n

Cao, Liyu (2016) *The influence of actions on auditory perception: cognitive and neural mechanisms.* PhD thesis.

http://theses.gla.ac.uk/7783/

The influence of actions on auditory perception: cognitive and neural mechanisms

Liyu Cao

Submitted in fulfilment of the requirements for the Degree of Ph.D.

School of Psychology
College of Science and Engineering
University of Glasgow
July, 2016

Summary

It is well known that self-generated stimuli are processed differently from externally generated stimuli. For example, many people have noticed since childhood that it is very difficult to make a self-tickling. In the auditory domain, self-generated sounds elicit smaller brain responses as compared to externally generated sounds, known as the sensory attenuation (SA) effect. SA is manifested in reduced amplitudes of evoked responses as measured through MEEG, decreased firing rates of neurons and a lower level of perceived loudness for self-generated sounds.

The predominant explanation for SA is based on the idea that self-generated stimuli are predicted (e.g., the forward model account). It is the nature of their predictability that is crucial for SA. On the contrary, the sensory gating account emphasizes a general suppressive effect of actions on sensory processing, regardless of the predictability of the stimuli. Both accounts have received empirical support, which suggests that both mechanisms may exist.

In chapter 2, three behavioural studies concerning the influence of motor activation on auditory perception were presented. Study 1 compared the effect of SA and attention in an auditory detection task and showed that SA was present even when substantial attention was paid to unpredictable stimuli. Study 2 compared the loudness perception of tones generated by others between Chinese and British participants. Compared to externally generated tones, a decrease in perceived loudness for others generated tones was found among Chinese but not among the British. In study 3, partial evidence was found that even when reading words that are related to action, auditory detection performance was impaired.

In chapter 3, the classic SA effect of M100 suppression was replicated with MEG in study 4. With time-frequency analysis, a potential neural information processing sequence was found in auditory cortex. Prior to the onset of self-generated tones, there was an increase of oscillatory power in the alpha band. After the stimulus onset, reduced gamma power and alpha/beta phase locking were found. The three temporally segregated oscillatory events correlated with each other and with SA effect, which may be the underlying neural implementation of SA.

In chapter 4, a TMS-MEG study was presented investigating the role of the cerebellum in adapting to delayed presentation of self-generated tones (study 5). It demonstrated that in sham stimulation condition, the brain can adapt to the delay (about 100 ms) within 300 trials of learning by showing a significant increase of SA effect in the suppression of M100, but not M200 component. Whereas after stimulating the cerebellum with a suppressive TMS protocol, the adaptation in M100 suppression disappeared and the pattern of M200 suppression reversed to M200 enhancement.

These data support the idea that the suppressive effect of actions on auditory processing is a consequence of both motor driven sensory predictions and general sensory gating. The results also demonstrate the importance of neural oscillations in implementing SA effect and the critical role of the cerebellum in learning sensory predictions under sensory perturbation.

Table of Contents

Acknowledgement

First and foremost, I would like to thank my supervisor Professor Joachim Gross. Under your supervision, I enjoy my Ph.D. study so much. You are always patient and supportive. I learned from you, explicitly and implicitly, how to become a good scientist and a decent man. I will pass on what I learned from you to my future students. Thank you again for being my supervisor.

I would also like to thank Dr. Katja Biermann-Ruben, Professor Klaus Kessler and Professor Gregor Thut for giving me so much help. I thank my officemates Drs Roberto Cecere and Christian Keitel. Over the daily conversation, I learned a lot from them.

I cannot forget the delicious (free) food from my landlord and friend Liran Liu. Thank you for being so nice.

A special thanks to my parents and sister. You have been giving me so much support to allow me to study somewhere so far away. I know that also means a lot of sacrifices. Words cannot express my gratitude. Now it is my turn to give you support.

There are many others who helped me a lot during the four years. Although not named here, they are in my mind.

Abbreviations

CI	Confidence interval
dB	Decibel
EEG	Electroencephalography
EQ	Empathy quotient
fMRI	Functional magnetic resonance imaging
ICA	Independent component analysis
MEEG	MEG and EEG
MEG	Magnetoencephalography
PDI	Peters et al. delusion inventory
PSE	Point of subjective equality
SA	Sensory attenuation
SD	Standard deviation
SOA	Stimulus onset asynchrony
SPL	Sound pressure level
SQ	Systemizing quotient
TMS	Transcranial magnetic stimulation

1 Introduction

This is a true story happened in my office during my PhD study. One day in the afternoon, I was doing some data analysis work in front of my computer, as concentrated as usual. I felt thirsty so I reached for my mug on the table and drank some tea. Then I put down my mug back on the table. Just after I put down my mug, my officemate Christian said ‘come in’ loudly towards the door as if someone just knocked at the door. I was quite puzzled because I did not hear any door knocking sound at all. Then I asked him and eventually we realized that it was the sound from putting down the mug on the table that he mistook for someone knocking at the door. Interestingly, as the generator of the sound, I interpreted the sound correctly without any confusion.

In the situation described above, the same physical auditory input (although it was closer in distance to me than to my colleague) triggered different reactions from people who were both concentrated on their work. As the generator of the sound, I knew that the sound resulted from my movement even without any cognitive processing reaching the consciousness level. If there had not been the confusion raised by my colleague, it may never have come into my consciousness that I just generated a sound through putting down my mug. However, my colleague received the sound stimulus from an external source while he was concentrated on his work. When the stimulus information reached his brain, the brain had to make sense of it. Since no prior information was available (he did not see me put down the mug), his brain interpreted it as someone knocking at the door. This is probably because that both the door and my table are made of wood so they sound alike and that the ‘knocking at the door’ interpretation is more accessible/common for my colleague’s brain. Overall, the story demonstrated how pre-stimulus states of the brain (e.g., the generator of the stimulus vs. the passive receiver of the stimulus) may have a profound influence on stimulus processing. In the particular instance considered above, the change of pre-stimulus brain state was brought about by actions. The influence of actions on perception is ubiquitous. A famous example is that it is very difficult to tickle oneself (Blakemore, Wolpert, & Frith, 1998; Weiskrantz, Elliott, & Darlington, 1971). That is to say, the ticklish sensation is attenuated when the somatosensory input is self-generated. This thesis deals with the similar phenomenon in the auditory domain, i.e., how actions may influence auditory perception and what the underlying neural mechanism is.

1.1 Sensory attenuation effect

Sensory attenuation (SA) is the effect that self-generated stimuli are attenuated as compared to externally generated stimuli. This is confirmed in auditory and somatosensory domains by both behavioural and electrophysiological studies (Blakemore et al., 1998; Hesse, Nishitani, Fink, Jousmaki, & Hari, 2009; Martikainen, Kaneko, & Hari, 2005; A. Sato, 2008; Schafer & Marcus, 1973; Shergill, Bays, Frith, & Wolpert, 2003; Weiskrantz et al., 1971). SA is also reported in the visual domain (Cardoso-Leite, Mamassian, Schuetz-Bosbach, & Waszak, 2010; Gentsch & Schutz-Bosbach, 2011), but it comes together with contradictory findings (Mifsud et al., 2016; see also Chapter 3). An introduction to auditory SA is given below.

1.1.1 Behavioural studies

Behavioural evidence for SA in the auditory domain comes from loudness perception studies. A. Sato (2008) compared the perceived loudness of a tone under different conditions. The tone was perceived to be softer when it was self-generated than when it was generated by the computer. This was later replicated by Weiss, Herwig, and Schuetz-Bosbach (2011), with a better control for the stimulus predictability in both conditions. When the stimulus was generated by the computer, a visual cue was included to allow participants to predict the timing of the stimulus. Stenner, Bauer, Sidarus, et al. (2014) provided more convincing evidence for a motor origin of this effect. In their study, tone-triggering motor responses were subliminally primed so that participants felt more control over their action consequences (i.e., the tone) in one condition than the other. In this case, the only difference between conditions was the manipulated motor processing through priming. A reduction in perceived loudness was found when participants felt more control over their action consequences after priming. These studies demonstrated that self-generated tones are perceived to be softer than externally generated tones, at least when the intensity of the tone (around 75 dB SPL (sound pressure level) in all the three studies) is well above the detection threshold. When the intensity of the tone is near detection threshold, however, sensory enhancement was found (Reznik, Henkin, Levy, & Mukamel, 2015).

1.1.2 Electrophysiological studies

The first report of auditory SA is from an EEG study by Schafer and Marcus (1973), which is later consistently replicated in EEG (electroencephalography) and MEG (magnetoencephalography) investigations (Aliu, Houde, & Nagarajan, 2009; Baess, Horváth, Jacobsen, & Schröger, 2011; Ford, Palzes, Roach, & Mathalon, 2013; Martikainen et al., 2005). In all these studies, auditory evoked responses are compared between two conditions: an active condition where the sound is generated by participants themselves (e.g., through pressing a button) and a passive condition where the sound presentation is controlled by the computer[1]. A consistent finding is that auditory evoked responses peaking around 100 ms after stimulus onset (M100 component in MEG; N100 component in EEG) are reduced in the active condition as compared to the passive condition (Figure 1.1). Dipole fitting source localization analysis showed that responses in auditory cortex were suppressed in the active condition (Aliu et al., 2009; Martikainen et al., 2005). Interestingly, when a sound was delayed for presentation (e.g., 100 ms) after a button press, SA effect was not so evident in the beginning but emerged within 300 trials of learning (Aliu et al., 2009).

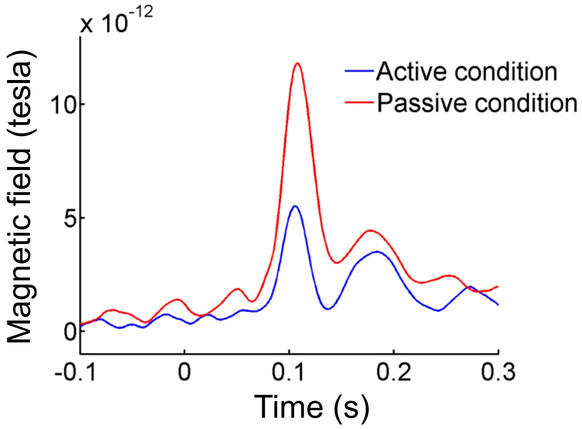


Figure 1.1 Illustration of sensory attenuation effect. When a sound is self-generated (active condition), the amplitude of evoked responses (e.g., M100 component) is smaller than the evoked responses from external sounds (passive condition).

SA is also studied with speech paradigm, in which a speech sound (e.g., ‘ah’) is compared between a speaking condition and a listening condition. Auditory evoked responses are significantly reduced in speaking condition (Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000; Houde, Nagarajan, Sekihara, & Merzenich, 2002; J. Wang et al., 2014). Using the speech paradigm, invasive recordings from animal studies and studies with neurosurgical

---

[1] Usually a motor only condition, which is the same to the active condition except that no sound is presented after participants’ movement, is included. This is for the aim of getting a clean auditory evoked response in the active condition that is free from the possible contamination of motor responses. This is done by subtracting evoked responses in the motor only condition from the active condition.

patients provide further support for a reduced neural response in the auditory cortex during speaking (Eliades & Wang, 2003, 2013; Greenlee et al., 2011). For example, Eliades and Wang (2003) showed that firing rates of the majority of neurons in the auditory cortex were reduced during vocalization in marmosets.

1.2 Mechanisms of SA

Four major different accounts have been put forward to account for SA: sensory prediction account, sensory gating account, attention account and preactivation account (for reviews, see Horváth, 2015; Hughes, Desantis, & Waszak, 2013b; Schröger, Marzecová, & SanMiguel, 2015). I will give a brief introduction to all accounts below.

1.2.1 Sensory prediction account: efference copy, corollary discharge and the forward model

Sensory prediction account of SA emphasises the importance of sensory prediction, i.e., the prediction of sensory consequences resulting from motor responses. It is posited that reduced responses to self-generated stimuli are a result of better prediction of stimuli (e.g., the onset timing and the identity of stimuli). For example, in the study by Hughes, Desantis, and Waszak (2013a), the authors manipulated the congruency between button presses and the frequency of induced sounds. They showed that congruent sounds following the button press elicited smaller amplitudes of N100 component than incongruent sounds. The prediction mechanism has been coined different terms by different researchers. These include, but not limited to, efference copy (von Holst & Mittelstaedt, 1950), corollary discharge (Sperry, 1950) and forward model (Wolpert & Ghahramani, 2000), which are still commonly seen and discussed in today’s papers. Below is a detailed introduction to these concepts.

1.2.1.1 Efference copy

The concept of efference copy has a long history.It revived in the 1950s through the introduction of von Holst and Mittelstaedt with their studies on fly Eristalis and fish (Grüsser, 1995; von Holst & Mittelstaedt, 1950). For example, when a fly Eristalis was put into a cylinder with black and white striped inner surface, it demonstrated the well-known optomotor reflex when the cylinder was rotated by the experimenter, i.e., the fly Eristalis rotated in the same direction as if it was trying to keep its visual scene constant. But when the insect moved voluntarily inside the cylinder (e.g., a clockwise rotation), it did not move back to its original position. This is strange because the optic flow on the retina during the voluntary movement was exactly the same as when the cylinder was rotated by the experimenter. The optomotor reflex did not show up during the voluntary movement suggesting that the visual scene must remain stationary in that case. von Holst and Mittelstaedt (1950) proposed that there is an efference copy signal sent to the retina during the voluntary movement so that fly Eristalis has a prediction about the incoming retinal stimulation. The prediction and the incoming stimulus then cancel each other out resulting in a stable visual scene. This proposition received support from the experiment in which the neck of fly Eristalis was rotated 180^o. After the neck rotation, a clockwise rotation of the cylinder would result in visual stimulation on the retina that was equivalent to the visual stimulation of an anticlockwise cylinder rotation under normal situation. In this case, if the fly Eristalis made a voluntary rotation, the prediction from the efference copy and the incoming stimulus would not be able to cancel each other out as the incoming stimulus was rotated. This is exactly what the experiment showed. The fly Eristalis ended up with continuous rotation in one direction or bidirectional rotations to and fro as if it was trying to reconcile the mismatch between the prediction and the incoming stimulus.

In what was called ‘the principle of reafference’ by von Holst and Mittelstaedt (1950), a general principle governing the interaction between the central nervous system and the peripheral organs was given (see Figure 1.2 a). A motor command is sent from the highest centre Zn to lower centres until Z1. Z1 then sends an efference signal to the effector (e.g., a muscle), which then generates a reafference signal that is sent back to Z1. The reafference signal denotes the input that results from self-movement contrasting the exafference signal which is input due to environmental perturbations. Importantly, an efference copy is generated along with the efference signal in Z1. The efference copy will be compared with the reafference signal and the mismatch between the two will be reported to higher centres. The reported mismatch functions as a feedforward signal enabling the higher centres to adjust the command until a match is achieved between the efference copy and the reafference.

The idea of efference copy also provides an elegant explanation of various perceptual phenomena in humans during or as a result of movements. We make saccades constantly in daily life but never experience that the visual scene moves or jumps during saccades.

Instead, we experience a stable world. Every saccade results in a displacement of the visual image on the retina, which is similar to the case when the eyeball is moved by an external force (e.g., by pressing the eyeball with fingers). In the latter case, we experience that the world jumps in the direction opposite to where the eyeball moves to. With the help of the efference copy idea, the stable visual scene during saccades is easy to explain. The visual perception during saccades is influenced by both efference copy signals and the visual input. When making a voluntary rightward saccade (Figure 1.2 b), the eye movement itself should result in a leftward jump of the visual scene, as in the case when the eyeball is pushed rightward by an external force. At the same time, an efference copy signal is sent to lower visual centres which alone should result in a rightward perception, as shown in patients with paralysis of eye muscles. The results from efference copy and visual image displacement on the retina then cancel each other out resulting in a stable environment (von Holst & Mittelstaedt, 1950).

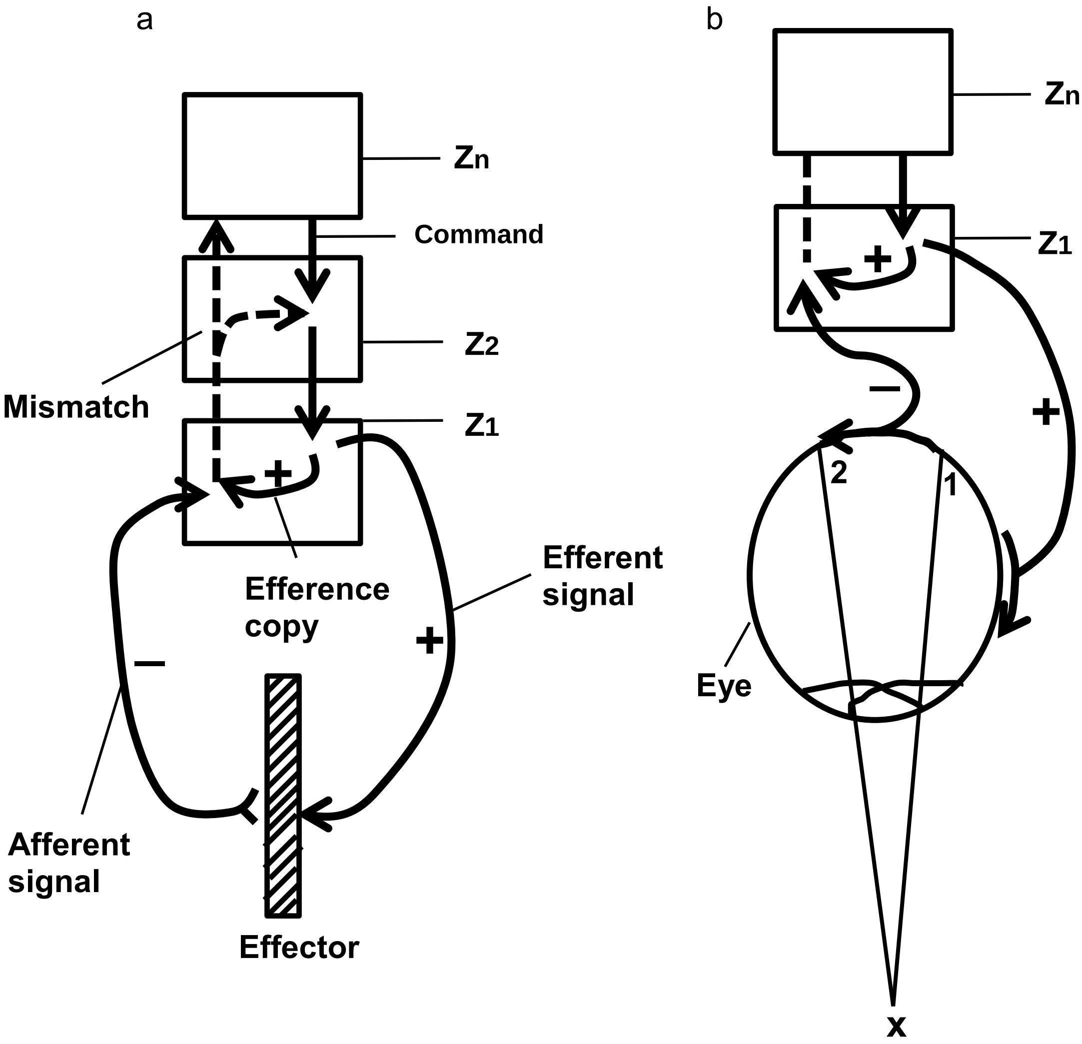


Figure 1.2 Illustration of the concept of efference copy. See text for detail. Adapted from von Holst and Mittelstaedt (1950).

1.2.1.2 Corollary discharge

Almost at the same time, Sperry (1950) independently put forward a similar idea to efference copy when studying the continuous turning behaviour of fish after eye rotation (similar to the rotation behaviour of fly Eristalis as introduced earlier). He used the term ‘corollary discharge’ instead of ‘efference copy’. Quoted from Sperry (1950; pp. 488):

> ‘Thus, any excitation pattern that normally results in a movement that will cause a displacement of the visual image on the retina may have a corollary discharge into the visual centers to compensate for the retinal displacement. This implies an anticipatory adjustment in the visual centers specific for each movement with regard to its direction and speed. A central adjuster factor of this kind would aid in maintaining stability of the visual field under normal conditions during the onset of sudden eye, head, and body movement. With the retinal field rotated 180 degrees, any such anticipatory adjustment would be in diametric disharmony with the retinal input, and would therefore cause accentuation rather than cancellation of the illusory outside movement.’

1.2.1.3 Forward model

In motor control studies, a forward model was put forward capturing the causal relationship between movement and its resulting consequences (e.g., hand velocity) (Miall & Wolpert, 1996; Wolpert & Ghahramani, 2000). The forward model is also widely referenced as an explanation of SA (Baess et al., 2011; Blakemore et al., 1998; Hughes et al., 2013a; Shergill et al., 2014). As Figure 1.3 shows, along with a motor command, an efference copy is generated allowing for the prediction of imminent sensory input. The predicted sensory input is then compared with real sensory input and SA can be observed when there is a match between the two.

The idea of forward model appears very similar to the concept of efference copy and corollary discharge. Proposers of the forward model explained that efference copy is merely a copy of the motor command thus it does not contain any sensory prediction. In their opinion, the corollary discharge represents the predicted sensory consequences (Blakemore et al., 1998; see also Pickering & Clark, 2014), albeit acknowledging the fact that the difference between efference copy and corollary discharge is not clear (Miall &

Wolpert, 1996). So efference copy can be viewed as input to the forward model, which then generates corollary discharge, i.e., predicted sensory consequences. However, there is evidence that efference copy includes the predicted sensory consequences. For example, when explaining why a rightward saccade does not result in a jumpy world, von Holst and Mittelstaedt (1950) explained that there is ‘an efference copy which on its own makes the visual scene move to the right, and an afferences which on its own makes it move to the left. Since, however, these two cancel each other out at the low level of Z1 no report ascends higher and we see neither movement; as witnessed by our everyday experience, the environment remains stable’ (see Figure 1.2 b for reference; translated by Peter C. Dodwell).

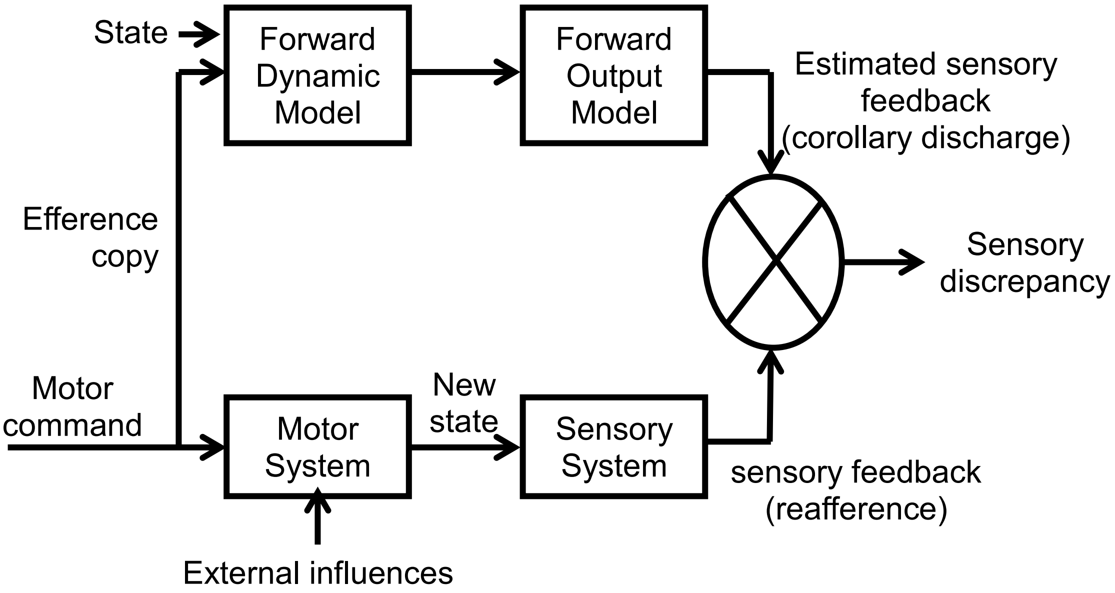


Figure 1.3 The forward model. See text for detail. Adapted from Miall & Wolpert (1996).

In summary, efference copy, corollary discharge and forward model acknowledge the same idea that self-generated sensory consequences are predicted and are processed differently from externally generated stimuli. In SA literature, they are the source of sensory prediction account of SA explanation. Throughout this thesis, these three concepts are used interchangeably to denote the same idea.

1.2.2 Sensory gating account

Sensory gating account asserts that movement imposes a general sensory gating towards incoming stimuli regardless of (or in addition to) prediction (Chapman, Bushnell, Miron, Duncan, & Lund, 1987; Horváth, Maess, Baess, & Tóth, 2012). For example, in the study by Horváth et al. (2012), participants performed a time interval generation task, in which they were asked to make successive button presses and to make sure that the interval between button presses formed a uniform distribution between 2 and 6 seconds. Brief tones were presented sporadically throughout the testing session. When irrelevant tones followed immediately after button presses, tone-evoked responses were smaller compared to when irrelevant tones were presented with a temporal delay (e.g., 1 second) after button presses. In this study, tones were not task-relevant and they did not necessarily follow a button press. SA was still observed. So the temporal proximity between a movement and a sensory effect might be enough to result in SA. Another study manipulated the predictability of the stimulus sound following a button press. Even when the frequency (ranging between 400 to 1990 Hz) and/or timing (ranging between 500 and 1000 ms after the button press) of the sound was unpredictable, SA was still present (Bäß, Jacobsen, & Schröger, 2008). Thus sensory gating account of SA also has empirical support.

1.2.3 Attention account

The argument of attention account for SA is that the attention directed towards self-generated stimuli is reduced, which contributes to the decreased amplitudes of evoked responses as seen in SA. When participants generate stimuli themselves (e.g., press a button to trigger a tone), attention may be diverted to the motor response task or other related sensory consequences (e.g., somatosensory feedback from the button press). Thus less attention is paid to the stimulus of interest (the tone) in the active condition as compared to the passive condition where the stimulus is passively presented (most SA studies have blocked design where different conditions are run in separate blocks). Since attention can boost sensory evoked responses (Hillyard, Hink, Schwent, & Picton, 1973; Lange, 2013; Schröger et al., 2015), it is not surprising that less attention in the active condition can lead to reduced responses, i.e., SA.

Several studies addressed this question. Baess et al. (2011) compared auditory evoked responses by bringing self-generated and externally generated tones within the same block so that the self-generated tone and the externally generated tone followed each other. In this way, the authors aimed to eliminate the possible attentional differences between conditions present in the blocked design. Interestingly, SA effect was found to be even larger in the mixed design than the traditional blocked design. Using similar experimental design, Timm, SanMiguel, Saupe, and Schröger (2013) took one step further by manipulating the attentional locus. In different conditions, participants were asked to count the number of tones, motor responses or visual cues on the screen. It was found that SA was present in all conditions and was independent of the attentional locus. Thus it is clear that a genuine SA effect exists independent of attention effect. However, it is highly likely that many SA studies were confounded (to some extent) by attention effect (Horvath, 2014; Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013).

1.2.4 Preactivation account

Preactivation account posits that before the onset of the action-induced stimulus, relevant sensory brain areas are activated by motor preparation, which can account for SA (Kuhn, Seurinck, Fias, & Waszak, 2010; Roussel, Hughes, & Waszak, 2013; Waszak, Cardoso-Leite, & Hughes, 2012). Since the activation of sensory brain areas is before the stimulus onset, it is called preactivation. The most direct support for this account comes from the study by Roussel et al. (2013), in which a visual contrast discrimination task was used. In the task, participants first learned the association between a motor response and a visual stimulus. The learned association then lead to decreased contrast discrimination sensitivity for the predicted visual stimulus (SA), which was accompanied by increased contrast ratings for the stimulus when the stimulus has low contrast (preactivation).

The major difference between preactivation account and the other three accounts (sensory prediction account, sensory gating account and attention account) is that preactivation account does not assume a genuine reduction in brain responses to self-generated stimuli as compared to externally generated stimuli, whereas the other three accounts do. The preactivation account explains the reduced evoked responses for self-generated stimuli as that, the evoked responses are always baseline corrected and the baseline level is enhanced for self-generated stimuli due to sensory preactivation. Thus amplitude differences between the post-stimulus period and the pre-stimulus period, i.e., evoked responses, get smaller. Preactivation account may be correct in the way that the sensory network for predicted stimuli is activated in the motor preparation period. But studies from single cell recordings clearly showed that the majority of recorded neurons in the auditory cortex have decreased firing rates during vocalizing compared to listening (Eliades & Wang, 2003, 2013), which argues for a genuine suppression effect in sensory cortices from motor responses.

It should be noted that the above mentioned four accounts are not necessarily mutually exclusive, except that a genuine SA effect should be acknowledged. The preactivation in the preactivation account may be a form of predictions in the prediction account. Prediction and general sensory gating may both exist for self-movements (Schröger et al., 2015). Attention may also play a role in determining the size of SA effect (Saupe et al., 2013).

1.3 Neural oscillations

Oscillatory activity is a prominent feature of electromagnetic brain signals. For example, alpha oscillation (8-12 Hz) can even be recognized with visual inspection from raw signals recorded with electroencephalography, as Hans Berger did in the 1920s (alpha oscillation is also known as Berger’s wave). Neural oscillations are categorized based on frequency. Apart from alpha oscillation, other well-studied oscillations in cognitive neuroscience include delta oscillation (1-3 Hz), theta oscillation (4-7 Hz), beta oscillation (13-30 Hz) and gamma oscillation (>30 Hz). Over the past decades, research in this field shows that neural oscillations play an important role in regulating brain functions. Oscillations in different frequency bands have been associated with different cognitive processes. For example, alpha oscillation is well-documented in visual spatial attention studies. In these studies, participants are usually asked to pay attention to one hemisphere of the visual field while simultaneously fixating on a cross in the centre of the screen. The attention diversion leads to an increase of alpha power (related to the amplitude of oscillations) in the ipsilateral hemisphere and a decrease of alpha power in the contralateral hemisphere over the occipital area (Foxe & Snyder, 2011; Jensen & Mazaheri, 2010; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). Besides the power, the phase of neural oscillations is also found to be important. For example, the phase of alpha oscillation is found to be associated with neural firing rates and visual perception (Busch, Dubois, & VanRullen, 2009; Haegens, Nacher, Luna, Romo, & Jensen, 2011; Mathewson, Gratton, Fabiani, Beck, & Ro, 2009).

Apart from influences on local neural processing, neural oscillations are also suggested to be a means of communication between neural populations. The distinction between two types of neural information propagation is apparent: feedforward (bottom-up) and feedback (top-down) information passage. Feedforward information goes from the periphery to the central nervous system and usually moves from lower hierarchies to higher hierarchies. It carries stimulus-related information. Whereas feedback information is considered to be related to internal states, such as attention and prediction. It is not surprising that feedback information goes from higher hierarchies to lower hierarchies. Some recent studies provide evidence that feedforward information is carried by gamma oscillation and that feedback information is carried by alpha/beta oscillation (Bastos et al., 2015; Michalareas et al., 2016; van Kerkoerle et al., 2014; X. J. Wang, 2010). For example, Michalareas et al. (2016) demonstrated that gamma oscillation and alpha/beta oscillation mediate feedforward and feedback influences along human visual hierarchies using MEG.

