The role of mechanical impact in

action-related auditory attenuation

János Horváth

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

Hungarian Academy of Sciences, Budapest, Hungary

Running head: Impact and action-related auditory attenuation

Address correspondence to:

János Horváth
Institute of Cognitive Neuroscience and Psychology, RCNS, HAS
P.O.B. 286,
H-1519
Budapest
HUNGARY
Phone: +36 1 382 6800
E-mail: horvath.janos@ttk.mta.hu

(46 pages, including 7 figures)

Abstract

A number of studies found that sounds temporally close to one's own finger-movements elicit lower amplitude auditory event-related potentials (ERPs) than the same tones when they are only listened to. In these studies actions involved making a mechanical contact with an object. In the present study, the role of mechanical contact with an object was investigated in action-related auditory attenuation. In three experiments, participants performed a time-interval production task. In each experiment, in one condition the action involved touching an object, but no mechanical contact was made in the other. The estimated tone-related ERP contributions to the action-tone coincidence ERP waveforms (calculated by subtracting the action-related ERP from the coincidence ERP) were more attenuated when the action involved moving the finger and making a mechanical contact at the end of the movement. However, when participants kept their finger on a piezoelectric element, and applied pressure impulses without moving their finger, the action did not result in stronger attenuation of the tone-related auditory ERP estimates. Although these results may suggest that auditory ERP attenuation is stronger for actions resulting in mechanical impact, they also imply that mechanical impact may confound and lead to the overestimation of auditory ERP attenuation in such paradigms because the impact may result in faint, but audible sounds.

Introduction

Actions are seldom performed for their own sake. Most of our actions are goal-directed, and result in direct interactions with the environment. Many studies show that stimulation caused by our own actions is processed differently from stimulation caused by external sources. In the auditory domain, it has been found that sounds occurring in temporal proximity to one's own actions are processed differently from those far from actions, even if the actions are not vocalizations. This is reflected in the attenuation of the N1 and P2 auditory event-related brain potentials (ERPs, e.g. Aliu, Houde, & Nagarajan, 2009; Baess, Horváth, Jacobsen, & Schröger, 2011; Ford, Gray, Faustman, Roach, & Mathalon, 2007; Knolle, Schröger, Baess, & Kotz, 2012; Martikainen, Kaneko, & Hari, 2005; McCarthy, & Donchin, 1976; Sowman, Kuusik, & Johnson, 2012; Schafer & Marcus, 1973), as well as subjective reports of lower loudness levels (Weiss, Herwig, & Schütz-Bosbach, 2011; Desantis, Weiss, Schütz-Bosbach, & Waszak, 2012). These attenuation effects are attributed to predictive (Hughes, Desantis, Waszak, 2013) or attentional mechanisms (Horváth, Maess, Baess, & Tóth, 2012) related to the voluntary performance of the given movement. One common feature of all studies investigating action-related auditory attenuation with non-speech actions is that the actions are not simply movements, but *interactions with objects* (typically button-presses). The success of the action depends not only on our intention and capability to act, but on the physical properties of the manipulated objects and our knowledge of these properties. For example, too short or too weak movements may not result in a successful button-press (although these movements are still actions). The goal of the present study was to investigate whether the mechanical interaction with an external object played a role in the auditory ERP attenuations reported in the literature.

Most studies investigating action-related attenuation of auditory ERPs administered protocols in which the sounds were *caused* by the task-relevant actions (*contingent* paradigm).

Because of this arrangement, it was hypothesized that these attenuations reflected the cancellation of sensory re-afference (Miall & Wolpert, 1996): It was assumed that causal action-sound relationships were represented in the form of *internal forward models*, and through the *efference copy* or *corollary discharge* (von Holst & Mittelstaedt, 1950, Sperry, 1950, for a recent summary see Crapse & Sommer, 2008a, 2008b) function, the auditory system was prepared for the consequences of the to-be performed action to prevent the redundant processing of the incoming sounds. Because predictions provided by internal forward models are used by various cognitive subsystems beyond those directly involved in the control of the given effector (Davidson & Wolpert, 2005), this hypothesis seems plausible. A couple of studies, however, demonstrated that the processing of sounds randomly coinciding with one's own button-presses is attenuated even in the absence of a causal action-sound relationship (*coincidence* paradigm, Hazemann, Audin, & Lille, 1975; Makeig, Mueller, & Rockstroh, 1996; Horváth, et al., 2012; Horváth, 2013a, 2013b). Because an internal forward model would be dysfunctional if no actual action-effect contingency existed (preparing for an effect over and over again even if the action never results in the given effect is hardly adaptive), these results are difficult to explain in the forward modeling framework. The most prominent alternative explanation is that these attenuations reflect a dynamic allocation of attention from the incoming sounds to the action. It is well-known that auditory N1 is sensitive to the allocation of attention: sounds in an attended channel elicit larger N1s than those presented in an unattended one (e.g. Hillyard, Hink, Schwent, & Picton, 1973, Okamoto, Stracke, Wolters, Schmael, & Pantev, 2007), and N1 and P2 elicited by sound events presented shortly after a distracting event are also reduced (Horváth & Winkler, 2010). That is, sounds in close temporal proximity to actions may elicit lower amplitude ERPs than those far away, because attention is drawn away from the sounds by the action-related activities.

Various specifications of the rather general notion of "attending the action" showed diverging results. Horváth (2013b) and Timm, SanMiguel, Saupe, and Schröger, (2013) attempted to influence the allocation of attention by manipulating the task-relevancy of the tones, and found that these manipulations did not influence the magnitude of attenuation. Saupe, Widmann, Trujillo-Barreto, & Schröger (2013), on the other hand, found evidence suggesting that although attention may not explain the phenomenon in its entirety, attention effects may contribute to the observed ERP attenuation. The present study investigated whether finger-movement-related auditory attenuation was caused by the allocation of input attention to the tactile stimulation occurring when a finger-movement was performed. Because previous studies used exclusively button-presses or -releases as actions, and participants' primary task was to perform a sequence of actions corresponding to a pre-defined timing schedule, it was hypothesized that these settings might compel participants to attend the tactile feedback when the action was performed to support maintaining the optimal interaction with device, or as an integral part of action preparation (Galazky, Schütze, Noesselt, Hopf, Heinze, & Schoenfeld, 2009; Brown, Friston, & Bestmann, 2011).

Experiment 1 and 2 were conducted to test this hypothesis. In both experiments participants produced a sequence of finger-movements which either resulted in mechanical contact with an object or not. Experiment 1 used a *coincidence paradigm*: during the performance of the action sequence, random, action-independent tone sequences were presented. In Experiment 2, a *contingent* arrangement was administered, in which the given actions always resulted in tones. It was hypothesized that if mechanical interaction played a role in action-related auditory attenuation, then the magnitude of N1- and P2-attenuation would be reduced when no mechanical interaction took place. As described in the Methods of Experiment 1, establishing a mechanical contact with an external object might have caused faint, but audible transient sounds, so in order to better specify the nature of the mechanical

5

contact-related effects found in these two experiments, in Experiment 3, in a coincidence arrangement, participants performed actions which did not produce such transient sounds: participants maintained continuous contact with an object and applied force impulses from time-to-time, or produced finger-movements without mechanical contact (Figure 1). As usual in the literature, in all three experiments, the contribution of the auditory ERP to the action-tone coincidence-related ERP was estimated by subtracting the action-related ERP registered when the action did not co-occur with a tone.

Experiment 1

Materials & Methods

Participants

19 young adults (nine women, aged 18-25 years, mean 21 years, 17 right-, two left-handed) participated in Experiment 1 for monetary compensation. They reported normal hearing and no history of neurological disorders, and gave written informed consent after the experimental procedures were explained to them.

Stimuli and procedures

Participants were seated in an armchair in a sound-proofed room during the experiment. In front of their dominant hand, a frame with an infrared light beam was positioned (Figure 1, left and middle column). Participants were instructed to interrupt the light beam with their index finger with a quick downward-upward movement (i.e. the resting state was above the light beam) at regular, 4 s intervals. The frame was positioned so that the light beam was close to the lowest point of the finger movement. To minimize eye-movements, participants were instructed to look at a fixation puppet placed in front of them during the task. Before the experiment, participants practiced the finger movements and their

6

pacing in 2-3 blocks with 30 actions. In these blocks, a histogram of the between-action intervals was presented on a screen with 1 s bins, which was updated for each action. During the experiment, such feedback was only provided at the end of each block.

There were two conditions: in the Tap condition, a plate was placed under the light beam, so that participants' fingers tapped on this plate when the light beam was interrupted. In the No Tap condition, the plate was removed, so no mechanical contact was made. Each condition was administered in seven blocks, each lasting 5 minutes. The blocks were presented in an interwoven order ("TNNTTNNT..."), with 9 participants starting with the Tap (denoted by "T"), 10 with the No Tap ("N") condition.

In both conditions, 1000 Hz pure tones of 50 ms duration (including 10-10 ms linear rise and fall times) were presented with pre-generated, random inter-tone-intervals sampled from a 2-6 s uniform distribution through open headphones (HD-600, Sennheiser, Wedemark, Germany). Tone intensity was individually adjusted to 60 dB above hearing threshold level. Similarly to that described in Horváth et al. (2012), the tone presentation schedule was adjusted when an action occurred: The tones following the action were shifted uniformly towards the action (i.e. earlier): if the next tone would have been delivered within 300 ms according to the original schedule, it was delivered immediately (an action-tone coincidence); if it was to be delivered between 300 and 600 ms, it was re-scheduled to 300 ms, and so on. Due to the constraints of the stimulation equipment, a constant action-tone timing difference of 7 ms was always present – this was taken into account during electroencephalogram (EEG) processing. If multiple actions preceded a tone, the adjustment was made only in reference to the last action. Due to this adjustment, the inter-tone-interval was shorter for tones preceded by actions than for tones preceded by another tone. Because shorter inter-tone-intervals result in lower N1 amplitudes (Näätänen & Picton, 1987), only ERPs to tones immediately preceded by actions (and not by tones) were analyzed in this experiment.

