Action-related auditory ERP attenuation: Paradigms and hypotheses

János Horváth

Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

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

32 pages

Abstract

A number studies have shown that the auditory N1 event-related potential (ERP) is attenuated when elicited by self-induced or self-generated sounds. Because N1 is a correlate of auditory feature- and event-detection, it was generally assumed that N1-attenuation reflected the *cancellation of auditory re-afference*, enabled by the internal forward modeling of the predictable sensory consequences of the given action. Focusing on paradigms utilizing non-speech actions, the present review summarizes recent progress on action-related auditory attenuation. Following a critical analysis of the most widely used, *contingent* paradigm, two further hypotheses on the possible causes of action-related auditory ERP attenuation are presented. The *attention hypothesis* suggests that auditory ERP attenuation is brought about by a temporary division of attention between the action and the auditory stimulation. The *pre-activation hypothesis* suggests that the attenuation is caused by the activation of a sensory template during the initiation of the action, which interferes with the incoming stimulation. Although each hypothesis can account for a number of findings, none of them can accommodate the whole spectrum of results. It is suggested that a better understanding of auditory ERP attenuation phenomena could be achieved by systematic investigations of the types of actions, the degree of action-effect contingency, and the temporal characteristics of action-effect contingency representation-buildup and -deactivation.

1. Introduction

Hearing provides a constant stream of information about the events of the environment. Filtering this flood for relevant pieces of information is enabled by various functions of the human cognitive system. A number of studies show that the auditory system maintains a neural model representing the regularities of the auditory environment (Winkler, 2007; Bendixen, SanMiguel, Schröger, 2012), which produces predictions on forthcoming auditory events. These sensory predictions contribute substantially to the information filtering capability of the auditory system by calling capacity-limited (attention and cognitive control) processes only for stimuli which are incompatible with these predictions (Schröger, 1997). Auditory processing can also be influenced voluntarily: one may establish various selective attention sets which allow performing a given auditory tasks more efficiently when the task-relevant sound is presented, while suppressing task-irrelevant auditory input (see e.g. Hillyard, Hink, Schwent, & Picton, 1973; Okamoto, Stracke, Wolters, Schmael, & Pantev, 2007). Sounds, however, are not only generated by external sources. We move around, handle objects, talk, and perform various actions, which result in predictable sound events. A number of recent studies suggest that voluntary actions may directly influence auditory processing if the actions result in consistent, predictable patterns of auditory stimulation. These influences are mainly reflected by the attenuation of auditory event-related potentials (ERPs) elicited by such self-induced or self-generated sounds. Results generally regarded as fundamental pieces of evidence for theories on speech production (Hickok, 2012), or understanding sensory deficits in schizophrenia (Ford & Mathalon, 2012). The goal of the present review is to summarize recent progress on the measurement and interpretation of action-related auditory ERP attenuation.

Specifically, this review focuses on studies utilizing non-speech-producing actions (mainly finger movements). The main question in this line of research is whether performing the action influences the processing of concurrently (or temporally closely) presented auditory stimuli, and if it does, then what stages of processing are affected and through which mechanisms. Following a technical, non-interpretative description of the paradigm mainly used to address these questions, the typical results and the various hypotheses put forward to explain these findings are presented. Experiments supporting or challenging these hypotheses are presented in the context of the respective hypothesis. The review is concluded by a delineation of outstanding questions.

2. Measuring auditory processing activity in the presence of on-going action

Measuring auditory processing activity as the participant performs an action is not trivial. In the simplest case, physiological or behavioral responses to sound probes recorded in two conditions are compared: in one condition, the probe is presented concurrently with, in the other in the absence of the given action. Response differences are interpreted as reflections of action-related sound-processing changes.

Depending on the selectivity of the method, such a direct comparison may or may not yield results reflecting only the processing of the probe. Methods with high spatial resolution (e.g. electrocorticography) may allow one to characterize auditory processing activity in the auditory cortex alone, and the relative "blindness" of MEG to certain sources of brain activity (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993) may also allow a characterization of (some parts of) auditory processing without much interference from non-auditory activities. The most widely used ERP measures, however, generally do not allow a clear separation of auditory activity in the two conditions: non-auditory contributions to the ERP elicited by the combined action-sound event are likely to contaminate the results.

2.1 The contingent paradigm

Relying on a general assumption that the given method reflects only the processing of the auditory stimulus when the stimulus occurs concurrently with an action, may often lead to confounds. In order to eliminate such confounds, the most widely used paradigm (referred to as *contingent paradigm* in the following) features three conditions (McCarthy & Donchin, 1976), which are administered in separate experimental blocks:

1) In the Motor-Auditory condition, the participant voluntarily repeats an action from time-to-time according to the instructions. In this condition, each action leads to the presentation of a probe sound, and probe sounds are only initiated by the given action (i.e., the action-tone relationship is fully contingent).

2) In the Auditory condition, a sound probe sequence is presented, and the participant listens to these without performing any actions.

3) In the Motor condition, the participant voluntarily repeats an action according to the instructions, which are the same as in the Motor-Auditory condition, but no sound is presented.

To estimate the auditory ERP contribution to the ERP recorded in the Motor-Auditory condition, the ERP recorded in the Motor condition is subtracted from that recorded in the Motor-Auditory condition. This *motor-corrected* waveform is then compared to the one obtained in the Auditory condition. Typically, the comparisons show that ERPs are attenuated in the motor-corrected waveform.

Simple as it seems, the contingent arrangement has a number of limitations.

First, because auditory N1 ERP amplitude (which is the primary focus of this research) increases with increasing inter-sound interval (Davis, Mast, Yoshie, & Zerlin, 1966), it is important to make sure that inter-sound intervals are not longer in the Auditory than in the Motor-Auditory condition, because this could result in an *apparent* N1-attenuation in the critical comparison. Because the inter-sound intervals in the Motor-Auditory condition are inherently variable, most studies (e.g. Schafer & Marcus, 1973; Martikainen, Kaneko, & Hari, 2005; Baess, Widmann, Roye, Schröger, & Jacobsen, 2009) presented *replays* of the probe sound sequences generated by the same participant in the Motor-Auditory experimental blocks. On one hand, this is the optimal design choice, because the sequence of sound probes (and inter-sound intervals) is the same in the two conditions. On the other hand, this introduces a presentation order confound – the amplitude differences in the critical comparison could be attributed to the presentation order of the experimental blocks, because the Auditory (replay) blocks always follow the corresponding Motor-Auditory blocks.

Second, it needs to be decided what instructions the participants receive in the three conditions. Obviously, it is not possible to give participants the same task in the three conditions: in the Motor and Motor-Auditory conditions a sequence of actions is performed (e.g. press a button isochronously), but participants obviously cannot perform this task in the Auditory condition. In most studies, the instruction in the Auditory condition is "passive listening", that is, participants are instructed not to perform any specific task regarding the sounds. Because this difference in the task set may cause differences in the attentional focus of the participants (e.g. more or less attention towards the sounds, see below), this complicates the interpretation of the results. A further problem with "passive listening" is that what participants do (in terms of cognitive operations) during these periods can be highly variable, and cannot be monitored by the experimenter. Some recent studies attempted to make the conditions more similar in terms of cognitive operation and attentional focus by instructing participants to perform various sound-related tasks: Horváth & Burgyán (2013) and Horváth (2014c) instructed participants to count the sounds produced by the actions in the

Motor-Auditory condition, because only the first 60-70 actions produced a sound. Similarly, participants had to count the sounds in the Auditory condition (also a random number of sounds were presented, the replay of the first 50-60 sounds from the Motor-Auditory condition). Analyses showed that more miscounting errors were made in the Motor-Auditory condition, probably because the Motor-Auditory condition had been a dual-task (maintain an isochronous action pace and count), whereas the Auditory condition had been a single-task situation (count only). That is, although the explicit tasks allow the monitoring of the participants' cognitive behavior, they also introduce a mental-workload confound. Saupe, Widmann, Trujillo-Bareto & Schröger (2013) attempted to reduce between-condition differences by instructing participants to produce a sequence of button-presses with highly variable between-sound (i.e. between-action) intervals within a 1.8-5.0 s range in the Motor-Auditory condition. When listening to the replay, participants had to detect too short (< 1.8 s) or too long (> 5.0 s) between-sound intervals. This instruction ensures that participants attend the same temporal feature of the sequence in both conditions, therefore, presumably, reduces between-condition differences.