The information relay feature of neural oscillations is particularly relevant to the idea of efference copy. In Figure 1.2 a, the efference copy (EK) is actually a feedback signal that is related to the internal state of the brain, i.e., prediction. The mismatch between efference copy and reafference (M) is a feedforward signal that needs to be relayed to higher processing hierarchies. Thus it appears likely that recent findings on neural information processing provide a neural basis for the model proposed by von Holst and Mittelstaedt in 1950 (Figure 1.2 a). Indeed, modern theories of brain information processing (e.g., predictive coding theory) are trying to relate the prediction process to alpha oscillation and the prediction error (mismatch between the predicted stimulus and the real incoming stimulus) to gamma oscillation (Bauer, Stenner, Friston, & Dolan, 2014; Stenner, Bauer, Haggard, Heinze, & Dolan, 2014).

1.4 Neural circuits of SA

The neural circuit underlying SA has been studied in the context of the forward model. Several brain areas, including but not exclusive to, medial prefrontal cortex (Müller, Leske, Hartmann, Szebenyi, & Weisz, 2014), inferior parietal lobule (Sirigu et al., 2004), inferior frontal gyrus (J. Wang et al., 2014)), motor cortex (Ford et al., 2013; Reznik, Ossmy, & Mukamel, 2015) and the cerebellum (Blakemore, Frith, & Wolpert, 2001; Knolle, Schröger, & Kotz, 2013), have been suggested to be involved in generating the forward model.

Of special interest is the cerebellum. Knolle, Schröger, Baess, and Kotz (2012) showed that auditory SA is absent in cerebellar lesion patients, which provided causal evidence that the cerebellum is involved in generating the forward model. Furthermore, there is also evidence showing that the cerebellum is involved in updating the outdated forward model when a perturbation is brought for self-generated stimuli. For example, in visuomotor tasks, participants controlled the movement of a cursor on the screen. When a discrepancy was introduced between the controlling movement and the real cursor movement, normal participants can adapt to this change after a few trials of learning. But cerebellar lesion patients were significantly impaired in adapting to the change, indicating a deficit in updating the forward model (Izawa, Criscimagna-Hemminger, & Shadmehr, 2012; Synofzik, Lindner, & Thier, 2008).

In chapter 4, the role of the cerebellum in updating the forward model was probed in a study by stimulating the cerebellum with an inhibitory TMS (transcranial magnetic stimulation) protocol and measuring the updating of the forward model with MEG.

1.5 SA and diseases

A deficit in SA has been reported in several patient groups, including schizophrenia (Blakemore, Smith, Steel, Johnstone, & Frith, 2000; Ford et al., 2013; Shergill et al., 2014), Parkinson (Macerollo et al., 2016) and stuttering (Daliri & Max, 2015). Smaller or even abolished SA was found in these patients. Dysfunctional SA in schizophrenic patients has been linked to a deficit in generating the efference copy signal, which was related to positive symptoms such as delusion and hallucination (Heinks-Maldonado et al., 2007; Shergill et al., 2014). Even among normal participants, a correlation between the magnitude of SA effect and schizotype scores measured by questionnaires has been found (Oestreich et al., 2015; Teufel, Kingdon, Ingram, Wolpert, & Fletcher, 2010).

1.6 Purpose of this thesis

The general aim of this thesis is to extend the understanding of the mechanism how actions influence auditory processing. Compared to the number of SA studies using neuroimaging methods (EEG, MEG, fMRI), only a limited number of studies addressed the issue behaviourally. So the first aim is to extend the understanding of this phenomenon in the behavioural domain. Three behavioural studies are presented in chapter 2. The first study compared the effect of SA and attention. The second study is a cross-cultural comparison of how watching others generate a tone may influence auditory perception. The third study investigated the modulation of auditory perception in the process of language comprehension. Most SA studies with MEEG so far focused on the comparison of the amplitudes of evoked responses. In chapter 3, the classic SA effect was replicated with

MEG and time-frequency analysis was performed to interrogate the role of neural oscillations in implementing SA (study 4). In chapter 4, the role of the cerebellum in updating the forward model was investigated by interrupting the normal cerebellar function with TMS (study 5). Some questionnaire scores were collected in study 2 and study 4 to study the relationship between psychological traits (e.g., delusional ideation) and the prediction related process as implemented in SA among normal participants.

2 The influence of motor activation on auditory perception: behavioural investigation

2.1 Overview

This chapter contains three behavioural studies investigating the influence of motor activation on auditory perception. In study 1, sensory attenuation effect (SA) was compared with attention effect in a tone detection task. SA leads to attenuated sensory processing, whereas attention generally enhances sensory processing. In this study, we found that SA was still observable when the stimulus was paid substantial attention to. Yet the attention effect only decreased by around 20% when SA was present (Cao & Gross, 2015a). In study 2, loudness perception of a tone was compared cross-culturally in three conditions: the tone was either triggered by the participant, the experimenter, or the computer. Previous studies from different cultural regions reported divergent effects in the loudness perception of sounds generated by others, which suggests that a cultural difference exists. Indeed, a cultural difference was found in perceiving others triggered tone in this study, with Chinese showing attenuated loudness perception whereas the British not (Cao & Gross, 2015b). The amount of this effect correlated with the independent self-construal score, which constitutes an important cultural difference in social cognition (Markus & Kitayama, 1991). In study 3, we demonstrated that reading verbs that are related to action could also lead to a decrease in an auditory detection task. This is probably due to the motor activation from reading verbs that are related to action, which then has a suppression effect on auditory cortex (Cao, Klepp, Schnitzler, Gross, & Biermann-Ruben, 2016).

2.2 Study 1

(Published in *Plos One (2015)*, 10(8), e0136585)

2.2.1 Introduction

Behavioural studies on the loudness perception of self-triggered tones only started very recently. In his seminal study, A. Sato (2008) compared the perceived loudness of a same tone under different conditions. The tone was perceived to be softer when it was self-triggered than when it was triggered by the computer. The forward model was used to explain this effect, i.e., the motor area sends predictive information about the incoming stimulus to the auditory area and SA results from a matching between the prediction and the reafferent information. This was later replicated by Weiss et al. (2011), with a better control for the stimulus predictability in both conditions. When the stimulus was triggered by the computer, a visual cue was presented to allow participants to predict the timing of the stimulus. Stenner, Bauer, Sidarus, et al. (2014) provided more convincing evidence for the involvement of the motor area in this effect. In their study, tone-triggering motor responses were subliminally primed so that participants felt more control over their action consequences (i.e., the tone) in one condition than the other. In this case, the only difference between conditions resulted from the manipulation of motor processing through priming. A reduction in perceived loudness was found when participants felt more control over their action consequences after priming.

In contrast to SA from actions, attention improves sensory processing. For example, Greenberg and Larkin (1968) investigated how attending to a particular frequency tone could improve its detection. They found a significant increase of detection rate for the attended tone than the unattended tone, which should be equally detectable without the intervention of attention. However, no studies have looked at the effect of attention in SA behaviourally. This question is not trivial due to the ubiquity of attentional influences on sensory processing. Dissociating attentional effects from SA is very important for a better understanding of the latter. In electrophysiological studies of SA, attention effect has been addressed in several studies (Baess et al., 2011; Jones, Hughes, & Waszak, 2013; Saupe et al., 2013; Timm et al., 2013). These studies acknowledged a genuine SA effect as explained by the forward model independent of attention. However, how attention may interact with SA is still not very clear (Schröger et al., 2015).

Since attention enhances sensory perception and SA suppresses sensory perception, an interesting question is whether SA can still be observed when the stimulus is sufficiently attended to. We studied this question using a behavioural auditory detection task. In the task, participants were required to detect a near-threshold sound, where attention and SA were manipulated orthogonally. With this design we can measure the effect of attention and SA at the same time thus enabling a direct comparison between the two.

2.2.2 Methods

2.2.2.1 Participants

28 participants (13 females; mean age = 25.1, SD = 7.0; two left handed) were recruited through poster advertising on campus. They gave written informed consent prior to the experiment and received £6/hour compensation. The experiment was conducted conforming to the ethical codes of the declaration of Helsinki and was approved by Ethics Committee of College of Science & Engineering, University of Glasgow.

2.2.2.2 Main task

An auditory attention paradigm adapted from Borra, Versnel, Kemner, van Opstal, and van Ee (2013) was used (Figure 2.1). In the task, participants were required to detect a near-threshold target tone. In the beginning of a trial, a cue tone (clearly audible) was first played. After 600 ms, a cross appeared in the centre of the screen. In ‘no SA’ condition, the cross was present for 400 ms, and a sound detection task started automatically 100 ms after its offset. In the sound detection task, a near threshold target tone was always presented in either the first or the second interval, which participants were required to make judgment of. In ‘SA’ condition, participants were told to press a button (number ‘2’ on numeric section of a standard keyboard) with right index finger when they saw the cross. The same detection task started 100 ms after the button press, which was supposed to activate the forward model and lead to SA effect. The cue tone was always 1000 Hz, and the target tones could be 880 Hz, 1000 Hz, or 1120 Hz. Participants were told to ‘search for’ the same target tone in the sound detection task as the cue tone, whereas being informed that the target tone could be different from the cue tone at the same time. Thus, in the sound detection task, participants’ attention was drawn to the 1000 Hz tone (‘attended’), but not to the 880 Hz or 1120 Hz tone (‘unattended’). Previously, an auditory attention effect has been reported and is reflected in improved detection performance for attended target tone compared to unattended target tone (Borra et al., 2013; Greenberg & Larkin, 1968). We expect to replicate this in ‘no SA’ condition, which is similar to a standard auditory attention task.

All tones had durations of 250 ms (8 ms rise/fall). Continuous background white noise (volume set to a comfortable level) was presented during the experiment to control for environmental noise. The cue tone was 12 dB lower than the background noise and it was clearly audible. The volume of target tones was determined through a pre-test phase (see below). Auditory stimuli were delivered through headphones (Model: Sony MDR-XD100). Each condition consisted of 128 trials and was divided into two blocks. In half of trials, the target tone was 1000 Hz; in the other half of trials, the target tone was 880 Hz or 1120 Hz (each 50%). The probability of target tone presentation was equally distributed across the two intervals (both intervals had the duration of 250 ms, which is the same as the duration of the tone). The correspondence between response buttons and intervals was counterbalanced among participants (half participants pressed ‘A’ for the first interval, ‘D’ for the second interval; the other half pressed ‘A’ for the second interval, ‘D’ for the first interval). The order of the two conditions was also counterbalanced. Participants were instructed to avoid any unnecessary movements during the experiment.

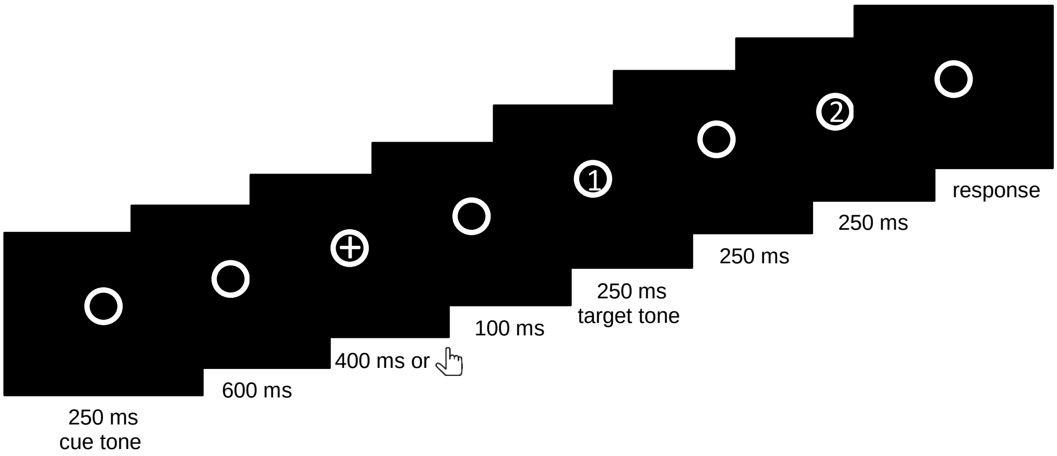


Figure 2.1 Timeline of one trial in the main task. Participants first heard a cue tone. 600 ms later, a cross appeared in the centre of the screen and then the sound detection task started either following participants’ button press or after a delay of 400 ms. The inter-trial interval was 1000 ms. In this example, the target tone is in the first interval.

2.2.2.3 Pre-test

An approximate detection threshold for each participant was established before the experiment. The task was similar to the main task, but no cue tone was played. Five 1000Hz tones with different intensities (compared to background noise; normally from -30.5 dB to -20.5 dB with 2.5 dB increment) were each presented 48 trials in a random order. Then the accuracy data for the five tones were fitted with a sigmoid function and 85% detection threshold for the 1000 Hz target tone was determined (Borra et al., 2013). The 85% threshold for the 880 Hz/1120 Hz target tone was attained by subtracting/adding 0.24 dB to the threshold for the 1000 Hz target tone according to the function between detection threshold and frequency reported by Green et al. (Green, McKey, & Licklider, 1959).

2.2.2.4 Data analysis

Corrected d′ as a measure of sensitivity between the two intervals (see formula 1) and response bias c as a measure of possible bias for certain responses (see formula 2) were calculated, respectively (Hautus, 1995). The data for 880 Hz and 1120 Hz target tones were concatenated because both tones were unattended. After concatenation, there were equal numbers of trials for attended (1000 Hz) and unattended tones (880 Hz and 1120 Hz). There were two factors in this study: tones could be ‘attended’ or ‘unattended’; ‘SA’ (with a button press) or ‘no SA’ (without a button press) was present in the sound detection task. Thus a within-subject 2 by 2 ANOVA was run with SPSS 19. Data from two participants were excluded due to chance level performance for all the target tones[2].

$$d' = z(\text{hit rate}) - z(\text{false alarm rate}) \qquad (1)$$

$$c = -0.5 \times [z(\text{hit rate}) + z(\text{false alarm rate})] \qquad (2)$$

Hit rate = (number of hits for first interval + 0.5) / (number of trials with target tone in first interval + 1);

False alarm rate = (number of false alarms for first interval + 0.5) / (number of trials with target tone in second interval + 1).

2.2.3 Results

A 2 (‘attended’ vs. ‘unattended’) by 2 (‘SA’ vs. ‘no SA’) ANOVA on d′ data revealed a main effect of attention ($F(1,25) = 94.99$, $p < .01$, $\eta^2 = .79$) and an interaction between

---

[2] Results do not change qualitatively even if data from these two participants are included.

attention and SA (F(1,25) = 9.89, *p* < .01, η^2 = .28). Sensitivity for the attended tone (mean = 1.45, SD = 0.62) was larger than the unattended (mean = 0.15, SD = 0.36), which is a replication of the classic auditory attention effect (Figure 2.2). Post-hoc analysis for the interaction effect suggested that sensitivity for attended tone in ’SA’ condition (mean = 1.33, SD = 0.73; mean hit rate: 71.45%, false alarm rate: 25.87%) was smaller than in ‘no SA’ condition (mean = 1.58, SD = 0.57; mean hit rate: 77.86%, mean false alarm rate: 24.59%) (t(25) = -2.86, *p* < .01). Attention effect still remained significant in ‘SA’ condition, so in this condition the performance for ‘attended’ tone was better than the ‘unattended’ tone (t(25) = 8.01, *p* < .01). There was no significant change for unattended tones between ‘no SA’ (mean = 0.12, SD = 0.42; mean hit rate: 51.52%, false alarm rate: 47.09%) and ‘SA’ condition (mean = 0.19, SD = 0.42; mean hit rate: 50.93%, false alarm rate: 43.82%)) (t(25) = -0.77, *p* = .45). Performance for unattended tones was at chance level and was therefore not further analysed or discussed. The main effect of SA was not significant (F(1,25) = 1.73, *p* = .20, η^2 = .07). The effect size for attention and SA effect in ‘attended’ condition was 2.02 and 0.56, respectively (Cohen’s d).

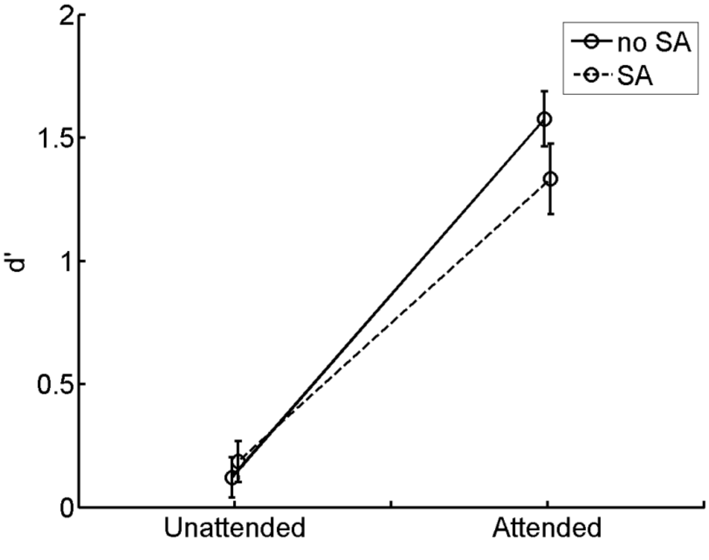


Figure 2.2 d’ results. d’ of behavioural performance for different conditions. Vertical bars stand for standard error. SA: sensory attenuation.

Response bias c was also subjected to ANOVA analysis, and no significant results were found indicating that differences between conditions were not due to changes in response bias (see Table 2.1).

Table 2.1 Response bias c for different conditions (with standard deviation in brackets). There are no significant main effects (attention: F(1,25) = 0.72, *p* = 0.41); sensory attenuation: F(1,25) = 2.33, *p* = 0.14) or interaction effect (F(1, 25) = 0.84, *p* = 0.37).

	Attended	Unattended
SA	0.05(0.25)	0.07(0.26)
no SA	-0.05(0.31)	0.02(0.32)

2.2.4 Discussion

We investigated the effects of SA and attention on auditory detection performance. The classic auditory attention effect was replicated, i.e., performance for the attended tone was better compared to the unattended tone. More interestingly, we found that in the presence of both SA and attention, auditory attention effect was significantly reduced but still maintained. Analysis of response bias confirmed that this result was not confounded by condition specific changes of response bias.

The predominant account for SA posits that SA results from correct prediction (with a motor origin through the forward model) of incoming sensory stimuli. To build up this prediction, some studies even included a pre-test phase where participants learned the association between the motor response and sensory stimuli (e.g., A. Sato (2008), Cardoso-Leite et al. (2010)). But this is not the case in our study. There is no motor prediction due to the design of our experiment, i.e., motor response does not necessarily produce a particular tone at a particular time. The data are consistent with the idea of a general suppressive effect of motor response to auditory cortex (maybe somatosensory cortex as well, see Walsh and Haggard (2007)) regardless of prediction (Horváth et al., 2012; Sanmiguel, Todd, & Schröger, 2013). A recent animal study also found indistinguishable modulatory effects on auditory cortex across many different movement patterns (Schneider, Nelson, & Mooney, 2014). This general suppression may be developed from daily experience that a motor response is always, if not all the time, associated with some kind of auditory feedback. Further studies are needed to confirm this hypothesis. Note that the general suppression effect is not mutually exclusive to motor prediction based SA, which may require some learning for association between motor response and corresponding sensory effect. Both of them may be developed from experience, but intensive learning during a short period of time makes the motor prediction based SA salient over the general suppression effect.

When the stimuli were attended, a clear SA effect was still observed, which is consistent with recent brain imaging studies (Schröger et al., 2015; Timm et al., 2013). An alternative explanation for the SA effect is that the decreased performance might result from attentional withdrawal from the discrimination task caused by the button press. As with conventional SA studies, this is difficult to address experimentally because a motor response (button press) is a prerequisite for SA, which can always be linked to attentional changes. In SA literature, the attentional withdrawal question has already been addressed and the conclusion is that attentional withdrawal cannot explain SA exclusively (Hughes et al., 2013b; Timm et al., 2013). For our data, an additional analysis was performed to indirectly address this question. We used coefficient variation (standard deviation divided by mean) of reaction time to the cross as a measure of attention allocated to the button press task. A small coefficient variation indicates a large attentional withdrawal from the discrimination task caused by the button press, thus should lead to a large performance decrease (SA). Therefore, if the reported effects were compatible with the attentional withdrawal account we would expect a negative correlation between coefficient variation and SA. However, the correlation is positive and not statistically significant (Spearman's rho = 0.37, $p > 0.05$). Thus attentional withdrawal due to button press is unlikely to explain the SA.

The attention effect (Cohen's d = 2.02) was much stronger than SA effect (Cohen's d = 0.56) with the current paradigm. The net effect on performance with attention and SA coexisting was still better for the attended than the unattended tone, i.e., attention outperforms SA. It is important to note that the absence of SA effect when the stimuli were not attended was due to the floor effect for the unattended tone, i.e., the detection performance for unattended tone was always at chance level. In the current study, there are not enough levels of attention to study the interaction between SA and attention, which might be an interesting topic for further behavioural exploration. An interesting question would be if SA can be completely suppressed with very strong attention, or if attention can be completely suppressed with very strong SA. With the current study paradigm, the unpredictability of the stimulus shouldn't have an effect on the attentional gain control process (Schröger et al., 2015), thus we predict that a very strong attention effect has the potential to completely supress SA and a very weak attention effect can be suppressed by SA. Van Hulle and colleagues (Van Hulle, Juravle, Spence, Crombez, & Van Damme,

2013) found that tactile suppression was decreased if attention was focused on the to be stimulated body part, thus suggesting that attention could cancel tactile suppression. Timm et al. (2013) reported an ERP study suggesting that SA is independent of attention. They compared SA effect in different attentional context and found that SA effect did not change in different attention conditions. While they did not report whether attention effect was still significant during SA, visual inspection of their results suggests that this is the case, which is similar to our finding. It is also interesting to compare our results with recent work on the interaction between attention and higher level prediction (Kok, Rahnev, Jehee, Lau, & de Lange, 2011). Higher level prediction has a similar effect to SA in reducing sensory signals, but that can be reversed by attention. In our study, SA was still significant in the presence of attention, which suggests that SA is different from higher level prediction (Hughes et al., 2013b).

2.3 Study 2

(Published in *Frontiers in Psychology* (2015), 6, 1865)

2.3.1 Introduction

In the section 1.1.1, we briefly reviewed behavioural studies on the loudness perception of self-triggered tones. Those studies all demonstrated sensory attenuation effect for tones that are self-triggered as compared to tones that are controlled by a computer (A. Sato, 2008; Weiss et al., 2011). In the same study by A. Sato (2008), a sensory attenuation effect for tones that were triggered by others (i.e., the experimenter) was also reported. Compared to the computer controlled tone, the tone generated by the experimenter was also perceived to be lower in loudness. However, the sensory attenuation for others effect was not replicated in the study by Weiss et al. (2011), even though a very similar testing paradigm was used as in the study by A. Sato (2008).

We noted that A. Sato (2008) study was from Japan and Weiss et al. (2011) study was from Germany. This raises the possibility that the sensory attenuation for others effect may be contingent on the cultural backgrounds of participants. Findings from cross cultural studies indicate that easterners (including but not limited to people from Asian countries like China and Japan) have more interdependent selves and westerners (including but not limited to people from European countries like UK and Germany, and North American countries like USA) have more independent selves (Markus & Kitayama, 1991). The difference between interdependent and independent selves manifests in many aspects. Generally speaking, people with more interdependent selves give more weight to their social (interpersonal) self than their personal self, which makes the influence from others more profound. On the contrary, people with more independent selves are influenced by others to a lesser degree (Markus & Kitayama, 1991). Numerous studies have demonstrated how self-construal has a profound influence on various aspects of social cognition (Cross, Hardin, & Gercek-Swing, 2011; Han & Northoff, 2008). For example, Chinese participants have less preference to their own face over others’ faces than British participants (Sui, Liu, & Han, 2009), and further studies showed that this difference may be modulated by self-construal priming (Sui, Hong, Liu, Humphreys, & Han, 2013). When interacting with others, people with more interdependent self-construal tended to unconsciously mimic other’s behaviour (e.g., face rubbing) more than people with more independent self-construal (van Baaren, Maddux, Chartrand, de Bouter, & van

Knippenberg, 2003). It was suggested that the difference in self-construal could be fundamental to cultural differences in cognition (Varnum, Grossmann, Kitayama, & Nisbett, 2010). The dominance of interdependency over independency may make the differentiation between self and others less salient among easterners than among westerners. Thus we hypothesize that easterners' perceptual experience towards a tone generated by others is similar to a self-triggered tone and they show sensory attenuation for both self and others. While for westerners, there is a clear differentiation between self and others so that others generated tone is just like an external tone. Thus we hypothesize that there is sensory attenuation for the self but not from others for westerners.

To summarize, we predict a cultural difference in sensory attenuation for other-generated tones, with people from collectivism-dominated cultural backgrounds (e.g., Chinese) showing this effect and people from individualism-dominated cultural backgrounds (e.g., British) not showing this effect. We hypothesise that there is no cultural difference with respect to sensory attenuation for the self, which is accounted for by the common forward model mechanism (Wolpert and Ghahramani (2000)). Further support for this hypothesis is that sensory attenuation for the self was shown in both A. Sato (2008) and Weiss et al. (2011) studies. In both studies, the perceived intensity of a standard tone was estimated through a volume comparison task between the standard tone and tones of differing intensities. Then the perceived intensity of the same tone was compared between different conditions, e.g., between the self-triggered tone and computer generated tone. The same method will be used in the current study, with participants from both cultural backgrounds being tested in the same experimental setting to reduce any possible confounding factors. We also included a battery of questionnaires, including Self Construal Scale (Singelis, 1994), short forms of Empathy Quotient (EQ-short), short forms of Systemizing Quotient (SQ-short) (Wakabayashi et al., 2006) and Peters et al. Delusion Inventory (PDI) (Peters, Joseph, Day, & Garety, 2004), to measure cultural differences and cognitive style (for more information, please see the methods section). Relevant to the focus of this study, we predict a correlation between sensory attenuation for others and self-construal, as self-construal plays an important role in cultural differences as outlined earlier.

2.3.2 Methods

2.3.2.1 Participants

30 Chinese (15 females, mean age = 22.4, SD = 1.7, 1 left-handed) and 30 British (15 females, mean age = 21.9, SD = 1.7, all right handed) participants, most of whom are students from the University of Glasgow, were recruited through a local subject pool. Chinese participants were all born and educated in China and they were tested within three months of their first arrival in UK. British participants were selected if their self-reported nationality information was UK and ethnicity information was British & White. All participants have self-reported normal hearing and normal or corrected to normal vision. Participants were debriefed and received a payment of £6/hour after the experiment. The experiment was conducted in accordance with ethical codes of conduct of APA, BPS, and declaration of Helsinki and was approved by the Ethics Committee of College of Science & Engineering, University of Glasgow. For each participant written consent was obtained prior to experiment.

2.3.2.2 Stimuli, Task and Procedure

Participants completed a sound comparison task, in which they were asked to compare which was louder between a standard tone and a comparison tone. The standard tone was 74 dB SPL in intensity and the comparison tone ranges from 71 dB to 77 dB with 1 dB increment. All tones (1000 Hz, 100 ms in duration, 10 ms rise/fall ramp, sampling rate at 48000 Hz) were generated with MATLAB (http://www.mathworks.com).

The experiment consisted of a pretesting phase and a testing phase. In the pretesting phase, participants pressed a button (number ‘2’ on numeric section of a standard keyboard) with their right index finger about once every 3 seconds. After a button press, they heard a standard tone immediately and received visual feedback on the screen whether the response was good, too slow (more than 3.8 seconds after last response) or too fast (less than 2.2 seconds after last response). There were 200 trials in total, and participants were allowed to take a break when needed. The purpose of including this pretesting phase was to keep the procedure the same as the procedure used in A. Sato (2008). Since the procedure was identical for participants from both cultural groups, the pre-testing phase was not crucial to the cultural difference question we were interested in.

In the testing phase (Figure 2.3), participants completed the sound comparison task. In each trial, participants first heard the standard tone. After a jittered interval of 800-1200 ms, they heard a comparison tone and then made a judgment which tone was louder by

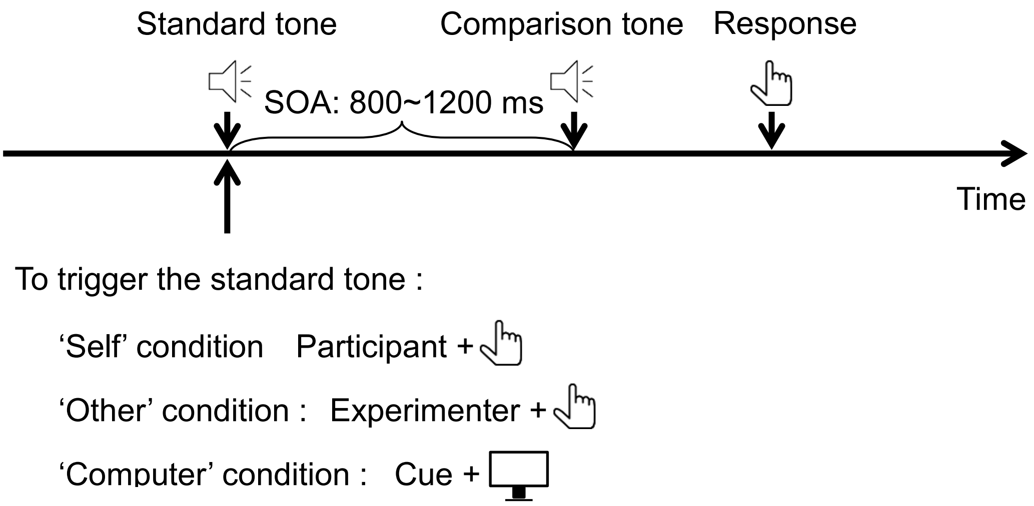


Figure 2.3 Schematic illustration of a typical trial. The standard tone is triggered by participant ('self' condition), experimenter('other' condition), or computer ('computer' condition). The comparison tone plays automatically after an SOA of 800–1200 ms. Participants are instructed to respond as accurately as possible which tone is louder. SOA: stimulus onset asynchrony.

pressing button 'F' (if the first one is louder) or 'J' (if the second one is louder) on the keyboard with their left hand. The intensity of comparison tones was randomized across trials. There were 3 conditions in the testing phase that differed in the way the standard tone was triggered. In 'self' condition, the standard tone followed immediately after participants pressing the button '2' with their right hand as in the pretesting phase. They were asked to press the button about once every 3 seconds after a response was made for the previous trial. No feedback of press latency was provided. In 'other' condition, the standard tone was triggered by the experimenter pressing the button '2' with right hand in the same way as participants did in 'self' condition. Participants were required to pay attention to the experimenter's hand in the whole process. In 'computer' condition, the computer controlled the presentation of the standard tone and participants received visual cues (from one second before the onset of standard tone, the cross in the screen centre gradually enlarged in size and then changed its colour from black to red just before the tone presentation) before the tone played. The onset of the standard tone was between 2.5 and 3.5 seconds after participants' response to the previous trial. The three conditions each contained 210 trials (30 x 7 comparison tones) and were presented in a random order. Each condition was presented in three mini blocks, each containing 70 trials, and participants were offered a break after each mini block.