Tapping on the plate inevitably causes sound, which may be audible. To give a rough estimate of the loudness of these sounds, after the experiment, the intensity of the presented tone (adjusted for an individual with 0 dB SPL hearing threshold level at 1000 Hz), and the sounds caused by tapping were recorded in a setup closely matching the original with microphones set in an artificial head (HSU III.2, without torso, Head Acoustics, Germany). To capture the power of the transient signals resulting from tapping, root-mean-squared (RMS) signal energy was calculated in 20 ms sliding windows across the whole recording. Although forceful hits on the plate (produced with the intention to make a sound as loud as possible by moving the whole hand from the wrist and hitting the plate with the index finger) could produce measurement maxima reaching -15 dB in comparison to the tone, the typical maximal power for taps (measured in a sequence of 40 taps delivered at a rate of once per second) was around -35 dB (standard deviation, SD: ±3 dB) in comparison to the tone. Note that these are rough estimates because the actual loudness also depends on a number of factors, most importantly on the individual threshold level for the tone used in the experiment and the individual movement characteristics.

EEG-recording and analysis

The EEG was recorded with 1000 Hz sampling rate and on-line, 200 Hz low-pass filtering (Synamp 2, Compumedics, Victoria, Australia) by 61 Ag/AgCl electrodes mounted on an elastic cap (EASYCAP, Herrsching, Germany) according to the 10% system (Nuwer et al., 1998). Additional electrodes were placed at the mastoids. The reference was placed on the tip of the nose, the ground on the forehead. Horizontal electrooculogram (EOG) was obtained by a bipolar setting between electrodes placed close to the outer canthi of the two eyes. Vertical EOG was calculated off-line by subtracting the signal of an electrode placed under the right eye from the signal recorded at Fp2. The EEG was off-line re-referenced to average reference, 1-20 Hz band-pass filtered, and segmented into epochs of 600 ms corresponding to

events described below, including a 200 ms long interval preceding the given event. Amplitude calculations were referred to the average signal measured in this interval. Epochs with a signal range exceeding 100 μV on any channel were discarded. In some blocks some actions were not registered because of improper hand positioning or technical problems. When such problems were detected on-line, the affected block was re-run with a different pre-generated tone-schedule. To minimize the influence of potentially undetected actions on the ERPs, time intervals in which between-action-intervals exceeded 6.5 s were rejected from analyses.

The auditory ERPs elicited by action-tone coincidences were estimated by subtracting the ERPs elicited by actions from the ERPs elicited by coincidences (only actions separated from any other events by at least 1 s were used for this purpose). To assess whether the auditory ERP was attenuated, this estimate (*corrected coincidence ERP*) was compared to the *tone ERP*, that is, the ERP elicited by tones following an action by at least 1200 ms (and no action following it within 400 ms). Peak latencies and topographical extrema of the N1, and P2 auditory ERPs were assessed in the group average ERPs to such tones. ERP amplitudes were then calculated as average signals in 20 ms long windows centered at these time points at the electrode with the maximal signal for N1 and P2. Amplitudes were analyzed in Condition (Tap vs. No Tap) × Event (corrected coincidence vs. tone) repeated-measures analyses of variance (ANOVAs). The hypothesis-relevant question was whether a significant interaction was present, that is, whether the tone-minus-corrected-coincidence differences were modulated by the Condition factor. The power of the experiment to reveal this effect in the N1 and P2 time ranges was assessed as suggested by Cohen (1988), post-hoc. Effect size (d) was determined by dividing the mean between-condition attenuation difference by its standard deviation, for a two-sided arrangement. To assess whether attenuation was present

despite the modulation by the Condition factor, corrected coincidence and tone ERPs elicited in the No Tap condition were compared by paired, two-tailed Student's t-tests.

Previous studies (Horváth et al., 2012, Horváth, 2013a) suggested that the topography of the coincidence-related N1-attenuation effect may be more posterior than that of the tone-related N1, suggesting that N1 attenuation was not a "pure" N1-effect, but it included an ERP of different origin as well. Because of this, the topography of the tone-related N1 and P2 was compared to the corresponding ERP effect or effects: that is, the modulatory effect of condition on the level of attenuation (i.e. the No Tap-minus-Tap differences of the attenuation effects), the main ERP attenuation effect (i.e. the mean of the attenuation effects across conditions), or the attenuation in the No Tap condition. For these analyses the vector-normalized amplitudes (as described by McCarthy & Wood, 1985) were submitted to Signal (tone-related ERP vs. the appropriate ERP-effect) × Electrode (Fz, Cz, Pz, Oz) ANOVAs for N1 and P2 separately. In these analyses, a significant interaction would mean that the ERP-effect was not a "pure" modulation of the underlying tone-related ERP, rather, that it was (in part) due to a different ERP component. That is, for example, an interaction in this analysis could signal that the topography of the modulation effect was different from that of the N1. In these analyses Greenhouse-Geisser correction was applied; Greenhouse-Geisser epsilon, unadjusted degrees of freedom, and adjusted p-levels are reported. The alpha-level was set to 0.05. Generalized eta-squared effect sizes (Olejnik and Algina, 2003; Bakeman, 2005) are reported. Amplitude values are reported in the form of amplitude ± standard deviation. Statistical calculations were performed in R (version 3.0.0, R Core Team, 2013).

Results

Participants complied with the instruction. The mean between-action interval was 3968 ± 194 ms in the Tap, and 4004 ± 137 ms in the No Tap condition (no significant

difference: t[18] = 0.755, p = 0.460; paired, two-tailed t-test). In the Tap, 5.7 ± 1.1 %; in the

No Tap condition 5.7 ± 0.9 % of the sounds were coincidences (no significant difference:

t[18] = 0.278, p = 0.784). The ratio of intervals excluded from the analyses due to potentially

undetected actions did not differ between conditions (t[18] = 1.439, p = 0.167); this affected

0.6 ± 1.1 % of all between-action intervals.

The number of epochs contributing to coincidence ERPs was 23 ± 6 in both

conditions. The number of epochs contributing to the tone ERPs was 231 ± 35 in the Tap and

223 ± 33 in the No Tap condition. Tones elicited the N1 and P2 waveforms (Figure 2) with

peaks occurring at the expected latencies and electrode positions documented in the literature

(Näätänen and Picton, 1987). For tones (at least 1200 ms after an action), N1 peaked at 107

ms at FCz in both conditions (Figure 2, the amplitudes - average signal in the 97-117 ms

range at FCz - were -4.343 ± 2.331 , and -4.521 ± 2.452 µV in the Tap and No Tap

conditions, respectively). The ANOVA of the amplitudes showed a significant Condition

main effect: $F(1,18) = 13.450$, $\eta_G^2 = 0.028$, $p = 0.002$; and a Condition × Event interaction:

$F(1,18) = 6.934$, $\eta_G^2 = 0.017$, $p = 0.017$, indicating stronger attenuation in the Tap, than in the

No Tap condition. The modulation of the attenuation-effect (i.e. the mean between-condition

difference of the attenuations) was 1.229 µV, yielding d = 0.604, and a (post-hoc) power of

0.702. The corrected coincidence and tone ERP amplitudes significantly differed in the Tap

(t[18] = 3.229, p = 0.005, 1.002 µV attenuation), but not in the No Tap condition (t[18] =

0.697, p = 0.494, -0.226 µV "attenuation", i.e. the amplitude was numerically more negative

for the coincidence).

P2 peaked at FCz at 185 ms and 184 ms in the Tap and No Tap conditions,

respectively (the amplitudes - average signal in the 174-194 ms range at FCz – were 3.971 ±

2.797, and 4.157 ± 2.716 µV in the Tap and No Tap conditions, respectively). The ANOVA

of the amplitudes showed a significant Condition main effect: $F(1,18) = 12.034$, $\eta_G^2 = 0.039$,

p = 0.003; a significant Event main effect: $F(1,18) = 21.966$, $\eta_G^2 = 0.099$, $p < 0.001$; and a

Condition × Event interaction: $F(1,18) = 12.489$, $\eta_G^2 = 0.027$, $p = 0.002$, indicating stronger

attenuation in the Tap, than in the No Tap condition. The modulation of the attenuation-effect

was 1.708 µV, yielding $d = 0.811$, and a (post-hoc) power of 0.916. The corrected

coincidence and tone ERPs amplitudes differed significantly in the No Tap condition: ($t[18] =$

2.125, $p = 0.048$, 0.856 µV attenuation), showing that despite the modulatory effect of the

Condition factor, a significant attenuation was still present.

The topographical distributions of the effects were visually similar to that of the

respective tone-related ERPs (Figure 3). Comparing the topographies of the mean tone ERP

and the attenuation-modulatory effect of the Condition factor (No Tap-minus-Tap attenuation

differences) showed no significant Signal × Electrode interactions (N1: $F[3,54] = 0.554$, $\varepsilon =$

0.535, $p = 0.543$; P2: $F[3,54] = 0.493$, $\varepsilon = 0.567$, $p = 0.585$) providing no evidence for a

different ERP component causing the modulatory effects of the Condition factor. Comparing

the mean tone ERP and the attenuation effect in the No Tap condition in the P2 time range

showed no significant interaction ($F[3,54] = 0.229$, $\varepsilon = 0.608$, $p = 0.777$).

Interim Discussion - Experiment 1

The results of Experiment 1 are on-a-par with previous studies utilizing button-press

actions: the estimated auditory N1 and P2 amplitudes for tones coinciding with an action in

the Tap condition were reduced in comparison to the N1 and P2 elicited by tones separated

from the preceding action by at least 1200 ms. Conforming to the hypothesis, when finger-

movements did not result in mechanical contact with an external object (No Tap condition),

the magnitude of attenuation was reduced. This suggests that mechanical impact played a

substantial role in the observed ERP attenuation pattern. Because tapping caused faint, but

probably audible transient sounds as evidenced by the measurements presented in the

Methods section, interpreting these effects is not trivial, as discussed in the General Discussion section. The results, however, also indicate that mechanical impact was not the only factor behind the observed action-related ERP attenuation: a significant P2 attenuation was still present in the No Tap condition.

