS amplitudes. The top shows the ERF distributions in the N1m and P2m time ranges.



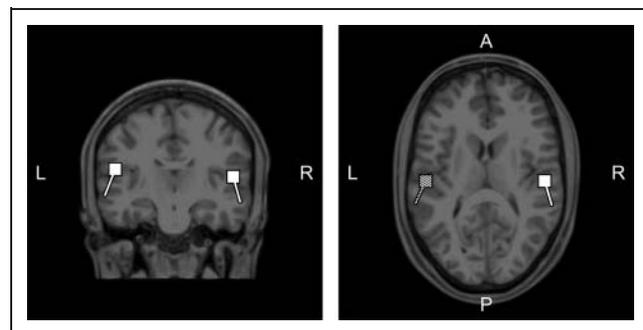
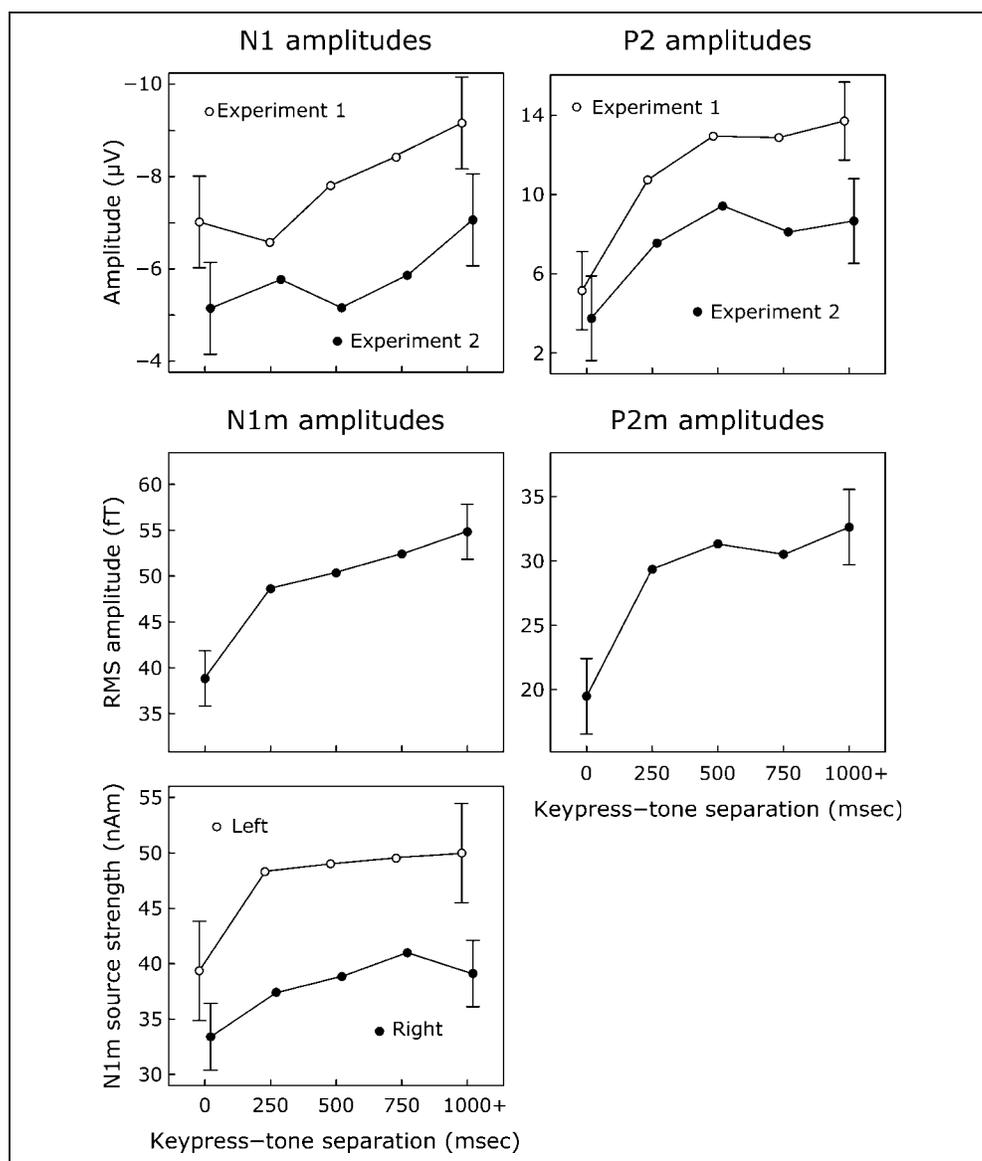