Third, when estimating the auditory ERP contribution to the Motor-Auditory ERP, it is assumed that the action-related activity is the same when the actions consistently lead to the presentation of the probe sound (Motor-Auditory condition), and when they do not (Motor condition). Although this may be true, currently, there is not much direct evidence for, or against this assumption. The most relevant result (Ford, Palzes, Roach, & Mathalon, 2014) is that the readiness potential preceding the action was more lateralized (De Jong, Wierda, Mulder, & Mulder, 1988; Gratton, Coles, Sirevaag, Eriksen, & Donchin, 1988; for a summary, see Eimer, 1998) when the button-pressing resulted in a tone than when it did not. Because the readiness potential reflects response-activation (or - rather - action-activation in the present context), this suggests that action-related activity probably differs between conditions.

Between-condition differences in action timing may also hint at action-related processing differences. In most studies, the instruction regarding action performance is the same in the two conditions, but only few reported statistics on action timing. Such analyses revealed diverging - probably instruction-dependent - patterns (no difference in the study by SanMiguel, Todd & Schröger, 2013; faster pace in the Motor-Auditory than in the Motor condition in the studies by Ford et al., 2014 – healthy control group; Horváth & Burgyán,

2013; and Horváth, 2014c; and slower average pace in the Motor-Auditory than in the Motor condition in the study by Saupe et al., 2013).

The fourth limitation is a potential technical confound (Horváth, 2014c). Most of the actions used in this line of research result in mechanical interaction with an object - most commonly the action is a button-press. Such transient interactions result in sound themselves. A quick finger movement touching a surface mostly produces faint, but audible sounds, and buttons are typically constructed to produce clicks to allow one an unambiguous feedback on whether the button was successfully pressed or not. If button-presses result in audible transient sounds, then the action-related ERPs elicited in the "Motor-only" condition may actually contain a low-amplitude auditory ERP contribution. Because the probe sounds are probably louder than these transient sounds, the transient is likely to be masked in the Motor-Auditory condition, and the elicited auditory ERP amplitude is likely to be less than the sum of the ERPs which would be elicited if the probe and the action-induced transient would be presented in isolation (that is, auditory ERPs are not additive). Therefore, subtracting the "Motor-only" waveform from the Motor-Auditory waveform leads to the subtraction of an auditory ERP, which may result in an *apparent* auditory ERP attenuation.

2.2 Typical findings in the contingent paradigm

Although Schafer & Marcus (1973) reported that essentially all ERP waveforms were reduced for self-initiated sounds, later studies focused on the attenuation of the auditory N1 waveform, even if differences in other waveforms (e.g. P2) were also observable (see for example, Ford & Mathalon, 2004; Baess, Horváth, Jacobsen, & Schröger, 2011). Recent studies, however started to explore changes in other waveforms, and started to differentiate N1 subcomponent-effects.

The N1 waveform comprises at least three subcomponents (Näätänen & Picton, 1987):

1) When sounds follow long (> 5-6 s) silent periods, the elicited N1 waveform is dominated by a centrally maximal (negative) subcomponent, the non-specific N1 (Hari, Kaila, Katila, Tuomisto, & Varpula, 1982, see also Näätanen, 1988), which is also elicited by non-auditory stimuli (Davis, Davis, Loomis, Harvey, Hobart, 1939; Davis & Zerlin, 1966).

2) When sounds are presented with short (< 4 s) inter-stimulus intervals, the N1 waveform shows a more frontal or fronto-central distribution. When the EEG is recorded with a nose reference, the fronto-centrally peaking negative waveform often shows a polarity inversion at mastoid sites (Vaughan & Ritter, 1970), suggesting that one of the

subcomponents originates from a supratemporal generator structure. Indeed, event-related magnetic field (ERF) counterparts of N1 (the N1m) recorded with MEG closely correspond to fields generated by tangentially oriented dipoles located in the supratemporal auditory cortex (see Elberling, Bak, Kofoed, Lebech, & Særmark, 1980; Hari, Aittoniemi, Järvinen, Katila, & Varpula, 1980).

3) The T-complex (Wolpaw & Penry, 1975) comprising a positivity at around 90-100 ms (Ta), and a negative peak at around 140-150 ms (Tb), which are most readily observable at the temporal (T3 and T4, in the 10-20 system, see Jasper, 1958) electrodes. The T-complex is thought to be generated in secondary auditory cortices (e.g. Scherg and von Cramon, 1985, 1986; Ponton, Eggermont, Khosla, Kwong, Don, 2002).

Although early studies often used N1-P2 peak-to-peak measurements (and in order to improve signal-to-noise ratio some studies still do), a few studies show that the N1 and P2 waveform reflect different processes (e.g. because of their differing sensitivity to temporal separation from preceding stimuli: Roth, Krainz, Ford, Tinklenberg, Rothbart, & Kopell, 1976; or differential impact of lesions on them: Knight, Hillyard, Woods, & Neville, 1980).

In general the supra-temporal N1 is regarded as a reflection of an auditory event- and feature-detection process (Näätänen & Winkler, 1999), while the non-specific N1 reflects processing leading to a conscious detection and orientation towards a (not necessarily auditory) sensory event (Näätänen & Picton, 1987). Whereas there is more-or-less clear consensus on these N1 components, a functional interpretation of the T-complex and P2 is missing (Crowley & Colrain, 2004). Some recent studies suggested that P2 reflects a process supporting auditory perceptual learning. Exposure to, and interaction with sounds in a perceptual task context leads to enhanced P2 ERPs even when the sounds are presented in an inattentive situation, and this enhancement is retained for months (Seppänen, Hämäläinen, Pesonen, & Tervaniemi, 2012; Tremblay, Ross, Inoue, McClannahan, & Collet, 2014).

Due to these unclarities regarding the functional interpretation of the ERP components, hypotheses on the cause of ERP attenuations observed in the contingent paradigm are often based on the generalizations of supra-temporal N1 effects. That is, current hypotheses mostly do not differentiate between N1-, Tb-, and P2- attenuations.

Because the central question of this line of research is whether performing an action leads to genuine sensory processing changes, the focus of inquiry is the supra-temporal N1 component, which originates from the auditory cortex. Due to the relative insensitivity of

MEG to non-superficial, non-tangential sources, attenuated N1m ERFs in contingent stimulation protocols are generally attributed to the attenuation of this N1 subcomponent (e.g. Aliu, Houde, & Nagarajan, 2009, Martikainen, et al., 2005). In the ERPs, a decreased positivity at the mastoid electrodes beside the decreased (less negative) fronto-central negativity is a good indicator that the experimental manipulation affected the supra-temporal component if the EEG was recorded with nose-reference. Unfortunately, studies often do not record EEG with such a reference, or do not specifically inspect whether a topographical inversion of the N1-effect occurred.