Because action-related auditory attenuation is often measured in a contingent stimulation arrangement, Experiment 2 was conducted with such an arrangement to investigate whether the effects observed in Experiment 1 were present in such a paradigm as well.

Experiment 2

Materials & Methods

Participants

19 young adults (fourteen women, aged 20-26 years, mean 22 years, all right-handed) participated in Experiment 2 for monetary compensation. They reported normal hearing and no history of neurological disorders, and gave written informed consent after the experimental procedures were explained to them.

Stimuli and procedures

The stimulus and the action-related measurement arrangements were largely similar to that in Experiment 1 (see also Figure 1, left and middle columns). Participants were explicitly instructed to move only their fingers and avoid forcefully hitting the plate. Furthermore, a curtain was hanged above the arm, between the participant's head and hand, which occluded the sight of the hand and the frame. The instructions and procedures closely followed that described in the study by Horváth & Burgyán (2013). As usual in *contingent* paradigms, there were three types of *stimulation arrangements*, which were administered in separate

experimental blocks: In Motor-Auditory blocks, participants' finger-movements triggered the presentation of a tone (with 9 ms constant delay, which was taken into account during EEG processing). Participants were instructed to keep a steady, one-movement-every-4 s pace, and more importantly, count the elicited tones, because "at the end of the block the finger-movement will no longer elicit the tone", and they will be required to report how many tones were elicited. At the end of the block, the experimenter provided feedback on accuracy, and also on movement timing. The number of tones was randomly sampled from a uniform distribution in the 61-70 range. In Motor blocks, participants were instructed to maintain a steady, one finger-movement every 4 s pace till the experimenter instructed them to stop. In Auditory blocks, participants listened to a sequence of tones and were instructed to count the tones, and report their number at the end of the block. As for Motor-Auditory blocks, the experimenter provided feedback on accuracy. Unknown to the participants, the sequence in each block was a replay of the first 51-60 tones (the sequence length determined randomly) from the tone-sequence produced in the preceding Motor-Auditory block. In contrast with paradigms in which no explicit instruction is given in these blocks, the instruction to count the tones and providing feedback on the accuracy allowed monitoring and motivating participants to attend the tones. The three types of stimulation arrangements were presented in *block triplets:* the first block was a Motor-Auditory, the second a Motor, and the third one was an Auditory block. Within each triplet, the condition (Tap or No Tap) was the same. For each participant two Tap and two No Tap triplets were administered. The order of the conditions was "TNNT" for ten, and "NTTN" for nine participants (where "T" stands for a Tap condition- and "N" for a No Tap condition triplet).

EEG-recording and analysis

Experimental blocks with improper hand positioning leading to unregistered movements were re-run immediately. The auditory ERPs elicited by action-tone events in the

14

Motor-Auditory arrangement were estimated by subtracting the ERPs elicited by actions in the Motor arrangement from the ERPs elicited by the action-tone events. To assess whether the auditory ERP was attenuated, this estimate (*corrected coincidence – Motor-Auditory ERP*) was compared to the *tone – Auditory - ERP*, that is, the ERP elicited by tones in the Auditory arrangement. Only the first fifty trials of each experimental block were used in these calculations (corresponding to the lowest number of trials-minus-one in the Auditory blocks; this was chosen to maximize between-arrangement similarity). Peak latencies and topographical extrema of the N1, and P2 auditory ERPs were assessed in the group average ERPs to such tones. ERP amplitudes were then calculated as average signals in 20 ms long windows centered at these time points at the electrode with the maximal signal for N1 and P2. Amplitudes were analyzed in Condition (Tap vs. No Tap) × Event (corrected coincidence vs. tone) repeated-measures ANOVAs. Potential topographical differences for N1- or P2 attenuation-modulatory effects (between-condition attenuation differences) were assessed as in Experiment 1. To assess whether the experiment had sufficient power to detect differences in N1 and P2 attenuations equaling the mean between-condition attenuation differences measured in Experiment 1, the respective attenuation-difference measured in Experiment 1 was divided by the respective standard deviation of the attenuation-difference measured in Experiment 2. The resulting effect-size for each component was then used to calculate power. Furthermore, attenuation difference magnitudes detectable with 95% power were also calculated.

Results

Participants complied with the instructions. The between-action interval in the Tap condition was 3745 ± 333 ms in the Motor-Auditory and 4130 ± 347 ms in the Motor arrangements. In the No Tap condition it was 3849 ± 325 ms in the Motor-Auditory and 4268 ± 388 ms in the Motor arrangements. The Condition (Tap vs. No Tap) × Stimulation

15

Arrangement (Motor-Auditory vs. Motor) ANOVA showed a significant Condition main effect: $F(1,18) = 4.864$, $\eta_G^2 = 0.030$, $p = 0.041$; and a Stimulation Arrangement main effect: $F(1,18) = 20.071$, $\eta_G^2 = 0.259$, $p < 0.001$. That is, participants kept a slower pace in No Tap than in the Tap conditions, and in the Motor than in the Motor-Auditory arrangements.

Most tone counting errors were off-by-one errors, with a couple of instances of larger deviations. The Condition (Tap vs. No Tap) × Stimulation Arrangement (Motor-Auditory vs. Auditory) ANOVA of the number of erroneous counts showed only a tendency for more errors in the Motor-Auditory than in the Auditory arrangement: $F(1,18) = 3.729$, $\eta_G^2 = 0.081$, $p = 0.069$ (overall, 29 counting errors occurred in the Motor-Auditory, and 14 in the Auditory arrangements out of a total of 76 experimental blocks).

The number of epochs contributing to the average ERPs were 85 ± 15, 86 ± 15, and 93 ± 7, in the Motor-Auditory, Motor, and Auditory arrangements, respectively, in the Tap; and 86 ± 19, 90 ± 12, and 93 ± 7, in the Motor-Auditory, Motor, and Auditory arrangements, respectively, in the No Tap condition.

Tones elicited the N1 and P2 waveforms (Figure 4) with peaks occurring at the expected latencies and electrode positions documented in the literature (Näätänen and Picton, 1987). For tones in the Auditory arrangement, N1 peaked at 95 ms at FCz in the Tap and at 96 ms in the No Tap condition (the amplitudes - average signal in the 86-106 ms range at FCz – were -4.435 ± 1.677, and 4.296 ± 1.979 μV in the Tap and No Tap conditions, respectively). The Condition (Tap vs. No Tap) × Event (corrected coincidence vs. tone) ANOVA of the N1 amplitudes showed a significant Event main effect: $F(1,18) = 34.417$, $\eta_G^2 = 0.114$, $p < 0.001$, indicating attenuated N1 response (by 1.273 μV) to the self-induced tones. (There was also a tendency for a Condition × Event interaction: $F(1,18) = 3.930$, $\eta_G^2 = 0.009$, $p = 0.063$, hinting at the possibility of stronger attenuation in the Tap, than in the No Tap condition). The effect

size corresponding to the between-condition attenuation-difference in Experiment 1 (1.229 μV) was d = 0.853, and the power to detect such a difference was 0.940. The mean attenuation difference detectable with 95% power was 1.261 μV in this arrangement.

P2 peaked at Cz at 170 ms in the Tap, and 168 ms in the No Tap condition (the amplitudes - average signal in the 159-179 ms range at Cz – were 4.294 ± 2.060, and 4.220 ± 1.699 μV in the Tap and No Tap conditions, respectively). The Condition × Event ANOVA of the amplitudes showed a significant Condition main effect: $F(1,18) = 5.270$, $\eta_G^2 = 0.015$, p = 0.034; a significant Event main effect: $F(1,18) = 29.640$, $\eta_G^2 = 0.162$, p < 0.001; and a Condition × Event interaction: $F(1,18) = 4.560$, $\eta_G^2 = 0.021$, p = 0.047, indicating stronger attenuation in the Tap, than in the No Tap condition (by 0.963 μV) . The effect size corresponding to the between-condition attenuation-difference in Experiment 1 (1.708 μV) was d = 0.850, and the power to detect such a difference was 0.938. The mean attenuation difference detectable with 95% power was 1.760 μV in this arrangement. The corrected coincidence and tone ERPs amplitudes significantly differed in the No Tap condition: $t(18) = 2.921$, p = 0.009, showing that despite the modulatory effect of the Condition factor, a significant 0.964 μV attenuation was still present.

The topographical comparison of the mean tone ERP and mean attenuation effect in the N1 range (Figure 5, top) showed no significant Signal × Electrode interaction: $F[3,54] = 1.135$, $\varepsilon = 0.620$, p = 0.330), providing no evidence for a different ERP component causing the ERP attenuation. In the P2 range (Figure 5, bottom), comparing the mean tone ERP with the difference of the attenuations showed no significant Signal × Electrode interaction: $F[3,54] = 0.198$, $\varepsilon = 0.621$, p = 0.806), providing no evidence for a different ERP component causing the modulatory effect of the Condition factor. Comparing the mean tone ERP and the attenuation effect in the No Tap condition in the P2 time interval, however, showed significant Signal × Electrode interaction: $F[3,54] = 4.494$, $\varepsilon = 0.588$, $\eta_G^2 = 0.072$, p = 0.023,

indicating that the attenuation-effect involved ERP component(s) other than the tone-related P2.