Tones were delivered to subjects through a set of headphones (Beyerdynamic, DT770 PRO Headset-250 OHM). A male experimenter (LC, Chinese) tested all the male participants and two female experimenters (one Ukrainian and one Swedish) tested all the female participants. Under this design, half of the participants from either cultural group were tested by a same race experimenter and the other half were tested by an other race experimenter. If there is an effect from the race of the experimenter, both groups should be equally affected. During the experiment, the experimenter sat next to the participants while in front of the stimulus computer all the time. Participants were told to try their best to avoid any unnecessary movements. Except in the 'other' condition, participants were asked to always fixate on the cross in the centre of the screen. Participants' behaviour was monitored online by the experimenter. Questionnaires were completed after the experiment. Chinese participants completed translated versions of Self-construal Scale (Y. Wang, Yuan, & Xu, 2008), EQ-short and SQ-short questionnaires (translated and back-translated). The experiment took between 2 and 2.5 hours for each participant.

2.3.2.3 Questionnaires

The included questionnaires were: Self Construal Scale (Singelis, 1994), short forms of Empathy Quotient (EQ-short), short forms of Systemizing Quotient (SQ-short) (Wakabayashi et al., 2006) and Peters et al. Delusion Inventory (PDI) (Peters et al., 2004). Self Construal Scale is intended to measure individual's self-construal pattern, which is the key measurement of cultural difference in the study. An independent self-construal score and a dependent self-construal score was provided. EQ-short and SQ-short questionnaires are developed based on Empathizing-Systemizing theory (Baron-Cohen, 2002). They are intended to be used to measure the general cognitive style (more social or more systematic). Since social interaction is involved in sensory attenuation for others, we suspect that the way other's behaviour is cognized could be very important. So it is an interesting question to see whether sensory attenuation for others is related to EQ or SQ. PDI measures individual delusional state, which is related to sensory attenuation for the self (see discussion section). The correlation between PDI and sensory attenuation for the self has been reported (Teufel et al., 2010). We hypothesise a replication of this correlation between PDI and sensory attenuation for the self.

2.3.2.4 Data analysis

Point of subjective equality (PSE) was used to assess the subjective perceptual intensity of the standard tone. To compute PSE, the percentage of comparison tones perceived as louder than the standard tone was computed for each of the seven intensities of the comparison tone and for each participant and condition. Data were fitted with a logistic function using the maximum-likelihood method (Wichmann & Hill, 2001). PSE is defined as the intensity where participants respond 50% of times that the comparison tone is louder. Three data sets from the British sample were excluded due to deviant PSE values (more than 2.5 standard deviations away from the mean), resulting in 30 data sets with Chinese sample and 27 data sets with British sample for the final analysis. Since group size is unbalanced, linear mixed-effects model was used to test the interaction and main effects of PSE (Baayen, Davidson, & Bates, 2008; Barr, 2013). The effects of sensory attenuation for the self and others were analysed separately. When analysing sensory attenuation for others, different linear mixed-effects models were built comprising of main and/or interaction effects from the 2 (cultural groups) by 2 (‘others’ and ‘computer’ conditions) design. For example, when testing the interaction effect, the model that includes both interaction effect and main effects was compared to the model that only includes main effects. A p value will be derived from this model comparison. All significant effects will be followed up by t-test of PSEs between conditions. Sensory attenuation for the self was analysed in the same way but including ‘self’ and ‘computer’ conditions instead of ‘other’ and ‘computer’ conditions. Besides PSE values, participants’ responses toward the 74 dB comparison tone were analysed in the same way to assess the sensory attenuation effect. This is the sensitive part of the task where the standard tone and the comparison tone are actually identical (some studies only used this part in their design, see Reznik, Henkin, et al. (2015); Stenner, Bauer, Sidarus, et al. (2014)). When the two tones are identical, sensory attenuation would lead to more responses for the comparison tone as louder. So we calculated the percentage of responding the comparison tone as louder in this case, and an increase of this percentage value would indicate sensory attenuation.

For the correlation analysis, we quantified sensory attenuation the self effect by subtracting the PSE value in ‘computer’ condition from the PSE value in ‘self’ condition (SA_{self}); sensory attenuation for others is measured by subtracting the PSE value in ‘computer’ condition from the PSE value in ‘other’ condition (SA_{other}). Then correlation analysis was performed between SA_{self}/SA_{other} and the four questionnaire scores with an open source Matlab toolbox that uses robust correlation (Pernet, Wilcox, & Rousselet, 2012). Pearson’s correlation was used. Whenever a significant correlation was found, it was then tested with percentage-bend correlation and skipped-correlation for robust correlation tests. Pearson percentage-bend correlation results were reported in the main text.

2.3.3 Results

2.3.3.1 Sensory attenuation for others

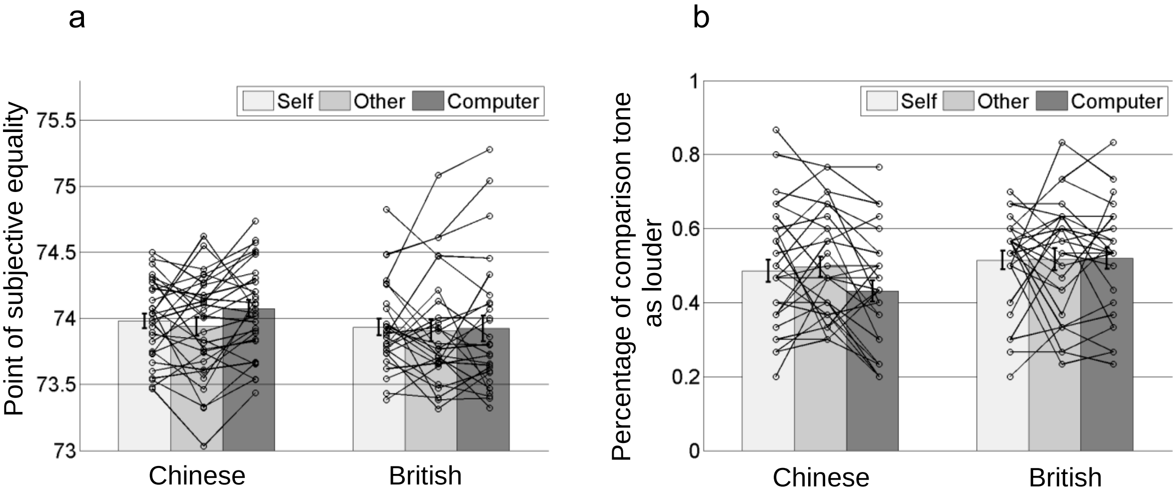


Figure 2.4 Main results. Point of subjective equality (a) and Percentage of comparison tone (74dB) as louder (b) in three conditions for Chinese and British. Vertical bars represent standard error. Each superimposed line represents the data of one participant.

A significant main effect was found for conditions (χ^2 (1) = 4.11, p < .05), with PSE values smaller in ‘other’ condition (mean = 73.92; SD = 0.40) than in ‘computer’ condition (mean = 74.0; SD = 0.43) (Figure 2.4 a). No other effects were significant (cultural groups: χ^2 (1) = 0.80, p = .37; interaction: χ^2 (1) = 2.58, p = .11). Based on our a priori hypothesis about the cultural difference, we performed within cultural group t-test as a follow-up. The within cultural group t-test showed that sensory attenuation for others among Chinese was significant with a smaller PSE value in ‘other’ condition (t(29) = -2.79, p < .01, 95% Confidence Interval (CI) = [-0.24, -0.04], Cohen’s d = -0.51) and no such effect was found among British (t(26) = -0.28, p = .78, CI = [-0.13, 0.10], Cohen’s d = -0.05). Analysis with participants’ responses toward the 74 dB comparison tone resulted in a main effect for conditions(χ^2 (1) = 4.28, p < .05), i.e., more responses for the comparison tone as louder in ‘other’ conditions as compared to ‘computer’ condition. And most importantly, a significant interaction effect emerged (χ^2 (1) = 4.95, p < .05) (Figure 2.4 b). Post-hoc analysis gave similar results to the analysis with PSE. Chinese showed significant sensory attenuation for others with a higher percentage of responding comparison tone as louder in ‘other’ condition (mean = 0.50, SD = 0.15) than in ‘computer’ condition (mean = 0.43, SD = 0.16) (t(29) = 3.20, *p* < .01, CI = [0.02, 0.11], Cohen’s d = 0.58), which was not the case for British participants (‘other’ conditions: mean = 0.52, SD = 0.16; ‘computer’ condition: mean = 0.52, SD = 0.15; t(26) = -0.11, *p* = .91, CI = [-0.05, 0.04], Cohen’s d = -0.02).

2.3.3.2 Sensory attenuation for the self

Surprisingly, no main effects (cultural groups: χ^2 (1) = 1.31, *p* = .25; conditions: χ^2 (1) = 0.85, *p* = .36) or interaction (χ^2 (1) = 1.29, *p* = .26) were found (Figure 2.4 a). Following up within group t-test showed that there was a trending significant sensory attenuation for the self among Chinese with a lower PSE in ‘self’ condition (mean = 73.98, SD = 0.30) than in ‘computer’ condition (mean = 74.08, SD = 0.34) (t(29) = -1.86, *p* = .07, CI = [-0.20, 0.01], Cohen’s d = -0.34), and no sign of sensory attenuation for the self among British (‘self’ conditions: mean = 73.93, SD = 0.33; ‘computer’ condition: mean = 73.92, SD = 0.50; t(26) = 0.12, *p* = .90, CI = [-0.15, 0.17], Cohen’s d = 0.02). Analysis with responses for 74 dB comparison tone gave similar results (for Chinese: (t(29) = 1.88, *p* = .07, CI = [-0.005, 0.11], Cohen’s d = 0.34; for British: (t(26) = -0.19, *p* = .85, CI = [-0.06, 0.05], Cohen’s d = -0.04; no interaction effect: χ^2 (1) = 2.29, *p* = .13) (Figure 2.4 b).

2.3.3.3 Correlations

Table 2.2 showed a summary of the correlations between SA_{self}/SA_{other} and questionnaire scores. SA_{self} is the PSE difference between ‘computer’ condition and ‘self’ condition, which is used to quantify sensory attenuation for the self effect. And SA_{other} is used to quantify sensory attenuation for others effect (see methods section). For SA_{other}, a significant correlation was found only with independent self-construal score (r = .40, *p* < .01, CI = [.14, .65]; Figure 2.5 a). Since negative values indicated positive SA_{other} effects, the correlation suggested that a higher independent self-construal score was associated with a reduced SA_{other}. The correlation is still significant within the Chinese sample (r = .48, *p* < .01, CI = [0.03, 0.80]), and the British sample (r = .41, *p* < .05, CI = [0.02, 0.74]). For SA_{self}, a significant correlation was found with PDI overall score among the British group (r = .44, *p* < .05, CI = [.02, 0.76]; Figure 2.5 b) but not among the Chinese group (r

= -.11, *p* = .56, CI = [-0.49, 0.24]). Among the British group, higher PDI overall scores (delusions) are associated with reduced SA_{self}. SA_{self} also has promising correlation with other PDI break-down measurements: yes/no score (r= .39, *p* = .04, CI = [-0.04, 0.72]), distress score (r = .38, *p* = .05, CI = [-0.07, 0.71]), preoccupation score (r = .44, *p* = .02, CI = [0.02, 0.76]), conviction score (r = .40, *p* = .04, CI = [0.01, 0.73]). All significant (or trending significant) results remained significant when tested with normal Pearson correlation or other robust statistics such as skipped-correlation (see Table 2.2 for results with normal Pearson correlation test).

Table 2.2 Summary of correlation analysis between SAself/SAother and 4 questionnaires scores using Pearson's correlation and Pearson bend correlation.

Measure		EQ	SQ	Ind. SC	Dep. SC	PDI (overall) (British only)
SA_{self}	(*Method*) Pearson	-.11 [-.36, .14]	0.01 [-0.26, 0.25]	.17 [-.07, .40]	0 [-.26, .30]	.37† [.02, .67]
	Pearson bend	-.07 [-.35, .20]	0 [-0.28, 0.28]	.17 [-.12, .43]	0 [-.29, .28]	.44* [.02, .76]
SA_{other}	Pearson	.01 [-.27, .30]	0 [-.28, .23]	.44* [.21, .64]	-.02 [-.31, .31]	.20 [-.23, .65]
	Pearson bend	.07 [-.22, .38]	-.04 [-.33, .23]	.40* [.14, .65]	.06 [-.23, .33]	.15 [-.28, .59]

† *p* = .06; * *p* < .05

Note: Data shown are Pearson's r values, and 95% confidence intervals are included in the brackets. Pearson skipped correlations gave similar results. Note that the p value for the correlation between SA_{self} and PDI overall is just above .05, but the bootstrapped 95% confidence interval is above 0 and Pearson bend correlation was significant. SA_{self}: measurement of sensory attenuation for the self effect; SA_{other}: measurement of sensory attenuation for others effect; EQ: empathy quotient; SQ: systemizing quotient; Ind. SC: independent self-construal; Dep. SC: dependent self-construal; PDI: Peters et al. delusion inventory.

2.3.4 Discussion

In this study, we report a cultural difference in perception of sensory consequences

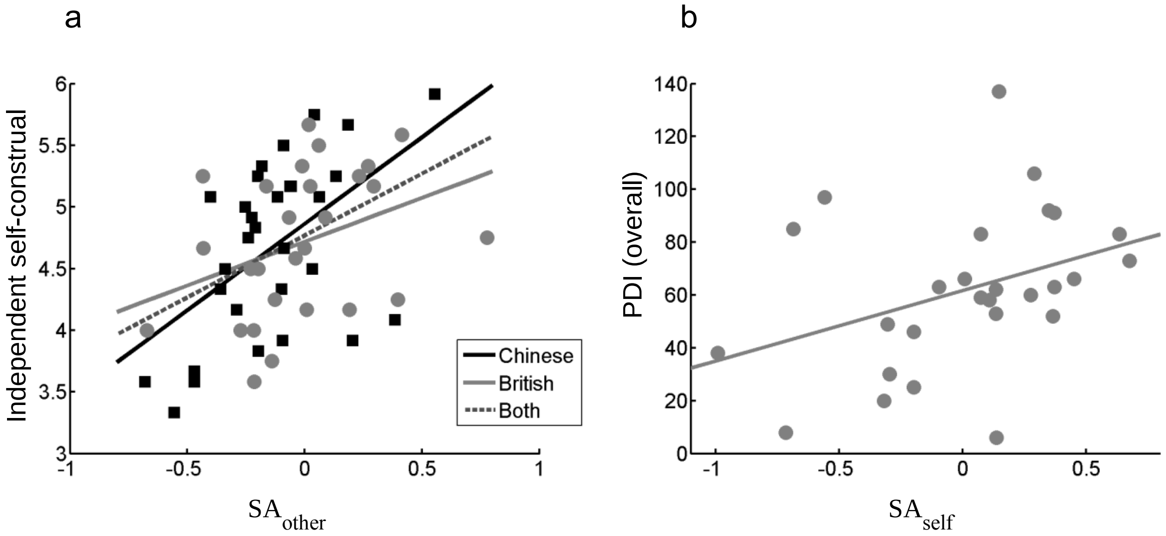


Figure 2.5 Correlation results. Scatter plot with SA_{other} on x-axis and Independent self-construal on y-axis (a), and SA_{self} on x-axis and PDI (overall) score on y-axis (b). Linear fit was shown with lines. Low $\text{SA}_{\text{self}}/\text{SA}_{\text{other}}$ values correspond to strong sensory attenuation effects. PDI: Peters et al.(2004) delusion inventory; SA_{self}: sensory attenuation for self-generated stimuli; SA_{other}: sensory attenuation for others generated stimuli.

generated by others. Chinese participants showed sensory attenuation to stimuli generated by others, but this effect was not observed in British participants. Moreover, SA_{other} (quantifying sensory attenuation for others effect) was significantly correlated with independent self-construal, not interdependent self-construal or emphasizing abilities. Higher independent self-construal scores were related to smaller sensory attenuation for others effects. We failed to replicate sensory attenuation for the self in either Chinese or British sample, though a trend towards significance was found among Chinese (p = .07). However, a significant correlation between SA_{self} (quantifying sensory attenuation for the self effect) and delusional ideation was found among British.

Sensory attenuation for the self is generally considered as a consequence of internal forward model, which posits that along with any action an efference copy is sent to sensory areas that allows prediction of the expected sensory consequence of an action. Sensory attenuation for the self occurs when the reafferent signal corresponds to the expectation (Sperry, 1950; von Holst & Mittelstaedt, 1950; Wolpert & Ghahramani, 2000). Its functional role was suggested to help individuals distinguish self-triggered stimuli from external stimuli, thus keeping the sense of agency (Blakemore, Frith, & Wolpert, 1999; A. Sato & Yasuda, 2005). Schizophrenic patients suffer from impaired sense of agency and indeed they show reduced or diminished sensory attenuation for the self (Ford, Gray, Faustman, Roach, & Mathalon, 2007; Shergill, Samson, Bays, Frith, & Wolpert, 2005). Specifically, the key factor responsible for the abnormality with sensory attenuation for the self was found to be delusional ideation (Heinks-Maldonado et al., 2007; Teufel et al., 2010). Supporting this, we found a significant correlation between SA_{self} measurement and delusional ideation among healthy subjects in the British sample (cf. Teufel et al., 2010). A closer examination of the correlation suggests that SA_{self} is related to almost all the aspects of delusional ideation as measured by the PDI break-down scores. Yes/no score (indicating the scope of delusional ideation) and distress score (indicating how distressing one feel about the delusional ideation) were either marginally significantly correlated with SA_{self} or the confidence interval for the correlation just contained 0. The correlations between SA_{self} and conviction score (indicating to what extent one is convinced of the delusional ideation), between SA_{self} and preoccupation score (indicating how much time one would spend on delusional ideation) were significant and had good confidence intervals suggesting that sensory attenuation for the self is more related to the intensity of delusional ideation than the scope or emotional consequences of delusional ideation. However, the correlation between SA_{self} and delusional ideation was not significant in the Chinese sample. We suspect that this may be because the PDI questionnaire is not suitable for Chinese population. This is evidenced by the surprisingly high PDI overall scores in our Chinese sample (mean = 102; median = 103). Peters et al. (2004) reported that in Western cultural context the mean scores for normal and deluded groups are 59 and 131, respectively. Since no norms are available for Chinese PDI scores, further studies are needed to clarify this question.

For unknown reasons we failed to obtain a significant sensory attenuation for the self effect, which has been previously shown (A. Sato, 2008; Weiss et al., 2011). Like previous studies, the standard tone in the current study was a unisensory stimulus in all conditions. The difference among conditions was what happened before the presentation of the tone. The ‘computer’ condition functioned to exclude the alternative explanation that the predictability of the stimulus onset explained the conditional differences. Similar manipulations for the control condition have been reported before (e.g., Ford et al. (2007)).Sensory attenuation for the self renders a 74 dB tone to be perceived about 0.4 dB (A. Sato, 2008) or 0.2 dB (Weiss et al., 2011) softer, which is a very small perceptual effect and may be difficult to be detected. In our study, the 74 dB tone was perceived as 0.1 dB softer and 0.01 dB louder when it was self-triggered for Chinese and British participants, respectively. Whereas the mean of SA_{self} may shift due to unknown reasons, the variance of SA_{self} across participants may not. This might be the reason why we can still find a significant correlation between SA_{self} (but not SA_{other}) and delusional ideation, which also argues for the validity of the data. It is also worth noting that it is unlikely that there is a cultural difference in sensory attenuation for the self, despite the numerical differences in SA_{self} found here. This is because that consistent cross-cultural results were reported on this topic (A. Sato, 2008; Weiss et al., 2011) and that the underlying theoretical forward model account applies universally (for example, it is true that self-tickling is less ticklish than being tickled by others for both easterners and westerners; Crapse & Sommer, 2008). However, this should be left as an open question for future studies that have more robust measurements of sensory attenuation for the self. Unlike the behavioural measurement of sensory attenuation for the self, the electrophysiological measurement of this effect is very robust (Hughes et al., 2013b; Schröger et al., 2015). Typically sound evoked responses are smaller when the sound is self-triggered (e.g., 'self' condition in the current study) as compared to when it is from external sources (e.g., 'computer' condition in the current study). However, the relationship between the suppression of evoked responses and behavioural measurements of sensory attenuation is not clear from the literature. Some recent studies from the visual domain suggest that both effects may reflect the consequences of a common underlying mechanism, i.e., internal forward model (Hughes, 2015; Stenner, Bauer, Haggard, et al., 2014). In the auditory domain, we speculate that the suppression of evoked responses is a more unambiguous and direct measure of sensory attenuation but that the perceptual intensity judgement changes are a likely consequence of the alleged underlying mechanism. Further studies are needed to clarify this point.

Interestingly, Chinese but not British showed sensory attenuation for tones generated by others. This is consistent with our prediction stemming from the self-construal difference between cultural groups. The dominance of interdependency over independency in Eastern culture may lead to fewer differences between sensory consequences generated by others and self, thus easterners show sensory attenuation for others just like sensory attenuation for the self as reported in the literature. For westerners, stronger independency may make the sensory consequences from others distinct from sensory consequences from the self, thus the sensory consequences from others may be of no difference to external sensory stimuli. So westerners did not show sensory attenuation for sounds caused by others. Further support for the above explanation comes from the strong correlation between

SA_{other} and the independent self-construal score. A large SA_{other} (large sensory attenuation for others effect) is associated with a small independent self-construal score. This reconciles the discrepant findings reported by A. Sato (2008) and Weiss et al. (2011). In the study by A. Sato (2008), participants and experimenter always pressed the same button with the same finger to trigger the same tone, whereas in the study by Weiss et al. (2011), participants and experimenter pressed different buttons to trigger a different (experiment 1) or same (experiment 2) tone. It is possible that the similarity between participants' and experimenter's response pattern leads to similar sensory attenuation effect following self and other's movement in A. Sato (2008). By testing both groups of participants in the very same experimental setting, this possibility was ruled out and we confirmed sensory attenuation for others as a cultural phenomenon. A recent EEG study showing attenuated N1 neural responses to a tone after watching a goal-directed button press video was reported from Australia, where individualism is more prominent (Poonian, McFadyen, Ogden, & Cunnington, 2015). This seems to be at odds with Weiss et al. (2011) study and our data here. However, a different behavioural task (time estimation) in their study may hinder a direct comparison of the results. As discussed earlier, this may also suggest that sensory attenuation measured from cortical responses and perceptual intensity judgements may not exactly be the same thing.

Social context can modulate sensory processing (Baess & Prinz, 2015; Desantis, Weiss, Schütz-Bosbach, & Waszak, 2012). For example, Desantis et al. (2012) showed that sensory attenuation effect can be modulated by authorship belief. We also found that SA_{other} was significantly correlated with independent self-construal both when analysing the two cultural groups together and separately, which suggests an influence of social orientation (general belief) on sensory processing. This is also direct evidence supporting the social orientation hypothesis on the origin of cultural differences in cognition (Varnum et al., 2010). According to the social orientation hypothesis, the fact that westerners are more independent and easterners are more interdependent is the origin to the various aspects of cultural differences in cognition. Independent/Interdependent self-construal is a key factor of the independency/Interdependency. SA_{other} is not correlated with EQ or SQ (see Table 2.2), which might suggest that empathizing-systemizing cognitive style is not related to sensory attenuation for others. What are the potential neural mechanisms underlying the sensory attenuation for others effect? One possible mechanism could be that other brain areas (possibly prefrontal cortex) modulate the neural responses in auditory cortex during this process (Müller et al., 2014). Another mechanism could be that the internal forward model still accounts for sensory attenuation for others, but it is activated by seeing other’s movement (Kilner, Friston, & Frith, 2007; Poonian et al., 2015). As discussed by A. Sato (2008) and Weiss et al. (2011), mirror neurons could be the mediator. However, that would also assume a cultural difference in mirror neurons, which is under debate (Cook, Bird, Catmur, Press, & Heyes, 2014).

Given that Chinese showed sensory attenuation for others and British did not, and the effect was correlated with independent self-construal, it is reasonable to predict that a difference in independent self-construal would be found between the two groups. However, this is not the case (mean for Chinese: 4.67; mean for British: 4.71; t(55) = -0.20, *p* = .84, CI = [-0.38 0.32]). The sample size could be too small to identify this difference. Another explanation is that the explicit self-construal measure in the Singelis self-construal scale is prone to situational influence (Cross et al., 2011). All the Chinese participants were tested shortly (mostly within one month) after their first arrival in UK, when they were trying to adapt to a new environment by themselves. That could promote their explicit sense of independence, which was reflected in Singelis self-construal scores. Independence as revealed through sensory attenuation for others can be more resistant to the influence from social environment within a short period of time, as it is more like an implicit measure of attitudes towards others. Whether or not it changes with more enculturation is an interesting question for further studies. Contrary to independent self-construal, we found a significant difference in interdependent self-construal with Chinese scoring higher (mean for Chinese: 5.25; mean for British: 4.72; t(55) = 2.87, *p* < .01, CI = [0.16 0.90]). Interestingly, interdependent self-construal was not correlated with SA_{other} (cf. Kitayama & Park, 2014). We view this as supporting evidence that independent and interdependent self-construals are two non-reducible aspects of self (Singelis, 1994) and may have independent influence on cognition and behaviour. For example, Kitayama and Park (2014) showed that the culture difference in self-centric motivation was mediated by interdependent self-construal but not independent self-construal.

2.4 Study 3

(Published in *Experimental Brain Research (2016),* 234(10), 3049-3057)

2.4.1 Introduction

According to theories of embodied cognition, language comprehension involves a distributed brain network, including sensorimotor areas - at least for concrete words (Barsalou, 2008; Pulvermüller, 1999, 2013). Accumulating evidence from neuroimaging studies supports this idea. For example, in an fMRI study Hauk, Johnsrude, and Pulvermüller (2004) showed a somatotopically organized activation in the motor and premotor cortex when participants passively read action-related verbs, i.e., leg-related action words lead to activations more medially than arm- or face-related action words. A recent MEG study replicated this finding and found that the activation in the motor cortex could be detected about 80 ms after the onset of the full word, thus ruling out the possibility that the observed activation in the motor cortex is a result of mental imagery and supporting the idea of embodied language comprehension (Shtyrov, Butorina, Nikolaeva, & Stroganova, 2014). In addition to neuroimaging studies, behavioural studies showed that motor responses (reaction time (RT), accuracy) were influenced after reading action-related words (Andres, Finocchiaro, Buiatti, & Piazza, 2015; Boulenger et al., 2008; Klepp, Niccolai, Buccino, Schnitzler, & Biermann-Ruben, 2015; Mirabella, Iaconelli, Spadacenta, Federico, & Gallese, 2012; M. Sato, Mengarelli, Riggio, Gallese, & Buccino, 2008). For example, Mirabella et al. (2012) found increased RTs and increased error rates in motor responses in a Go-NoGo task when the same effector was involved as in the presented verbs.

Besides the findings from the motor area, domain-specific activation during language comprehension in olfactory, gustatory, and auditory brain areas was also found (Barrós-Loscertales et al., 2002; González et al., 2006; Kiefer, Sim, Herrnberger, Grothe, & Hoenig, 2008). Kiefer et al. (2008) found stronger activation in the auditory brain area when participants read words that were sound related compared to words that were not sound related. In a single case report, a patient with focal lesions in the left auditory area demonstrated deficits in processing words depicting sound-related everyday objects, e.g., a bell (Trumpp, Kliese, Hoenig, Haarmeier, & Kiefer, 2013). However, unlike in the motor domain, behavioural effects in sensory domains are rarely reported. The aim of this study is to investigate whether auditory perception is modulated when participants read verbs related to sound, to sound plus action, or to neither sound nor action.

To assess auditory perception, a sound detection task was used, in which participants were required to indicate whether a near threshold tone was presented or not. We hypothesize a priming effect on auditory perception after reading sound-related words (i.e., an enhanced detection performance), drawing on the finding that the auditory area is involved in the comprehension of visually presented sound-related words (Kiefer et al., 2008). For words that are related to both sound and action, we predict a cross domain effect on auditory perception from the motor involvement, which adds on the auditory priming effect hypothesized above. The rationale of this hypothesis is that a suppressive pathway from the motor cortex to the auditory cortex was reported in human and animal studies. Martikainen et al. (2005) showed that brain responses to a simple tone were smaller when the tone was triggered by a button press of the respective participant than when the tone was externally controlled by the computer. Recent animal studies suggested a motor origin of this suppression effect (Schneider et al., 2014). We predict that covert activation of motor/premotor brain areas generated by reading (sound-plus-) action-related words should have a comparable suppression effect on auditory perception. This would result in a reduced detection performance as compared to purely sound-related verbs.

Additionally, behavioural studies of action-related words on motor responses suggested that the stimulus onset asynchrony (SOA) between prime verb and Go signal has an effect on behavioural results, with a short SOA (100 ms) leading to an interference effect and a long SOA (350 ms) leading to a facilitation effect (de Vega, Moreno, & Castillo, 2013). We included a shorter (50 ms) and a longer (300 ms) latency between the offset of the stimulus word and the onset of the sound detection task to explore this factor in the auditory domain.

2.4.2 Methods

2.4.2.1 Participants

30 participants (mean age: 26.4; age range: 18 – 44; 17 females; 1 left-handed) were recruited from university campus. All participants were native German speakers and reported normal hearing. Informed consent was obtained prior to the experiment. Participants were debriefed and received monetary compensation after the experiment. The

experiment was conducted in accordance with the declaration of Helsinki and was approved by Düsseldorf University ethics committee (study number 3400).

2.4.2.2 Word stimuli

First we collected 54 purely sound related verbs, 96 sound plus mouth action related verbs and 48 abstract verbs with neither sound nor action content. Choosing mouth action verbs instead of hand action verbs allowed us to compare manual RT between different verb conditions without interference between verb content and response effector. These verbs will further be labelled sound plus action verbs. This initial selection of verbs for the three categories was assessed in a multistep rating, exclusion and matching procedure to define suitable experimental stimuli. To this end, 35 participants were asked to rate sound relatedness, action relatedness and familiarity of all verbs in a randomized online questionnaire (35 participants completed the sound relatedness judgment; 28 the action relatedness judgment; 28 the familiarity judgment. None of them participated in the main behavioural study). Participants rated from 1 (very weak) to 6 (very strong), indicating to what extent they associate the word with a sound (sound relatedness), to what extent the action depicted by the word is executed with physical strength and/or amplitude (action relatedness), and to what extent they are familiar with the word (familiarity). For action relatedness, the scale included the option “not an action”, indicating that the verb is not associated with any kind of action that a person can execute. ‘Not an action’ was coded as 0 in addition to the options 1 to 6 for sound verbs and sound plus action verbs. All abstract verbs were taken from former studies (Klepp et al., 2014; Niccolai et al., 2014) where the ratings did not include this option since all verbs were executable actions – either abstract or concrete. The data were collected online with SoSci Survey (https://www.soscisurvey.de/). Finally we selected 16 sound verbs that were related to sound (mean rating = 4.73; SD = 0.70) but not to action (mean rating = 1.10; SD = 0.41; ‘not an action’ ratio: 0.64), 16 sound plus action verbs that were related to both sound (mean rating = 4.83; SD = 0.50) and action (mean rating = 3.34; SD = 0.59; ‘not an action’ ratio: 0), and 16 abstract verbs that were not related to sound (mean rating = 1.19; SD = 0.10) or action (mean rating = 1.26; SD = 0.17; see Appendix 1 for a list of all the words used). We failed to find enough words that are (mouth) action related but not sound related. Independent sample t-tests showed that sound plus action verbs had higher action relatedness than both sound verbs (t(30) = 12.08, p <0.001) and abstract verbs (t(30) = 13.18, p <0.001, p < 0.001). Moreover, sound plus action verbs (t(30) = 27.39, p <0.001) and sound verbs (t(30) = 19.44, *p* <0.001) had higher sound relatedness than abstract verbs. No other t-tests reached significance. For each word, we also obtained word frequency (Biemann, Heyer, Quasthoff, & Richter, 2007), bigram frequency and trigram frequency (Baayen, Piepenbrock, & Gulikers, 1995), the latter two of which were calculated as the mean frequency of all relevant units in the word (e.g., bigrams in the word 'surren' include 'su', 'ur', 'rr', 're' and 'en') from the database. Word frequency was different among verbs (F(2,45) = 5.00; *p* = 0.01). Independent sample t-tests showed that sound plus action verbs (mean = 16.50; SD = 1.41) have a lower frequency than abstract verbs (mean = 14.63; SD = 1.54) (t(30) = 3.48, *p* = 0.002) (note that a higher frequency class value indicates a lower word frequency). No differences were present in frequency between sound verbs (mean = 15.69; SD = 1.89) and sound plus action verbs (t(30) = -1.33, *p* = 0.19) or between sound verbs and abstract verbs (t(30) = 1.69, *p* = 0.10). Despite the differences in frequency, all three selected categories of words were matched for familiarity (F(2,45) = 1.88; *p* = 0.16), word length (F(2,45) = 1.57; *p* = 0.22), bigram frequency (F(2,45) = 0.57; *p* = 0.57) and trigram frequency (F(2,45) = 1.63; *p* = 0.21). Table 2.3 provides a summary of the above mentioned word parameters.