SanMiguel et al. (2013) found that ERP amplitudes measured in time windows corresponding to the earlier supra-temporal subcomponent were attenuated by 14-19 %, whereas attenuation was about 54% in the later interval corresponding to the non-specific subcomponent (which was observable only for the longest - 3.6 s inter-sound interval). Based on this mean difference, SanMiguel et al. suggested that most ERP studies utilizing relatively long (typically 2-6 sec) between-action intervals mainly reflect the attenuation of this later subcomponent.

Interestingly, some recent studies also found an attenuation of the Tb component (SanMiguel, et al., 2013; Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013) of the T-complex in contingent arrangements.

The attenuation of the P2 was also inspected in some recent studies, with many reporting significant attenuation (Knolle, Schröger, Baess, & Kotz, 2012; Horváth & Burgyán, 2013; Knolle, Schröger, & Kotz, 2013; SanMiguel et al., 2013, Saupe, et al., 2013; Van Elk, Salomon, Kannape, & Blanke, 2014; Horváth 2014c), and some reporting no attenuation (Sowman, Kuusik, & Johnson, 2012; and Ford et al., 2014).

N1, P2, and the T-complex are not the earliest waveforms that are attenuated in a contingent paradigm. Baess, et al. (2009) found that in such a paradigm the middle latency response and the evoked 40-Hz response was also attenuated.

At this point, there seems to be no obvious patterns in the reported ERP waveform attenuations. Whether the various ERP attenuations are brought about by a single cause or by multiple causes is largely unexplored. One may speculate for example, that results suggesting multiple causes show a significant attenuation of one ERP, but not of the other. It is suggested, for example, that N1- and P2-attenuations may be differentially affected by the presence of cerebellar lesions (Knolle et al., 2012), or by the type of the effector (hand vs.

foot movements, Van Elk et al., 2014). Because there is no reason to assume that a single manipulation would have the same magnitude of effect on two ERP components, a convincing dissociation would have to show effects of different signs, that is, the manipulation would result in higher amplitude for one of the ERPs, and lower amplitude for the other. (If the polarity of the ERPs differ, like for the N1 and P2, it has to be kept in mind that such a pattern could also be caused by an ERP shift – slower, low-frequency ERP - overlapping the time range of the two ERPs).

3. Interpreting action-related ERP attenuations

When interpreting the ERP-effects found in contingent paradigms, it has to be kept in mind that some (parts) of the effects may be brought about by the four types of confounds described above. Nonetheless, if the studies manipulate experimental variables for which it can be safely assumed that they are independent from (i.e. they do not modulate) these confounding effects, then the resulting ERP modulations can be readily interpreted as action-related ERP effects.

The common ground of all explanations is that the attenuation-effects are not simply caused by movement, but by voluntary, goal-directed actions. This was recently confirmed by Timm, SanMiguel, Keil, Schröger, & Schönwiesner (2014), who compared ERP responses to sounds initiated by voluntary tapping movements and those induced by transcranial magnetic stimulation, and found auditory ERP attenuation to voluntary movements only.

The explanations also agree on that finger-movement-related auditory (ERP) attenuation reflects central (and not peripheral) processes. Because speech-related auditory attenuation may be caused (in part) by a peripheral effect as described below, this assumption is not self-evident. Although the experimenter may present identical sounds, the participant's behavior may change sound transmission before it reaches the auditory system. One way in which such a change may occur is the contraction of the middle ear stapedius muscle, which leads to the effective attenuation of incoming sounds. Typically the contraction can be observed when vocalization, chewing, or even limb movements are initiated. Because the contraction starts 50–100 ms before the actual movement, sounds (Carmel & Starr, 1963; Salomon & Starr, 1963; Simmons, 1964) co-occurring with these movements are attenuated before reaching the auditory system. On the basis of this, Horváth & Burgyán (2013) suggested that auditory ERP attenuation may not reflect the modulation of the *neural signals*, rather the modulation of the *sound signal*. Since softer sounds elicit lower-amplitude auditory

ERPs (Picton, Woods, Baribeau-Braun, & Healey, 1977; Bak, Lebech, & Saermark, 1985; Lütkenhöner & Klein, 2007), ERP attenuations may reflect the attenuation of the sound itself, and not the attenuation of its processing. Because the stapedius contraction reduces sound transmission efficiency only for frequencies lower than about 2 kHz (Borg & Zakrisson, 1974; Zakrisson & Borg, 1974) measuring attenuation magnitudes for tones below and above this frequency allows for testing this hypothesis. Administering the contingent arrangement with button-press actions, Horváth & Burgyán (2013) however, found that this mechanism probably did not substantially contribute to the ERP effect measured in the button-press version of the contingent paradigm.

The three interpretations presented in the following suggest that auditory ERP attenuations are brought about by our ability to represent action-effect mappings: they all suggest that such representations allow us to adjust sensory processing in advance (or just in time) to prepare for the sensory consequences of a voluntary action. Although the explanations rely on the concept of forward modeling (for a summary, see Miall & Wolpert, 1996), this general principle can be implemented in various ways (Crapse & Sommer, 2008a, 2008b) and the three hypotheses seem to disagree on how such internal forward models are implemented in the human neuro-cognitive system.

3.1. Attention

The *attention hypothesis* is built on the results of research utilizing ERPs to understand how auditory processing is adapted to the task-relevant and irrelevant auditory input when participants perform sound-related perceptual tasks.

N1 does not only reflect the physical stimulation parameters, it also reflects whether the participant's attention is focused on the sounds or not. In most studies investigating this phenomenon, task-instructions are used to manipulate the allocation of input attention, that is the way in which the processing of a subset of stimuli or stimulus features is prioritized. Hillyard et al. (1973) found that sounds presented in the attended ear elicited higher-amplitude N1 waveforms than those presented in the unattended one (but found no P2 effect; see also Rif, Hari, Hämäläinen, & Sams, 1991; Woldorff & Hillyard, 1991; Ozaki, Jin, Suzuki, Baba, Matsunaga, & Hashimoto, 2004). Attention sets can not only be established for different modalities, input channels or moments in time (for a summary, see Lange, 2013) but for task-relevant sound features as well. Selective N1(m) enhancements have been found for example, when the task required attending a given tone frequency (Kauramäki, Jääskeläinen, & Sams, 2007; Kauramäki, Jääskeläinen, Hänninen, Auranen, Nummenmaa, Lampinen, &

Sams, 2012; Okamoto, et al., 2007). It has been also demonstrated that unexpected, task-irrelevant auditory events presented in the same channel as the task-relevant auditory events may disrupt the task-relevant attention set, which leads to the removal of the attentional N1- (and possibly P2) -enhancement for closely following task-relevant events (Horváth & Winkler, 2010; Horváth, 2014a, 2014b). That is, establishing a stimulus-specific attention sets is reflected in the enhancement of the N1 (and possibly P2) waveform.