Interim Discussion - Experiment 2

As in previous studies utilizing button-press-contingent stimulation protocols, the estimated auditory N1 and P2 amplitudes were reduced for tones coinciding with (and also elicited by) the finger-movements in comparison to those elicited by tones only listened to. Conforming to the hypothesis, when finger-movements did not result in mechanical contact with an external object (No Tap condition), the magnitude of attenuation significantly decreased for P2, and there was a tendency for a similar N1 attenuation-decrease as well. Similarly to the results of Experiment 1, this suggests that mechanical impact played a substantial role in the observed ERP attenuation pattern. Similarly to Experiment 1, the observed action-related ERP attenuation was not exclusively related to the mechanical impact, because a significant N1 and P2 attenuation was still present in the No Tap condition. The topographical difference between the tone-related P2 and the attenuation effect in the P2 time range supports the notion that ERP attenuation effects measured in contingent paradigms reflect more, probably task dependent (Saupe et al., 2013) ERP effects, possibly superimposed on genuine N1 or P2 modulation.

Although participants were instructed to refrain from forcefully hitting the plate, tapping may have still caused faint, but audible transient sounds as evidenced by the intensity measurements presented in the Methods section of Experiment 1. That the N1 attenuation-decrease in the No Tap condition did not reach significance in contrast with Experiment 1, might have been caused by this instruction. Because of the potential presence of such transients, interpreting these effects is not trivial, as discussed in the General Discussion section.

The goal of the present study was to assess whether mechanical interaction resulting from the action played a role in action-related auditory (ERP) attenuation. The results of Experiment 1 and 2 suggest that mechanical impact does play a role, but due to the potentially confounding effect of transient sounds generated by tapping, this effect cannot be unequivocally attributed to attention. To eliminate this confound, in Experiment 3 participants maintained continuous contact with an object and applied force impulses from time-to-time (Press condition), or produced finger-movements without mechanical contact (No Tap condition identical to that in Experiment 1 and 2, see Figure 1). Because applying a force impulse did not result in displacement, no action-related sound was generated. It was hypothesized, as in Experiment 1 and 2, that the availability of tactile feedback (Press condition) would compel participants to allocate attention to the tactile stimulation as the action was performed, which would lead to a stronger action-related reduction of the auditory ERPs than in the No Tap condition, in which tactile feedback was not available.

Experiment 3

Materials & Methods

Participants

21 young adults participated in Experiment 3, one of whom got tired during the experiment and could not perform the task, so the final sample included data from 20 participants (nine women, aged 18-28 years, mean 23 years, all right-handed) for monetary compensation. They reported normal hearing and no history of neurological disorders, and gave written informed consent after the experimental procedures were explained to them.

Stimuli and procedures

The experiment was very similar to Experiment 1. In the following, only the differing details are described. There was a No Tap condition identical to that in Experiments 1 and 2 (Figure 1, left column); and there was a Press condition, in which participants were instructed to put their index finger on a piezoelectric element and maintain contact throughout the experimental block (Figure 1, right column). The center of the piezoelectric element was marked with a small bump. They were instructed to apply short pressure impulses without moving their fingers once every 4 s. The pressure impulses did not produce sound. Each condition was administered in six blocks, each lasting 5 minutes. The blocks were presented in an interwoven ("PNNPPNN…", in which "P" denotes a block from the Press condition, and "N" denotes a block from the No Tap condition) order, with the starting condition counterbalanced between participants. As in Experiment 2, a curtain was hanged above the arm, between the participant's head and hand, which occluded the sight of the hand and the frame or piezoelectric element. The temporal adjustment (which was 300 ms in Experiment 1) was set to 500 ms. That is, tones scheduled to be delivered within 500 ms after an action were delivered immediately (an action-tone coincidence); tones scheduled to be delivered between 500 and 1000 ms were presented at 500 ms, tones scheduled to be delivered between 1000 and 1500 ms were presented at 1000 ms, and so on. This change in temporal adjustment would result in higher coincidence rates, and thus allow the recording of more coincidence epochs than in Experiment 1, which would result in enhanced signal-to-noise ratio, and more power to the detect attenuation- and attenuation-modulatory effects in the repeated-measures design. The ratio of coincidences would be lower than 10%, that is, coincidences would still be relatively rare.

EEG-recording and analysis

EEG recording and analysis was identical to that of Experiment 1, but in order to further enhance signal-to-noise ratio, eye-movement correction was applied on the basis of a

two-minutes-long EEG recording (registered at the beginning of the session) featuring various eye-movements as described by Schlögl, Keinrath, Zimmermann, Scherer, Leeb, and Pfurtscheller (2007). To accommodate eye-movement correction into the processing chain, the EEG was off-line 1 Hz high pass filtered before eye-movement correction, and then 20 Hz low pass filtered. Also, because of the 500 ms adjustment window, tone ERPs were calculated as the ERP elicited by tones following an action by at least 1500 ms (and no action following within 400 ms).

Power calculations as described in the Methods section for Experiment 2 were performed to assess whether Experiment 3 had sufficient power to detect differences in N1 and P2 attenuation equaling the mean between-condition difference measured in Experiment 1. Attenuation-differences detectable with 95% power are also reported.

Because of its similarity to Experiment 1, further between-experiment comparisons were conducted separately for the N1 and P2 amplitudes measured in the Tap condition in Experiment 1 and the Press condition in Experiment 3, that is, amplitudes were analyzed in Experiment (Experiment 1 Tap vs. Experiment 3 Press, as between groups factor) × Event (corrected coincidence vs. tone, as repeated measures factor) ANOVAs. The main question in these analyses was whether an interaction would be found, which would suggest that although both tapping and applying more force involved a tactile change, differences in some other aspects of the two actions played a substantial role in the measured attenuations as well.

Because no significant N1 attenuation was found in the No Tap condition of Experiment 1, but a significant N1 attenuation was found in Experiment 3 (see below), between-experiment comparisons (Experiment × Event ANOVAs, as described above) were also conducted separately for the N1 and P2 amplitudes measured in the No Tap conditions of Experiment 1 and 3.

Results

Participants complied with the instruction. The mean between-action interval was 3808 ± 494 ms in the Press, and 3925 ± 417 ms in the No Tap condition (no significant difference: $t[19] = 1.140$, $p = 0.268$). The sound-coincidence rate was lower in the No Tap (8.7 ± 1.2 %) than in the Press (9.4 ± 1.1 %) condition: $t[18] = 2.171$, $p = 0.043$. The ratio of intervals excluded from the analyses due to potentially undetected actions did not differ between conditions ($t[18] = 1.052$, $p = 0.306$); this affected 3.1 ± 8.3 % of all between-action intervals.

The number of epochs contributing to coincidence ERPs was 35 ± 7 in the No Tap, and 39 ± 7 in the Press condition. The number of epochs contributing to the tone ERPs was 169 ± 31 in the No Tap and 176 ± 30 in the Press condition.

For tones (at least 1500 ms after an action), N1 peaked at 105 ms in the No Tap and at 107 ms in the Press condition, (both) at Cz (Figure 6, the amplitudes - average signal in the 96-116 ms range at Cz – were 4.330 ± 1.938, and 4.372 ± 1.866 µV in the Press and No Tap conditions, respectively). The ANOVA of the amplitudes showed a significant Event main effect only: $F(1,19) = 9.985$, $\eta_G^2 = 0.022$, $p = 0.002$ (for all other effects $F < 1$), indicating 0.557 µV lower (less negative) ERP amplitude for corrected coincidence than for tones. The effect size corresponding to the between-condition difference in attenuation measured in Experiment 1 (1.229 µV) was $d = 1.426$, and the power to detect such a difference was >0.999. The mean attenuation difference detectable with 95% power was 0.733 µV in this arrangement. The corrected coincidence and tone ERPs amplitudes were compared in both conditions, and significant differences were found: $t(19) = 2.859$, $p = 0.010$, attenuation of 0.602 µV; and $t(19) = 2.683$, $p = 0.015$, attenuation of 0.511 µV, respectively in the No Tap and Press conditions.

For tones, P2 peaked at 185 ms at FCz in the No Tap, and at 182 ms at Cz in the Press condition (the amplitudes - average signal in the 174-194 ms range at Cz - were 3.925 ± 2.120, and 3.776 ± 2.156 µV in the Press and No Tap conditions, respectively). The ANOVA of the amplitudes showed a significant Event main effect only: $F(1,19) = 6.504$, $\eta_G^2 = 0.020$, $p = 0.026$ (for all other effects $F < 1$), indicating 0.618 µV lower (less positive) ERP amplitude for corrected coincidences than for tones. The effect size corresponding to the between-condition attenuation-difference in Experiment 1 (1.708 µV) was $d = 1.594$, and the power to detect such a difference was >0.999. The mean attenuation difference detectable with 95% power was 0.911 µV in this arrangement. The corrected coincidence and tone ERP amplitudes were compared in both conditions. A significant difference was only found in the Press condition ($t[19] = 3.250$, $p = 0.004$, attenuation of 0.777 µV; in the No Tap condition: $t[19] = 1.540$, $p = 0.140$, attenuation of 0.460 µV).

The topographical comparison of the mean tone ERPs and mean attenuation effects showed no significant Signal × Electrode interactions (N1: $F[3,57] = 1.140$, $\varepsilon = 0.561$, $p = 0.324$; P2: $F[3,57] = 0.962$, $\varepsilon = 0.527$, $p = 0.375$) providing no evidence for a different ERP component causing the attenuation effects of the Event factor. An unexpected finding observable at the topography of the attenuation effects (Figure 7) is the attenuation of the positive aspect of the N1 at the temporal (T3/T4) sites (the mean amplitudes were 2.866 ± 1.571; and 2.994 ± 1.353 µV, in the Press and No Tap conditions, respectively). The post-hoc Condition (No Tap vs. Press) × Event (corrected coincidence vs. tone) × Electrode Side (T3 or T4) repeated-measures ANOVA showed a significant main effect of Event only: $F(1,19) = 13.998$, $\eta_G^2 = 0.025$, $p = 0.001$ (for all other effects $F < 1$), indicating a 0.538 µV lower (less positive) ERP amplitude for corrected coincidences than for tones. The topographical comparison of the mean tone- and mean corrected-coincidence ERPs using the T3, T4, TP7, TP8, and the signals from the left and right mastoids (Lm and Rm respectively) showed

significant Signal × Electrode interaction: $F(5,95) = 3.297$, $\varepsilon = 0.555$, $\eta_G^2 = 0.037$, $p = 0.023$, indicating that the tone N1 and the attenuation-effect topographies differed in their shapes.