Table 2.3 Means and standard deviation (in brackets) of related word parameters. Note that for word frequency, higher values indicate lower frequencies.

	Sound plus action verbs	Sound verbs	Abstract verbs
Sound relatedness	4.83 (0.50)	4.73 (0.70)	1.19 (0.10)
Action relatedness	3.34 (0.59)	1.10 (0.41)	1.26 (0.17)
Familiarity	5.68 (0.20)	5.51 (0.30)	5.66 (0.30)
Word length	7.69 (1.25)	7.19 (1.38)	6.94 (1.00)
Word frequency	16.50 (1.41)	15.69 (1.89)	14.63 (1.54)
Bigram frequency	12940.91 (1913.91)	12398.25 (2255.83)	12025.23 (3034.00)
Trigram frequency	4531.27 (1859.08)	4022.69 (2443.22)	3098.41 (2478.61)

For the lexical decision task, 48 pseudowords were generated as counterparts to the selected word stimuli using Wuggy, a multilingual pseudoword generator (Keuleers & Brysbaert, 2010). Given a German word as input, it automatically generates a list of counterpart pseudowords that are matched for subsyllabic structure and transition frequencies.

2.4.2.3 Procedure

In a pre-testing phase, the 75% detection threshold for a 1000 Hz tone (100 ms duration; 5 ms rise/fall) was individually identified for each participant, specifying the stimulus intensity to be used in the main task. To control for environmental noise, a background noise set at a comfortable level was presented throughout the whole experiment (including the main task). The test tone was presented in five different intensities (-29, -26, -23, -20, -17 dB in reference to background noise) and participants were required to judge if they heard a tone or not. Each intensity was presented 36 times in random order.

In the main task (Figure 2.6), participants sat in front of a computer screen at a viewing distance of approximately 50 cm. A trial began with the presentation of a fixation cross in the centre of the screen for 500 ms. After that, a word stimulus was centrally presented for 300 ms. Participants were required to silently read the word in order to judge if it was a real word or a pseudoword at the end of the trial. After a short (50 ms) or long latency (300 ms) during which the screen remained blank, a circle appeared in the centre of the screen. In half of the trials, the circle was accompanied by the test tone with the intensity determined in the pre-test. Participants were required to respond as accurately and as quickly as possible whether they had heard a tone. After the response, they were given the lexical decision task, judging if the word presented at the beginning of the trial was a real word or a pseudoword. This was aimed to make participants pay attention to the word stimuli. After the participant’s response the next trial started with a jittered interval between 1000 and 1500 ms.

Each real word stimulus was presented 8 times in total, equally divided between short and long latency, sound stimulus present and sound stimulus absent. The pseudowords were presented once each, resulting in 432 trials in total. All (pseudo-) word stimuli were presented in random order. The experiment was run in 4 blocks, each containing 108 trials. Participants used a keyboard to respond. The assignments of responses and buttons were counterbalanced between participants in the sound detection task: half of the participants pressed the left arrow button for tone present, the right arrow button for tone absent and vice versa. The lexical decision was always reported with button ‘A’ for word and button ‘D’ for pseudoword. All sound stimuli were delivered through a pair of headphones (Sony MDR-XD100) and the experiment was run with Psychtoolbox-3 on Matlab (The MathWorks, Inc., Natick, MA, USA). The duration of the study was about 70 minutes for each participant including pre-test and main experiment.

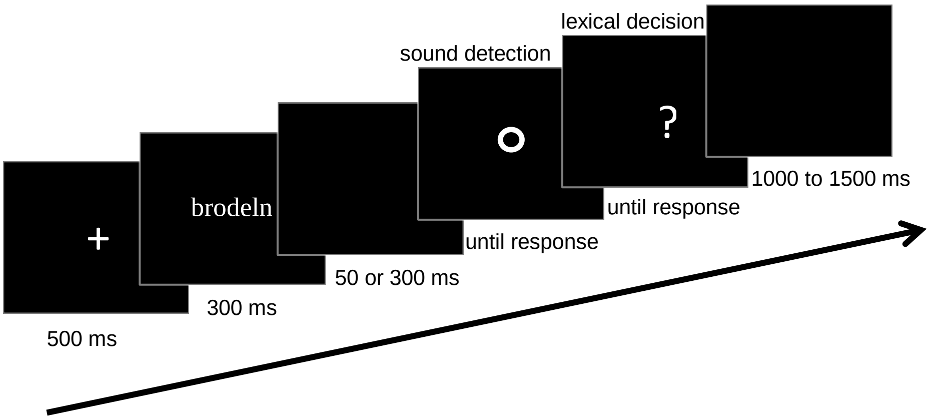


Figure 2.6 Main task. Each trial starts with a fixation cross for 500 ms, which is followed by a word stimulus for 300 ms (in this example, ‘brodeln’, ‘to seethe’ in English). After a short (50 ms) or long (300 ms) latency, a circle is presented. In half of the trials, the circle is presented together with a test tone which the participants are required to detect as quickly and as accurately as possible (sound detection task). A question mark follows after the sound detection response prompting the participant to respond whether the word shown in the beginning of the trial is a real word or not (lexical decision task). This is followed by a random inter-trial interval between 1000 and 1500 ms.

2.4.2.4 Data analysis

For the pre-testing phase data, we obtained a detection probability for each of the five tone intensities. Then the five data points were plotted on a Cartesian coordinate system with tone intensity on the x-axis and detection probability on the y-axis. At last, a sigmoid curve was fitted to the five data points and 75% detection threshold was determined (the same method as used in Borra et al. (2013)).

For the main task data, one participant had to be excluded from further analysis due to very low accuracy (0.58) in the lexical decision task. For all other participants, trials starting with pseudowords were not analysed. Trials with incorrect answers in the lexical decision task were also excluded. In these trials, either no real verbs were presented or participants failed to recognize the verbs thus excluding a conclusive sensory modulation causally related to verb processing. Trials that were followed by RTs exceeding 3 standard deviations of the individual mean RT in the sound detection task were excluded. 83.1% (SD = 3.8%) trials remained for the final analysis. The accuracy and RT data in the sound detection task were submitted to a 2 (latency: long and short) by 3 (word category: sound, sound plus action and abstract verbs) repeated measures ANOVA. In addition, planned comparisons were carried out by means of paired t-tests to follow up significant main effects or interactions. We also grouped the participants into two subgroups based on their performance in the lexical decision task and then performed a 2 (group: HP and LP) by 2 (latency: long and short) by 3 (word category: sound, sound plus action and abstract verbs) mixed-design ANOVA analysis with the data. The rationale is that participants with higher accuracy in the lexical decision task were more attentive to the word stimulus and thus might have processed the word at deeper levels. Level of processing has been shown to be important in a language-motor interaction study (M. Sato et al., 2008). The modulation effect in the sound detection task, if present, should be more likely to be observed in participants with better lexical decision performance. The 29 participants were divided into a High Performance (HP) group and a Low Performance (LP) group. The HP group consists of the first 15 participants in the lexical decision accuracy ranking. Their mean accuracy resulted in 0.98 (SD = 0.01). The LP group consists of the other 14 participants according to the accuracy ranking. Their mean accuracy resulted in 0.91 (SD = 0.05). Additionally, a sound detection performance modulation effect was calculated. Since a higher tone detection performance was predicted for sound verbs than both sound plus action verbs and abstract verbs, the detection accuracy following both verbs was subtracted from the detection accuracy following sound verbs separately, for both short and long latency conditions. The average of the resulting four (2 verb categories x 2 latencies) values was then taken as the modulation effect, indexing the strength of the modulation of auditory perception after reading sound words. A positive modulation effect emerges in case a participant on average benefits from sound verbs as opposed to the other verb categories with regard to tone detection accuracy. A negative modulation effect indicates that on average a participant’s performance declines after sound verbs presentation. We performed a correlation analysis between the modulation effect and accuracy in the lexical decision task across participants who showed a positive modulation effect. The AVOVA analysis was conducted with SPSS 19 and the correlation analysis was performed with the robust correlation toolbox (Pernet et al., 2012) implemented in Matlab.

2.4.3 Results

2.4.3.1 Accuracy of tone detection

The repeated measures ANOVA analysis with all 29 participants revealed no significant effects (word category: F(2,56) = 1.64, $p = 0.21$; latency: F(1,28) = 0.36, $p = 0.55$; interaction: F(2,56) = 2.07, $p = 0.14$). When participants were grouped based on the lexical decision performance, a 2 (group) by 2 (latency) by 3 (word category) mixed-design ANOVA analysis revealed a significant interaction effect between group and word category (F(2,54) = 3.41, $p = 0.04$; sphericity assumed). Post hoc analysis showed that there was a significant modulation of the tone detection accuracy across word categories in the lexical decision High Performance group (F(2,28) = 4.53, $p = 0.02$; Figure 2.7 a), with the accuracy after reading sound verbs being higher than the accuracy after reading sound plus action verbs (t(14) = 2.42, $p = 0.03$) and abstract verbs (t(14) = 2.33, $p = 0.04$). No such modulation was found in the Low Performance group (F(2,26) = 1.21, $p = 0.31$ Figure 2.7 b). No other effects from the mixed-design ANOVA analysis reached statistical significance (group: F(1,27) = 0.02, $p = 0.88$; latency: F(1,27) = 0.32, $p = 0.58$; word category: F(2,54) = 1.71, $p = 0.19$; group vs. latency: F(1,27) = 0.98, $p = 0.33$; latency vs. word category: F(2,54) = 2.05, $p = 0.14$; group vs. latency vs. word category: F(2,54) = 0.28, $p = 0.76$).

A control analysis was performed to compare lexical decision performance among word categories with a 2 (group) by 3 (word category) mixed-design ANOVA. This revealed significant main effects of group (F(1,27) = 29.78, $p < 0.001$) and word category (F(2,54) = 8.38, $p = 0.002$; Greenhouse-Geisser correction) and also a significant interaction effect (F(2,54) = 4.24, $p = 0.03$; Greenhouse-Geisser correction). The main effect of group with lexical decision performance is not surprising given the group categorization procedure. The main effect of word category was followed by a post hoc analysis, which showed that the lexical decision performance was significantly lower following sound verbs than following sound plus action verbs (t(28) = -2.83, $p = 0.01$) and abstract verbs (t(28) = -2.95, $p = 0.01$). However, the significant interaction effect showed that the modulation of lexical decision performance across word categories was only true for the LP group (F(2,26) = 6.33, $p = 0.01$), but not for the HP group (F(2,28) = 2.09, $p = 0.16$) (see Table 2.4 for the lexical decision accuracy data for each group). Post hoc paired t-tests within the LP group showed that sound verbs led to significantly lower lexical decision accuracy than sound plus actions verbs (t(13) = -3.00, $p = 0.01$) and abstract verbs (t(13) = -2.71, $p = 0.02$).

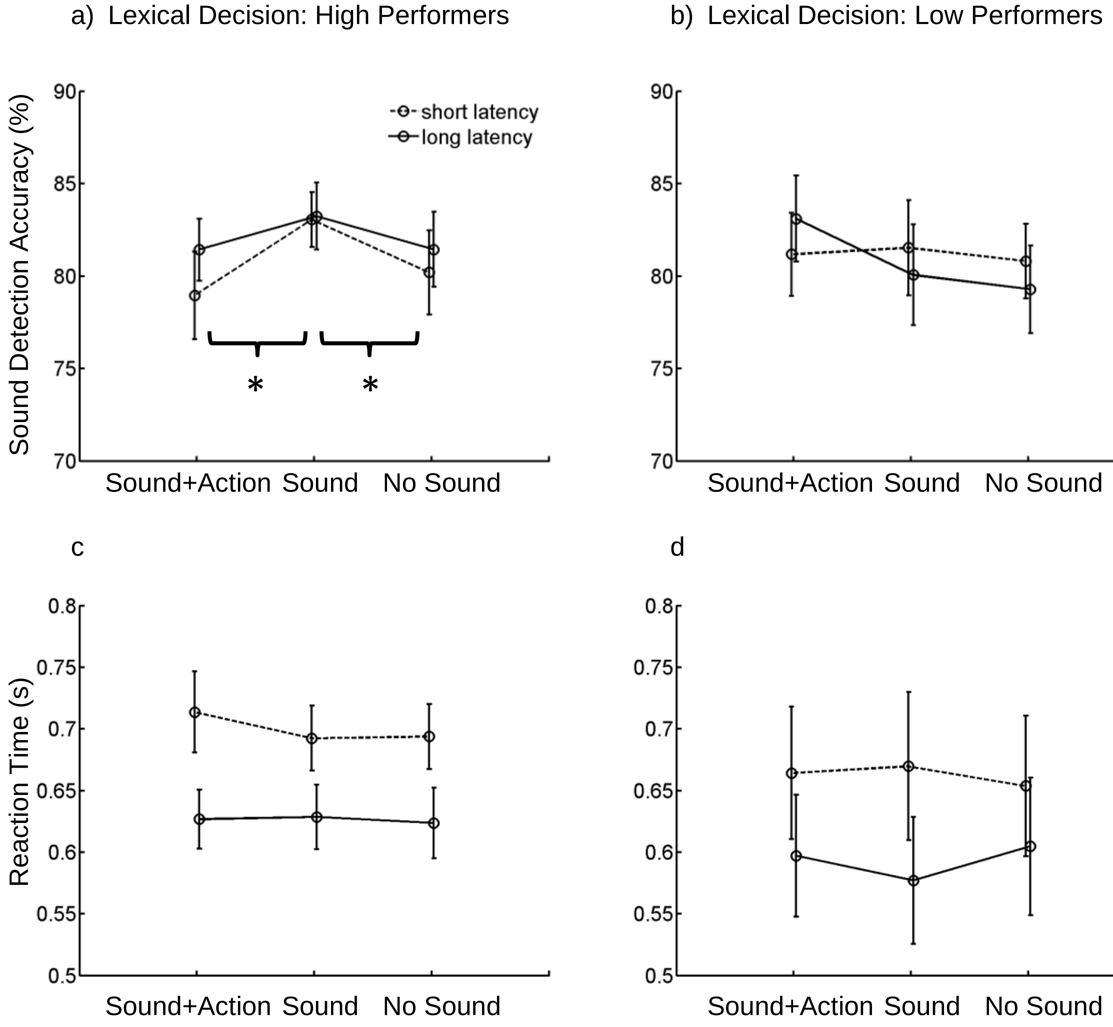


Figure 2.7 Accuracy (upper row) and reaction times (lower row) in the sound detection task (mean ± standard error of mean) separated for participants with high (>96 % correct, mean = 98 %; left column) and low (<96 % correct, mean = 91 %; right column) lexical decision performance. For high lexical decision performers, accuracy for sound verbs is significantly higher than for sound/action and abstract (no sound) verbs. The pattern of results is comparable but accentuated for short as compared to long latency condition. This effect is not seen in the low lexical decision performers (upper right panel). Reaction time is not modulated by word category in either participant group.

2.4.3.2 RT of tone detection

RT data were analysed similarly with a 2 (group) by 2 (latency) by 3 (word category) mixed-design ANOVA as for the accuracy data. There was a significant main effect of latency (F(1,27) = 58.01, $p < 0.001$), with participants responding faster in the long latency condition (mean = 610 ms, SD = 151) than in the short latency condition (mean = 682 ms, SD = 161). No other effects reached statistical significance (group: F(1,27) = 0.38, p = 0.55; word category: F(2,54) = 0.69, p = 0.45 (Greenhouse-Geisser correction); group vs. latency: F(1,27) = 0.05, p = 0.83; group vs. word category: F(2,54) = 0.24, p = 0.79; latency vs. word category: F(2,54) = 0.85, p = 0.43; group vs. latency vs. word category: F(2,54) = 1.58, p = 0.22).

Table 2.4 Lexical decision accuracy across word categories for both HP and LP groups (with standard deviation in brackets). A mixed-design ANOVA analysis shows a main effect that the HP group has higher accuracy than the LP group. An interaction effect indicates that, in the LP group sound verbs lead to significantly lower accuracy than the other two word categories, the pattern of which is not present in the HP group.

	Sound plus action verbs	Sound verbs	Abstract verbs
HP group	98.23% (1.3%)	97.66% (2.2%)	98.80% (0.9%)
LP group	93.30% (4.7%)	88.45% (7.9%)	94.42% (2.8%)

2.4.3.3 Correlation between modulation effect and lexical decision accuracy

For each participant we calculated a sound detection performance modulation effect indexing to what extent the detection performance benefitted from sound verb reading (see Methods section). A positive value is associated with an enhancement in sound detection performance after reading sound verbs and vice versa. Modulation effects ranged from -0.10 to 0.08 and 19 out of 29 participants had positive modulation effect values. From the high lexical decision performance group, 11 participants showed positive modulation effects. We further reasoned that within these 19 participants, better lexical decision accuracy should be associated with larger modulation effects because better lexical decision accuracy may indicate deeper word processing. This is indeed the case, as the modulation effect was significantly correlated with lexical decision accuracy (Spearman correlation coefficient = 0.47, p < 0.05, Figure 2.8). Two bivariate outlier data points, i.e., the two participants with a lexical decision accuracy of less than 90% shown on Figure 2.8, were identified with the robust correlation analysis toolbox (Pernet et al., 2012). Leaving these two participants out, the correlation was still significant (skipped Spearman correlation coefficient = 0.47, bootstrapped 95% Confidence Interval = [0.01 0.78]).

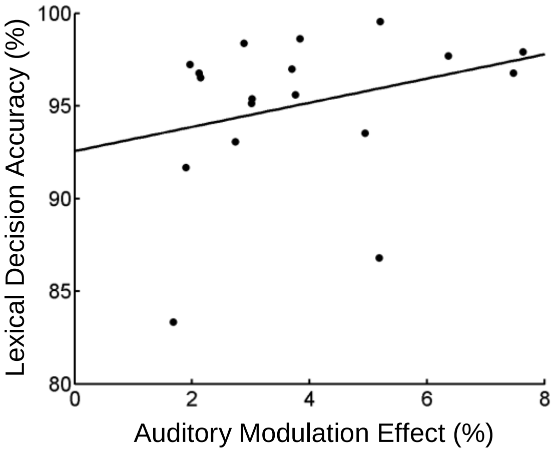


Figure 2.8 Significant correlation between auditory modulation effect and lexical decision accuracy (r = 0.47, $p < 0.05$) indicates higher modulation effect values coinciding with higher lexical decision performance. Robust correlation analysis (skipped Spearman correlation) showed that the correlation is still significant when the two outlier data points (the two dots below the data points cluster) are excluded from the analysis. For detail and the calculation of the modulation effect, please see the Methods section. The solid line shows a linear fit to the data.

2.4.4 Discussion

We investigated the modulation of auditory perception during language comprehension using a sound detection task. The data showed that there in fact was modulation of auditory perception, but the effect depended on the participants' performance in a lexical decision task. In the High Performance group, the group consisting of participants with a high accuracy in the lexical decision task, a clear modulation emerged. Sound detection accuracy was significantly higher after reading sound verbs compared to reading sound plus action verbs and abstract verbs. This modulation effect was not observed in the Low Performance lexical decision group. RT as dependent measure did only show an effect of latency with shorter RTs in the long latency condition. A control analysis with lexical performance across word categories showed an interesting modulation effect in the LP group but not in the HP group.

Word frequency differences between word categories need to be taken into account, because they could contribute to the result. However, here word frequency is rather unlikely to contribute to the auditory modulation effect, as the only difference occurred was that sound plus action verbs have a lower frequency than abstract words. Critically, there are no differences between sound verbs and sound plus action verbs or between sound verbs and abstract verbs, where auditory modulation effect was found. Furthermore, bigram frequency, trigram frequency, word length and subjective familiarity were controlled between word categories, which allows for an equal difficulty in the lexical decision task. These variables did not differ between conditions.

For both the short and the long latency conditions, we found a similar facilitation effect on sound detection performance from sound verb reading in the HP group. It should be noted that this does not contradict with the findings from motor studies, where an interference effect was found for short SOAs (100 ms) and a facilitation effect was found for long SOAs (350 ms) (de Vega et al., 2013). The short (50 ms) and long latencies (300 ms) in our study correspond to an SOA of 350 ms and 600 ms between the word onset and the response cue onset, respectively. Thus both the short and long latencies should be viewed as long SOAs in the context of the study by de Vega et al. (2013). Very short SOAs (e.g., 100 ms) are not appropriate for the current study as in such case the stimulus presentation time is not long enough for participants to make meaningful lexical judgments. In the study by de Vega et al. (2013) the stimulus word was presented in the context of a sentence in which case prior predictions were available. Most importantly, the stimulus word continued to be present after the response cue onset. Future studies may combine this knowledge to investigate the effect of sound verbs on auditory performance in a shorter time scale.

While no significant interaction was found between latency (of tone presentation onset) and word category, the modulation effect was numerically stronger in the short latency condition, when the tone was presented 50 ms after the offset of the word (Figure 2.7 a). Interestingly, even in the LP group exhibiting no significant modulation effect, the pattern of accuracy values in the short latency condition resembled the pattern of the HP group (Figure 2.7 b). This may suggest that the modulation effect occurs with or just after reading words and decays gradually. This supports the idea that the modulation effect reflects language comprehension but not consecutive mental imagery (Kiefer et al., 2008; Klepp et al., 2014; Pulvermüller, 1999; Trumpp, Traub, & Kiefer, 2013).

Word processing depth may be critical for the tone detection modulation effect. First, the modulation effect was only significant in the HP group but not in the LP group. The HP group participants may have been more attentive to the lexical decision task and thus processed the stimulus word in a deeper level. Second, the modulation effect was correlated with lexical decision accuracy across participants who showed a positive modulation effect. This may indicate that the deeper the word processing (as indexed by lexical decision accuracy), the stronger the modulation effect by word processing. We are aware that the above interpretation is speculative given the fact that we did not manipulate the word processing depth in the task. We propose that a semantic instead of a lexical decision task as we used here could produce an even stronger modulation effect than the one observed here. This would parallel the findings from language-motor interference studies (M. Sato et al., 2008).

In the case of abstract verbs, we interpret our findings as that the activation of auditory brain areas by sound verbs enhanced cortical excitability, thus leading to better perceptual performance compared with non-sound abstract verbs. In line with our prediction, sound detection accuracy after reading sound plus action verbs, which were both action and sound related, was lower than accuracy after reading sound verbs, which were sound related only. According to our hypothesis of combined effects of sound- and action-relatedness we interpret this pattern of result as follows: sound relatedness of verbs, present in both sound and sound plus action verbs, enhances auditory performance. In the case of sound plus action verbs, the additional effect of action relatedness possibly inhibits auditory excitability and decreases performance in comparison to the sound verbs - in this case to a level comparable with the abstract verbs. Motor induced suppression of auditory perception is a ubiquitous phenomenon across species (Crapse & Sommer, 2008). In humans, it can even be observed during silent lip reading (Kauramaki et al., 2010) or imagining speaking (Tian & Poeppel, 2014). The finding here suggests that the comprehension of action related verbs involves a simulation process including the activation of both motor and (auditory) sensory brain areas. One limitation of the current study is that we did not study the exclusive effect of action relatedness. This was because we were not able to collect a sufficient number of well-matched purely action-related mouth verbs that are not associated with sound. Studies using other action- and sound-related verbs with the capability to disentangle motor and sound effects completely are needed to substantiate this finding.

Interestingly, our control analysis showed that in the LP group, participants were specifically impaired in processing sound verbs (lower lexical decision accuracy with sound verbs than with the other two verb categories). This pattern was not observed in the HP group. While the original idea of the study was to test the influence of word processing on auditory perception, the real scenario could be a bidirectional influence, i.e., word processing and auditory perception can influence each other (we thank anonymous reviewers for igniting the suggestion). The directionality of the influence may be related to participants' strategy used in the task. When participants put more weight on the lexical decision task, the involvement of auditory cortex may have a positive effect on auditory perception (HP group). This also corresponds to our word processing depth hypothesis put forward earlier. In contrast, when participants put more weight on the auditory detection task, the involvement of the auditory cortex may have a negative effect in processing sound related verbs (LP group) (Trumpp, Kliese, et al., 2013). The RT data lend partial support for the idea that LP group participants put more weight on the sound detection task and that HP group participants put more weight on the lexical decision task. In the sound detection task, the LP group (mean = 628 ms; SD = 202 ms; averaged across all conditions) had numerically lower RTs than the HP group (mean = 663 ms; SD = 94 ms; averaged across all conditions) (see Figure 2.7 c & d). The RT in the lexical decision task was not recorded but can be estimated since the whole trial duration was recorded. The estimated RT data showed the opposite pattern to the RT data in the sound detection task, i.e., the HP group (mean = 722 ms; SD = 242 ms; averaged across all conditions) had numerically lower RTs than the LP group (mean = 855 ms; SD = 180 ms; averaged across all conditions). Unfortunately, a 2 (group) by 2 (task) mixed-design ANOVA analysis with the RT data only led to a significant main effect of task (task: $F(1,27) = 7.36$, $p = 0.01$; group: $F(1,27) = 1.14$, $p = 0.30$; interaction: $F(1,27) = 2.53$, $p = 0.12$). Albeit interesting, the RT in the lexical decision task was not timely recognized as a potential interest when planning the study - participants were explicitly told that they can take their time for the lexical decision task. Nevertheless, the reversal pattern of RT was reported having future studies in mind.

In conclusion, our study supports the view of embodied language comprehension, focussing on the auditory system and auditory contents of verbs. We have shown convergent evidence of auditory perception modulation after word reading in terms of enhancing behavioural performance. In addition to the finding that sound relatedness in words facilitates auditory perception, the study also suggests an interference effect on auditory perception from action relatedness in words. This extends our understanding of embodiment in language processing, taking into account different modalities. Unexpected (interesting) results from the lexical decision accuracy data further complement the embodied language comprehension view, suggesting a negative effect on sound verb processing from the auditory task.

2.5 General discussion

In a series of three experiments, we investigated motor modulation over auditory perception from different perspectives. In study 1, SA was compared with attention effect. SA was still present for attended stimuli but attention effect was much larger than SA. In study 2, we showed that watching others generate a tone can lead to SA among Chinese but not among the British, which reconciled existing controversial findings in this field. Moreover, the effect of SA from watching others generate a tone was correlated with the independent self-construal score. Although we failed to replicate SA for self-generated tones in this study, a significant correlation between the amount SA effect and delusional ideation was found. In study 3, we showed evidence that reading verbs that are movement related could also lead to SA, possibly due to the motor activation from verbs reading.

2.5.1 Behavioural investigation of auditory SA

A. Sato (2008) reported the first evidence of behavioural SA in the auditory domain, which received confirmatory evidence later (Stenner, Bauer, Sidarus, et al., 2014; Weiss et al., 2011). However, we failed to find this effect in study 2 using a similar paradigm to Sato's. Different reasons may account for this replication failure. SA effect for a 74 dB tone, which was used in study 2, is very small. In the study by Sato and Weiss et al., a reduction of 0.4 dB and 0.2 dB in the perceived loudness for the self-triggered 74 dB tone was found, respectively. In study 2, a reduction of 0.1 dB and 0.01 dB was found for Chinese and British participants, respectively. Although the perceived loudness for the self-triggered tone was still lower than for the computer triggered tone in our study, the effect was much smaller than previous studies. The source of this difference is unknown. The significant correlation between SA and delusional ideation tendency among British participants supports the validity the dataset (Teufel et al., 2010). It also suggests that the SA effect may be shifted downwards by unknown factors whereas the inner variance is still kept among participants. We conclude that the SA replication failure in study 2 should not be counted as evidence against SA.

In study 1, SA was found using an auditory detection task. A button press prior to the detection task leads to a significant decrease in the detection performance. The results support the idea of an unspecific sensory gating mechanism from movements (Chapman et al., 1987; Horváth et al., 2012; Schröger et al., 2015). In the task, the frequency and the timing of the stimulus were not predictable, thus excluding an explanation based on the prediction from the forward model. Study 3 provided further evidence for this in showing that tone detection accuracy was lower after reading verbs that were related to both sound and action than reading verbs that are only related to sound. In this case, the motor activation resulted from reading action related verb, so no prediction from the forward model was available.

Interestingly, when low intensity sounds could be predicted from motor responses, sensory enhancement was found (Reznik, Henkin, et al., 2015; Reznik, Henkin, Schadel, & Mukamel, 2014). The authors explained it as that the efference copy carries context-dependent information so it serves to enhance loudness perception under relevant situations. This sensory enhancement effect is in contrast to the SA as seen in study 1. It again suggests that there are two mechanisms underlying the influences of motor responses over auditory perception: one is the unspecific action-related sensory gating mechanism and the other is the forward model based mechanism. Reznik et al. studies also suggest that the forward model based mechanism is more influential than the unspecific sensory gating mechanism as sensory enhancement won the battle against SA when the sound intensity was low.

Despite the claim of an existence of auditory sensory attenuation measured from behavioural tests, it should be noted that there is no sufficient evidence that self-generated sounds are really perceptually softer. For example, the just noticeable difference for a 74 dB tone in Weiss et al. (2011) study was above 0.6 dB and was 0.79 dB in study 2, which was clearly more than the largest reduction (0.4 dB) found in perceived loudness for self-generated tones. It is not clear, at this stage, that if (and to what extent) the decision making process contributes to the reported behavioural measurements of auditory sensory attenuation. Future studies should clarify this point. This is in a drastic contrast to the clear perceptual effect for self-generated somatosensory inputs (i.e., it is difficult to tickle oneself), which makes a strong argument for different modulation effects from actions on sensory perception in different modalities.

2.5.2 SA and attention

One drawback of the behavioural SA studies is that they generally don’t have a control of attention level, which potentially opens the door for alternative explanations of the results.