The attentional enhancement of the N1 waveform is brought about by the modulation of the supra-temporal N1, and a slightly later, overlapping ERP of different origin (Alho, Paavilainen, Reinikainen, Sams, & Näätänen, 1986; Knight, et al., 1981) - termed *processing negativity* (PN, Näätänen, Gaillard & Mäntysalo, 1978) or *negative difference* (Nd, Hansen & Hillyard, 1980). The auditory N1 modulation reflects the attentional amplification of the task-relevant aspects of auditory input processing on the sensory level, and PN (or Nd) is thought to reflect additional, voluntary, task-relevant processing beyond the registration of the auditory event (Näätänen & Michie, 1979), possibly related to matching the event to a voluntarily maintained stimulus template (attentional trace, Näätänen, 1982, 1990). The lack of attentional P2 enhancement in some experiments manipulating sound-focused attention may be due to PN, which may overlap and cancel the positive attentional P2-enhancement.

Attention may play a role in action-related auditory ERP attenuation because pressing a button may include allocating attention to the action itself. That is, auditory attenuation may result from a difference in the distribution of attention between the action and the sound. It is assumed that participants attend the sounds in the Auditory condition (spontaneously, in the interest of maintaining vigilance, for example, or because they are instructed to) which leads to the enhancement of the auditory N1 (and P2). But this enhancement is removed when the participant performs the action in the Motor-Auditory condition because the participant focuses primarily on the action at the moment when it is initiated or performed, which leads to an apparent ERP attenuation (Horváth, Maess, Baess, & Tóth, 2012). Whether the enhanced N1 waveform includes a PN contribution probably depends on the experiment settings. Studies with explicit tone-related tasks are likely to include a PN contribution, but "passive listening" instructions may also compel some participants to occupy themselves with a tone-related task (e.g. count the tones) to maintain vigilance. Because PN emerges slightly later than the auditory N1, a PN contribution could be manifested in an attenuation effect which is delayed in comparison to the N1 waveform observable in the Auditory condition (see e.g. Bäß, Jacobsen, and Schröger, 2008)

Saupe et al. (2013) found evidence suggesting that although the allocation of attention may not explain the phenomenon in its entirety, attention effects may contribute to N1 and P2 attenuation measured in contingent paradigms. In a contingent setting, they instructed participants to produce a sequence of button-presses with highly variable between-sound (i.e. between-action) intervals within a 1.8-5.0 s range in the Motor-Auditory condition. They administered two types of Auditory conditions: In the Active Listening condition participants had to detect too short (< 1.8 s) or too long (> 5.0 s) between-sound intervals in the replay of the sound sequence produced in the Motor-Auditory condition, whereas no task was given in the Passive Listening condition. Comparisons of the corrected Motor-Auditory ERPs to those obtained in the two Listening conditions showed that N1 was attenuated. The difference was larger for the comparison with the Active Listening condition, suggesting that attentional differences may contribute to the attenuation effect. Importantly, however, a topographical difference in the attention-, and self-induction-related effects was found, which suggested that the origins of the effects were different: self-induction mainly affected the Tb component of the T-complex, while attention mainly affected the supratemporal N1. (Because of a potential PN overlap it is difficult to assess whether P2 was differentially affected or not).

3.2 Cancellation of auditory re-afference

The *cancellation of auditory re-afference hypothesis* of action-related auditory ERP attenuation is based on research uncovering the neural bases of how stimulation resulting from external and internal sources can be separated in the neural system during sensory processing.

The most widely accepted explanation of auditory ERP attenuation observable in contingent paradigms suggests that the attenuations reflect the cancellation of sensory re-afference. This forward modeling explanation is based on the concept of *corollary discharge* (Sperry, 1950, see also efference copy, von Holst, & Mittelstaedt, 1950). The corollary discharge is a direct neural signal originating from motor- or motor-planning-related neural centers (brain areas) to sensory systems, which adjusts sensory processing concurrently with the initiation of a movement. Such signals allow for example, the cancellation of proprioceptive re-afference resulting from muscle movements (Proske & Gandevia, 2012), the stabilization of the visual field despite eye-movements (Duhamel, Colby, Goldberg, 1992), and the cancellation of re-afference in the vestibular system resulting from head movement (Cullen, Brooks, Sadeghi, 2009). Because a neural circuitry enabling direct motor-induced suppression of auditory processing has been delineated in recent studies (Nelson, Schneider,

Takatoh, Sakurai, Wang, & Mooney, 2013; Schneider, Nelson, & Mooney, 2014)[1], it seems plausible that action-related auditory ERP attenuations observed in contingent paradigms may reflect a "macroscopic" aspect of the cancellation of auditory re-afference by a corollary discharge-based mechanism.

Evidence for this hypothesis is, however, rather scarce. It is mainly supported by studies showing that in comparison to healthy controls, the magnitude of N1 attenuation is reduced in patients with schizophrenia (e.g. Ford et al., 2007, and Ford et al., 2014; for a recent review see Ford & Mathalon, 2012), whose forward modeling capabilities may be impaired (Frith, Blakemore, & Wolpert, 2000). It has been also found that participants with larger readiness potential lateralization (toward the action-contra-lateral side), exhibited stronger N1- attenuation (Ford et al., 2014), which is compatible with a notion that more pronounced action-initiation activities give rise to a stronger cancellation of the auditory re-afference.

The promise of this line of research is that action-related auditory ERP attenuation may reflect functions related to agency attribution, that is, the attribution of sensory events to our own actions or to external sources. Although the argumentation is somewhat circular (in the sense that attenuation is regarded both as a result of prediction and an indicator that the sensory events were self-produced), some studies show results compatible with this notion. Extending the Motor-Auditory condition of the contingent paradigm, Baess et al. (2011) interspersed externally initiated tones with the tone sequence produced by the participants' own button presses. It was found that N1 attenuation for the self-induced sounds was stronger in this "mixed" condition than in the condition in which only tones initiated by the participants' actions were present. The stronger attenuation may reflect a mechanism enhancing the discriminability of self-induced and external stimulus events.

The foundation of this hypothesis is that action-related auditory ERP attenuation depends on the existence of a contingent action-sound relationship, however, evidence for this is missing[2]. On the contrary, some studies show that action-related auditory ERP attenuation

---

[1] The suppression effects shown in the two rodent studies were contingent on movement, but were not specific in terms of movement-types or the affected sound frequencies. Whether direct motor- and stimulus-specific suppression patterns exist is an open question.
[2] Note that showing the buildup of an ERP attenuation effect for an action-effect contingency would lend support for this hypothesis. Although often interpreted as a reflection of such a

occurs in the absence of an action-sound contingency. Makeig, Müller and Rockstroh (1996) showed that the amplitude and phase of the auditory steady state response was affected by concurrent, voluntary finger movements. Moreover, auditory N1 and P2 ERPs were attenuated when the sounds coincided with an action in the absence of a contingent action-sound relationship (Hazemann, Audin, & Lille, 1975; Horváth et al., 2012; Horváth 2013a, 2013b, 2014c).

The *coincidence* paradigm (Horváth et al., 2012) administered in these studies requires participants to produce a sequence of actions just like the *contingent* paradigm does. Concurrently, but independently from the actions, a sound-sequence with random between-sound intervals is presented. Because the action- and sound-sequence is independent, the action-, sound-, and combined action-sound-events occurring in separate blocks in the contingent paradigm occur in this paradigm by chance within the same block, as governed by the probability distribution of the sounds and the actions. As in the contingent paradigm, the auditory contribution to the ERP elicited by the action-sound coincidences is estimated by subtracting the ERP to action events temporally separated from sounds. The corrected waveform is compared to that elicited by sounds temporally separated from actions. Since all the relevant events occur within the same experimental blocks, potential confounds due to between-block differences are eliminated: there are no potential block-order-related differences, no task differences, and because action-sound coincidences occur by chance, confounds related to differences in action-preparation and –execution can be ruled out.