Figure 5. ECD positions (squares) and orientations (bars) fitted at the N1m, superimposed on the T1-weighted MRI (coronal: left, axial: right) of one representative participant. The axial slice is positioned at the height of the right dipole. The shaded square shows the vertical projection of the left dipole, which was about 6 mm above the slice. Slice orientations are indicated by letters: A = anterior, P = posterior, L = left, and R = right.

Figure 6. Group-mean corrected N1 and P2 amplitudes elicited at the Cz electrode in Experiments 1 and 2 (top left and top right, respectively). N1m and P2m RMS amplitudes (center left and center right, respectively) and N1m source strengths (bottom left) in Experiment 3 for tones following keypresses by 0 (coincidences), 250, 500, 750, and at least 1000 msec are shown. The whiskers represent 95% confidence intervals for the pairwise comparison of coincidence with 1000+ msec post-keypress amplitudes (calculated according to Masson & Loftus, 2003).



The group-average RMS amplitude peaked at 96 msec (N1m) and 156 msec (P2m) in the corrected 1000+ msec post-keypress tone ERF response. The RMS amplitude in the N1m and P2m time ranges (Figure 4, center) were significantly lower in the corrected coincidence tone ERF than in the corrected 1000+ msec post-keypress tone ERF response [$t(19) = 7.82, p < .001$ for N1m and $t(19) = 6.60, p < .001$ for P2m]. The trend analyses of the RMS amplitudes (Figure 6, center) showed significant linear [$F(1, 19) = 61.61, p < .001$] and quadratic [$F(1, 19) = 7.13, p < .01$] trends for N1m. For the P2m, significant linear [$F(1, 19) = 31.24, p < .001$], quadratic [$F(1, 19) = 9.91, p < .01$], and cubic [$F(1, 19) = 4.84, p < .05$] trends were found. This indicates that RMS amplitude in the N1m and P2m time ranges increased with increasing keypress–tone separation.

The Component \times Keypress–tone separation ANOVA of the z-transformed RMS amplitudes showed no significant interaction.

DISCUSSION

The results of the three experiments show that the N1(m) and P2(m) auditory event-related responses are consistently suppressed when the eliciting sound coincides with an action, even when no contingent action–stimulus relationship exists. In contrast with previous research, the present results are not contaminated by between-block differences or condition order because the relevant keypress, sound, and coincidence events occurred within the same experimental blocks.

Whereas the N1m suppression observed in Experiment 3 shows that auditory processing was attenuated for tones coinciding with keypresses, the topography of the ERP suppression effect in Experiments 1 and 2 was more posterior than expected (i.e., larger at the parietal than at the frontal leads). Moreover, keypress–tone coincidences even increased the positive aspect of the N1 at the mastoids in Experiment 1. This indicates that, besides the attenuation of auditory sensory processing, further processing changes take place when an action–sound coincidence occurs. This may be the attenuation of the widely distributed nonspecific N1 component, which, in turn, would make the polarity reversal of the stimulus-specific N1 more obvious, but it is also possible that coincidences result in a further, parietally positive ERP component overlapping the N1.

Whereas P2(m) was attenuated in all three experiments, the two methods showed somewhat different aspects of the P2. P2m ERF peaked earlier, and its sensitivity to the keypress–tone separation was not different from that of the N1m ERF. The P2 ERP amplitude, on the other hand, showed a different dependence on keypress–tone separation than the N1 (in Experiment 1, where amplitudes were higher because of their relative loudness). This suggests that the N1 and P2 ERP waveforms include subcomponents that reflect functionally different aspects of processing (see also Grimm, Schröger, Baess, & Kotz, in press; Ford et al., 2001).