Take the Reznik, Henkin, et al. (2015) study as an example. One may argue that more attention is drawn to self-triggered tones when tones have very low intensities, which would lead to sensory enhancement effect for low intensity tones. Indeed, a recent study showed that attention can be synchronized with movement to synergistically improve sensory processing (Morillon, Schroeder, & Wyart, 2014).

A modelling study suggested that the cause of SA is the withdrawal of attention from the self-triggered sensory consequences (Brown, Adams, Parees, Edwards, & Friston, 2013). This view is compatible with behavioural and neurophysiological findings that SA can still be observed even if a considerable amount of attention is paid to the stimulus (Cao & Gross, 2015a; Timm et al., 2013). This is because the button press will always take some attention away from the stimulus. Thus the attention paid to the stimulus will be lower in the active condition as compared to the passive condition. An interesting prediction from this modelling study is that SA should disappear or at least get smaller when little attention is paid to the stimulus, as the attention paid to the stimulus should be similar between conditions in this case. Unfortunately this prediction cannot be tested with our behavioural data in study 1 as a floor effect easily explained the unattended condition data. This remains an open question for future studies.

However, Morillon et al. (2014) showed that movement can serve as a reference to align attention for sensory processing, which seems contradictory to the account put forward by Brown et al. (2013) that attention should be driven away from the incoming stimulus for making a movement. The contradiction might result from differences in testing paradigms, the predictability of the stimulus etc. For example, the task in Morillon et al. (2014) study is very challenging which may be a source of enhanced attention along with movements. It’s difficult to draw a firm conclusion as very few studies are available on this topic for comparison. As noted earlier, this is only the beginning of SA and attention studies. More studies are needed to unravel their interplay.

2.5.3 Attenuated auditory perception from indirect motor activation

Study 2 and study 3 investigated the effect of indirect motor activation on auditory perception. Indirect motor activation denotes that the activation of motor cortex is not induced by a direct movement from participants but by, for example, watching others’ movements (study 2) or word reading (study 3). Although we did not measure the brain activity in study 2 or study 3, previous studies confirmed an activation of motor cortex when observing others’ movements (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Hari et al., 1998; Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010) and when reading action-related words (Hauk et al., 2004; Pulvermüller, 1999). The fact that indirect motor activation also influences auditory perception, in which case there is no explicit association between motor activation and sound presentation (study 3) suggests an automatic activation of the motor-auditory neural pathway (Reznik, Ossmy, et al., 2015; Schneider et al., 2014). However, in study 3 we did not compare between verbs that are only related to actions and abstract verbs due to a failure in finding a sufficient number of such verbs that are only related to (mouth) actions but not sound. Future studies can address this issue by selecting verbs that are related to other actions or from a different language.

The perception of others generated tone is malleable, with participants from Eastern culture showing SA_{other} whereas participants from Western culture not showing SA_{other} (Cao & Gross, 2015b; A. Sato, 2008; Weiss et al., 2011). It is interesting to compare this with the cultural difference in visual perspective taking performance (Kessler, Cao, O'Shea, & Wang, 2014; Wu & Keysar, 2007). In visual perspective taking task, participants were asked to take the visual perspective of others, e.g., if others can see something or not. It was found that easterners have better performance than westerners in this task. Recent developments in this area suggest that visual perspective taking process is an off-line run of forward model (Mast & Ellis, 2015). In study 2, if the task in ‘other’ condition (listening to the experimenter generated tone) is viewed as an ‘auditory perspective taking’ process, it would fit very nicely with the visual perspective taking studies. With this idea in mind, the cultural difference in SA_{other} can be interpreted as that easterners are better than westerners in auditory perspective taking, as they are in visual perspective taking. This would also mean that the difference might be a result of differential degrees of the involvement of the forward model process.

3 Neural mechanisms of sensory attenuation

(Partly published in *NeuroImage* (2016), doi: 10.1016/j.neuroimage.2016.11.001)

3.1 Introduction

3.1.1 Electrophysiological sensory attenuation studies

In our interactions with the environment, action and perception are tightly linked. Voluntary motor actions typically lead to predictable sensory consequences. For example, knocking on a door results in a predictable sensory input to the auditory and somatosensory systems. It is well established that these self-generated sensory stimuli elicit smaller brain responses than externally generated stimuli (Martikainen et al., 2005; Schafer & Marcus, 1973) - a phenomenon known as sensory attenuation (SA). For example, an MEG study showed a reduced auditory M100 component when the sound was generated by participants pressing a button compared to when the sound was passively presented. With source localization analysis, the authors first showed that SA was localized in the auditory cortex (Martikainen et al., 2005).

Whereas SA is well-established in auditory and somatosensory modalities (Hesse et al., 2009; Hughes et al., 2013b; Shergill et al., 2003), it is not so clear whether SA is also present in the visual domain. Several EEG studies investigated this topic, with both SA and sensory enhancement being reported. Schafer and Marcus (1973) reported SA of evoked responses in the vertex when comparing self-triggered flash to the flash sequence played back by the computer. But the information of the evoked components was not given. Gentsch and Schutz-Bosbach (2011) found SA with visual evoked N100 component (see also Gentsch, Kathmann, and Schutz-Bosbach (2012)). Interestingly, they also found that the N100 component was lower when participants felt more control over the visual stimulus through subliminally priming, thus suggesting that the N100 attenuation was related to self-agency. SA was also found in even later visual evoked components from 150 to 450 ms after the stimulus onset (Hughes & Waszak, 2011). Both visual SA reports from Gentsch and Schutz-Bosbach (2011) and Hughes and Waszak (2011) were obtained from frontocentral sensors. However, Hughes and Waszak (2011) also reported a sensory enhancement effect of P100 component in the occipital area, which was echoed by a recent finding from Mifsud et al. (2016) showing a sensory enhancement with visual N145 component in the occipital area. Thus existing studies showed mixed results with regards to the influence of actions on visual processing, with both attenuated and enhanced responses from self-generated stimuli being reported.

3.1.2 Mechanisms of SA

A forward model, initially developed from motor control studies, has been proposed as a possible account for SA (Blakemore, Frith, et al., 1999; Ramnani, 2006; Wolpert & Ghahramani, 2000). The model posits that along with a motor command, an efference copy (von Holst & Mittelstaedt, 1950) is sent that allows the computation of the predicted, imminent sensory consequences. The predicted sensory signal is then compared to the actual incoming sensory signal leading to a modulation of the brain responses depending on the match between the real sensory signal and the prediction (attenuated when matching). A detailed conceptual explanation can be derived from the predictive coding theory (Friston, 2005). In this framework, the evoked response is an expression of prediction error, which is the discrepancy between the predicted sensory consequence and the actual sensory input. Accurately predicted stimuli lead to smaller prediction errors, which are reflected in decreased evoked responses. In addition, it has been suggested that predictions and prediction errors are communicated along cortical hierarchies in distinct frequency bands. More specifically, recent evidence suggests that predictions are communicated along anatomical feedback connections via alpha/beta oscillations and prediction errors are communicated via feedforward connections via gamma oscillations (Bastos et al., 2015; Michalareas et al., 2016; X. J. Wang, 2010).

The neural circuit of the efference copy signal is largely under dispute. Several brain regions, including but not exclusive to medial prefrontal cortex (Müller et al., 2014), inferior frontal gyrus (J. Wang et al., 2014)) and motor cortex (Ford et al., 2013; Reznik, Ossmy, et al., 2015), have been suggested to be involved in generating the efference copy signal. The involvement of motor cortex seems natural as the efference copy is originally proposed as a copy of motor command (von Holst & Mittelstaedt, 1950). Recent animal studies demonstrated a neural circuit that connects the motor area and the auditory area, which provides a potential neural basis for the efference copy signal (Schneider et al., 2014).

Abnormal SA has been found in several patient groups, including schizophrenia (Blakemore et al., 2000; Ford et al., 2013; Shergill et al., 2014). Among normal subjects, a correlation between the magnitude of SA effect and the schizotypal traits (especially with delusional state) was also found (Oestreich et al., 2015; Teufel et al., 2010). Higher delusional state was associated with a smaller SA effect, which suggests that delusions may be associated with deficits in predictions (Malassis, Del Cul, & Collins, 2015).

3.1.3 Aims and hypotheses of the present study

In this study, we investigated the effect of actions on both auditory and visual stimuli. In the auditory domain, we expected to replicate the classic SA effect. In the visual domain, we had no prior expectations of the result given the mixed findings from previous EEG studies as reviewed above.

For SA mechanisms, we sought to investigate how brain oscillations orchestrate SA effect, with an emphasis on auditory cortex activity. Neural oscillations in lower frequency bands are likely candidates for the implementation of SA as they are tightly linked to excitability changes in neural populations (Jensen & Mazaheri, 2010; Thut, Miniussi, & Gross, 2012; Weisz, Hartmann, Muller, Lorenz, & Obleser, 2011), and therefore may mediate gain control mechanisms to incoming sensory information. A number of studies provided converging evidence that low frequency oscillations particularly in the 10 Hz range (alpha band) support active inhibition. An increase in alpha power is typically associated with a decrease in perceptual performance (Frey et al., 2014; Thut et al., 2006; van Dijk, Schoffelen, Oostenveld, & Jensen, 2008). Moreover, the phase of low frequency oscillations (including alpha) was also shown to modulate neural excitability, so that near-threshold stimuli are more likely to be perceived or neural responses to be enhanced if stimulus presentation is aligned to a certain phase of the ongoing oscillations (Arnal & Giraud, 2012; Busch et al., 2009; Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007; Mathewson et al., 2009). We therefore hypothesized that pre-stimulus changes in low frequency oscillations may be a candidate reflecting prediction related processes, which is generated by the forward model to implement the suppression of post-stimulus responses for SA. Indeed, some studies already provided evidence that pre-stimulus alpha power is higher in the sensory cortex when speech/visual stimuli are self-induced by movement (Müller et al., 2014; Stenner, Bauer, Haggard, et al., 2014).

The prediction error was shown to be reflected in gamma oscillations (Bauer et al., 2014; Behroozmand et al., 2016). This is in line with findings showing that gamma oscillations relay feedforward information (e.g., Michalareas et al. (2016)). In the context of SA, intracranial recordings from neurosurgical participants showed that gamma power (70-150 Hz) was suppressed in response to speech stimuli during speaking as compared to during listening (Flinker et al., 2010). Thus reduced gamma power may indicate decreased prediction errors when the stimulus is better predicted during speaking. Therefore, we predict a replication of gamma reduction for self-generated stimuli in the present study.

Post-stimulus gamma power decrease does not seem to contribute to SA in trial-averaged evoked responses (e.g., reflected in attenuated M100 component) as a low pass filter at around 40 Hz was applied in many SA studies (e.g., Baess et al. (2011); Martikainen et al. (2005); Müller et al. (2014)). Our understanding about the underlying changes responsible for SA at the level of single trials is still incomplete. A reduced amplitude of evoked responses after averaging across trials during SA could result from single-trial amplitude reduction, increased single trial phase jitter or a combination of both. Since sensory evoked responses are primarily reflected in an increase in the power and phase locking of theta oscillation, one may expect that a reduction of power and/or phase locking in the same frequency band contributes to SA.

To investigate the involvement of motor cortex in the generation of the efferency copy signal, Granger causality was used to measure the information flow between motor cortex and auditory cortex. Granger causality has the advantage of capturing the neural information exchange between neural assemblies on single-trial level and it has never been employed in the context of SA.

A set of questionnaire data, including 21-item Peters et al. Delusions Inventory (PDI) (Peters et al., 2004), Self Construal Scale (Singelis, 1994), short form Empathy Quotient (EQ-short) and Systemizing Quotient (SQ-short) (Wakabayashi et al., 2006)), was obtained from participants to study the relationship between SA and psychological traits. Although the correlation between SA and delusional ideation has been addressed in behavioural SA studies and electrophysiological SA studies using a speech paradigm (Cao & Gross, 2015b; Oestreich et al., 2015; Teufel et al., 2010), it has never been investigated under the current SA paradigm with self-generated tones in MEG. Furthermore, the EQ-short and SQ-short are questionnaires measuring autistic traits. To the best of our knowledge, SA has never been studied together with autism, whereas the current trend points out that a deficit in prediction may be a central theme in autism studies (Pellicano & Burr, 2012; Sinha et al., 2014).

In summary, the present study addresses the following questions: 1) Replicating SA in the auditory domain and exploring the effect of actions on visual stimulus processing. 2) What is the role of neural oscillations in SA? 3) How can Granger causality analysis inform us about the neural information exchange between motor and auditory cortex? 4) Exploring the relationship between SA (the prediction process) and psychological traits (especially schizotypal traits and autistic traits). To address these questions, an MEG experiment was conducted using a well-established SA paradigm in both auditory and visual domains, in which neural responses from self-generated stimuli and passive stimuli were compared (Baess et al., 2011; Schafer & Marcus, 1973).

3.2 Methods

3.2.1 Participants, Procedure and Recording

14 healthy, right-handed volunteers (6 males, mean age = 22.6, SD = 1.8) were recruited from a local participant pool. Participants gave written informed consent prior to the experiment and received monetary compensation after the experiment. The study was approved by the local ethics committee (Ethics Committee of College of Science and Engineering, University of Glasgow) and was conducted in accordance with the Declaration of Helsinki.

A 248-magnetometers whole-head MEG system (MAGNES 3600 WH, 4-D Neuroimaging) was used for data recording with a sampling rate of 1,017Hz. Auditory stimulus was a pure tone (1000 Hz, 50 ms in duration, 90 dB sound pressure level) delivered through a plastic tube. The visual stimulus was a black and white image (a trigram; 100 ms in duration) that was projected to a white board in front of participants. The distance between the visual stimulus and eyes is about 108 cm (visual angle 18.3^o x 18.3^o). There were four conditions for each modality (Figure 3.1). In active condition, the stimulus was presented immediately after an index finger lifting movement that the participants were asked to perform about once every three seconds without inner counting. We used a light sensor (instead of a response box) to record the finger movement without noise associated with the movement. Every movement unblocked the beam from the light sensor (placed next to participant’s right index finger), which then generated a sound stimulus. Motor only condition was identical to active condition except that no stimulus was presented after a finger lift. In passive periodic condition, the stimulus was controlled by the computer and was presented once every three seconds. In passive jittered condition, the stimulus was controlled by computer and was presented with a jittered interval between 2000 and 4000 ms. Each condition consisted of 100 trials. The order of the two modalities was counterbalanced among participants with the four conditions within each modality being presented in a random order. Participants were also asked to close their eyes during the testing. Before the start of the experiment, participants received 50 trials of practice to familiarize themselves with the light sensor and the rate of finger movements. During this practice, they were asked to move the finger about once every three seconds without inner counting and they received visual feedback for their timing performance after each trial. No such feedback was provided in the real data collection. Participants were encouraged to take a break between testing blocks if they want.

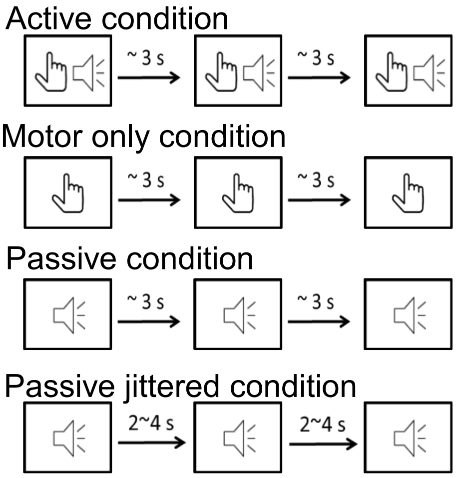


Figure 3.1 Illustration of the 4 conditions used in the auditory domain. The same 4 conditions are also used in the visual domain with the stimulus being a visual image.

A set of questionnaire data was also obtained from 11 of the participants. The questionnaires included 21-item Peters et al. Delusions Inventory (PDI) (Peters et al., 2004), Self Construal Scale, short form Empathy Quotient (EQ-short) and Systemizing Quotient (SQ-short) (Wakabayashi et al., 2006). These questionnaire data were used for correlation analysis with SA.

3.2.2 Data analysis

Data analysis was performed with Matlab using FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011) and in-house codes in accord with current MEG guidelines (Gross et al., 2013). Trials with very short inter-trial intervals (less than 1500 ms) were discarded. Then MEG signals were denoised using ft_denoise_pca which removes artefact components measured by the MEG reference sensors. Trials with artifacts were removed following visual inspection with ft_rejectvisual. On average, 8.1% trials were excluded from each condition (SD = 2.3%). Eye movement and heart artefacts were rejected using ICA (mostly 2 or 3 components).

3.2.3 Evoked responses in sensor space

MEG signals were low-pass filtered with 30 Hz cut-off frequency. Original magnetometer signals were converted to planar gradient representation. Visual evoked responses were computed from the average of 6 sensors in the occipital area that showed strongest activation at the latency of visual M100 component (90-130 ms post-stimulus) in passive periodic condition (Figure 3.2 c). Similarly, three sensors from each hemisphere of temporal area that were predominantly responding at the latency of the auditory M100 component (95-120 ms post-stimulus) in passive periodic condition were selected for auditory analysis (Figure 3.3 c). Event related fields aligned to the onset of stimuli were computed for each condition with baseline (-500 to -100 ms) correction. The M100 component was statistically compared between conditions using a paired t-test with a Monte Carlo randomization with 1000 permutations.

3.2.4 Time-frequency analysis

In sensor space analysis, the time-series data from three interesting regions (left motor area, left and right auditory areas) were subjected to time-frequency analysis, separately for each participant and experimental condition. Sensor space time-frequency analysis was exploratory. Therefore no statistical p value was calculated. Three sensors that showed the largest activation in relevant components (motor or auditory) were selected for analysis from each region. We performed time-frequency analysis with a temporal resolution of 20 ms and a spectral resolution of 1 Hz on 500 ms long sliding windows with a Hanning taper for frequencies below 40 Hz. To test the differences in oscillations between conditions, oscillatory power was first log transformed and compared between conditions with a between-subject t-test on each selected sensor for each participant. The resulting t values were then averaged among sensors of each interesting region for each participant, which were then subjected to a t-test against 0 to detect differences between conditions. For oscillatory phase, phase locking value (PLV) was first computed for each participant in each condition as an index of phase consistency over trials:

$$\mathrm{PLV}_{t,f} = \frac{1}{N}\left|\sum_{n=1}^{N} \exp(i * \theta_{t,f}(n))\right|$$

where $\theta_{t,f}(n)$is the oscillatory phase at time-frequency point (t,f) of trial n and N is the total trial number in the condition. PLV was compared between conditions with paired t-tests to detect differences between conditions.

The time-frequency analysis in source space was based on one representative voxel from right auditory cortex (see section 3.2.5). The analysis was performed similarly as for sensor space analysis except that that a temporal resolution of 10 ms was used for frequencies below 40 Hz and that frequencies from 40 to 120 Hz were also analyzed. For oscillatory power above 40 Hz, the time-frequency analysis was performed with a temporal resolution of 10 ms and a spectral resolution of 1 Hz on 200 ms long sliding windows with the multi-taper approach with 10 Hz smoothing. Oscillatory phase above 40 Hz was analyzed similarly but with a Hanning taper. For source space statistics, individual time-frequency maps were subjected to dependent-sample t-test between active and passive periodic condition. Oscillatory power was log transformed with reference to the mean power from -700 to 700 ms. The null distribution was estimated using 1000 randomizations and multiple comparison correction was performed using the cluster method (Maris & Oostenveld, 2007). Only significant results ($p < 0.05$, cluster corrected) are reported.

3.2.5 Source localization

T1-weighted structural magnetic resonance images (MRIs) of each participant were co-registered to the MEG coordinate system using a semi-automatic procedure. Anatomical landmarks (nasion, left and right pre-auricular points) were manually identified in the individual's MRI. Initial alignment of both coordinate systems was based on these three points. Subsequently, numerical optimization was achieved by using the ICP algorithm (Besl & McKay, 1992).

Individual head models were created from anatomical MRIs using segmentation routines in FieldTrip/SPM5. Leadfield computation was based on a single shell volume conductor model (Nolte, 2003) using a 10 mm grid defined on the template (MNI) brain. The template grid was transformed into individual head space by linear spatial transformation. The localization of auditory and visual evoked components was based on eLoreta algorithm as implemented in Fieldtrip (covariance matrix from -500 ms to -100 ms; normalization 0.6; regularization 7%). All other analyses used LCMV filters computed based on a covariance matrix from -500 ms to 500ms with a regularisation of 7% of the mean across eigenvalues of the covariance matrix.

All further analyses of auditory data in source space were based on a representative voxel from the right primary auditory cortex that was anatomically defined and showed clear reconstructed evoked responses. We focused the analysis on the right auditory cortex because activity estimates for left auditory voxels can be contaminated by activity from the left primary motor cortex related to the movement of the right hand finger. For the selected voxel, we computed an LCMV filter along the orientation of maximal power (across all experimental conditions) and extracted the single-trial time series separately for each experimental condition.

3.2.6 Correlation analysis

Various cross-participant correlations were performed. For correlation analysis between SA and questionnaire data, SA (sensor) was calculated as the average magnitude difference of M100 component (95-120 ms) between active condition and passive periodic condition from auditory sensors (‘active condition’ minus ‘passive periodic condition’). Similar correlation analysis was done between SA (sensor) and beta rebound, which was measured from sensors showing strongest activation by motor responses.

For correlation analysis in the source space, first, SA was used for correlation with the pre-stimulus alpha power increase, post-stimulus alpha phase locking decrease and post-stimulus gamma power decrease. SA (source) was calculated by taking the relative change of evoked responses (70-160 ms) between active and passive periodic condition using the voxel reconstructed time series data (i.e., the amplitude difference between active and passive periodic condition divided by the amplitude in passive periodic condition). Other components used for correlation analysis were derived from clusters showing significant differences between active and passive periodic conditions (see results). The pre-stimulus alpha power increase (10 Hz, -400 to 0 ms), post-stimulus alpha phase locking decrease (9-10 Hz, 0-150 ms), and post-stimulus gamma power decrease (high gamma: 85-104 Hz, 90-

120 ms; low gamma: 57-62 Hz, 30-80 ms) refer to relative changes that were calculated in the same way as for SA. Second, correlation analysis was performed among the oscillatory changes between conditions: the pre-stimulus alpha power increase (9-12 Hz, -400 to -60 ms) and the post-stimulus gamma power decrease (93-106 Hz, 90-120 ms); the post-stimulus alpha phase locking decrease (9-11 Hz, 0-150 ms) and the post-stimulus gamma power decrease (85-104 Hz, 80-110 ms).

Correlations between a vector (e.g., SA measured from sensor space) and another vector (e.g., PDI questionnaire score) were performed with the robust correlation toolbox (Pernet et al., 2012) except where specified differently. Correlations between a vector (e.g., SA measured from sensor space) and a matrix (e.g., the time-frequency representation of beta rebound) were performed with Spearman correlation as implemented in ft_statfun_correlationT.

3.2.7 Analysis of passive jittered condition data

This part of analysis aims at corroborating the existence of oscillatory neural information processing sequence (Figure 3.19) from the between-condition comparison analysis by testing them in a single condition setting. In single condition data, single-trial analysis within participants can provide additional evidence for the proposed sequence of events. In the passive jittered condition, single trial pre-stimulus alpha power (8-12 Hz, -300 to 0 ms) was extracted per participant and correlated with the single trial gamma power (absolute baseline correction; baseline: -300 to 0 ms). Pearson correlation coefficients were Fisher z transformed before being subjected to t-tests against 0 with cluster correction for multiple comparisons. For analyzing the relationship between post-stimulus gamma power and post-stimulus alpha/beta phase, we computed the phase deviation as the absolute angular difference of a single trial phase to the mean phase across trials. Then we used the phase deviation in the time-frequency window that showed a significant difference between conditions (Figure 3.10 b; 12 to 14 Hz, 70 to 160 ms) to correlate with single trial gamma power (with circ_corrcl from CircStat toolbox; Berens (2009)). The correlation coefficients within the first 100 ms after the stimulus onset were statistically compared to mean correlation coefficients in a baseline period (-300 ms to 0; paired t-test with cluster correction) after Fisher z transformation. Next, we computed the correlation between post-stimulus phase deviation (1-20 Hz; 70-160 ms) and post-stimulus gamma power (40-120 Hz; 0-40) ms for each time-frequency combination and the average over the time doamin was compared to the same correlation but calculated with a baseline gamma power (-300 to 0 ms). Fisher z transform was applied before paired t-tests with cluster correction. The result of this test can reveal the frequency specificity of the correlation between post-stimulus gamma power and post-stimulus low frequency phase deviation. Lastly, we correlated the gamma power (99-106 Hz, 0-40 ms; Figure 3.13 b) from the time window where significant correlations were found with the alpha/beta (12-14 Hz) phase deviation over time (from -750 ms to 750 ms) to reveal the temporal relationship between them.

3.2.8 Granger causality analysis

The computation of Granger causality is typically based on autoregressive modeling of neural time series data. If adding signal A to signal B reduces residual variances of the autoregressive model with only signal B, it indicates causal influence of signal A to signal B (Granger, 1969). Here, we computed Granger causality using an alternative implementation - nonparametric spectral matrix factorization of the cross-spectral density matrix (Dhamala, Rangarajan, & Ding, 2008) as implemented in FieldTrip. It is conceptually equivalent to the traditional implementation based on autoregressive models. However, since it relies on the factorization of the cross spectral density matrix it does not require the selection of a model-order which can be difficult to determine. The cross-spectral density matrix was computed using the same parameters as for source space time-frequency analysis below 40 Hz. In this way, the Granger causality result at each time point represents the information flow between two signals in the 500 ms time window centred at that time point. Granger causality was computed between 6 representative voxels from left motor cortex and 2 representative auditory voxels (one left and one right, which was also the one used for previous analysis) for each of the 12 (6 x 2) combinations. The results for the same auditory voxel were averaged across the 6 motor voxels. To statistically test Granger causality results, a bootstrapping procedure (100 randomization) was performed by calculating the Granger causality with trial orders of both motor and auditory signals randomized. The 95% percentile of bootstrapped Granger causality results was taken as the statistical significance cut-off.

3.3 Results

3.3.1 Sensor space analysis

3.3.1.1 Visual evoked responses

In all conditions where visual stimuli were presented, clear visual evoked responses were found (Figure 3.2). As expected, no visual evoked responses were found in movement only condition. No significant differences in the amplitude of visual evoked M100 component were found between active condition and passive periodic condition (t(13) = -0.91, p = 0.20), or between active condition and passive jittered condition (t(13) = -0.09, p = 0.46) (Figure 3.4 a). Therefore, no SA or sensory enhancement was found with early visual evoked responses in sensor space data. Therefore, further analyses focused on the auditory data.

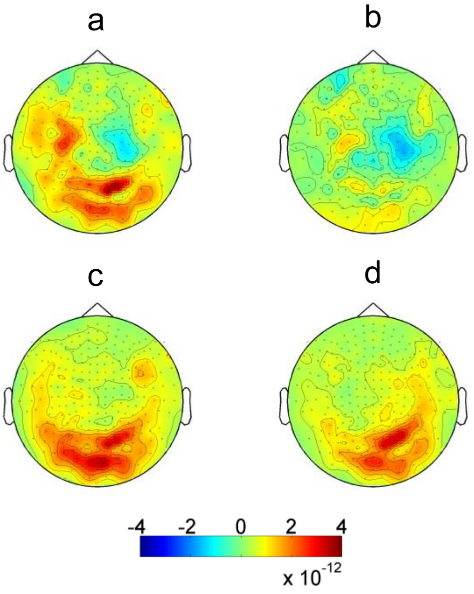


Figure 3.2 Topographies of visual evoked responses (90-130 ms) in (a) active condition, (b) motor only condition, (c) passive periodic condition and (d) passive jittered condition. Unit: tesla.

3.3.1.2 Auditory evoked responses

In all conditions where auditory stimuli were presented, clear auditory evoked responses were found. No auditory evoked responses were found in movement only condition (Figure 3.3). Right auditory area had clearly stronger responses as compared to left auditory area (paired t-test; t(13) = 5.02, $p < 0.001$). The reason of this lateralization is not clear. The last participant reported that she did not hear anything from the right ear and later checks confirmed that the right ear plug was blocked. But it is not clear how many participants had the same problem.

The classic SA effect with auditory evoked responses was replicated. There was a significant decrease in the amplitude of sound evoked M100 component in the active condition as compared to the passive periodic condition (left sensors: t(13) = -3.67, $p < 0.01$; right sensors: t(13) = -3.99, $p < 0.01$) and passive jittered condition (left sensors: t(13) = -3.13, $p < 0.01$; right sensors: t(13) = -5.14, $p < 0.01$) (Figure 3.4 b & c). There were no significant differences between passive periodic and passive jittered condition (left sensors: t(13) = 0.22, $p = 0.43$; right sensors: t(13) = 0.33, $p = 0.37$).

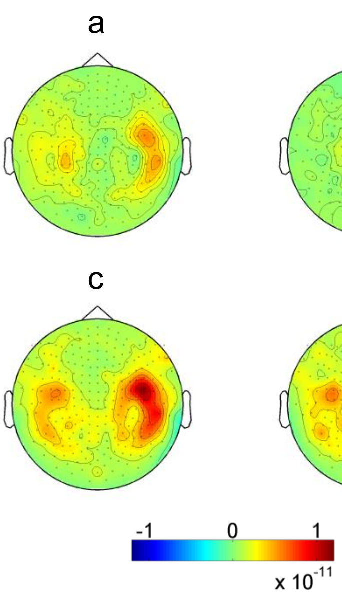


Figure 3.3 Topographies of auditory evoked responses (95-120 ms) in (a) active condition, (b) motor only condition, (c) passive periodic condition and (d) passive jittered condition. Unit: tesla.

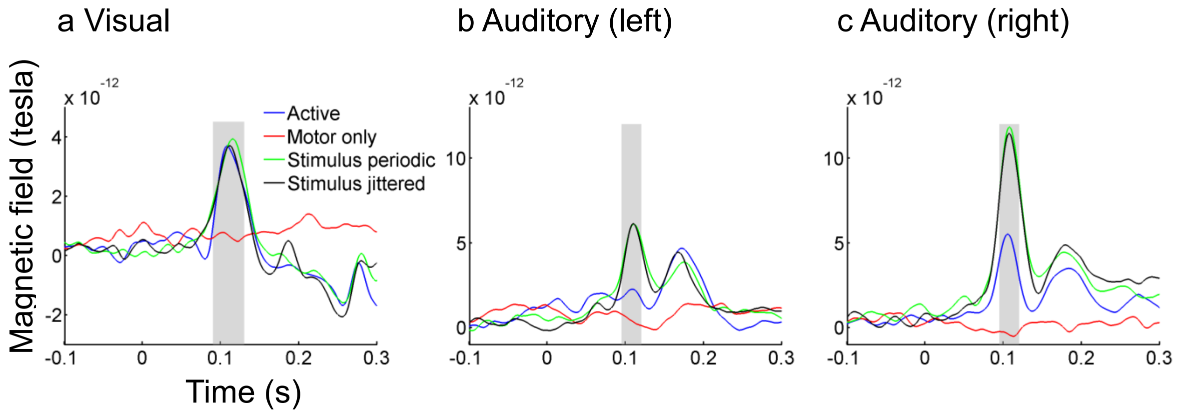


Figure 3.4 Visual (a) and auditory (b: left auditory sensors; c: right auditory sensors) evoked responses in all four conditions. The shaded area shows the selected time window for statistical comparisons (90-130 ms for visual evoked responses and 95-120 ms for auditory evoked responses).