The studies administering the coincidence paradigm yielded similar ERP patterns as those administering the contingent paradigm: N1(m) and P2 waveforms were consistently attenuated. The estimated auditory ERPs and ERFs for sounds *following* an action in a short time showed that the N1- and P2-attenuation effects fell off quickly (in couple hundred ms) as a function of action-sound separation (Horváth et al., 2012). Similarly to recent studies reporting Tb-attenuations in contingent settings, variations of the coincidence paradigm also showed Tb-attenuations (Horváth, 2013a; the retrospective inspection of previous data showed significant Tb-attenuations in Experiment 2, but not in Exp. 1 in the study by Horváth et al., 2012, and in Exp. 1 in the study by Horváth 2013b). Interestingly, Horváth (2014c) found (post-hoc) a Ta-attenuation.

---

buildup, the results in the study by Aliu and colleagues (2009) do not convincingly demonstrate such an effect.

Although these results do not rule out the possibility that action-related auditory ERP attenuations are caused at least in part by contingent action-sound relationships, they show, however, that conclusions regarding this matter cannot be based solely on the presence of attenuation-effects. That is, the coincidence paradigm provides a baseline condition for the contingent paradigm. Whether contingent action-sound relationships give rise to ERP attenuations beyond that caused by coincidence, is an open question.

The mechanism causing the coincidence-effect is unclear. As for the contingent arrangement, one hypothesis is that the effect is caused by the division of attention, that is, when the actions are performed, attention is drawn away from the auditory stimulation toward the action, which results in N1 and P2 differences when the corrected coincidence- and the tone-related ERP waveforms are compared. Timm, et al. (2013) investigated this hypothesis by manipulating the task-relevancy of experimental events in a stimulation arrangement resembling those in a coincidence paradigm (i.e. button presses resulted in tones 50% the time, and tones and visual stimuli were also presented independently from button-presses). Participants were instructed to count either the tones, button-presses or the visual stimuli, thus manipulating the focus of attention, and found that these manipulations did not influence the magnitude of N1 attenuation. Horváth (2013b) also manipulated the task-relevancy of the tones, by instructing participants to try to "hit" a tone: that is, to create coincidences - press the button exactly at the time when a tone was presented. Despite paying a bonus for each "hit" (coincidence), an N1-attenuation comparable to that in a "non-motivated" condition was still present. That is, task-relevance-based manipulations of the attentional focus did not substantially influenced the coincidence-effect. Beside the task-relevancy, a selective attention set may also be manipulated by the presentation of a single or multiple tone frequencies in a given sequence. Presenting a sequence of random-frequency tones, however, also did not substantially influence the coincidence-effect (Horváth, 2013b).

Manipulating the type of action (button-press vs. –release) also did not substantially modulate the coincidence-effect (Horváth, 2013a). Although manipulating whether a finger-movement resulting in mechanical contact with an external object did modulate the coincidence-effect (i.e. making contact resulted in stronger attenuation, Horváth, 2014c), due to the potential confound caused by soft but audible tapping sounds described above, the cause of this modulation is unclear.

3.3. Pre-activation

The *pre-activation hypothesis* is mainly based on experimental psychological research on voluntary actions.

As described earlier, N1-attenuation is often interpreted as a reflection of attenuated auditory processing. Forward-modeling-based interpretations, however, not only suggest that actions causing sounds *generally* influence auditory processing, rather, that *identity prediction* (Hughes, Desantis, & Waszak, 2013a) occurs, that is, processing of the action-contingent sounds will be *specifically, selectively* influenced.

A number of studies suggest that during task-performance, actions are represented by their sensory consequences in the cognitive system, and the activation of such representations, and allocating attention to the task-relevant sensory consequence form an integral part of action preparation (Galazky, Schütze, Noesselt, Hopf, Heinze, & Schoenfeld, 2009; Brown, Friston, & Bestmann, 2011). The Theory of Event Coding (TEC, Hommel, Müsseler, Aschersleben, & Prinz, 2001) suggests that actions are encoded primarily (but not exclusively) in their distal, task-relevant consequences. When multiple action-consequences are available, then the one corresponding to the intentional reference frame is dominantly used (Sutter, Sülzenbrück, Rieger, & Müsseler, 2013). Intentional coding allows for considerable (but not unlimited) freedom in representing a task. For example, Hommel (1993) modified the Simon-paradigm (Simon & Rudell, 1967), in which participants had to respond with left or right key-presses to the pitch of a target tone. The tone, however, was presented on the left or the right. Reaction times typically show a spatial interference pattern: responses are slower if the response is to be given on the side opposite to the (task-irrelevant) side of the tone. Hommel showed that if responses are coupled with an action-effect on the action-opposite side (a light emitting diode flash, i.e. pressing the left button results in a flash on the right and vice versa), the spatial interference-effectcould be reversed if the instructions emphasized the effect-side instead of the button-press side (for a summary see, Hommel, 2011).

It is important to note that although actions are encoded primarily in their task-relevant sensory consequences, irrelevant sensory consequences are also automatically (that is, without intention) acquired and represented together with the action, and when the action is executed, the associated irrelevant sensory representations are activated as well. It is hypothesized (Hommel, et al., 2001) that the automatic acquisition of task irrelevant effects makes it possible to initiate actions without a known task-relevant effect. Elsner and Hommel (2001) demonstrated that task-irrelevant action consequences (key-presses resulting in tones of

different frequencies) are quickly coupled to the respective actions. Being exposed to such irrelevant action-consequences only about a hundred times in about 9-10 minutes produced robust action-selection interference effects in the tasks following the initial exposure to these contingencies. That is, action-effect representations are formed without intention to do so, even for artificial, arbitrary action-effect contingencies in the typical timeframe used in experiments on action-related auditory ERP attenuation. In the present context, it is also important that stimuli appearing during the execution of an action are more difficult to discriminate or detect if the action and the given stimulus are coupled in a concurrent task (Müsseler and Hommel, 1997a, 1997b).

Stimulus-specific action-related auditory ERP changes were observed in the contingent paradigm by Bäß et al. (2008). They investigated ERP attenuation to self-induced tones with predictable (1000 Hz) and unpredictable (random, 400-1990 Hz) frequencies and predictable (immediate) and unpredictable (500-1000 ms delayed) onsets. N1-attenuation was present in all cases, with stronger attenuation when frequency or onset was predictable. This suggests that a stronger action-stimulus coupling leads to stronger attenuation.

Going beyond the contingent paradigm, instead of comparing the processing of self-induced sounds to that of the same sounds only listened to, one may also compare responses to action-congruent and action-incongruent sounds, that is, sounds which have been previously associated with one action or with another action. In such studies, the main question is whether the action-related ERP effect is association-specific or not. In the study by Hughes, Desantis, & Waszak (2013b) participants learned and maintained action-sound (hand and finger movements to tone-frequencies) associations during the experiment. It was found that action-congruent sounds elicited lower amplitude N1s than action-incongruent ones (with no P2-effect visible in the reported waveforms). Hughes et al. argued that it was probably the congruency and not the incongruency that caused this difference.