The Experiment (Experiment 1 Tap vs. Experiment 3 Press) × Event (corrected coincidence vs. tone) ANOVA for the N1 amplitudes showed a significant Event main effect only: $F(1,37) = 17.384$, $\eta_G^2 = 0.033$, $p < 0.001$, showing lower (less negative) ERP amplitude for corrected coincidences than for tones (Experiment main effect: $F<1$; Experiment × Event interaction: $F[1,37] = 1.860$, $\eta_G^2 = 0.004$, $p = 0.181$).

For the P2 amplitudes the same type of ANOVA yielded a significant Event main effect: $F(1,37) = 40.309$, $\eta_G^2 = 0.136$, $p < 0.001$, and an Experiment × Event interaction: $F(1,37) = 11.854$, $\eta_G^2 = 0.044$, $p = 0.001$, indicating a stronger P2 attenuation in the Tap condition of Experiment 1 than in the Press condition of Experiment 3 (the Experiment main effect was not significant: $F[1,37] = 1.799$, $\eta_G^2 = 0.040$, $p = 0.188$).

The Experiment (Experiment 1 No Tap vs. Experiment 3 No Tap) × Event (corrected coincidence vs. tone) ANOVA for the N1 amplitudes showed only a significant interaction: $F(1,37) = 4.682$, $\eta_G^2 = 0.009$, $p = 0.037$, showing that N1-attenuation was stronger in Experiment 3 than in Experiment 1 (Experiment main effect: $F<1$; Event main effect: $F[1,37] = 1.077$, $\eta_G^2 = 0.002$, $p = 0.306$).

For the P2 amplitudes the same type of ANOVA yielded a significant Event main effect only: $F(1,37) = 4.878$, $\eta_G^2 = 0.018$, $p = 0.012$, showing lower (less positive) P2 amplitude for corrected coincidences than for tones ($F<1$ for both the Experiment main effect, and the interaction).

Interim Discussion - Experiment 3

The estimated auditory N1 and P2 amplitudes were reduced for tones coinciding with finger-movements or pressure impulses in comparison to tones following such actions by at least 1500 ms. No significant between-condition difference in the action-related attenuations were found, despite the increased power resulting from the introduction of the eye-movement correction and the stronger temporal adjustment. This result does not support the hypothesis that the action-related attenuation of auditory ERPs is brought about by allocating attention to the tactile feedback related to performing the action. The comparison of the N1 and P2 attenuations observed in the Tap condition of Experiment 1 and the Press condition of Experiment 3 showed that P2 attenuation was significantly stronger in the Tap condition of Experiment 1. The comparison of the N1 and P2 attenuations in the No Tap conditions of Experiment 1 and 3 showed that N1-attenuation was stronger in Experiment 3.

General Discussion

The goal of the present study was to explore the potential role of mechanical interaction in the action-related auditory ERP attenuation reported in the literature. It was hypothesized that having a reliable mechanical interaction with an external object as part of the given action compelled participants to allocate attention to the tactile feedback as the action was performed, which would draw processing resources away from the simultaneously presented auditory stimuli, and lead to attenuated N1 and P2 auditory ERPs. The contribution of the auditory ERP to the action-tone coincidence-related ERP was estimated by subtracting the action-related ERP registered when the action did not co-occur with the tone.

In Experiments 1 (coincidence paradigm) and 2 (contingent paradigm), mechanical contact was established by tapping on an external object at the endpoint of a finger-movement. Both the coincidence and the contingent arrangements replicated the results of previous studies: auditory processing activity was attenuated when the tones coincided with a

25

finger-movement. Moreover, this attenuation as reflected by the auditory N1 and P2 ERPs was reduced when finger-movements did not result in mechanical contact with an external object in Experiment 1, and a reduced P2 attenuation (and a hint at a reduced N1 attenuation) was obtained in Experiment 2. These results clearly show that mechanical interaction played a substantial role in the finger-movement-related auditory ERP attenuation reported in the literature. Because mechanical impact inevitably causes faint, but audible transient sounds (see below for a detailed discussion), in Experiment 3 participants continuously maintained contact with the external object, and applied pressure impulses from time to time, which did not result in displacement and transient sounds. Although auditory ERPs were attenuated as in Experiment 1 and 2, applying pressure impulses did not result in stronger N1 or P2 attenuations in comparison to the condition in which no mechanical contact was made.

Whereas the pattern of results shows that mechanical impact, but not the mere presence of mechanical contact played an important role in finger-action-related auditory ERP attenuation, connecting this technical description to the hypothesis motivating the present study, that is, that auditory ERP attenuation was caused by shifting attention to the action-contingent tactile stimulation, is not trivial. Indeed, as discussed in the following, the manipulation of action-contingent mechanical interaction in the present paradigm allows for a number of interpretations regarding the cause of this pattern.

The first two interpretations are based on the fact that movements resulting in mechanical impact lead to the elicitation of potentially audible transient sounds. As evidenced by the measurements described in the Methods of Experiment 1, tapping may have produced faint, but probably audible sounds.

(1) The first interpretation suggests that the ERP attenuation effects are ERP artifacts stemming from this methodological shortcoming. One of the basic difficulties in measuring

action-related auditory ERP attenuation is to estimate the auditory ERP contribution within the action-tone coincidence ERP waveform. The most widely used approach – which was used in the present study as well - is to subtract an "action-only" waveform from the coincidence waveform. Whereas this is a plausible approach, it may cause problems if the actions generate auditory impact-sound transients, because the "action only" waveform will actually be a superposition of the motor- and the impact-sound-ERP (for clarity, in the following *probe sound* refers to the tone presented by the experimenter, and *impact-sound* refers to the faint sound produced by the mechanical interaction with the device). When the action coincides with the probe sound, the contribution of the impact-sound to the ERP may be negligible due to masking, because the probe is much louder. In contrast, when the impact-sounds occur during periods of silence (far away from the probe sounds, or in a different experimental block), low-amplitude auditory ERPs may be elicited by the unmasked impact-sound (superimposed on the motor ERP). When estimating the auditory contribution to the coincidence ERP, the subtraction of this low-amplitude impact-sound ERP may artificially decrease the auditory ERP amplitude for the probe. Note that the presence or absence of the superimposed impact-sound related ERP to the motor ERP might not be visually salient in the ERP waveform at all. Obviously, this suggests that experiments in which actions produce faint sounds because of the nature of interaction with the response-device (e.g. tapping, or button clicks) overestimate the magnitude of auditory ERP attenuation. The finding that the estimated P2 attenuation was stronger in the Tap condition of Experiment 1 than in the Press condition of Experiment 3 fits this interpretation well (but note that no such difference was found for the N1). Note that whereas some studies (e.g. SanMiguel et al., 2013) used noise-attenuation techniques to counteract this type of confound, no study reported direct measurements whether these were in fact effective or not (the present study being no

exception: the effectiveness of the instruction to refrain from forcefully hitting the plate in Experiment 2 was not formally assessed).

Certain stimulation arrangements may be suitable for the elimination of the ERP artifact described above. For example, in contingent paradigms, introducing a delay between the action and the probe sound may allow one to temporally separate the probe-related ERP from the impact-related waveform. In such arrangements, however, one would have to control for refractoriness effects, which are known to affect the N1 (Näätänen & Picton, 1987). That is, the probe-related N1 might be elicited with lower amplitude in the Motor-Auditory arrangement than in the Auditory one because of the refractoriness caused by the impact sound.

(2) The second interpretation suggests that although the results do not reflect an ERP artifact as described above, they reflect prediction- or attention-based "filtering" related to the production of the impact-sound. This hypothesis suggests that action-related auditory attenuation (which may not only be reflected in ERP, but other measures as well) is "genuine" in the sense that it reflects a preparation for the *reliable* auditory consequence of the action, that is, the impact-sound. If there is a faint but audible sound produced by the mechanical interaction with the response device, then in the typical contingent or coincidence-based experimental designs, this is the most reliable auditory consequence of the action. In terms of the internal forward modeling framework the auditory attenuation would reflect an internal forward model predicting this action-contingent auditory stimulation, which would also lead to the attenuation of the probe tones. This would also readily explain why auditory attenuation occurs for probe-action coincidences in the absence of a contingent action-probe relationship (but see below). In terms of the attention-based framework, it could be speculated that participants adopt an attention set which is optimal for detection of the impact-sound (even

against the background of other sounds), which leads to attenuated activity when a different sound – the probe – also occurs.

The most important methodological implication of these two lines of thought is that controlling the auditory consequences of the actions must be adopted as a standard procedure in research on action-related auditory attenuation. Whereas one may attempt to acoustically isolate the participant from the action-induced auditory transients, controlling whether the isolation of these extremely short, wide-spectrum sounds was successful might require extensive acoustical measurements specific to the given settings. It seems plausible that the best practice involves the use of devices designed for silent operation (like the infrared light beam- or piezoelectric element-based arrangements used in the present study).

The next two speculations suggest that despite the potential confounds, the observed ERP-modulations nonetheless reflect attention- or forward-model-based cognitive processes. That is, although one *could* cause impact-sounds with the given action (i.e. potentially hit the given response-device with a momentum sufficient to produce an audible sound), typically this does not happen, and the measured attenuation-differences reflect other, genuine cognitive effects.