3.3.1.3 Time-frequency analysis on auditory data

Auditory and motor induced oscillations

To illustrate the oscillatory components associated with auditory and motor responses, the power and PLV of right auditory sensors in passive period condition and motor sensors in motor only condition, respectively, were shown in Figure 3.5. An auditory input led to an increase of both power and PLV, which is most prominent in the theta range (Figure 3.5 a & b). Motor responses were characterized by increased power and PLV at around 3-4 Hz at the time of movement onset (Figure 3.5 c & d). The typical beta desynchronization and beta synchronization associated with movements were also evident (Figure 3.5 c, see also Supplementary Figure S3.10 a & b).

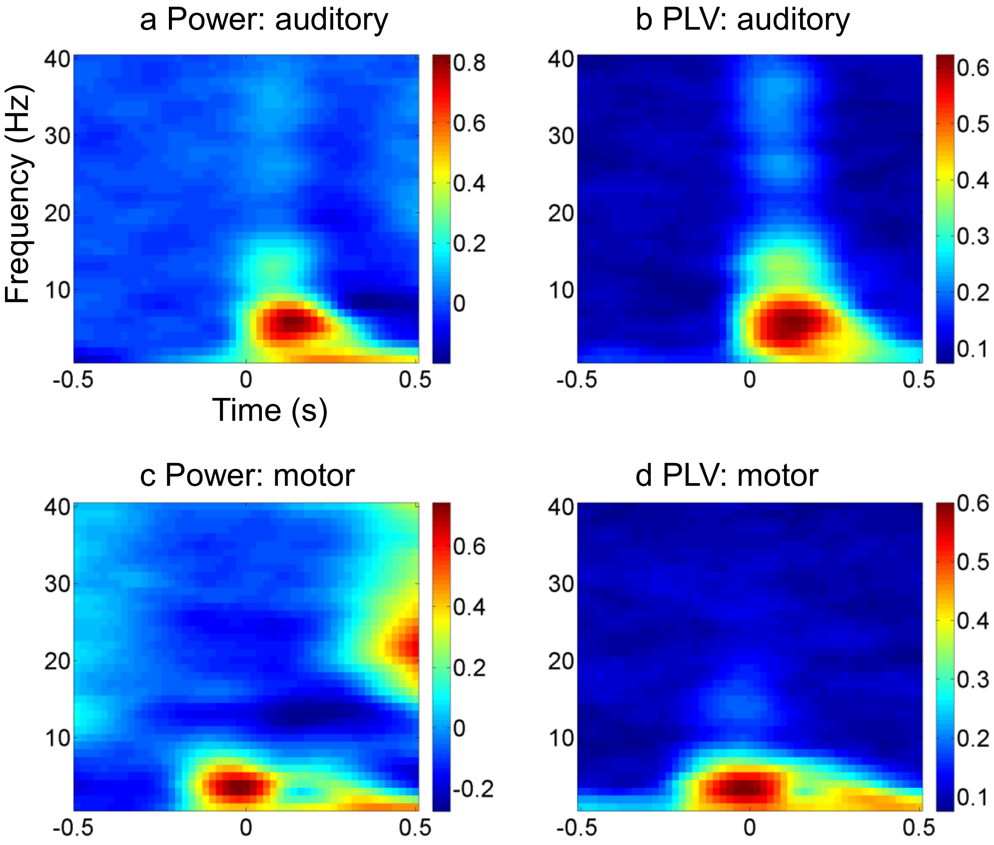


Figure 3.5 Grand average oscillatory components associated with auditory input (a & b; time 0: tone onset) and motor responses (c & d; time 0: action onset). Power data are baseline corrected (absolute correction with baseline from -750 to 0 ms).

Comparing oscillations between conditions

As introduced above, we hypothesised that in the SA auditory evoked responses are modulated by anticipatory prediction mechanisms, possibly reflected in the alpha band oscillation. Therefore, we tested for differences in oscillations between active and passive periodic condition in the pre-stimulus period. The onset timing of tones was predictable in both conditions, so the comparison can reveal specific effects associated with motor related predictions.

In the right auditory sensors, there was a clear increase of pre-stimulus alpha power in active condition as compared to passive periodic condition (Figure 3.6 a). Importantly, no obvious differences in the pre-stimulus alpha power were found between motor only condition and passive periodic condition (Figure 3.6 b). The power comparisons in the left auditory sensors did not reveal clear pre-stimulus alpha power changes, possible due to the contamination from motor related signals (Figure 3.6 c & d; see also Supplementary Figure S3.1 and Figure S3.2 for comparisons between active and passive jittered condition and comparisons made with motor sensors).

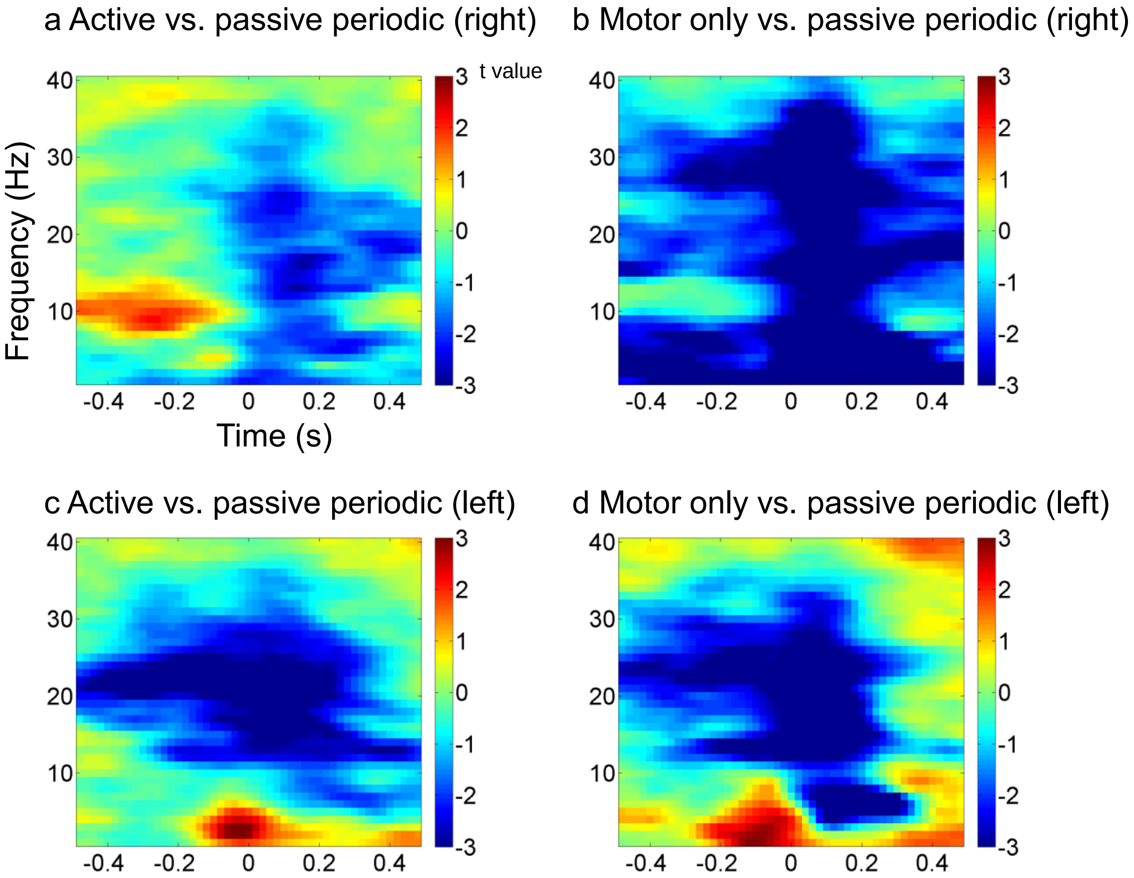


Figure 3.6 Power comparisons between active and passive periodic condition in the right (a) and left (c) auditory sensors, between motor only and passive periodic condition in the right (b) and left (d) auditory sensors.

Pre-stimulus theta band PLV seems to be higher in active condition as compared to passive periodic condition in both left and right auditory sensors. However, the same pattern was observed when comparing motor only condition and passive conditions in both auditory and motor sensors and it was stronger in left auditory sensors and motor sensors (Figure 3.7; Supplementary Figure S3.3 and Figure S3.4). So the pre-stimulus theta phase locking increase might be an artefact from motor preparation. To examine how SA is expressed in single-trials across different frequencies of neural activities, we focused next on oscillatory activities in the post-stimulus time window that overlaps with the M100 component. Both power and phase locking decreases were found in the post-stimulus time window (Figure 3.6 & 3.7). However, it seemed that prominent decreases were in the alpha and beta band, but not in the theta band in right auditory sensors when comparing active condition and passive periodic condition. It has been shown earlier that auditory evoked components were most strongly represented in the theta band for both power and PLV (Figure 3.5 a & b). Interestingly, differences between conditions seemed to predominantly occur at higher frequencies.

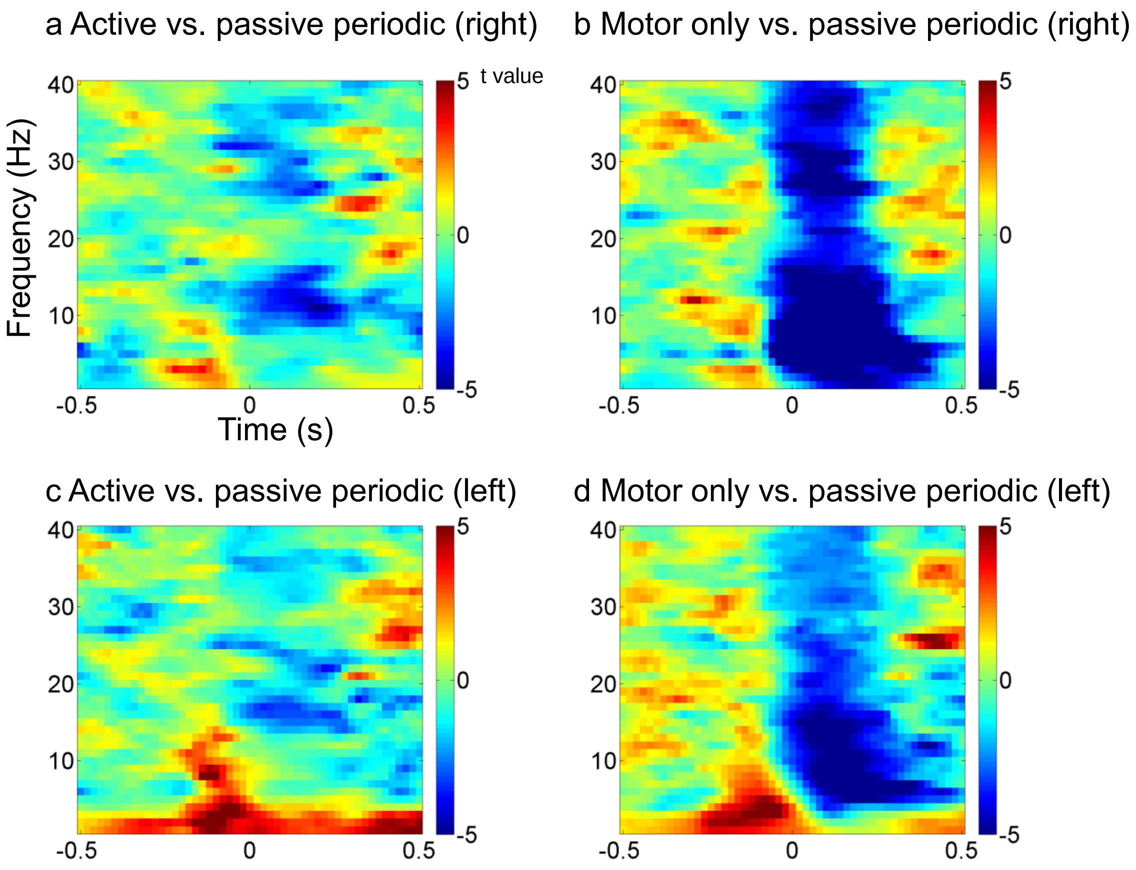


Figure 3.7 Phase locking value comparisons between active and passive periodic condition in the right (a) and left (c) auditory sensors, between motor only and passive periodic condition in the right (b) and left (d) auditory sensors.

3.3.2 Interim summary with sensory space analysis

In the visual domain, no significant modulations on the amplitude of M100 component by actions were found. In the auditory domain, the classic SA effect was replicated. Comparisons of oscillatory activities showed that an increase of pre-stimulus alpha power is a likely candidate of motor prediction induced changes in auditory cortex. In the post-stimulus time window, decreases in power and PLV of alpha/beta band oscillations are likely neural representations of SA as seen in evoked responses. In the source space analysis below, we compared visual evoked responses again and focused on activities in right auditory cortex in both active and passive periodic conditions where prominent effects were found in the sensor space analysis.

3.3.3 Source space analysis

3.3.3.1 Visual evoked responses in source space

A whole-brain t-test was performed comparing early visual evoked responses (90-160 ms; to validate the source analysis, source localization was performed for the active condition and showed, as expected, the activation of left primary motor cortex and visual cortex, see supplementary Figure S3.5). No significant differences were found between action condition and passive periodic condition (Figure 3.8 a), or between action condition and passive jittered condition (Figure 3.8 b) in visual cortex. Additionally, the evoked responses from the voxel that showed the strongest activation in visual cortex in passive periodic condition were selected for analysis. This single voxel analysis led to similar results (Figure 3.8 c), i.e., no significant differences in visual evoked responses were found between action condition and passive periodic condition (t(13) = 0.69, $p = 0.29$), or between action condition and passive jittered condition (t(13) = 0.91, $p = 0.21$). Therefore, further analyses focused on the auditory data. As a side note, in the post-stimulus time window, there were increased responses in active condition compared to passive conditions in motor cortex (see supplementary Figure S3.5).

3.3.3.2 Localizing SA effect

Source localization analysis demonstrated the maximum SA effect in the auditory cortex, confirming a significant reduction of primary auditory cortex response amplitude for self-initiated sounds compared to external sounds (Figure 3.9). In the voxel selected for further analysis below, significant SA from reconstructed time series data was confirmed (t(13) = -2.87, *p* = 0.01) (see Supplementary Figure S3.11 for the reconstructed time series data in source space).

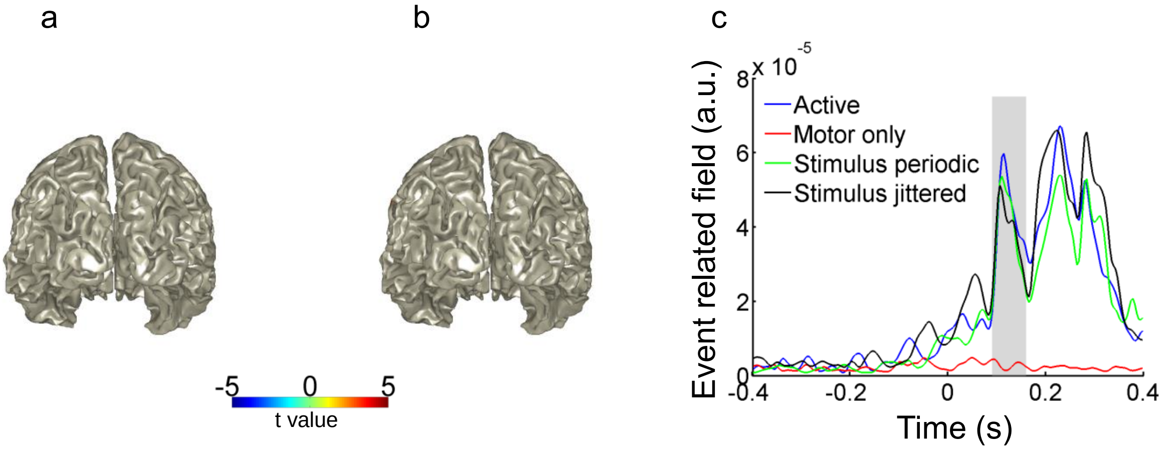


Figure 3.8 Source space comparisons of visual evoked responses. There is no significant different between active condition and passive periodic condition (a) or between active condition and passive jittered condition (b) for the whole-brain analysis (cluster correction). Similarly, for the selected visual voxel, no differences are found (c). The shaded area (90 ms to 160 ms) shows the time window selected for analysis.

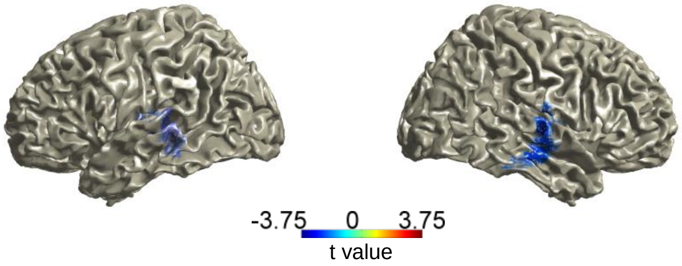


Figure 3.9 SA effect is localized in auditory cortex.

3.3.3.3 Neural activity preceding sensory attenuation in auditory cortex

A significant alpha power increase was found in active condition compared to passive periodic condition starting around 400 ms before stimulus onset (Figure 3.10 a), which is consistent with results from sensor space analysis. Testing for a relationship between this pre-stimulus alpha power increase and the magnitude of SA across participants revealed a significant correlation: increased alpha power was associated with increased SA (Spearman's rho = -0.74, *p* = 0.003, 95% CI = [-0.92 -0.33]; Figure 3.11 a). No significant changes in PLV were found in the pre-stimulus time window (Figure 3.10 b).

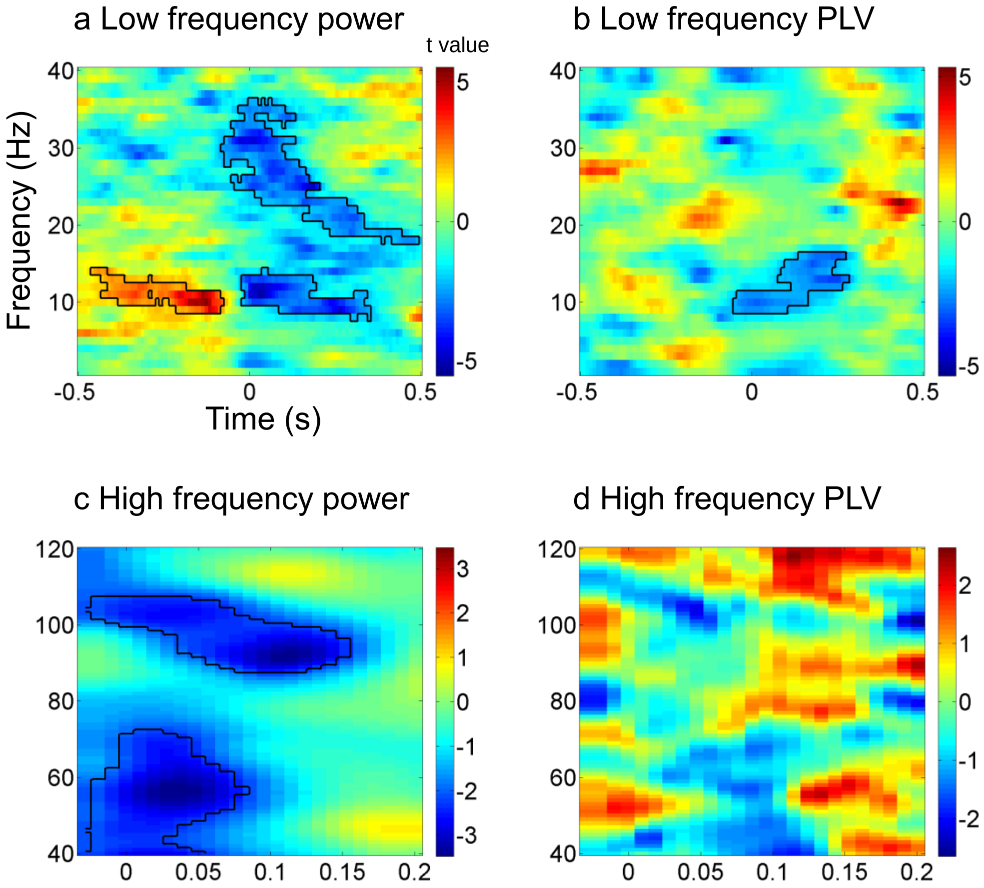


Figure 3.10 Power and phase locking value comparisons between the active and passive periodic condition. In the pre-stimulus time window, a clear alpha power increase is shown (a). In the post-stimulus time window, broadband power decreases coincide with sensory attenuation from the evoked fields analysis (a & c). Of particular interest is the post-stimulus gamma power decrease. But there are no changes to theta band oscillations. Post-stimulus phase locking is decreased in the alpha/beta range (b). No difference is found in the gamma range phase locking (d).

3.3.3.4 Neural representation of sensory attenuation in auditory cortex

Time-frequency analysis revealed a significant decrease of oscillatory power at frequencies in the alpha/low beta (9-15) and higher beta band (20-35 Hz) as well as in the gamma band (40-70 Hz and 90-110 Hz, Figure 3.10 a & c) for the active as compared to the passive condition. These broadband changes overlapped in time with the SA effect. In parallel, phase locking to stimulus onset was significantly reduced in a limited frequency band, spanning the alpha/low beta frequency (9-15 Hz) in the same time window, for the active as compared to the passive condition (Figure 3.10 b). Again, this analysis confirmed that post-stimulus oscillatory changes only occurred at frequencies higher than the theta band, where evoked responses were prominently represented. No significant differences were found in gamma band phase locking (Figure 3.10 d).

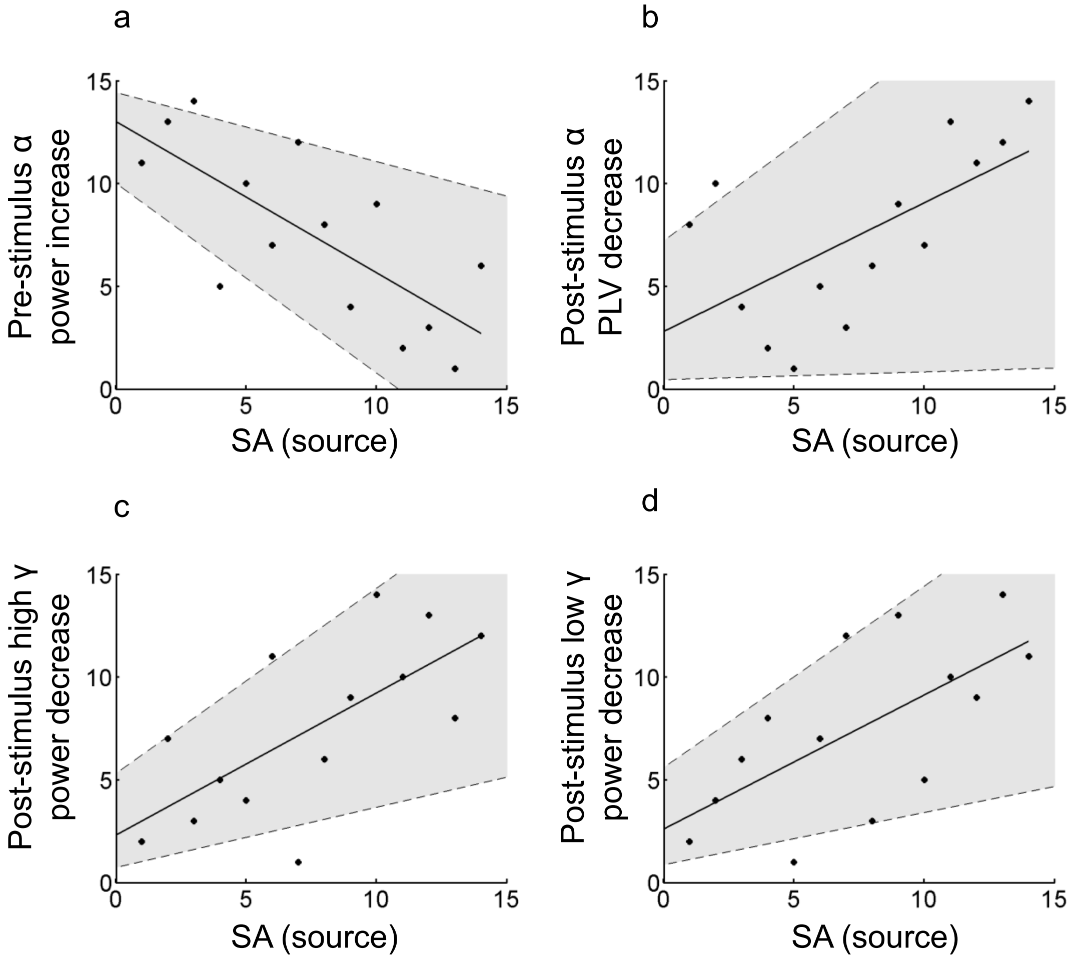


Figure 3.11 Scatter plots for the correlations between SA in source space and significant power/phase locking changes between conditions. (a) SA is negatively correlated with pre-stimulus alpha power increase (r = -0.74, *p* = 0.003, CI = [-0.92 -0.33]). (b) SA is positively correlated with post-stimulus alpha phase locking decrease (r = 0.63, *p* = 0.02, CI = [0.05 0.94]). (c) SA is positively correlated with post-stimulus high gamma power (85-104 Hz) decrease (r = 0.69, *p* = 0.006, CI = [0.29 0.90]). (d) same with C, but with lower gamma (57-62 Hz) (r = 0.65, *p* = 0.01, CI = [0.24 0.88]). The solid line indicates a linear fitting to the data points and the shaded area indicates the 95% confidence interval of the correlation. Note that the data being plotted are ranks of related measures.

When examining the relationship between these post-stimulus oscillatory events and SA, we found the SA effect to significantly correlate with three post-stimulus oscillatory events, namely alpha phase locking decrease (Spearman's rho = 0.63, *p* = 0.02, 95% CI = [0.05 0.94], Figure 3.11 b), high gamma power decrease (Spearman's rho = 0.69, *p* = 0.006, 95% CI = [0.29 0.90], Figure 3.11 c) and low gamma power decrease (Spearman's rho = 0.65, *p* = 0.01, 95% CI = [0.24 0.88], Figure 3.11 d). All the above correlations were resistant to influences from outliers as the correlations remained significant when eliminating outliers using the Spearman skipped correlation (see Appendix 2). SA was not significantly correlated with alpha/beta power changes (supplementary Figure S3.6 a). For a full correlation map, see supplementary Figure S3.6.

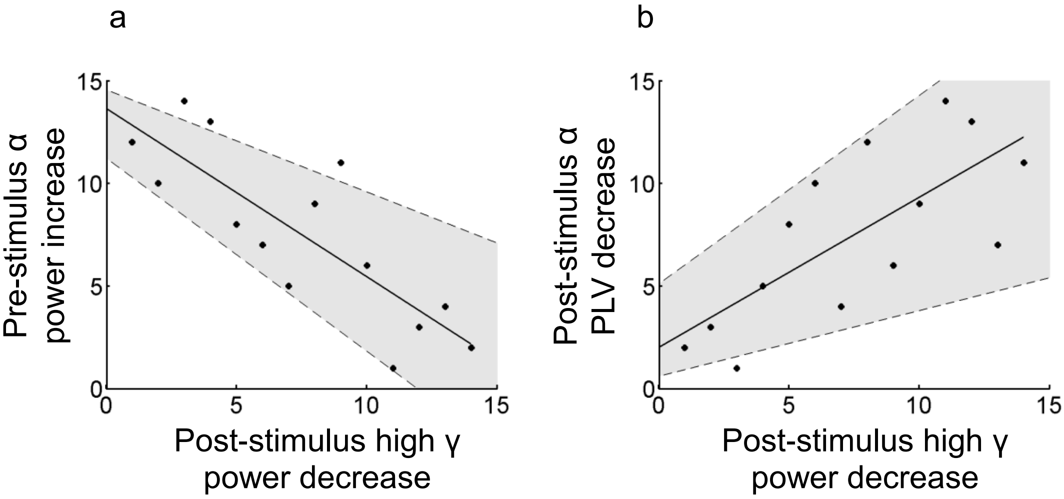


Figure 3.12 Scatter plots for cross-participant correlations between pre-stimulus alpha power increase and post-stimulus gamma power decrease (a), and between post-stimulus alpha phase locking decrease and post-stimulus gamma power decrease (b). An increase in the pre-stimulus alpha power is associated with a decrease in the post-stimulus gamma power (Spearman's rho = -0.82, *p* = 0.0003, 95% CI = [-0.94 -0.50]), and a decrease in the post-stimulus gamma power is associated with a decrease in the post-stimulus alpha phase locking (Spearman's rho = 0.73, *p* = 0.003, 95% CI = [0.31 0.92]). The solid line indicates a linear fit to the data points and the shaded area indicates the 95% confidence interval of the correlation. Note that the data being plotted are ranks of related measures.

3.3.3.5 Neuronal implementation of sensory attenuation in auditory cortex

Our results so far demonstrate that changes in pre-stimulus alpha power, post-stimulus alpha phase locking and post-stimulus gamma power were most relevant to SA, as evidenced by both, significant between-condition differences and correlations across participants. Because the pre-stimulus low frequency power change (especially in the alpha band) and post-stimulus oscillatory changes are possible candidates for mediating the SA, we investigated if these oscillatory components are correlated amongst each other. Across participants, there was a significant negative correlation between the pre-stimulus alpha power increase and the post-stimulus high gamma power decrease (Spearman's rho =

-0.82, p = 0.0003, 95% CI = [-0.94 -0.50]; Figure 3.12 a). Moreover, a significant positive correlation was found between the post-stimulus gamma power decrease and the post-stimulus alpha phase locking decrease (Spearman's rho = 0.73, p = 0.003, 95% CI = [0.31 0.92]; Figure 3.12 b). Both correlations were resistant to possible influences of outliers (see Appendix 2) and remained significant after Holm-Bonferroni correction. For a full correlation map, see Supplementary Figure S3.7 a & b.

3.3.3.6 Co-variation of the auditory pre-stimulus and post-stimulus oscillatory components related to sensory attenuation in the absence of predictive cues

Next, we tested whether the sequence of events described above (pre-stimulus alpha power relating to post-stimulus gamma power and alpha/beta phase reset) is also present during auditory sensory stimulus processing when no explicit predictions can be formed. Therefore, we tested for the presence of the same correlations within participants in a cross-trial analysis of the passive listening condition in which the inter-trial interval was randomly jittered. The jittered interval makes the exact onset of the stimulus unpredictable, thus leading to a variation in the participant's preparedness towards the stimulus.

We correlated pre-stimulus alpha power (8-12 Hz; -300 ms to 0 ms) with post-stimulus gamma power across trials and subjected the individual correlation maps to group statistics. Consistent with our analysis across participants, we found a significant correlation between pre-stimulus alpha power and early post-stimulus gamma power (Figure 3.13 a). The negative sign of the correlation indicated that a high pre-stimulus alpha power was associated with a low post-stimulus gamma power. The pre-stimulus alpha power was also correlated with gamma power (around 85 Hz) starting around 100 ms after stimulus onset.