Based on visual experiments, Waszak, Cardoso-Leite & Hughes (2012; see also Roussel et al., 2013) suggested that learned action-effect associations lead to the pre-activation of the sensory consequences of one's actions, which, in turn, lead to differences in stimulus detection performance, and the corresponding sensory ERPs (Roussel, Hughes, & Waszak, 2014). Interestingly, recent ERP studies in the auditory domain found compatible results. SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger (2013) found that when a mostly reliable, contingent action-sound relationship was present (i.e., actions elicited the same sound 88% of the time), rare omissions of the action-associated sound lead to the

elicitation of an ERP resembling part of the T-complex. Such an omission-response was not elicited when the actions induced sounds only 50% of the time, thus suggesting that the action-effect association was a pre-requisite of this effect, and the action in itself did not result in auditory activation. Moreover, this effect seemed to be sound-specific: if the action-induced sounds were randomly chosen on each button-press, the effect was no longer observable (SanMiguel, Saupe, & Schröger, 2013).

These results are compatible with the notion that contingent action-sound associations lead to the rapid formation of sensory effect representations (sensory templates), which are activated when the actions are voluntarily produced. Explaining why such pre-activated sensory templates lead to sensory ERP *attenuations,* is not self-evident. Whereas the corollary-discharge-based explanation essentially suggests that forward modeling results in a *negative sensory image,* which, superimposed on the (predicted) actual sensory input leads to cancellation, the presence of *positive* sensory (ERP) activation requires further considerations. Translating results (Cardoso-Leite, Mamassian, Schutz-Bosbach, & Waszak, 2010) reporting decreased detection sensitivity for action-congruent visual stimuli, Waszak et al. (2012) and Roussel et al. (2013) suggested that pre-activation increased the level of internal noise for the congruent stimuli, but the relation of this somewhat counter-intuitive explanation to action-related auditory ERP attenuations remains unclear.

Whereas one could speculate that the N1 ERP-differences between action-congruent and –incongruent stimuli could reflect the formation of selective attention sets based on the expectable sensory outcome of an action, the results of the studies by SanMiguel and colleagues on the omission-response referred to above would be difficult to explain in this framework. An initial study (Jones, Hughes & Waszak, 2013) attempting to separate attention- and motor-prediction-related contributions was uninformative regarding these ERP-effects.

4. Summary and outstanding questions

The studies published in the last couple of years lead to significant advances in the research on action-related auditory ERP attenuation. A number of basic assumptions were confirmed (or at least not rejected), and the set of affected ERP waveforms (and components) was extended.

In addition to the attenuation of the N1 waveform, recent studies also reported action-related attenuations of the P2, T-complex, middle latency-, and induced 40 Hz-responses. The

results, however, do not show patterns which would unequivocally correspond to experimental manipulations or hypotheses. Since these are relatively new developments, future studies should make it a priority to report whether such ERP effects are observable or not. At this point, it seems likely that these ERP phenomena do not reflect a single cause.

Beside the widely accepted *cancellation of auditory re-afference* hypothesis, two further hypotheses, the *attention* and the *pre-activation* hypotheses were put forward to explain the action-related auditory ERP attenuation. Although each hypothesis is compatible with a number of studies, none of them can accommodate all the findings on their own in their present form. Two findings especially difficult to reconcile with some of these hypotheses are the coincidence- (Hazemann et al., 1975; Horváth et al., 2012) and the sound omission-effects (SanMiguel, et al., 2013; SanMiguel et al, 2013). Similarly to the diverging results for the different ERP components, this suggests that action-related auditory ERP attenuation is a collection of separate ERP effects, probably with different underlying causes. It seems very likely that the mechanisms outlined in the three hypotheses shape auditory information processing together (see also Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010), and different experimental settings, instructions, tasks and contrasts may put different mechanisms into the foreground (e.g. a tone-related task in the contingent paradigm may lead to the dominance of attention-related processes in the given experimental contrast, see Saupe et al., 2013). Because of this, although studies investigating action-related auditory ERP changes mostly reported ERP attenuations, it seems possible that certain task settings may also result in action-related (attentional) activity enhancements as well. A recent functional magnetic resonance imaging study by Reznik, Henkin, Schadel, & Mukamel (2014) showing an action-related auditory activity enhancement may reflect the dominance of the attentional mechanism as well.

Currently, there are numerous outstanding questions in this line of research, answering which could reconcile the hypotheses and the conflicting evidence.

1) Although some studies already started to explore potential differences between sensory ERP attenuation patterns elicited by different actions, most of the research utilized only two types of actions: vocalizations, and finger-movements resulting in button-presses. Although it is generally assumed that sensory attenuations measured in paradigms using these two types of actions reflect the same mechanisms, evidence for this is scarce. Some studies find similar differences in auditory attenuation patterns between normal participants and patients with schizophrenia for speech-production and button-press actions (i.e. that the

magnitude of attenuation is reduced for the patients, Ford et al., 2007, and Ford et al., 2014), but direct evidence for this is yet unavailable. Since speech is probably be *the* best example for a behavior producing consistent patterns of auditory self-stimulation, it seems plausible that cancellation of sensory re-afference would play a role during speech production, and speech-related N1-attenuation effects reported in the literature may well reflect a corollary discharge function. For settings with arbitrary, non-speech-related actions and contingent (speech or non-speech) sounds, which are associated only for short periods (typically for about 5-20 minutes during the experiments), this seems less plausible. I am unaware of studies directly testing whether vocalization and finger-movement-based paradigms reflect the same mechanisms. On one hand, it would seem highly redundant if the highly similar ERP effects occurring in similar action-stimulus contexts would be produced by different subsystems, on the other hand, speech may have highly specialized subsystems, which are not readily available for the processing of other types of actions and stimuli (see e.g. Alho, Rinne, Herron, & Woods, 2014).

2) Our knowledge is rather limited on the temporal characteristics of auditory ERP attenuation. Although the results of studies investigating stimulus-specific attenuation patterns suggest a rapid (under 5-10 minutes) build-up of action-effect contingency representations, their deactivation times, and their long-term changes are unexplored. One speculation attempting to accommodate the coincidence-effect in an action-effect contingency-based explanatory framework (Horváth, et al. 2012) is that ERP attenuations observed in contingent paradigms may reflect a rapidly adapting action-effect contingency representation, whereas the coincidence- and speech-related-attenuation effects would reflect slowly adapting ones, presumably corollary discharge-based forward models.

3) Although of fundamental importance in this line of research, the effects of contingency have not been systematically investigated. The commonly used paradigms almost exclusively administer conditions in which only extreme action-effect contingencies occur: On one extreme, the Auditory-Motor condition of the contingent paradigm is 100% contingent (i.e. actions result in sounds all the time, and sounds are only generated by the actions). The other extreme is the coincidence paradigm in which there is no action-sound contingency at all. The importance of systematically exploring this avenue is emphasized by the success of two recent studies implementing contingency-manipulations. The ERP attenuation effects observed by Baess et al. (2011), may allow to relate action-related attenuation phenomena to agency attribution, while action-induced auditory activation

(SanMiguel et al., 2013) provides direct evidence for the existence of action-induced sensory templates.

Acknowledgements

References

Alho, K., Paavilainen, P., Reinikainen, K., Sams, M., & Näätänen, R. (1986). Separability of different negative components of the event-related potential associated with auditory stimulus processing. Psychophysiology, 23(6), 613–623. doi:10.1111/j.1469-8986.1986.tb00680.x

Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations in the auditory cortex: A meta-analysis of fMRI studies. Hearing Research, 307, 29–41. doi:10.1016/j.heares.2013.08.001

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. Journal of Cognitive Neuroscience, 21(4), 791–802.