Taken together with the results of Makeig et al. (1996) and Hazemann et al. (1975), this study provides strong evidence that performing an action leads to a suppression of concurrent auditory processing, that is, action–sound contiguity is a sufficient condition for N1 suppression. The suppression effects in this study are very similar to those routinely found in paradigms in which actions and sounds have a contingent relationship. Whereas previous studies assumed that contingency was a necessary condition for N1 suppression, the present design provides a baseline condition for the assessment of potential contingency-related N1 suppression effects; and the results show the necessity to directly test whether action–stimulus contingency and its forward modeling contributes to N1 suppression or not. It is important to emphasize that, based on the present results, it cannot be determined whether action–sound contiguity is the sole cause of auditory suppression or not. It seems possible that, when a contingent action–stimulus relationship exists, this might be represented by a

forward model, which could bring about further suppression. In comparison with the present experiments, in which actions had no auditory consequences, presenting action-contingent stimulation is likely to result in an explicit expectancy of the given contingent sensory event, which might contribute to N1 suppression as well. Indeed, some studies show suppression effects that would be difficult to explain solely on the basis of action–stimulus contiguity. For example, Baess et al. (2011) found that N1 suppression is stronger when action-independent sounds are mixed into a self-induced (contingent) action–sound sequence.

We started out by suggesting that, whereas it seemed plausible that speech-related N1 suppression reflects the workings of an internal forward model, N1 suppression measured in settings with arbitrary non-speech-related actions and contingent (speech or nonspeech) sounds may not reflect the workings of a forward model. Because this study does not provide direct information on the processing of speech or similarly “natural” action–stimulus relationships, one may argue that the present results cannot be generalized to these. On one hand, it would seem highly redundant if the highly similar event-related response effects occurring in similar contiguous action–stimulus contexts would be produced by different subsystems; on the other hand, however, speech may have highly specialized subsystems, which are not readily available for the processing of other types of actions and stimuli. At this point, because of the lack of empirical evidence, this issue cannot be convincingly resolved.

Interpreting the present N1 suppression effect in functional terms requires further research. At this point, we can offer three types of interpretations, which may independently influence N1 suppression effects. The first interpretation suggests that N1 suppression (in the present and other experiments) does not reflect the workings of a forward model rather that it results from a dynamic change in the distribution of attentional resources. It is well known that auditory N1 is enhanced when sounds are in the focus of attention (Hillyard, Hink, Schwent, & Picton, 1973) and is attenuated when attentional focusing is disrupted (Horváth & Winkler, 2010). It seems possible that pressing a button or performing an action draws attention away from task-irrelevant auditory stimulation for a short period, which results in attenuated N1 for tones close to keypresses. This explanation is on a par with that offered by Makeig et al. (1996) for the auditory steady-state response.

The second interpretation suggests that N1 suppression does not result from the cancellation of sensory reafference, rather, it reflects a process subserving the formation of a forward model or other contingency representation. Because contiguity is one of the cues that may allow the inference of a causal relationship between events (Hume, 1739/1896), the detection of action–contiguous sound events may be necessary for the formation of an action–sound contingency representation. N1 suppression might simply reflect a process that “flags”

such sounds and thereby provides a signal that could serve as a basis for the formation of an action–sound contingency representation.

A third type of interpretation suggests that, despite the absence of a contingent action–stimulus relationship, the suppression effect nonetheless reflects the cancellation of sensory refference, that is, it still results from the workings of an internal forward model. Whereas forward modeling allowed more efficient interactions with the environment in most previous studies, this interpretation suggests that N1 suppression reflects the workings of a dysfunctional contingency representation in the present experiments (after all, such a model would produce invalid predictions about the occurrence of self-initiated sounds). In the following, we elaborate this line of thought, which delineates some key questions regarding this interpretational framework.

Assuming that there is an internal forward model representing a keypress–sound contingency in the present experiments, there are two questions that should be answered: (1) How is this representation created and (2) how (and why) does such a representation get preserved?