(3) One may speculate (post-hoc) that although tapping (Experiment 1 and 2) and applying pressure (Experiment 3) on a surface both result in tactile changes related to the performance of the action, the information provided by these sensory changes is markedly different in the two contexts . Whereas tapping establishes a new mechanical contact and thereby allows one to confirm the presence of the object and its position in relation to the hand (which was essential for the performing the task in Experiment 1 and 2), changing the force applied to a continuously touched object does not provide more information on these task-relevant aspects of the situation. That is, tapping may compel participants to shift

attention to the tactile feedback when the action is performed, because that is how the optimal interaction with the device can be monitored and maintained. In contrast, applying more force to the already touched rigid object does not allow for substantial information gain on the circumstances of the interaction. In different words, making contact with an object strongly reduces uncertainties regarding the (potential) interaction, whereas applying more pressure on an already touched rigid object does not. In similar vein, one may speculate that there might be a movement-, or movement-planning-related difference between moving the index finger in the air and tapping with it on a surface, which may result in the observed ERP attenuation patterns.

(4) A further speculative account suggests that although no audible transient sounds are generated on most trials, participants might still be aware of the possibility (or opportunity) to generate such sounds, and this "awareness" might be sufficient to produce a generalized expectation (see also Desantis, Hughes, & Waszak, 2012; Hughes, Desantis, & Waszak, 2013). By attention- or forward model-based mechanisms, this expectation may activate sensory representations of the producible sounds, which results in stronger attenuation than in arrangements in which the action (pressure application, Press condition of Experiment 3) or the lack of object (No Tap conditions) does not warrant such an expectation.

Each of these interpretations is compatible with the pattern of results, still, these lines of thought have some plausible methodological implications, and offer some ideas for future research. Although the methodological implications of the results are severe, they do not invalidate previous research, because they also demonstrate that the ERP attenuation effects reported in the literature are not (entirely) due to the action-related mechanical impact-sound confound. Actions with silent response devices resulted in ERP attenuations: The contingent arrangement in Experiment 2 resulted in ERP attenuations in the N1 and P2 time range (although the different topography of the attenuation effect in the P2 range suggested that the

P2 effect was not a pure P2-modulation). The coincidence paradigm in Experiment 3 showed significant N1 and P2-attenuations , and a significant P2-attenuation was found in Experiment 1. Interestingly, the between-experiment comparison showed that N1-attenuation was stronger in the No Tap condition of Experiment 3 than that of Experiment 1. Although this might be due to the better signal-to-noise ratio in Experiment 3 (brought about by the longer adjustment window and the eye-movement correction), or the difference in coincidence rate (5.7 vs. 8.7 %), one may also speculate that this difference may be related to the different contexts provided by the other conditions of the experiments. Although some evidence suggests that context may influence action-related ERP attenuation (i.e. Baess et al., 2011 found that N1-attenuation was stronger when the tones were self-induced among external tones), at this point it is not clear which aspect of the other conditions could cause such a modulation-difference.

Although the present study mainly focused at the fronto-central aspect of the N1 ERP, it has to be noted that the N1 waveform elicited by tones is not a unitary phenomenon, but the superposition of a number of sub-components (for an overview, see Näätanen & Picton, 1987). Unfortunately, only few studies speculated on the component-specificity of action-related N1-attenuation (Horváth et al., 2012, SanMiguel et al., 2013, Saupe et al., 2013), and a systematic investigation on sub-component attenuations has not been conducted yet. Studies utilizing the coincidence paradigm (e.g. Horváth et al., 2012) did not report significant N1 attenuation-effects at electrode sites where the N1 polarity is typically reversed for pure tones (that is, below the Sylvan fissure with a maximum at the mastoids when the EEG is recorded with nose-, or average reference, suggesting a supra-temporal source, Vaughan & Ritter, 1970). The lack of such a polarity reversal also observed in some studies using contingent stimulation arrangements was interpreted as a sign that action-related attenuation might mainly (but not exclusively) affect the non-specific N1 subcomponent (SanMiguel et al., 2013), which shows no such reversal. Interestingly and unexpectedly, in Experiment 3 of the

present study, the N1-attenuation did show such a reversal. The effect, however, was not maximal at the mastoids, but at the T3 and T4 sites as evidenced by post-hoc analyses. This reversed attenuation effect therefore might not reflect the attenuation of the supra-temporal N1, but the attenuation of the positive Ta subcomponent of the so-called T-complex (Wolpaw & Penry, 1975). Whereas action-related attenuation of the Tb subcomponent (a negative peak observable typically at around 140 ms at temporal sites) have been observed (Horváth, 2013a; SanMiguel, et al., 2013; Saupe, et al., 2013), the results of Experiment 3 suggest that Ta might also be attenuated. Naturally, being a post-hoc finding, this interpretation should be handled with caution.

In summary, the present results demonstrate that using actions potentially resulting in mechanical impact sounds may bias the measurement of action-related auditory (ERP) attenuation. In the worst case scenario, the measured ERP attenuations may be artifacts resulting from the improper estimation of the auditory contribution to action-sound coincidence ERPs. Even in the absence of a direct ERP confound, measured action-related sensory attenuations (measured by ERPs or other methods) may reflect processing related to the impact sound instead of the probe sound presented by the experimenter. Importantly, however, the present study also demonstrated that when such confounds were eliminated by the use of silent response devices, tones coinciding with actions still elicited attenuated auditory ERPs .

Acknowledgements

References

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. Psychophysiology, 48(9), 1276–1283. doi:10.1111/j.1469-8986.2011.01196.x

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. Behavior Research Methods, 37(3), 379–384. doi:10.3758/BF03192707

Brown, H., Friston, K., & Bestmann, S. (2011). Active Inference, Attention, and Motor Preparation. Frontiers in Psychology, 2. doi:10.3389/fpsyg.2011.00218

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, N.J: L. Erlbaum Associates.

Crapse, T. B., & Sommer, M. A. (2008a). Corollary discharge across the animal kingdom. Nature Reviews Neuroscience, 9(8), 587–600. doi:10.1038/nrn2457

Crapse, T. B., & Sommer, M. A. (2008b). Corollary discharge circuits in the primate brain. Current Opinion in Neurobiology, 18(6), 552–557. doi:10.1016/j.conb.2008.09.017

Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: a short review. Journal of Neural Engineering, 2(3), S313–S319. doi:10.1088/1741-2560/2/3/S11

Desantis, A., Hughes, G., & Waszak, F. (2012). Intentional binding is driven by the mere presence of an action and not by motor prediction. PLoS ONE, 7(1), e29557. doi:10.1371/journal.pone.0029557

Desantis, A., Weiss, C., Schütz-Bosbach, S., & Waszak, F. (2012). Believing and perceiving: Authorship belief modulates sensory attenuation. PLoS ONE, 7(5), e37959. doi:10.1371/journal.pone.0037959

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. Psychophysiology, 44(4), 522–529. doi:10.1111/j.1469-8986.2007.00533.x

Galazky, I., Schütze, H., Noesselt, T., Hopf, J.-M., Heinze, H.-J., & Schoenfeld, M. A. (2009). Attention to somatosensory events is directly linked to the preparation for action. Journal of the Neurological Sciences, 279(1-2), 93–98. doi:10.1016/j.jns.2008.12.006

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(3), 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(4108), 177–180. doi:10.1126/science.182.4108.177

Holst, E., & Mittelstaedt, H. (1950). Das reafferenzprinzip. Naturwissenschaften, 37(20), 464–476.

Horváth, J. (2013a). Action-sound coincidence-related attenuation of auditory ERPs is not modulated by affordance compatibility. Biological Psychology, 93(1), 81–87. doi:10.1016/j.biopsycho.2012.12.008

Horváth, J. (2013b). 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

Horváth, J., & Burgyán, A. (2013). No evidence for peripheral mechanism attenuating auditory ERPs to self-induced tones. Psychophysiology, 50(6), 563–569. doi:10.1111/psyp.12041

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(9), 1919–1931. doi:10.1162/jocn_a_00215

Horváth, J., & Winkler, I. (2010). Distraction in a continuous-stimulation detection task. Biological Psychology, 83(3), 229–238. doi:10.1016/j.biopsycho.2010.01.004

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. Psychological Bulletin, 139(1), 133–151. doi:10.1037/a0028566

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(3), 698–706. doi:10.1162/jocn_a_00167

Makeig, S., Mueller, M. M., & Rockstroh, B. (1996). Effects of voluntary movements on early auditory brain responses. Experimental Brain Research, 110(3), 487–492.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. Cerebral Cortex, 15(3), 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(6), 581–590. doi:10.1111/j.1469-8986.1976.tb00885.x

McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 62(3), 203–208. doi:10.1016/0168-5597(85)90015-2

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. Neural Networks, 9(8), 1265–1279.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. Psychophysiology, 24(4), 375–425. doi:10.1111/j.1469-8986.1987.tb00311.x

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: a problem of significance. Nature Neuroscience, 14(9), 1105–1107. doi:10.1038/nn.2886

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J.-M., Hinrichs, H., … Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. Electroencephalography and Clinical Neurophysiology, 106(3), 259–261. doi:10.1016/S0013-4694(97)00106-5

Okamoto, H., Stracke, H., Wolters, C. H., Schmael, F., & Pantev, C. (2007). Attention improves population-level frequency tuning in human auditory cortex. Journal of Neuroscience, 27(39), 10383–10390. doi:10.1523/JNEUROSCI.2963-07.2007

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. Psychological Methods, 8(4), 434–447. doi:10.1037/1082-989X.8.4.434

R Core Team. (2013). R: A language and environment for statistical computing. (Version

    3.0.0). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from

    http://www.R-project.org/

SanMiguel, I., Todd, J., & Schröger, E. (2013). Sensory suppression effects to self-initiated

    sounds reflect the attenuation of the unspecific N1 component of the auditory ERP.

    Psychophysiology, 50, 334–343. doi:10.1111/psyp.12024

Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression

    of self-generated sounds and its dependence on attention. International Journal of

    Psychophysiology, 90(3), 300–310. doi:10.1016/j.ijpsycho.2013.09.006

Schafer, E. W. P., & Marcus, M. M. (1973). Self-Stimulation Alters Human Sensory Brain

    Responses. Science, 181(4095), 175–177. doi:10.1126/science.181.4095.175

Schlögl, A., Keinrath, C., Zimmermann, D., Scherer, R., Leeb, R., & Pfurtscheller, G. (2007).