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

Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. European Journal of Neuroscience, 29(7), 1514–1521. doi:10.1111/j.1460-9568.2009.06683.x

Bak, C. K., Lebech, J., & Saermark, K. (1985). Dependence of the auditory evoked magnetic field (100 msec signal) of the human brain on the intensity of the stimulus. Electroencephalography and Clinical Neurophysiology, 61(2), 141–149.

Bäß, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. International Journal of Psychophysiology, 70(2), 137–143. doi:10.1016/j.ijpsycho.2008.06.005

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. International Journal of Psychophysiology, 83(2), 120–131. doi:10.1016/j.ijpsycho.2011.08.003

Borg, E., & Zakrisson, J. E. (1974). Stapedius reflex and monaural masking. Acta Oto-Laryngologica, 78(3-4), 155–161.

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

Cardoso-Leite, P., Mamassian, P., Schutz-Bosbach, S., & Waszak, F. (2010). A New Look at Sensory Attenuation: Action-Effect Anticipation Affects Sensitivity, Not Response Bias. Psychological Science, 21(12), 1740–1745. doi:10.1177/0956797610389187

Carmel, P. W., & Starr, A. (1963). Acoustic and nonacoustic factors modifying middle-ear muscle activity in waking cats. Journal of Neurophysiology, 26, 598–616.

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

Crowley, K. E., & Colrain, I. M. (2004a). A review of the evidence for P2 being an independent component process: age, sleep and modality. Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology, 115(4), 732–744. doi:10.1016/j.clinph.2003.11.021

Crowley, K. E., & Colrain, I. M. (2004b). A review of the evidence for P2 being an independent component process: age, sleep and modality. Clinical Neurophysiology, 115(4), 732–744. doi:10.1016/j.clinph.2003.11.021

Cullen, K. E., Brooks, J. X., & Sadeghi, S. G. (2009). How actions alter sensory processing: reafference in the vestibular system. Annals of the New York Academy of Sciences, 1164, 29–36. doi:10.1111/j.1749-6632.2009.03866.x

Davis, H., Davis, P. A, Loomis, A. L, Harvey, E. N, & Hobart, G. (1939). Electrical reactions of the human brain to auditory stimulation during sleep. Journal of Neurophysiology, 2(6), 500–514.

Davis, H., Mast, T., Yoshie, N., & Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: recovery process. Electroencephalography and Clinical Neurophysiology, 21(2), 105–113.

Davis, H., & Zerlin, S. (1966). Acoustic Relations of the Human Vertex Potential. The Journal of the Acoustical Society of America, 39(1), 109. doi:10.1121/1.1909858

De Jong, R., Wierda, M., Mulder, G., & Mulder, L. J. (1988). Use of partial stimulus information in response processing. Journal of Experimental Psychology: Human Perception and Performance, 14(4), 682.

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(5040), 90–92.

Eimer, M. (1998). The lateralized readiness potential as an on-line measure of central response activation processes. Behavior Research Methods, Instruments, & Computers, 30(1), 146–156.

Elberling, C., Bak, C., Kofoed, B., Lebech, J., & Særmark, K. (1980). Magnetic Auditory Responses from the Human Brain A Preliminary Report. Scandinavian Audiology, 9(3), 185–190. doi:10.3109/01050398009076353

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. Journal of Experimental Psychology: Human Perception and Performance, 27(1), 229.

Ford, J. M., & Mathalon, D. H. (2004). Electrophysiological evidence of corollary discharge dysfunction in schizophrenia during talking and thinking. Journal of Psychiatric Research, 38(1), 37–46. doi:10.1016/S0022-3956(03)00095-5

Ford, J. M., & Mathalon, D. H. (2012). Anticipating the future: Automatic prediction failures in schizophrenia. International Journal of Psychophysiology, 83(2), 232–239. doi:10.1016/j.ijpsycho.2011.09.004

Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2014). Did I Do That? Abnormal Predictive Processes in Schizophrenia When Button Pressing to Deliver a Tone. Schizophrenia Bulletin, 40(4), 804–812. doi:10.1093/schbul/sbt072

Frith, C., Blakemore, S.-J., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. Brain Research Reviews, 31(2-3), 357–363. doi:10.1016/S0165-0173(99)00052-1

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

Gratton, G., Coles, M. G., Sirevaag, E. J., Eriksen, C. W., & Donchin, E. (1988). Pre-and poststimulus activation of response channels: a psychophysiological analysis. Journal of Experimental Psychology: Human Perception and Performance, 14(3), 331.

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(2), 413–497. doi:10.1103/RevModPhys.65.413

Hansen, J. C., & Hillyard, S. A. (1980). Endogeneous brain potentials associated with selective auditory attention. Electroencephalography and Clinical Neurophysiology, 49(3), 277–290.

Hari, R., Aittoniemi, K., Järvinen, M.-L., Katila, T., & Varpula, T. (1980). Auditory evoked transient and sustained magnetic fields of the human brain localization of neural generators. Experimental Brain Research, 40(2), 237–240.

Hari, R., Kaila, K., Katila, T., Tuomisto, T., & Varpula, T. (1982). Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: implications for their neural generation. Electroencephalography and Clinical Neurophysiology, 54(5), 561–569.

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

Hickok, G. (2012). Computational neuroanatomy of speech production. Nature Reviews Neuroscience, 13(2), 135–145.

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.

Hommel, B. (1993). Inverting the Simon effect by intention. Psychological Research, 55(4), 270–279.

Hommel, B. (2011). The Simon effect as tool and heuristic. Acta Psychologica, 136(2), 189–202. doi:10.1016/j.actpsy.2010.04.011

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. Behavioral and Brain Sciences, 24(5), 849–878; discussion 878–937.

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. (2014a). Probing the sensory effects of involuntary attention change by ERPs to auditory transients: Probing the sensory impact of distraction. Psychophysiology, 51(5), 489–497. doi:10.1111/psyp.12187

Horváth, J. (2014b). Sensory ERP effects in auditory distraction: did we miss the main event? Psychological Research, 78(3), 339–348. doi:10.1007/s00426-013-0507-7

Horváth, J. (2014c). The role of mechanical impact in action-related auditory attenuation. Cognitive, Affective, & Behavioral Neuroscience, 14, 1392-1406 doi:10.3758/s13415-014-0283-x

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. (2013a). Attenuation of auditory N1 results from identity-specific action-effect prediction. European Journal of Neuroscience, 37(7), 1152–1158. doi:10.1111/ejn.12120

Hughes, G., Desantis, A., & Waszak, F. (2013b). 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

Jasper, H., H. (1958). Report of the committee on methods of clinical examination in electroencephalography. Electroencephalography and Clinical Neurophysiology, 10(2), 370–375. doi:10.1016/0013-4694(58)90053-1

Jones, A., Hughes, G., & Waszak, F. (2013). The interaction between attention and motor prediction. An ERP study. NeuroImage, 83, 533–541. doi:10.1016/j.neuroimage.2013.07.004

Kauramäki, J., Jääskeläinen, I. P., Hänninen, J. L., Auranen, T., Nummenmaa, A., Lampinen, J., & Sams, M. (2012). Two-Stage processing of sounds explains behavioral performance variations due to changes in stimulus contrast and selective attention: An MEG study. PLoS ONE, 7(10), e46872. doi:10.1371/journal.pone.0046872

Kauramäki, J., Jääskeläinen, I. P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the guman auditory cortex. PLoS ONE, 2(9), e909. doi:10.1371/journal.pone.0000909