First, it seems possible that the hypothetical contingency representation is not built up at the beginning of or over the course of the experiment (see also Lange, 2011), but it already exists: It is a general “expectation” that our actions should generate some kind of a sensory event in the environment. Whereas it is an intriguing possibility that N1 suppression reflects an innate readiness for capturing contingent action–effect relationships during interactions with the environment, in this study, such a “readiness” could also be brought about by extensive training through the widespread use of keypress-based interfaces in everyday devices. Long-term training creates strong associations between actions and their sensory consequences, which influence perception even if the actions do not take place in their usual context: Repp and Knoblich (2007) showed that, when pianists performed movements that would generate a rising or descending tone pair on a piano, this induced a corresponding bias in the perception of an ambiguous pitch change, whereas for nonpianists, no bias was observable. In the present case, the long-term use of keypress-based interfaces might give rise to a general action–effect association in which the effect can be a large class of sensory events (including a tone). The generality of the effect would also explain how it is possible that an arbitrary contingency between a keypress and an artificial tone is represented by an internal forward model similarly to that hypothesized to exist between speech production and speech sounds.

A second possibility is that a contingency representation is built up (rapidly) during the experiment, but the build up of the representation does not depend on actual action–stimulus contingency, rather, it is based on the instances when the action and stimulus events were temporally contiguous. That is, coincidences give rise to an

“illusory contingency” in the present case despite the absence of an actual contingent relationship. It seems even possible that the cognitive system might not only take “real” coincidences as evidence for the establishment of such a representation but also stimulus events that follow actions within a sensitive period. Elsner and Hommel (2004) found that contiguity played a role in forming action–effect associations even if (contingent) task-irrelevant sounds followed actions by 1 sec. If stimuli following actions within 1 sec were interpreted as evidence for a contingent action–effect relationship, then in the present experiments, the relevant coincidence rate would be around 25% (whereas the “real” coincidence rate was 5%). Note, nonetheless, that the present results show a decrease of the suppression effect with growing keypress–tone separation over 1 sec, which suggests that the duration of such a sensitive period may be much shorter (on the order of a couple hundred of milliseconds).

The hypotheses outlined above suggest ways in which a contingency representation might be created despite the absence of actual contingency. These hypotheses, however, do not reflect on why or how such representations are preserved in these situations. Adaptation to changed contingencies in the sensory-motor system is usually investigated in paradigms in which an established action–stimulus contingency is abruptly changed to a different contingency. Adaptation to the new contingency takes place on multiple time scales and is based on learning from prediction errors (for a summary, see Shadmehr et al., 2010). On the basis of this, two speculations on the preservation of the hypothetical keypress–sound contingency representation can be put forward: First, similar to the acquisition of a contingency representation, the characteristic time for changing such a representation may simply be too long compared with the typical duration of an experiment. That is, if it takes long-term training to build up such representations, then changing them might take long time as well. Second, it is possible that representations are changed when the action starts to lead to novel consequences; however, when the associated effect is simply absent, the representation may remain unchanged because a prediction error cannot be calculated when no effects are present. In different terms, changes may be induced by the interference between the associated and actual consequences, whereas the lack of a consequence might not produce such interference, thereby not affecting the model itself.

In conclusion, we demonstrated in three experiments that the auditory N1 and P2 responses are attenuated when the eliciting tone is closely preceded by a keypress, even if no contingent relationship between keypresses and tones existed. These results are highly similar to those obtained in studies using action–contingent auditory stimulation and provide a baseline to further studies investigating the role of action–stimulus contingency and forward modeling in N1 and P2 suppression.

Acknowledgments

We thank Judit Rochéné Farkas and Yvonne Wolff for assistance in data collection. This work was supported by the European Community's Seventh Framework Programme (PERG04-GA-2008-239393), the German Academic Exchange Service (Deutscher Akademischer Austauschdienst, DAAD, Project 50345549), and the Hungarian Scholarship Board (Magyar Ösztöndíj Bizottság, MÖB, P/853).