In order to reveal the possible relationship between post-stimulus gamma power and post-stimulus alpha/beta phase across trials (in analogy to the cross participant analysis above), we calculated the phase deviation as the absolute angular difference of a single trial phase to the mean phase across trials for each individual at alpha/beta frequency shortly after stimulus onset. We then correlated this phase deviation with single-trial power across time and frequency. Post-stimulus gamma power was significantly correlated with alpha/beta band phase deviation shortly after stimulus onset (Figure 3.13 b). This correlation was only significant for post-stimulus phase from 11 to 14 Hz (Figure 3.13 c). Importantly, recalculating the correlation by taking into account only the very early gamma power data (first 40 ms of significant post-stimulus correlations) revealed its correlation with alpha/beta phase deviation to peak at a later time point (at around 150 ms, Figure 3.13 d), indicating that gamma power precedes alpha/beta phase resetting.

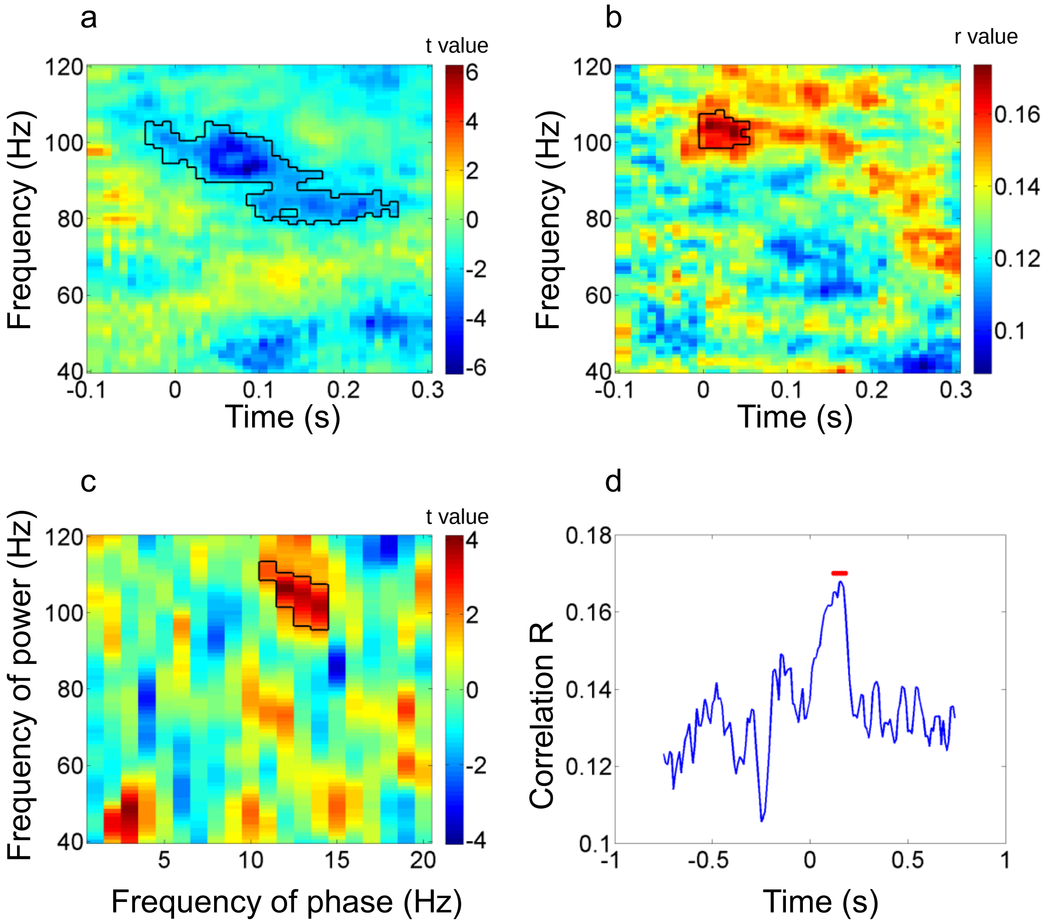


Figure 3.13 Results from the cross-trial analysis and schematic summary. In the passive jittered condition, pre-stimulus alpha power is correlated with post-stimulus gamma power (a), and post- stimulus alpha/beta phase deviation is correlated with the gamma power at comparable frequency bands and time points (b). (c) shows that the post-stimulus gamma power is only significantly correlated with post-stimulus alpha/beta phase deviation. Taking the first 40 ms post-stimulus gamma power where significant correlations were found in (b), the temporal dynamics of its correlation with post-stimulus alpha/beta phase deviation is shown in (d). There is a clear peak around 150 ms after the stimulus onset. The red line indicates post-stimulus points where there are significant higher correlations than the baseline period (-750 ms to 0; paired t test without multiple comparison correction; $p <$ 0.05).

3.3.3.7 Granger causality between motor cortex and auditory cortex

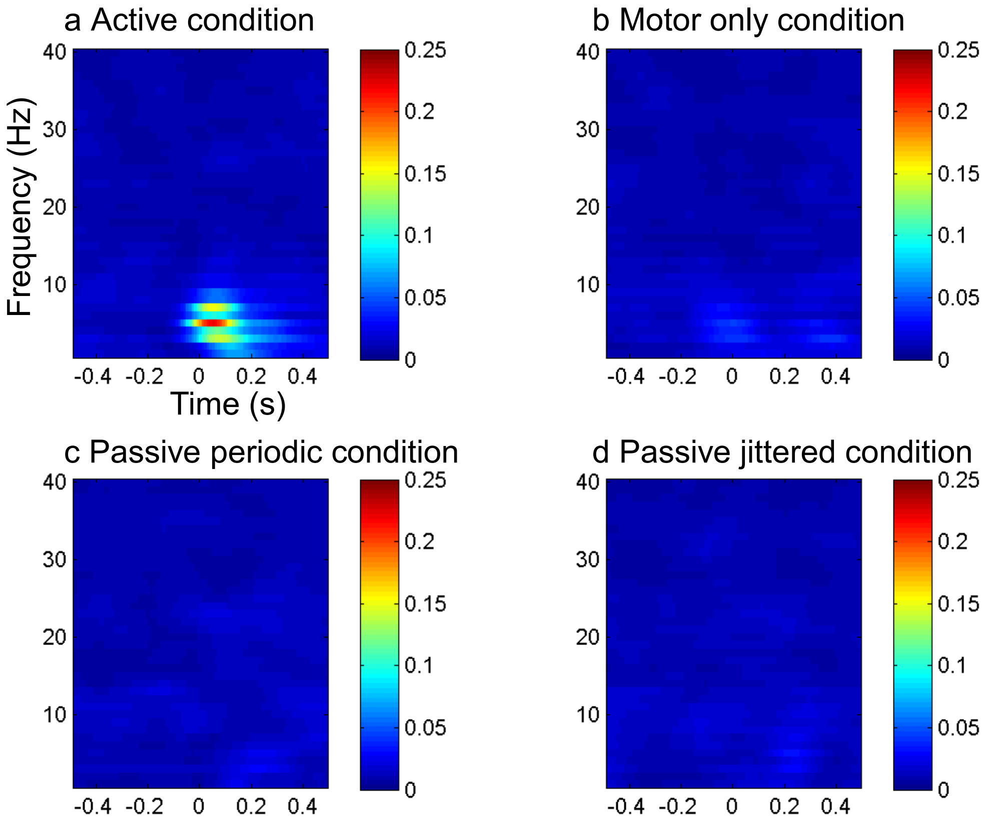


Figure 3.14 Granger causality results in the time frequency domain showing the causal influence of motor cortex over right auditory cortex in all the 4 conditions.

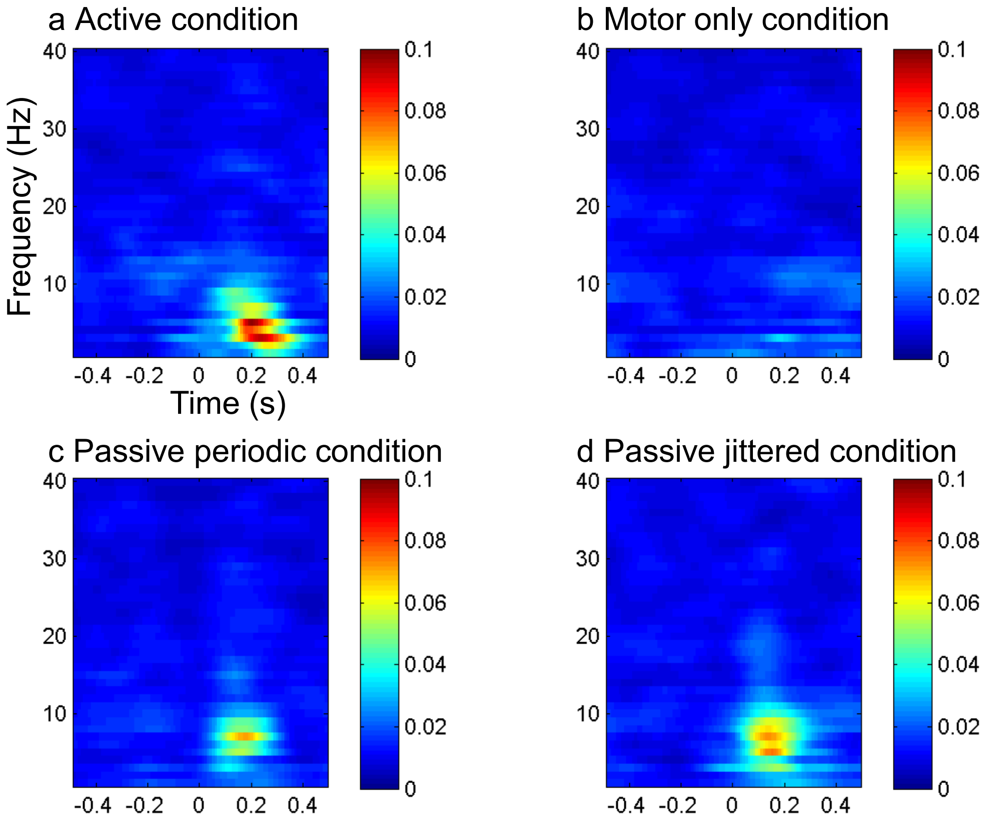


Figure 3.15 Granger causality results in the time frequency domain showing the causal influence of right auditory cortex over motor cortex in all the 4 conditions.

In Figure 3.14, each point indicates the causal influence of motor cortex over right auditory cortex in the specified time and frequency (Granger causality value indicated by color; time 0 aligned to stimulus onset). Similarly, Figure 3.15 shows the causal influence of right auditory cortex over motor cortex. The results showed that there was increased information flow from motor cortex to both right (Figure 3.14) and left auditory cortex (Supplementary Figure S3.8) in the theta band for both active condition and motor only condition, which was most prominent within the first 100 ms after stimulus onset. For the other direction of information flow, i.e., from auditory cortex to motor cortex, there was an increase in active condition, passive periodic condition and passive jittered condition in the theta band, which appeared around 200 ms (see Figure 3.15 for results of Granger causality from right auditory cortex to motor cortex and Supplementary Figure S3.9 for Granger causality from left auditory cortex to motor cortex). However, only the information flow from motor cortex to left auditory cortex in motor only condition and from left auditory cortex to

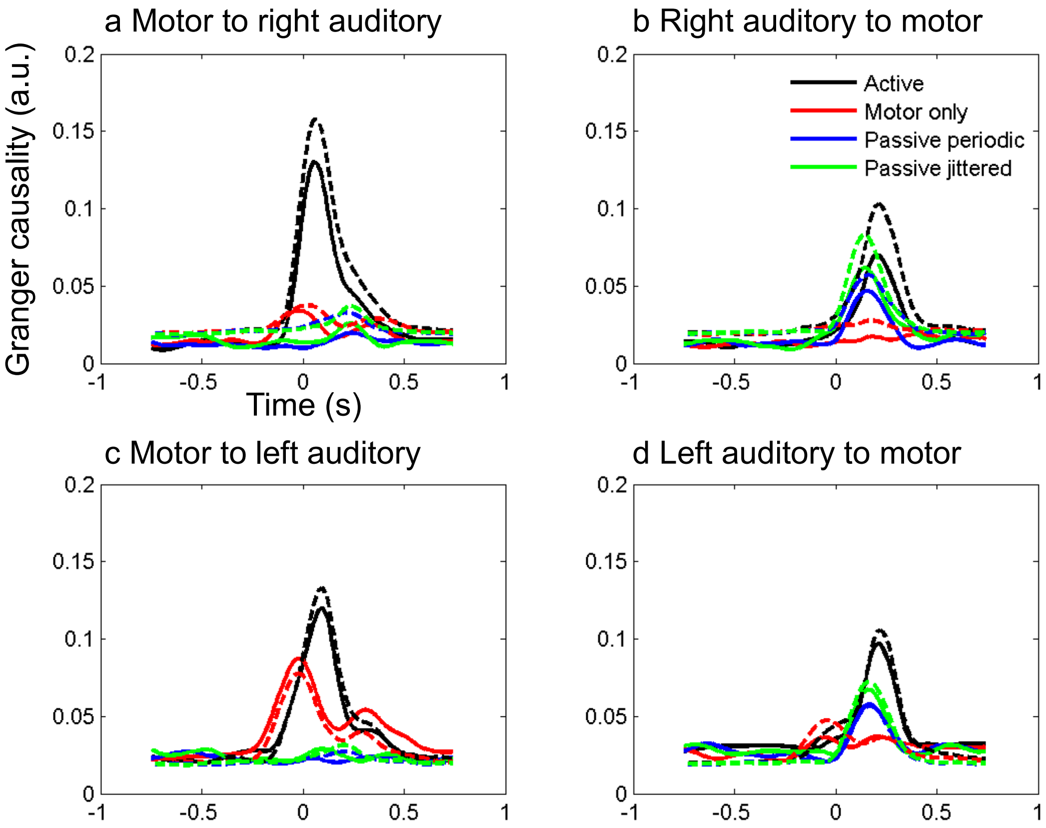


Figure 3.16 Granger causality results over time averaged between 3 Hz and 8 Hz of all conditions. The solid line shows the real granger causality data. The dashed line shows the 95% cut-off of bootstrapped granger causality results. (a) shows the information flow from motor cortex to right auditory cortex; (b) shows the information flow from right auditory cortex to motor cortex; (c) shows the information flow from motor cortex to left auditory cortex; (d) shows the information flow from left auditory cortex to motor cortex.

motor cortex in passive periodic condition were above statistical cut-off (Figure 3.16).

3.3.3.8 Correlations between SA and questionnaire data

Most measures from PDI were significantly correlated with SA from right auditory sensors with Pearson's correlation (yes/no score: Pearson's r = 0.63, $p = 0.04$; distress score: Pearson's r = 0.70, $p = 0.02$; preoccupation score: Pearson's r = 0.63, $p = 0.04$; conviction score: Pearson's r = 0.58, $p = 0.06$; overall score: Pearson's r = 0.66, $p = 0.03$; Figure 3.17 a). A high PDI score was associated with a small SA effect (note the negative sign of SA effect). PDI scores were correlated with amplitudes of evoked response in active condition (yes/no score: Pearson's r = 0.54, $p = 0.09$; distress score: Pearson's r = 0.63, $p = 0.04$; preoccupation score: Pearson's r = 0.77, $p = 0.01$; conviction score: Pearson's r = 0.61, $p = 0.046$; overall score: Pearson's r = 0.68, $p = 0.02$; Figure 3.17 b), but not in passive periodic condition. A high PDI score was associated with high amplitude of M100 component. No significant correlations were found with left auditory sensors or when using Spearman's correlation (Table 3.1).

PDI scores were also correlated with beta rebound from the motor sensors in both active condition and motor only condition (Table 3.2; see also Supplementary Figure S3.10 c & d). Figure 3.18 showed the correlation between beta rebound in active condition and PDI overall score, and the pattern looked similar with other PDI scores. A high PDI score was associated with high beta rebound.

For SQ-short questionnaire, an interesting correlation was found with SA effect from both left (Spearman's rho = -0.95, $p < 0.001$; Figure 3.17 c) and right auditory sensors (Spearman's rho = -0.78, $p < 0.01$). A high SQ score was associated with a big SA effect. With left auditory sensors, SQ score was correlated with the amplitude of M100 component in passive periodic condition (Spearman's rho = 0.78, $p = 0.005$; Figure 3.17 d) and in passive jittered condition (Spearman's rho = 0.63, $p = 0.04$), but not with the amplitude of M100 component in active condition (Spearman's rho = 0.08, $p = 0.81$). A high SQ score was associated with high amplitude of M100 component. SQ score was not significantly correlated with amplitudes of M100 component in passive periodic (Spearman's rho = 0.31, $p = 0.36$), passive jittered (Spearman's rho = 0.31, $p = 0.36$) or active condition (Spearman's rho = -0.08, $p = 0.81$) with right auditory sensors. For all correlations with evoked responses, see Table 3.1

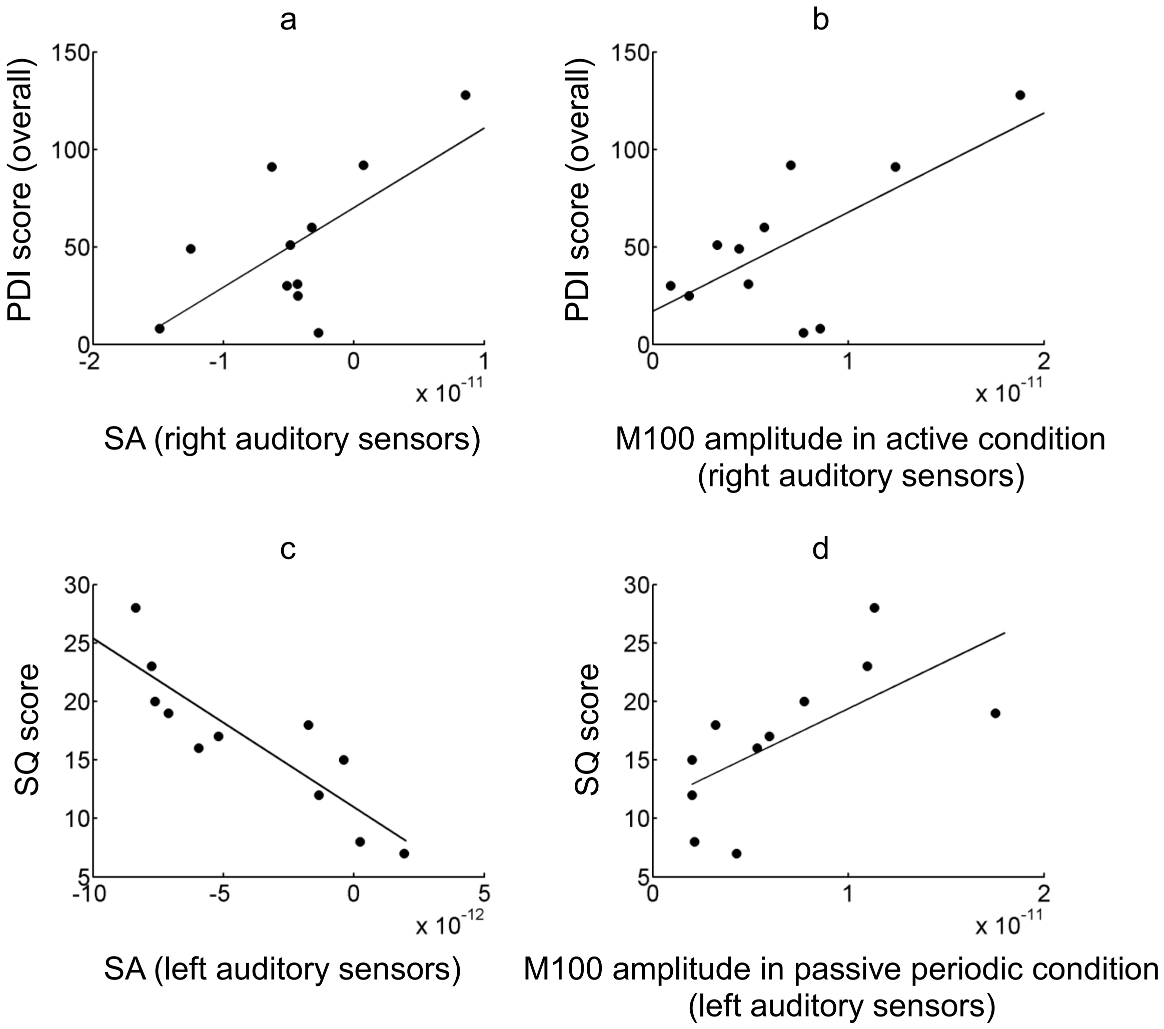


Figure 3.17 Correlations between questionnaire scores and SA measures. (a) and (b) show that PDI overall score is correlated with SA effect from right auditory sensors and the amplitude of M100 component in active condition, respectively. (c) and (d) show that SQ score is correlated with SA effect from left auditory sensors and the amplitude of M100 component in passive periodic condition, respectively.

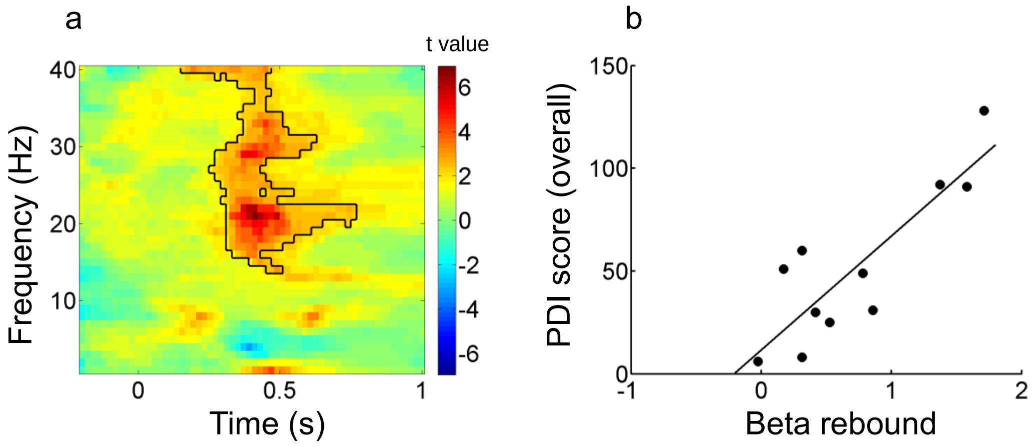


Figure 3.18 Correlations between PDI overall score and beta rebound in active condition. (a) shows that PDI overall score is strongly correlated with early beta rebound. (b) shows the scatter plot taking the beta rebound data from 15 to 25 Hz, 400 to 600 ms.

Table 3.1 Correlations between questionnaire scores and amplitude of evoked responses. 95% confidence interval of the correlation coefficient (using skipped correlation in Pernet et al. (2012)) is shown in brackets. EQ, empathy quotient; SQ, systemizing quotient; PDI, Peters et al. delusion inventory; SA, sensory attenuation; L, left; R, right.

Measure	(*Method*)	EQ	SQ	PDI (overall)
SA (L)	Spearman	0.24 [-0.51 0.91]	-0.95* [-1 -0.72]	0.21 [-0.53 0.88]
	Pearson	0.17 [-0.48 0.82]	-0.87* [-0.97 -0.68]	0.27 [-0.58 0.82]
M100 (active; L)	Spearman	-0.17 [-0.71 0.65]	0.08 [-0.65 0.69]	0.31 [-0.41 0.78]
	Pearson	-0.39 [-0.73 0.33]	-0.02 [-0.71 0.63]	0.56† [-0.33 0.84]
M100 (passive periodic; L)	Spearman	-0.30 [-0.87 0.49]	0.78* [0.35 0.97]	0.20 [-0.61 0.78]
	Pearson	-0.38 [-0.83 0.47]	0.65* [0.43 0.94]	0.16 [-0.53 0.77]
M100 (passive jittered; L)	Spearman	-0.15 [-0.77 0.66]	0.63* [0.01 0.89]	0.37 [-0.39 0.83]
	Pearson	-0.28 [-0.72 0.57]	0.44 [0.19 0.87]	0.39 [-0.22 0.86]
SA (R)	Spearman	0.04 [-0.69 0.83]	-0.78* [-0.99 -0.22]	0.37 [-0.46 0.92]
	Pearson	-0.11 [-0.71 0.65]	-0.80* [-0.97 -0.25]	0.66* [-0.23 0.92]
M100 (active; R)	Spearman	-0.07 [-0.70 0.56]	-0.08 [-0.74 0.72]	0.35 [-0.45 0.94]
	Pearson	-0.25 [-0.72 0.26]	-0.23 [-0.81 0.67]	0.68* [-0.31 0.96]
M100 (passive periodic; R)	Spearman	-0.17 [-0.86 0.58]	0.31 [-0.53 0.84]	-0.06 [-0.64 0.66]
	Pearson	-0.10 [-0.87 0.63]	0.65* [-0.17 0.91]	-0.10 [-0.60 0.62]
M100 (passive jittered; R)	Spearman	-0.07 [-0.66 0.86]	0.31 [-0.44 0.75]	0.23 [-0.43 0.72]
	Pearson	-0.33 [-0.75 0.67]	0.31 [-0.32 0.73]	0.35 [-0.32 0.74]

$^{\dagger} p = 0.08$; * $p < 0.05$

Table 3.2 Correlations between PDI (Peters et al. delusion inventory) scores and beta rebound in active condition (15 to 25 Hz, 400 to 600 ms), motor only condition (24-26 Hz, 650 to 900 ms). 95% confidence interval of the correlation coefficient (using skipped Spearman correlation in Pernet et al. (2012)) is shown in brackets.

Measure	(*Method*)	Beta rebound (active)	Beta rebound (motor only)
yes/no score	Spearman	0.67* [0.05 0.93]	0.61* [-0.08 0.94]
distress score	Spearman	0.72* [0.05 0.99]	$0.58^{\dagger}$ [-0.22 0.89]
preoccupation score	Spearman	0.69* [-0.22 0.99]	$0.58^{\dagger}$ [-0.34 0.91]
conviction score	Spearman	0.77* [0.21 0.97]	0.70* [0.08 0.94]
overall score	Spearman	0.71* [0.11 0.97]	0.62* [-0.05 0.94]

$^{\dagger}p = 0.06$; * $p < 0.05$

3.4 Discussion

First, in this study, the influence of actions on both auditory and visual processing was investigated with MEG by comparing brain responses to self-generated and external stimuli. The classic SA effect in the auditory domain (i.e., self-generated sound elicited a smaller amplitude of M100 component compared to externally generated sound) was replicated and was localized in the auditory cortex, whereas no SA or sensory enhancement was found in the visual domain.

Second, although SA has been typically studied in the auditory system, the neural mechanisms underlying this phenomenon are likely part of a neural functional architecture that acts along the different stages of sensory processing pathways. In fact, this well-studied effect has been linked to the predictive coding framework that postulates the importance of predictive neural models for general information processing in brain networks (Brown et al., 2013; Friston, 2005). In this Bayesian framework the brain generates predictions about the environment that are constantly compared to and updated by incoming sensory evidence. The resulting prediction errors are communicated to the next level in the processing hierarchy. An integral part of this theory is the control of gain of these prediction errors that is adjusted according to their expected precision (Friston, Bastos, Pinotsis, & Litvak, 2015). It has been argued that sensory attenuation originates from reduced precision of self-generated sensory information (Brown et al., 2013). Interestingly, brain oscillations provide efficient mechanisms for gain control and are ideal candidates for the neural mechanisms underlying SA. Our findings in the source space support this hypothesis. First, they show how pre-stimulus changes of auditory alpha band oscillatory power affect auditory stimulus processing and lead to SA. Second, we demonstrate that at the level of single trials SA is reflected in, both, reduced broadband power (including gamma) and reduced alpha/beta phase locking when comparing the active and passive condition. Third, we find a significant relationship between pre-stimulus alpha power changes, post-stimulus gamma power changes and post-stimulus alpha/beta phase changes, which may represent a functional sequence of neural information processing steps around the time of stimulus presentation (Figure 3.19). This receives further support from the single trial correlation analysis performed on the passive jittered condition data.

Third, the Granger causality analysis between motor cortex and auditory cortex revealed interesting results by showing increased information flow from motor cortex to auditory cortex in both active and motor only conditions. It also shows increased information flow from auditory cortex to motor cortex, in a later time window, in active condition, passive periodic condition and passive jittered condition. However, few of these passed statistical significance.

Fourth, correlations between questionnaire scores and MEG data resulted in several interesting findings. The correlation between PDI scores and SA was replicated, although it only reached statistical significance when using Pearson correlation. The correlations between PDI scores and beta rebound, between SQ scores and SA were robust and novel. The above four points are discussed more in detail below.

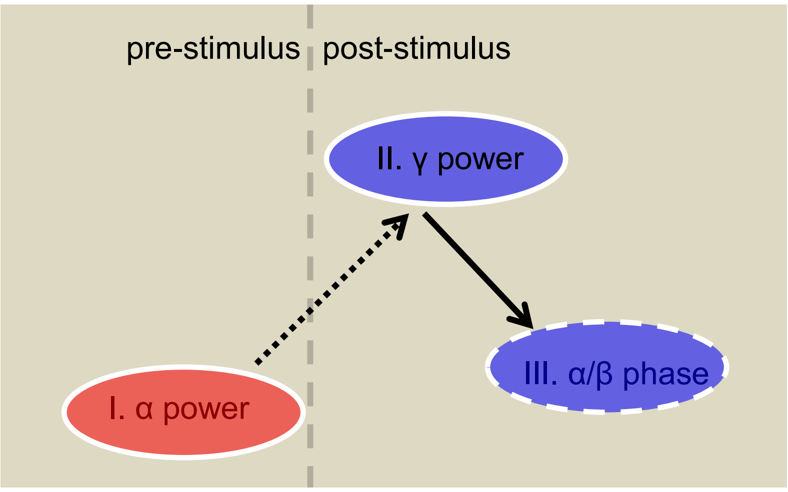


Figure 3.19 Schematic illustration of the relationship found among oscillatory changes between conditions. The increase of pre-stimulus alpha power is negatively correlated with the decrease of post-stimulus gamma power, which in turn is positively correlated with the decrease of post-stimulus alpha/beta phase locking. This may constitute a sequence of neural information processing (from I to III) which receives further support from the single-trial analysis with the passive jittered condition data. Red colour indicates a relative signal increase in the active condition and blue indicates a decrease. Solid ellipse edge indicates power and dashed line indicates phase. Dashed arrow indicates a negative correlation and solid arrow indicates a positive correlation.

3.4.1 Action effects on sensory processing

Despite the unequivocal SA effect of actions on auditory and somatosensory processing, which is replicated here in the auditory domain, the effect of actions on visual processing seems quite controversial. Only a handful of human EEG studies addressed this question, with SA being reported in frontocentral sensors (Gentsch & Schutz-Bosbach, 2011; Hughes & Waszak, 2011; Schafer & Marcus, 1973) and sensory enhancement being reported in occipital sensors (Hughes & Waszak, 2011; Mifsud et al., 2016) in visual studies. Here, we found neither SA nor sensory enhancement in processing self-generated visual stimuli as compared to externally generated visual stimuli in the occipital area.