Knight, R. T., Hillyard, S. A., Woods, D. L., & Neville, H. J. (1980). The effects of frontal and temporal-parietal lesions on the auditory evoked potential in man. Electroencephalography and Clinical Neurophysiology, 50(1-2), 112–124. doi:10.1016/0013-4694(80)90328-4

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

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Prediction errors in self- and externally-generated deviants. Biological Psychology, 92(2), 410–416. doi:10.1016/j.biopsycho.2012.11.017

Lange, K. (2013). The ups and downs of temporal orienting: a review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. Frontiers in Human Neuroscience, 7. doi:10.3389/fnhum.2013.00263

Lütkenhöner, B., & Klein, J.-S. (2007). Auditory evoked field at threshold. Hearing Research, 228(1-2), 188–200. doi:10.1016/j.heares.2007.02.011

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

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

Müsseler, J., & Hommel, B. (1997a). Blindness to response-compatible stimuli. Journal of Experimental Psychology: Human Perception and Performance, 23(3), 861–972. doi:10.1037/0096-1523.23.3.861

Müsseler, J., & Hommel, B. (1997b). Detecting and identifying response-compatible stimuli. Psychonomic Bulletin & Review, 4(1), 125–129. doi: 10.3758/BF03210785

Näätänen, R. (1982). Processing negativity: An evoked-potential reflection of selective attention. Psychological Bulletin, 92(3), 605–640.

Näätänen, R. (1988). Implications of ERP data for psychological theories of attention. Biological Psychology, 26(1), 117–163.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. Behavioral and Brain Sciences, 13(02), 201–233. doi:10.1017/S0140525X00078407

Näätänen, R., Gaillard, A. W., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. Acta Psychologica, 42(4), 313–329. doi:10.1016/0001-6918(78)90006-9

Näätänen, R., & Michie, P. T. (1979). Early selective-attention effects on the evoked potential: a critical review and reinterpretation. Biological Psychology, 8(2), 81–136.

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

Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. Psychological Bulletin, 125(6), 826–859.

Nelson, A., Schneider, D. M., Takatoh, J., Sakurai, K., Wang, F., & Mooney, R. (2013). A Circuit for Motor Cortical Modulation of Auditory Cortical Activity. Journal of Neuroscience, 33(36), 14342–14353. doi:10.1523/JNEUROSCI.2275-13.2013

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

Ozaki, I., Jin, C. Y., Suzuki, Y., Baba, M., Matsunaga, M., & Hashimoto, I. (2004). Rapid change of tonotopic maps in the human auditory cortex during pitch discrimination. Clinical Neurophysiology, 115(7), 1592–1604. doi:10.1016/j.clinph.2004.02.011

Picton, T. W., Woods, D. L., Baribeau-Braun, J., & Healey, T. M. (1977). Evoked potential audiometry. The Journal of Otolaryngology, 6(2), 90–119.

Ponton, C., Eggermont, J. J., Khosla, D., Kwong, B., & Don, M. (2002). Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. Clinical Neurophysiology, 113(3), 407–420. doi:10.1016/S1388-2457(01)00733-7

Proske, U., & Gandevia, S. C. (2012). The Proprioceptive Senses: Their Roles in Signaling Body Shape, Body Position and Movement, and Muscle Force. Physiological Reviews, 92(4), 1651–1697. doi:10.1152/physrev.00048.2011

Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. Nature Communications, 5. doi:10.1038/ncomms5059

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

Roth, W. T., Krainz, P. L., Ford, J. M., Tinklenberg, J. R., Rothbart, R. M., & Kopell, B. S. (1976). Parameters of temporal recovery of the human auditory evoked potential. Electroencephalography and Clinical Neurophysiology, 40(6), 623–632.

Roussel, C., Hughes, G., & Waszak, F. (2013). A preactivation account of sensory attenuation. Neuropsychologia, 51(5), 922–929. doi:10.1016/j.neuropsychologia.2013.02.005

Roussel, C., Hughes, G., & Waszak, F. (2014). Action prediction modulates both neurophysiological and psychophysical indices of sensory attenuation. Frontiers in Human Neuroscience, 8. doi:10.3389/fnhum.2014.00115

Salomon, G., & Starr, A. (1963). ELECTROMYOGRAPHY OF MIDDLE EAR MUSCLES IN MAN DURING MOTOR ACTIVITIES. Acta Neurologica Scandinavica, 39(2), 161–168. doi:10.1111/j.1600-0404.1963.tb05317.x

SanMiguel, I., Saupe, K., & Schröger, E. (2013). I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted "what" but not "when". Frontiers in Human Neuroscience, 7. doi:10.3389/fnhum.2013.00407

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

Scherg, M., & Von Cramon, D. (1985). Two bilateral sources of the late AEP as identified by a spatio-temporal dipole model. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 62(1), 32–44. doi:10.1016/0168-5597(85)90033-4

Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature, 513(7517), 189–194. doi:10.1038/nature13724

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. Current Opinion in Neurobiology, 20(2), 172–176. doi:10.1016/j.conb.2010.02.010

Seppänen, M., Hämäläinen, J., Pesonen, A.-K., & Tervaniemi, M. (2012). Music Training Enhances Rapid Neural Plasticity of N1 and P2 Source Activation for Unattended Sounds. Frontiers in Human Neuroscience, 6. doi:10.3389/fnhum.2012.00043

Simmons, F. B. (1964). PERCEPTUAL THEORIES OF MIDDLE EAR MUSCLE FUNCTION. The Annals of Otology, Rhinology, and Laryngology, 73, 724–739.

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. Journal of Applied Psychology, 51(3), 300–304. doi:10.1037/h0020586

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.

Sutter, C., Sülzenbrück, S., Rieger, M., & Müsseler, J. (2013). Limitations of distal effect anticipation when using tools. New Ideas in Psychology, 31(3), 247–257. doi:10.1016/j.newideapsych.2012.12.001

Timm, J., SanMiguel, I., Keil, J., Schröger, E., & Schönwiesner, M. (2014). Motor Intention Determines Sensory Attenuation of Brain Responses to Self-initiated Sounds. Journal of Cognitive Neuroscience, 26(7), 1481–1489. doi:10.1162/jocn_a_00552

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

Tremblay, K. L., Ross, B., Inoue, K., McClannahan, K., & Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? Frontiers in Systems Neuroscience, 8. doi:10.3389/fnsys.2014.00028

Van Elk, M., Salomon, R., Kannape, O., & Blanke, O. (2014). Suppression of the N1 auditory evoked potential for sounds generated by the upper and lower limbs. Biological Psychology, 102, 108–117. doi:10.1016/j.biopsycho.2014.06.007

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.

Waszak, F., Cardoso-Leite, P., & Hughes, G. (2012). Action effect anticipation: Neurophysiological basis and functional consequences. Neuroscience & Biobehavioral Reviews, 36(2), 943–959. doi:10.1016/j.neubiorev.2011.11.004

Winkler, I. (2007). Interpreting the Mismatch Negativity. Journal of Psychophysiology, 21(3), 147–163. doi:10.1027/0269-8803.21.34.147

Woldorff, M. G., & Hillyard, S. A. (1991). Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalography and Clinical Neurophysiology, 79(3), 170–191.

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

Zakrisson, J. E., & Borg, E. (1974). Stapedius reflex and auditory fatigue. Audiology: Official Organ of the International Society of Audiology, 13(3), 231–235.