    A fully automated correction method of EOG artifacts in EEG recordings. Clinical

    Neurophysiology, 118(1), 98–104. doi:10.1016/j.clinph.2006.09.003

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

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by

    visual inversion. Journal of Comparative & Physiological Psychology, 43(6), 482–489.

Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for

    self-initiated sounds is independent of attention. BMC Neuroscience, 14(1), 2.

    doi:10.1186/1471-2202-14-2

Vaughan Jr, H. G., & Ritter, W. (1970). The sources of auditory evoked responses recorded

from the human scalp. Electroencephalography and Clinical Neurophysiology, 28(4), 360–

367.

Weiss, C., Herwig, A., & Schütz-Bosbach, S. (2011). The self in action effects: Selective

attenuation of self-generated sounds. Cognition, 121(2), 207–218.

doi:10.1016/j.cognition.2011.06.011

Wolpaw, J. R., & Penry, J. K. (1975). A temporal component of the auditory evoked

response. Electroencephalography and clinical neurophysiology, 39(6), 609–620.

doi:10.1016/0013-4694(75)90073-5
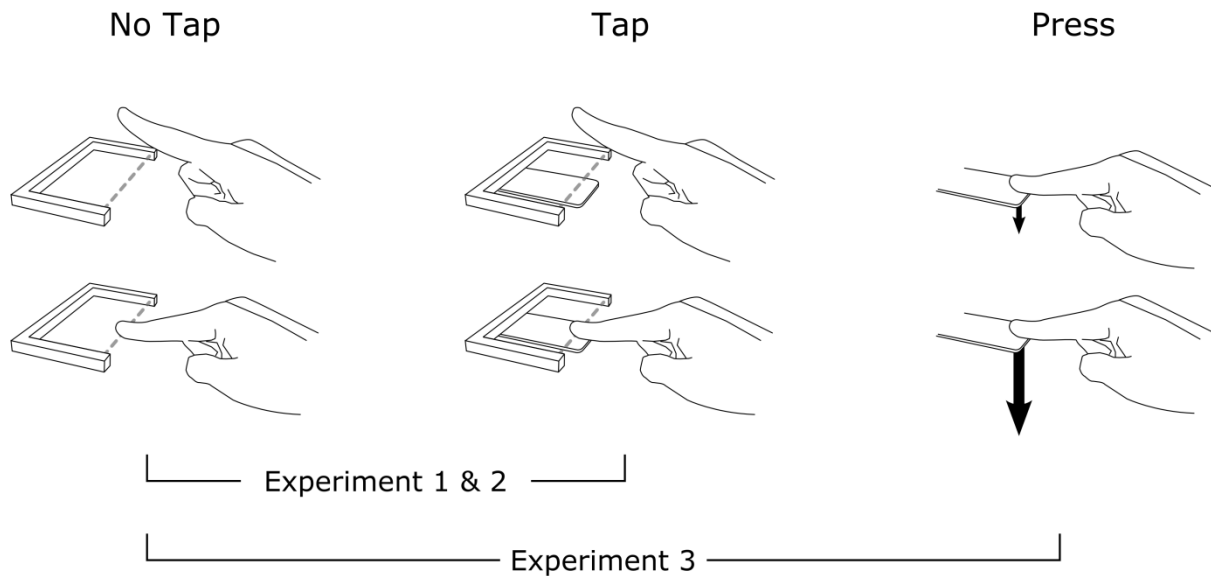
Figure 1. Action arrangements in the No Tap (left), Tap (middle), and Press (right) conditions. The top row shows the rest-, the bottom row the action phase. Participants interrupted the infrared (invisible) light beam by making a sharp downward-upward finger-movement in the No Tap and Tap conditions. In the Tap condition, the finger tapped on a plate positioned under the light beam. In the Press condition, participants maintained contact with a piezoelectric element and applied a short force impulse as an action. In Experiments 1 and 2 the No Tap and Tap conditions, in Experiment 3, the No Tap and Press conditions were administered.
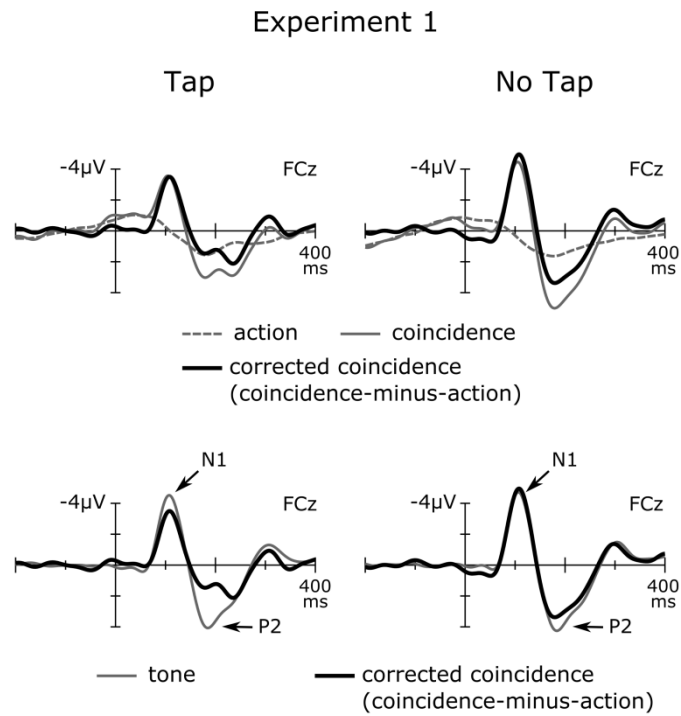
Figure 2. Group-average (N=19) ERP waveforms recorded in Experiment 1 at FCz in the Tap (left) and No Tap (right) conditions elicited by the respective actions, action-tone coincidences and the corresponding corrected coincidence waveforms (coincidence-minus-action difference waveforms; top row). Tone-related ERPs and corrected coincidence waveforms are contrasted in the bottom row.
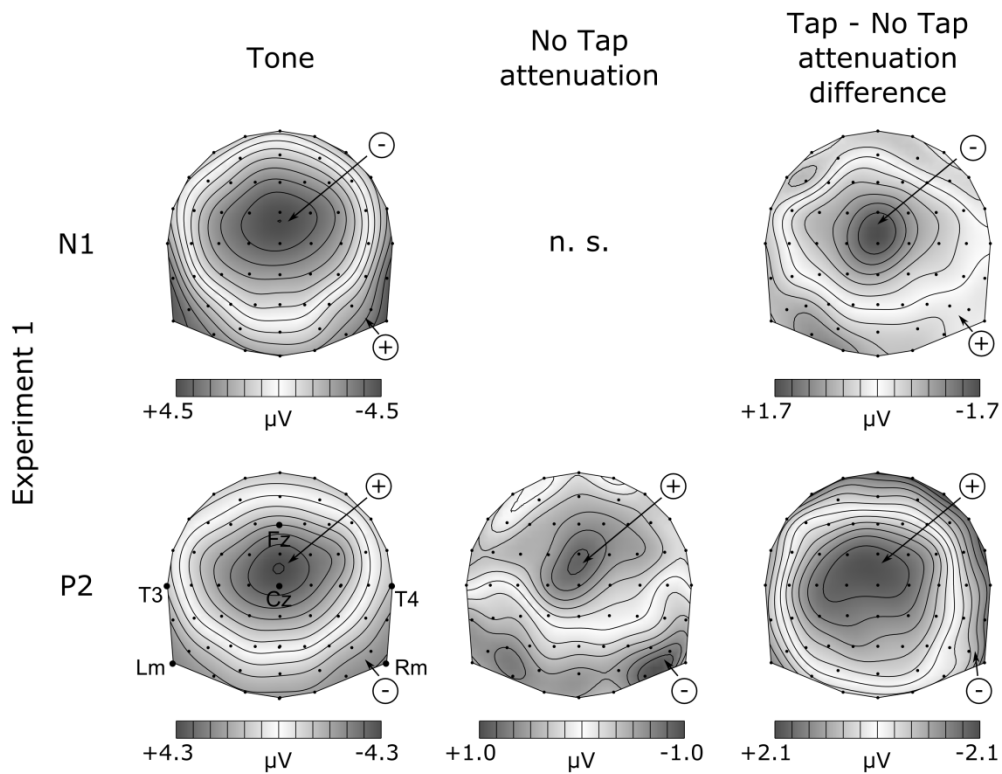
Figure 3. Group-average (N=19) topographical distributions of the N1 (top row) and P2 (bottom row) in Experiment 1. The left column shows topographies of ERPs elicited by tones (average signals in 20 ms long windows centered at the group-average peak latencies, averaged across conditions). The right column shows the corresponding attenuation-modulatory effect of the Condition factor, that is, the Tap-minus-No Tap attenuation differences. The middle column shows the attenuation effect (if significant) in the No Tap condition. Signal ranges differ between the individual topographical distributions in order to emphasize similarities or differences in shape. Positive and negative polarity areas are indicated by arrows.