Reprint requests should be sent to János Horváth, Institute for Psychology, Hungarian Academy of Sciences, P.O.B. 398, Szondi u 83/85, H-1394 Budapest, Hungary, or via e-mail: horvath@cogpsyphy.hu.

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.
- Ariff, G., Donchin, O., Nanayakkara, T., & Shadmehr, R. (2002). A real-time state predictor in motor control: Study of saccadic eye movements during unseen reaching movements. *Journal of Neuroscience*, *22*, 7721–7729.
- 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.
- Baess, P., Jacobsen, T., & Schröger, E. (2008). Suppression of unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*, 137–143.
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*, 635–640.
- Cullen, K. E. (2004). Sensory signals during active versus passive movement. *Current Opinion in Neurobiology*, *14*, 698–706.
- 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*, 183–191.
- Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: A short review. *Journal of Neural Engineering*, *2*, S313–S319.
- Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science*, *255*, 90–92.
- Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological Research*, *68*, 138–154.
- 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.
- 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.
- 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.
- Gentili, R., Cahouet, V., Ballay, Y., & Papaxanthis, C. (2004). Inertial properties of the arm are accurately predicted during motor imagery. *Behavioral Brain Research*, *155*, 231–239.
- Grimm, F., Schröger, E., Baess, P., & Kotz, S. A. (in press). The cerebellum generates sensorimotor-to-auditory predictions: ERP lesion data. *Journal of Cognitive Neuroscience*.
- Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*, 413–497.
- 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.
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*, 180–190.
- Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *NeuroReport*, *17*, 1375–1379.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Horváth, J., & Winkler, I. (2010). Distraction in a continuous-stimulation detection task. *Biological Psychology*, *83*, 229–238.
- 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.
- Hume, D. (1739/1896). *A treatise of human nature*. Oxford, UK: Clarendon. (Reprinted from the original edition in three volumes and edited, with an analytical index, by L. A. Selby-Bigge). Retrieved on Feb. 6, 2012, from oll.libertyfund.org/title/342.
- Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.
- Lange, K. (2011). The reduced N1 to self-generated tones: An effect of temporal predictability? *Psychophysiology*, *48*, 1088–1095.
- Makeig, S., Müller, M., & Rockstroh, B. (1996). Effects of voluntary movements on early auditory brain responses. *Experimental Brain Research*, *110*, 487–492.
- Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299–302.
- Masson, M., & Loftus, G. R. (2003). Using confidence intervals for graphically based interpretation. *Canadian Journal of Experimental Psychology*, *57*, 203–220.
- 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.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279.
- Näätänen, R. (1988). Implications of ERP data for psychological theories of attention. *Biological Psychology*, *26*, 117–163.
- 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.
- 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.
- Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, *18*, 6–7.
- Roy, J. E., & Cullen, K. E. (2004). Dissociating self-generated from passively applied head motion: Neural mechanisms in the vestibular nuclei. *Journal of Neuroscience*, *24*, 2102–2111.
- Schäfer, E. W., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, *181*, 175–177.
- Shadmehr, R., Smith, M. A., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in

- motor control. *Annual Review of Neuroscience*, *33*, 89–108.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*, 482–489.
- Taulu, S., Kajola, M., & Simola, J. (2004). Suppression of interference and artifacts by the signal space separation method. *Brain Topography*, *16*, 269–275.
- Thura, D., Hadj-Bouziane, F., Meunier, M., & Boussaoud, D. (2011). Hand modulation of visual, preparatory, and saccadic activity in the monkey frontal eye field. *Cerebral Cortex*, *21*, 853–864.
- Vaughan, H. G., Jr., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology*, *28*, 360–367.
- Vaziri, S., Diedrichsen, J., & Shadmehr, R. (2006). Why does the brain predict sensory consequences of oculomotor commands? Optimal integration of the predicted and the actual sensory feedback. *Journal of Neuroscience*, *26*, 4188–4197.
- Ventura, M. I., Nagarajan, S. S., & Houde, J. F. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neuroscience*, *10*, 58.
- von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*, 464–476.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*, 1880–1882.