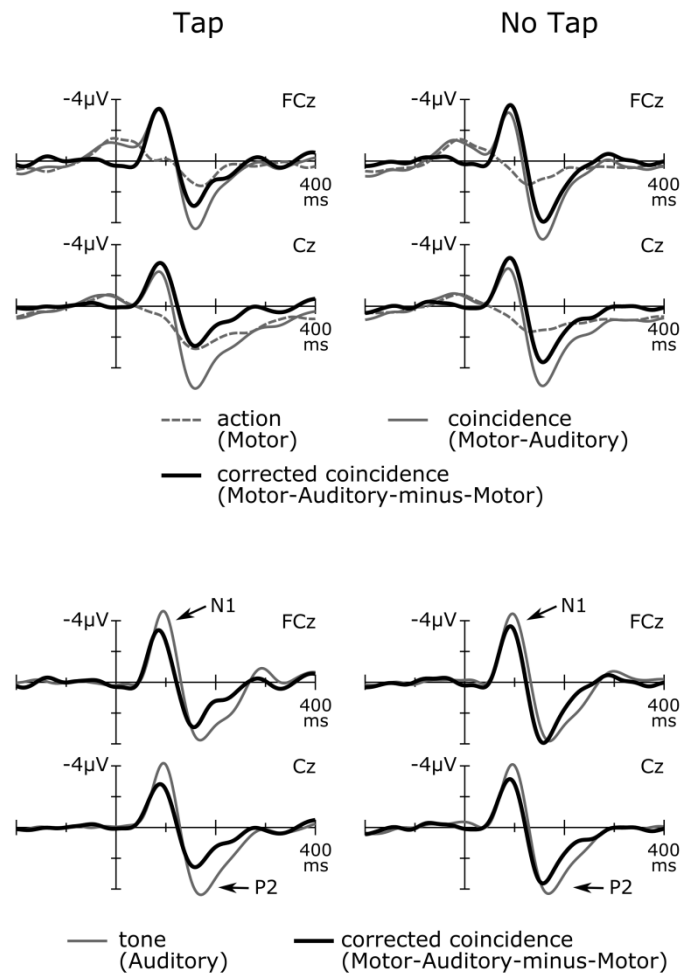
Figure 4. Group-average (N=19) ERP waveforms in Experiment 2 in the Tap (left) and No Tap (right) conditions elicited by the respective actions (obtained in the Motor arrangement), action-tone coincidences (obtained in the Motor-Auditory arrangement) and the corresponding corrected coincidence waveforms (coincidence-minus-action, that is, Motor-Auditory-minus-Motor difference waveforms; top panel). Tone-related ERPs (obtained in the Auditory arrangement) and corrected coincidence waveforms are contrasted in the bottom panel. Because N1 elicited by tones (in the Auditory arrangement) peaked at FCz, and P2 peaked at Cz, both signals are shown in separate rows.
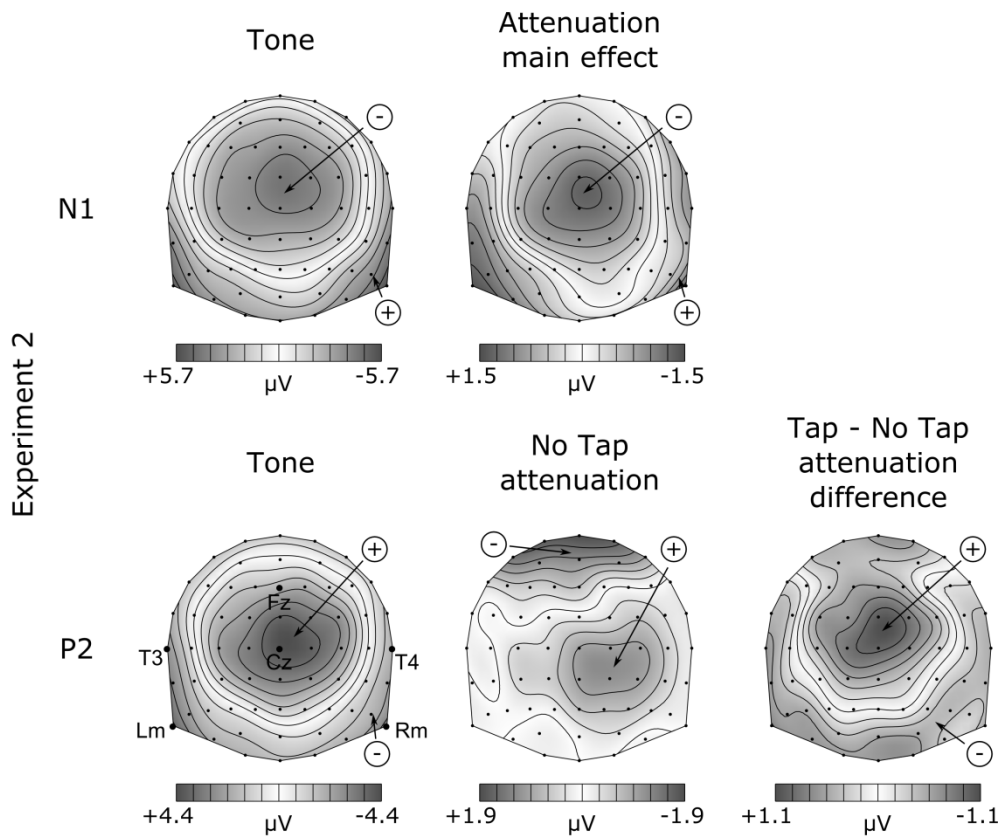
Figure 5. Group-average (N=19) topographical distributions of the N1 (top row) and P2 (bottom row) in Experiment 2. The left column shows topographies of ERPs elicited by tones (average signals in 20 ms long windows centered at the group-average peak latencies, averaged across conditions). In the top row, the middle column shows that attenuation main effect (tone-minus-corrected coincidence waveform averaged across conditions). In the bottom row the right column shows the attenuation-modulatory effect of the Condition factor (i.e., the Tap-minus-No Tap attenuation difference); the middle column shows the attenuation effect in the No Tap condition. Signal ranges differ between the individual topographical distributions in order to emphasize similarities or differences in shape. Positive and negative polarity areas are indicated by arrows.
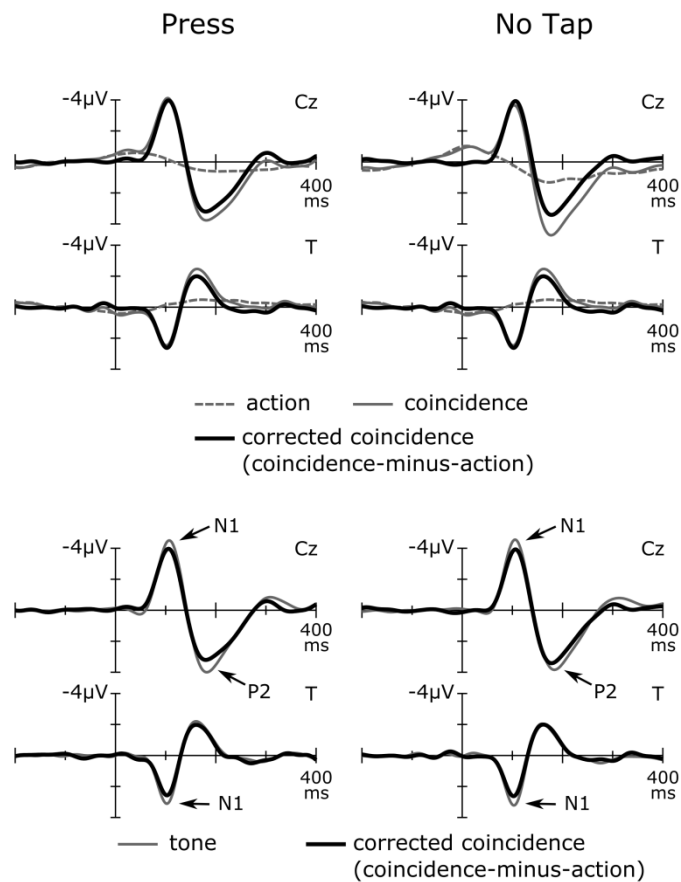
Figure 6. Group-average (N=20) ERP waveforms in Experiment 3 in the Press (left) and No Tap (right) conditions elicited by the respective actions, action-tone coincidences and the corresponding corrected coincidence waveforms (coincidence-minus-action difference waveforms; top panel). Tone-related ERPs and corrected coincidence waveforms are contrasted in the bottom panel. Because the N1 attenuation effect was also observable at the T3 and T4 electrodes (see Fig. 8) the average of these signals are shown in a separate row (marked by T) on both panels.
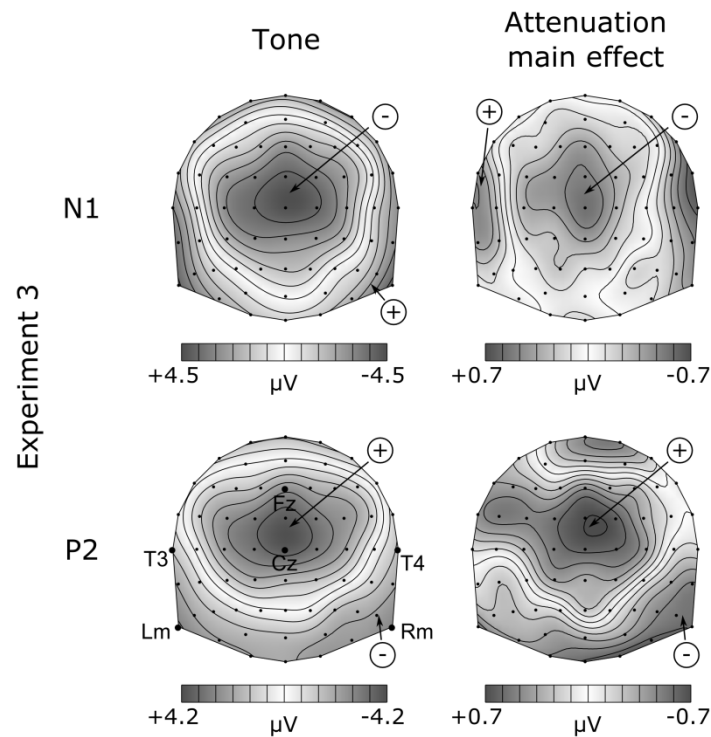
Figure 7. Group-average (N=20) topographical distributions of the N1 (top row) and P2 (bottom row) in Experiment 3 (average signals in 20 ms long windows centered at the tone group-average peak latencies). The left column shows topographies of ERPs elicited by tones (averaged across conditions). The right column shows the attenuation main effects (tone-minus-corrected coincidence waveform averaged across conditions). Signal ranges differ between the individual topographical distributions in order to emphasize similarities or differences in shape. Positive and negative polarity areas are indicated by arrows.