So what do actions do to visual processing? It may well be that there is a sensory enhancement in the visual domain. First, in the two sensory enhancement reports, occipital sensors were selected for analysis (Hughes & Waszak, 2011; Mifsud et al., 2016). Whereas in the studies where SA was reported, frontocentral sensors were used (Gentsch & Schutz-

Bosbach, 2011; Hughes & Waszak, 2011; Schafer & Marcus, 1973). Thus the visual signals from sensory enhancement studies are more reliable than from SA studies, whose ‘visual signals’ may be contaminated, e.g., by movement related signals. Second, ample evidence from animal studies shows that actions lead to enhanced visual responses (Bennett, Arroyo, & Hestrin, 2013; Niell & Stryker, 2010). For example, Niell and Stryker (2010) showed an enhancement of visual evoked firing rates in the visual cortex of mice from the state of standing still to the state of running. At the same time, animal studies also showed suppressive effects from actions in the auditory and somatosensory domain (Chapman, Jiang, & Lamarre, 1988; Schneider et al., 2014; Song & Francis, 2013; Zhou et al., 2014). This sensory enhancement effect in the visual domain might be smaller than the SA effect in the auditory domain in the scalp recorded electromagnetic signals and thus requires either a big sample size (40 participants in Mifsud et al. (2016)) or a large number of trials for active condition (240 trials in Hughes and Waszak (2011)). In the present study, 100 trials in action condition were collected from 14 participants. Neither the number of trials or the number of pariticipants may be big enough.

The difference in the effects of actions on different sensory modalities is an interesting question. Is efference copy signal enough to explain the effects across different modalities? If so, why does sensory enhancement effect seem to appear in the visual domain instead of SA? If different mechanisms exist, what are they? Mifsud et al. (2016) failed to find a correlation between auditory SA effect and visual sensory enhancement effect across participants in their study, which hints a different mechanism in the visual domain. In a similar vein, Williamson, Hancock, Shinn-Cunningham, and Polley (2015) reported differential modulatory effects in the auditory and visual regions of the thalamus with mice in locomotion. Very reccently, Zhang et al. (2016) found that both auditory and somatosensory cortices of mice are connected to motor cortex, whereas visual cortex is predominantly connected to anterior cingulate cortex cortex, which again highlights different brain circuits in top-down modulation of sensory processing. However, it is also possible that the brain is equipped with many different efference copy circuits (with a motor origin), which are not related to each other. These questions await further studies.

3.4.2 Pre-stimulus predictors of SA

We tested the hypothesis that modulation of low-frequency auditory oscillations are involved in the implementation of SA in the auditory domain. Indeed, our results demonstrate an enhancement of auditory alpha oscillations (~10Hz) for the active condition compared to the passive condition before stimulus onset. Low frequency oscillations particularly in the alpha range have previously been related to the suppression of task-irrelevant stimuli or stimulus features (Foxe & Snyder, 2011; Jensen & Mazaheri, 2010) in sensory systems such as visual cortex (Capilla, Schoffelen, Paterson, Thut, & Gross, 2012), auditory cortex (Weisz et al., 2011) and somatosensory cortex (Baumgarten, Schnitzler, & Lange, 2014; Haegens, Händel, & Jensen, 2011).

Therefore the upregulation of alpha oscillations is a viable mechanism for suppression of stimulus evoked activity in auditory cortex when stimulus presentation is self-generated. Indeed, our results are compatible with recent reports of enhanced alpha oscillations for self-uttered sound (Müller et al., 2014) and self-initiated visual stimuli (Stenner, Bauer, Haggard, et al., 2014). Interestingly, our analysis revealed a significant correlation between alpha power changes and SA, i.e., increased pre-stimulus alpha power changes were associated with increased SA (i.e., more attenuated M100 response) across participants. This provides further evidence for a close relationship between pre-stimulus alpha power modulation and SA and is consistent with reports that pre-stimulus alpha power correlates with early evoked responses (Ploner, Gross, Timmermann, Pollok, & Schnitzler, 2006). This finding is also consistent with results from a recent study that modulation of alpha power reflects the precision of predictions about upcoming stimuli (Bauer et al., 2014). This suggests that individual differences in the ability to predict the sensory consequences of one’s actions are expressed in differences in the modulation of alpha power.

Overall, the significantly increased alpha power for the active compared to the passive condition speaks in favour of an active inhibition of auditory areas at the time of motor preparation as a result of top-down mediated predictions in anticipation of the self-generated (predicted) sensory stimulus.

3.4.3 Post-stimulus representations of SA

The sound evoked response is characterized by an increase of both, oscillatory power and phase locking that is strongest in the theta (4-7 Hz) band (Figure 3.5 a & b). However, this activity is not modulated by the experimental conditions (active vs. passive). Instead, we show here that SA was associated with a significant decrease of power and phase locking in auditory cortex at higher frequencies. This suggests that the mechanisms responsible for

SA spare the low-frequency theta component and instead modulate alpha/beta and gamma components. A possible interpretation is that the low-frequency theta component reflects the physical stimulus properties that are unchanged between active and passive conditions whereas higher frequency components reflect more subjective properties of the stimulus that are discussed below in more detail (Gross, Schnitzler, Timmermann, & Ploner, 2007; Iversen, Repp, & Patel, 2009). The fact that SA was associated with a transient broadband power decrease in alpha and beta frequency bands up to almost 40 Hz is in line with a recent study that looked at the top-down modulation of brain responses to simple auditory rhythms (Iversen et al., 2009). Strongest top-down effects were observed in the beta range (while the alpha band was not studied).

Interestingly, high gamma band power was also reduced in the active condition compared to the passive condition, which is consistent with intracranial recordings from patients (Flinker et al., 2010). Animal studies have demonstrated that gamma rhythms are most prominent in supragranular layers suggesting that they mediate feedforward processes (Bastos et al., 2015; Michalareas et al., 2016; van Kerkoerle et al., 2014). In the framework of predictive coding, gamma band activity was suggested to reflect prediction errors for feedforward information processing (Arnal & Giraud, 2012; Bauer et al., 2014; Friston et al., 2015). A predicted stimulus (in active condition) is associated with smaller prediction errors thus leading to reduced gamma power.

Another significant difference between active and passive condition shortly after stimulus onset emerged from the phase locking analysis. Alpha/beta phase locking was weaker for active compared to passive condition. This difference is caused by a higher variability of single-trial phase across trials in the active condition. There are different possible explanations for this effect. It could be a byproduct of the reduced alpha/beta post-stimulus power for the active condition. In this scenario the reduced signal-to-noise ratio (SNR) due to the reduced power leads to an artificially reduced (less precise) estimate of single-trial phase. However, this scenario is unlikely for three reasons. First, post-stimulus power is also significantly reduced for high beta and gamma frequency bands without a difference in phase locking. Second, the correlation of post-stimulus gamma modulation to post-stimulus phase locking speaks against a simple SNR-induced effect, especially because post-stimulus gamma modulation correlates with alpha/beta phase but not alpha/beta power (Supplementary Figure S3.7 c & d). Third, the single-trial correlation of gamma power and alpha/beta phase favours a different interpretation where alpha/beta phase is causally linked to gamma amplitude.

3.4.4 Implications for neural information processing

Both, single-trial analysis and statistical contrasts between conditions revealed a functional neural information processing sequence among pre-stimulus alpha power, post-stimulus gamma power and post-stimulus alpha/beta phase. Prior to the stimulus onset, alpha power controls the gain of local neuronal populations reflecting the precision of the prediction about the incoming stimulus. Mechanistically, this may be implemented by modulating local neuronal excitability levels, known to be indexed by alpha activity (Romei et al., 2008). The pre-stimulus alpha power in the passive jittered condition fluctuated from trial to trial creating differential levels of precision over the incoming stimulus, with low alpha power corresponding to high levels of precision. When the stimulus arrives, any incongruency between the prediction and the actual incoming stimulus (prediction error) is fed forward for further processing through gamma oscillations. Since prediction error is weighted by precision, a negative correlation between pre-stimulus alpha power and post-stimulus gamma power is predicted. This is exactly what we observed (Figure 3.13 a). Next, brain areas from the higher hierarchy processing prediction errors provide feedback to the lower hierarchy through alpha/beta oscillations, which is captured by the significant correlation between post-stimulus gamma power and post-stimulus alpha/beta phase deviation (Figure 3.13 b). The idea of alpha/beta phase acting as top-down signals to resolve the bottom-up prediction error has received support from previous auditory studies (Arnal, Wyart, & Giraud, 2011; Fontolan, Morillon, Liegeois-Chauvel, & Giraud, 2014). In a recent study, Fontolan et al. (2014) showed that gamma power in A1 was modulated by alpha/beta phase in auditory association cortex suggesting the top-down origin of the latter. Our analysis provides further evidence for this by showing that an early gamma power led to a late alpha/beta phase resetting. This temporal asymmetry is an important step for establishing a causal role of gamma power in resetting alpha/beta phase. All these results fit very well with recent findings suggesting that high frequency band oscillation (e.g., gamma) relays feedforward information and that low frequency band (e.g., alpha and beta) oscillation relays feedback information (Bastos et al., 2015; Michalareas et al., 2016; van Kerkoerle et al., 2014).

3.4.5 Information flow between motor cortex and auditory cortex

Many empirical studies support the view that motor cortex is involved in generating efference copy signal (Christensen et al., 2007; Ford et al., 2013; Mock, Foundas, & Golob, 2011; Voss, Bays, Rothwell, & Wolpert, 2007). However, fMRI studies (e.g., Christensen et al. (2007)) or EEG studies based on correlations between evoked components (e.g., Ford et al. (2013)) can not reveal the temporal course of information exchange between relevant brain areas with a fine temporal resolution. Here, we tried to overcome this problem by using Granger causality. Granger causality results did show increased information flow from motor cortex to both left and right auditory cortices in both active and motor only condition (Figure 3.15). In addition, increased information flow from auditory cortex to motor cortex was found in all conditions except for motor only condition. However, most of the Granger causality results did not pass statistical tests. One possibility is that all the Granger causality results shown earlier are spurious as they merely reflect the fact that there are evoked responses in both motor and auditory cortices. Evoked responses from different brain areas that are close in time will lead to an increase in Granger causality measures, even if there are no true information exchanges between different brain areas. This account may fit with the results observed in active condition, but it cannot easily explain the results observed in other conditions. In motor only condition, obvious evoked responses were found in motor cortex but not in auditory cortices (Suppplementary Figure S3.11 b). However, increased information flow from motor cortex to auditory cortices was found with the information flow from motor cortex to left auditory cortex exceeding statistical significance cut-off. Similarly, in the passive listening conditions, there were obvious evoked responses in auditory cortices but not in motor cortex (Supplementary Figure Figure S3.11 c & d). However, increased information flow from auditory cortices to motor cortex was still found with the information flow from left auditory cortex to motor cortex in passive periodic condition being significant. Therefore, a more reasonable explanation is that the current analysis reveals true connectivity between motor and auditory cortices to some extent, but not enough to reach statistical significance.

3.4.6 Correlations with questionnaire scores

3.4.6.1 Prediction and schizophrenic/autistic traits

The correlation between delusional states and SA has been reported by several studies using different paradigms of SA (Heinks-Maldonado et al., 2007; Shergill et al., 2014; Teufel et al., 2010). Here, we have evidence that partially support this correlation from another SA measurement paradigm. The correlation between PDI scores and SA is only significant when using Pearson correlation, which is not ideal for this study with a relatively small sample size. However, due to the robustness of this correlation found in the literature, we tend to believe that the correlation here is not spurious and should still persist if more participants are tested. Large SA was associated with low PDI scores (low tendency to delusions). Interestingly, we also found that PDI scores are correlated with the amplitude of M100 component in active condition where the efference copy related prediction is present, but not in passive listening conditions.

For the first time we showed a significant correlation between SA and SQ scores, which measures the systemizing traits related to autism. Large SA effects are associated with high SQ scores (strong autistic traits). This suggests a link between SA effect and autism. Importantly, the source of this correlation seems to be that SQ scores are correlated with the amplitude of M100 component in passive periodic condition, but not in active condition.

Our correlation results suggest that both schizotypal traits and autistic traits are reflected in SA, but schizotypal traits lead to reduced SA while autistic traits lead to enhanced SA. Reduced SA among schizophrenics was suggested to be due to a failure in prediction (Ford et al., 2013; Shergill et al., 2005). When voluntarily generating a sound (in active condition), one also predicts the sound effect. Correctly predicted stimuli lead to reduced neural responses. Thus a failure of prediction prevents reduced neural responses in active action leading to reduced SA. A significant positive correlation between PDI score and M100 amplitude in active condition supports this point. For SQ score, however, it is positively correlated with M100 amplitude in passive periodic condition (Although SA score is also positively correlate with the amplitude of M100 component in passive jittered condition, the correlation coefficient is smaller; see Table 3.1). Note that in passive periodic condition the stimulus is presented once every three seconds, so it is predictable. One possible explanation is that autistic traits are associated with a deficiency in the prediction of periodic stimulus, which poses a difficulty to suppress the M100 component. So high SQ scores are associated with large M100 amplitude in passive condition. SQ measures the tendency to analyse a system with a mechanical view, and it has been shown to be correlated with local-focus attention in the visual domain (Billington, Baron-Cohen, & Bor, 2008; Walter, Dassonville, & Bochsler, 2009). This local focus processing feature may also apply to auditory domain, in which people with high autistic traits are less likely to detect the global feature of the stimulus train (in our case, the global feature is the temporal periodicity of the stimulus).

In principle, the correlation results support the Bayesian approach of understanding sensory anomaly in schizophrenia and autism (Fletcher & Frith, 2009; Pellicano & Burr, 2012). A deficiency with priors (or generating priors) can possibly explain the correlations found here. The analysis showed an opposite effect on SA from schizotypal traits and autistic traits. This opposite effect can be explained under the same Bayesian model, with a prior deficiency in active condition (predicting self-generated stimulus) for schizophrenia and a prior deficiency in passive condition (predicting externally generated periodic stimulus) for autism. This may be helpful for understanding the relationship between schizophrenia and autism (Craddock & Owen, 2010) and also points out new directions for studying aberrant perception under certain mental disorders. That is, the deficiency may lie under forming correct priors in certain situations.

3.4.6.2 Beta rebound and PDI score

Another interesting feature from this dataset is the significant correlation between PDI scores and beta rebound. As one of the most robust electromagnetic signals measured from the scalp, beta rebound is still not well understood for its functions. Some recent progress suggested that this sensorimotor originated rhythm might be related to the forward model, by representing the reliability of the forward model (Tan, Wade, & Brown, 2016) or participating in the forward model updating when the reafference is mispredicted (Cao & Hu, 2016). Although the exact function of beta rebound is still under debate, the correlation reported here between PDI scores and beta rebound echoes the latest advance in beta rebound study, as the delusional state measured by PDI scores was suggested to be a consequence of dysfunctional forward model (Heinks-Maldonado et al., 2007; Shergill et al., 2014). Interestingly, the correlation is not only significant in active condition, but also in motor only condition where no sound stimulus was played. In motor only condition, it might be the somatosensory input (participant’s finger touching the sensor) that acts as the reafference to the forward model. Large PDI scores are associated with high beta rebound, suggesting a negative effect of high beta rebound on normal brain functioning. However, one study showed reduced beta rebound in schizophrenic patients (Robson et al., 2015), which seems to be at odds with the correlation found here among normal participants. Despite this controversy, a clear point is emerging that beta rebound has a close relationship with the forward model and schizophrenia. Given the accidental observation of the interesting correlation and the small sample size in the present study, we look forward to further studies investigating the role of beta rebound in the forward model/mental diseases such as schizophrenia.

3.5 Conclusion

In summary, the present study replicates the classic SA effect in the auditory domain and fails to find SA or sensory enhancement in the visual domain. Subsequent analysis in the source space focusing on auditory cortex reveals an involvement of low-frequency oscillations in mediating SA, which is consistent with a predictive coding account of SA that rests on auditory oscillations for gain control of sensory evidence. The results also corroborate recent findings by providing evidence for hierarchical information processing in the brain mediated by gamma (bottom-up) and alpha/beta (top-down) oscillations. Granger causality analysis shows promising effects of information flow between motor cortex and auditory cortex, but further studies are needed to confirm the validity of the effects. Correlation analysis with questionnaire scores corroborates the tight link between SA and delusional ideation and also reveals interesting correlations (between SA and systemizing quotients, between beta rebound and delusional ideation) that call for further elaboration.

4 The cerebellum is involved in learning delayed action effects

4.1 Introduction

Self-generated stimuli are ubiquitous in everyday life. When I am typing these words, every strike on the keyboard generates, at least, predictable visual (i.e., the character), somatosensory (i.e., the fingertip tap) and auditory (i.e., the keyboard click) inputs. The sensory input resulting from actions elicits smaller brain responses as compared to the same sensory input when it is externally generated, a phenomenon known as sensory attenuation (SA) effect (Hughes et al., 2013b; Martikainen et al., 2005; Schafer & Marcus, 1973). For example, the amplitude of evoked responses in auditory cortex is smaller for self-generated tones than external tones in electrophysiological recordings (Martikainen et al., 2005; Schafer & Marcus, 1973). When a perturbation is introduced between actions and the sensory input (e.g., there is a delay between actions and the onset of stimuli), SA effect is reduced or even abolished in the case of a large perturbation (Aliu et al., 2009; Blakemore, Frith, et al., 1999). Yet, the brain can adapt to small perturbations if they are consistent so that SA effect re-emerges after learning. Aliu et al. (2009) showed that when a tone was delivered after a button press action with a 100 ms delay, auditory SA was initially absent but emerged again within 300 trials.

SA is explained as a result of prediction for self-generated stimuli, which is usually formulated in the context of the forward model (Schröger et al., 2015; Wolpert & Ghahramani, 2000). The forward model posits that predictions for self-generated stimuli are formed along with the motor command that will trigger these stimuli. When the signal of self-generated stimuli reaches the brain, it is compared to the prediction and SA will be observed if the real sensory input and the prediction match (Blakemore et al., 1998). With this in mind, the process of SA re-emerging after perturbation can be seen as a process of correcting previous predictions to account for the perturbation or the updating of the outdated forward model.

The cerebellum has been suggested to have a role in predicting self-generated stimuli and this received supporting evidence from functional magnetic resonance imaging (fMRI) studies using the SA paradigm in the somatosensory domain (Blakemore et al., 1998;

Blakemore, Wolpert, & Frith, 1999). Recently, the involvement of the cerebellum in SA received further support from studies with focal cerebellar lesion patients, with whom the classic auditory SA effect cannot be replicated (Knolle et al., 2012; Knolle et al., 2013). Moreover, there is evidence showing that the cerebellum is vital in updating the outdated predictions (or forward models). Specifically, cerebellar lesion patients were found to show deficits in predicting the position of self-controlled cursor on the screen when a discrepancy was introduced between the real cursor position and controlling movement (Izawa et al., 2012; Synofzik et al., 2008).

Taken together, fMRI and lesion studies emphasise the importance of the cerebellum for predicting sensory consequences of our actions. But very little is known about how the cerebellum may be involved in the forward model related processes. Here we introduced a delay between actions and action triggered tones to investigate the process of forward model updating using an auditory SA paradigm with magnetoencephalography (MEG) recordings. By transiently perturbing the normal cerebellar function using transcranial magnetic stimulation (TMS; with an inhibitory offline 1Hz-rTMS protocol), we interrogated how interfering with cerebellar function might affect the updating of forward model in healthy young participants.

4.2 Methods

4.2.1 Participants

10 healthy, right-handed volunteers (including LC; 5 males; mean age = 23.0, SD = 2.7) were recruited from a local participants’ pool. Participants gave written informed consent prior to the experiment and received monetary compensation after the experiment. The study was approved by the local ethics committee (Ethics Committee of College of Science and Engineering, University of Glasgow) and was conducted in accordance with the Declaration of Helsinki.

4.2.2 Equipment and Procedure

A double-cone coil TMS system (The Magstim Company, Whitland, Wales) was used for stimulating the cerebellum. This type of coil was used as a recent study showed that compared to other coil types (figure-of-eight and batwings), the double-cone coil was most effective for cerebellar stimulation (Hardwick, Lesage, & Miall, 2014). A 248-magnetometers whole-head MEG system (MAGNES 3600 WH, 4-D Neuroimaging) was used for data recording with a sampling rate of 1,017Hz.

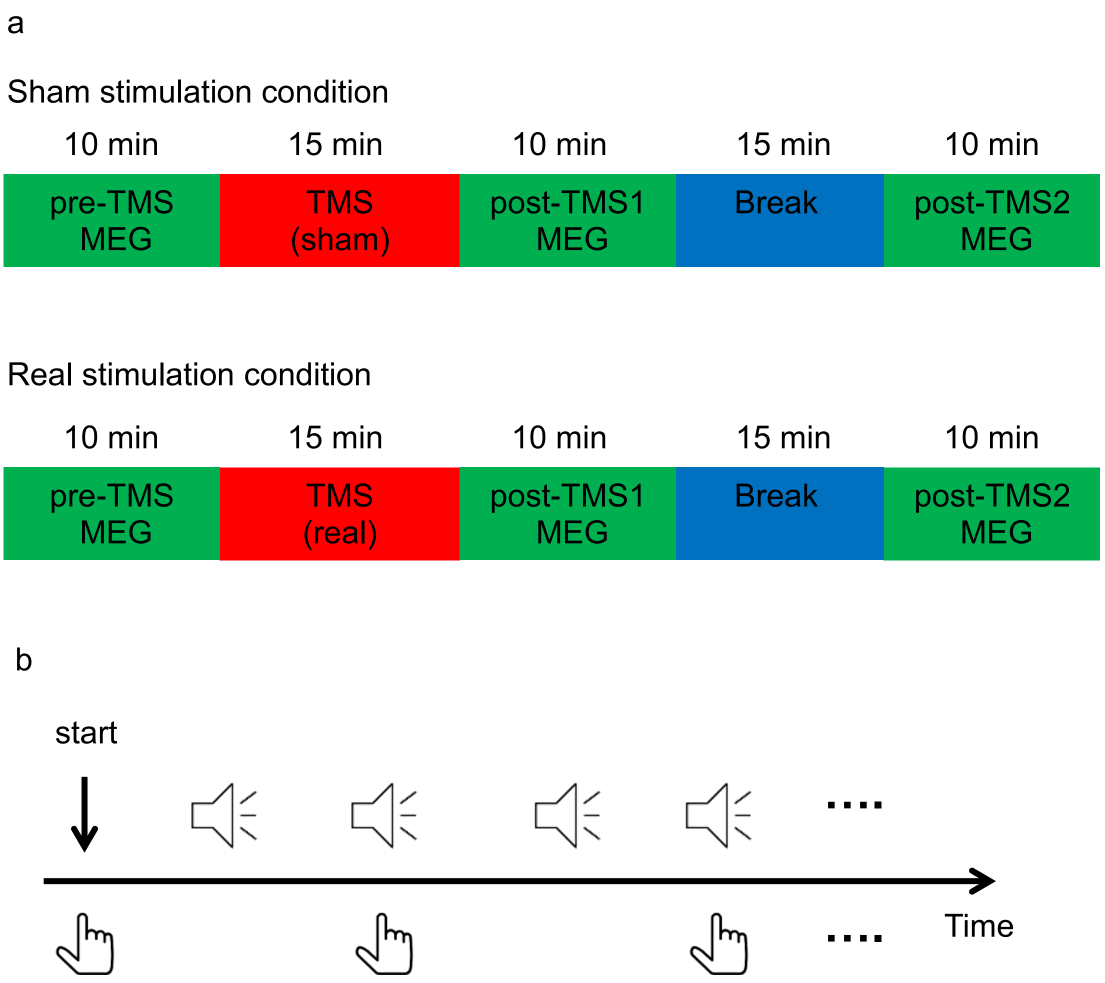


Figure 4.1 Testing procedure. (a) shows the testing procedure for both real and sham stimulation conditions (with at least 6 days apart). (b) shows the MEG testing paradigm, which is followed in all the MEG testing sessions. A finger lift marks the beginning of testing, which is followed by an external tone. After hearing the external tone, participants make a finger lift to trigger a self-generated tone. This procedure goes on until 200 tones are played. The inter-tone interval is around 3 seconds (see text for detail).

Each participant was tested in a real stimulation and a sham stimulation condition on different days (at least 6 days apart). In each stimulation condition (Figure 4.1 a), SA effect was measured three times using the same procedure (see below) and the stimulation was performed after the first SA measurement. After the first SA measurement (pre-TMS), participants were encouraged to take a break, after which the TMS stimulation was performed outside the magnetically shielded room. The second SA measurement (post-TMS1) started just after the stimulation. The delay between the end of TMS and the start of second SA measurement was comparable (t(9) = -0.97, *p* = 0.36) between the real (mean = 2.60; SD = 0.38; in minutes) and sham stimulation condition (mean = 2.92; SD = 1.09; in minutes). The delay between the end of the second SA measurement and the start of the third SA measurement (post-TMS2) was always about 15 minutes (about 25 minutes after TMS stimulation) for each condition and each participant.

For stimulation, a 15-minute inhibitory offline 1 Hz repetitive stimulation protocol (1Hz-rTMS) was used. The right cerebellum was stimulated in the real stimulation condition. The stimulation location was determined individually to make sure that it was right above the cerebellum and relatively away from the occipital cortex. The scalp point that is 1cm below and 3cm to the right of the Inion was first identified, which was then checked visually with Brainsight neuronavigation system (Rogue Research). In Briansight neuronavigation system, the selected point can be superimposed onto individual’s structural MRI (obtained at least 6 days before any MEG recordings) by placing the tip of a stylus pen right above the selected point. The tip of the stylus pen was coregistered with the MRI in the same coordinate system. The final stimulation point was moved down 0.5 cm from the initially selected point if the latter was close to the occipital cortex. The final stimulation point was determined prior to the first SA measurement. The stimulation intensity was also determined individually (mean intensity = 46.5% of maximal power output; ranging from 35% to 50%). This was because the stimulation point was close to neck muscles and very strong facial twitches can be induced in some participants if 50% stimulation intensity was applied. The stimulation intensity was determined so that participants did not feel any pain and were confident that they can receive the stimulation continuously for 15 minutes. In sham stimulation condition, the coil was placed right above the inion with the wings facing against the head. The stimulation intensity was always 50% of maximal power output. Three participants received real stimulation first and the remaining received sham stimulation first.

To measure SA effect, 100 self-generated tones and 100 external tones (computer-controlled) were presented in alternating order one by one in the same SA testing block (120 trials were used for LC, who was the first participant) (Baess et al., 2011). Both tones were 1000 Hz, 100 ms in duration and were set to be at a comfortable volume level. Tones were delivered through a plastic ear tube. The testing block started with an external tone, and then participants waited about 3 seconds to make a self-generated tone by briskly lifting their right index finger (Figure 4.1 b). The finger lift was detected by a laser sensor which served the function of a response box. Using the laser sensor has the advantage to avoid the noise associated with normal response box so that the only auditory input was the tone. Critically, a delay (mean = 92.0 ms; SD = 4.3 ms) was introduced between the finger lift and the tone output from the ear tube. After the self-generated tone, the next external tone was presented after a random interval between 1500 ms and 4500 ms. Then it was another self-generated tone and so on until 200 tones were played. Participants received a few trials of practice before the first SA measurement to get familiarized with the paradigm. Participants were asked to close their eyes during the SA measurement.

4.2.3 Data analysis

Data analysis was performed with Matlab using FieldTrip toolbox (Oostenveld et al., 2011) and in-house scripts conforming to the recent MEG data analysis guidelines (Gross et al., 2013). MEG signals in all testing sessions were high-pass filtered at 0.75 Hz to mitigate TMS induced artefacts. Trials with very short inter-trial intervals (less than 1500 ms for self-generated tones) were first discarded. Then very noisy trials and channels from visual inspection were rejected with ft_rejectvisual, followed by the denoise procedure using ft_denoise_pca. Rejected bad channels were repaired with interpolation methods using ft_channelrepair. Then MEG signals were visually inspected again and noisy trials were discarded. Eye movement and heart artefacts were rejected using ICA. After this step, 98.2 (SD = 6.8) trials and 98.7 (SD = 6.7) trials were left for self-generated tones and external tones, respectively.

4.2.3.1 Evoked responses

MEG signals were low-pass filtered with 40 Hz cut-off frequency. Event related fields aligned to the tone onset were computed for each testing session with baseline (-600 to -200 ms) correction. The M100 component was defined in a post-stimulus time window between 70 and 100 ms and M200 component was between 125 and 155 ms. To test the existence of the M100/M200 suppression effect for self-generated tones, paired t-tests were performed on the amplitudes of evoked responses between self-generated tones and external tones in the pre-TMS session of both stimulation conditions. Topographies of M100 and M200 components were illustrated with the evoked responses of external tones in the pre-TMS session.

4.2.3.2 ANOVA analysis

Changes of M100 suppression effect over testing sessions were tested with within-subject ANOVA analysis (implemented in Fieldtrip with ft_statfun_depsamplesFunivariate), separately for sham and real stimulation condition. Two significant auditory sensors with largest F values that best represent the M100 activation from this analysis were selected for a similar within-subject ANOVA analysis which was performed with the average data of the two selected sensors at each time point from 0 to 400 ms. This time point by time point ANOVA analysis can reveal whether the detected changes over testing sessions were specific to M100 suppression effect. To test effects associated with stimulation conditions (sham vs. real), a 2 (sham vs. real stimulation) by 2 (post-TMS1 vs. post-TMS2) within-subject ANOVA analysis was performed with baseline referenced M100 suppression effects, which was obtained by subtracting the pre-TMS M100 suppression effect from post-TMS M100 suppression effects. A paired t-test against 0 was also performed for each of the baseline referenced M100 suppression effects. Lastly, amplitudes of M100 components for self-generated and external tones in sham stimulation condition, where significant M100 suppression changes were found, were subjected to a 2 (self-generated vs. external) by 3 (testing sessions) within-subject ANOVA analysis. The M200 suppression effect related analysis was performed similarly as did with M100 suppression effect. See results section for details.

4.2.3.3 Correlation analysis

For each participant, the M100 (M200) suppression effect was averaged across all the 6 testing sessions to get an average M100 (M200) suppression effect. A Spearman correlation analysis was performed between the average M100 suppression effect and the average M200 suppression effect. To get a robust estimation of the confidence interval (CI) of the correlation, possible outliers were first identified with the median absolute deviation-median rule (Hall & Welsh, 1985) and then a bootstrapping approach (resampling with replacement; 1000 repetitions) was applied to the remaining data points (7 remained) with Spearman correlation. Values between 2.5% and 97.5% percentile of the resulting 1000 Spearman’s rho values were taken as the 95% CI of the correlation. A similar analysis was performed between the average magnitude of M100 component (averaged across both testing tones in all testing sessions) and the average magnitude of M200 component (averaged across both testing tones in all testing sessions). The robust correlation analysis was performed with the robust correlation Matlab toolbox (Pernet et al., 2012).

4.3 Results

4.3.1 Replication of M100 suppression and M200 suppression effects

In EEG SA studies, both N100 component suppression and P200 component suppression have been well documented (Schafer & Marcus, 1973; Schröger et al., 2015; Timm, Schönwiesner, Schröger, & SanMiguel, 2016). Therefore, here we focused on their MEG counterparts M100 and M200 components. Significant M100 suppression and M200 suppression effects were identified in the pre-TMS session of both stimulation conditions. Both effects were replicated at or close to sensors showing strong evoked responses. For example, sensors showing significant M100 suppression effect in the pre-TMS session of sham stimulation condition and sensors showing significant M200 suppression effect in the pre-TMS session of real stimulation condition are marked with ‘+’ in Figure 4.2 a and Figure 4.2 d, respectively, superimposed on the topography of evoked responses in the related testing session.

4.3.2 M100 suppression effect increased over testing sessions in sham stimulation condition

The within-subject ANOVA analysis of M100 suppression effect revealed a significant change in sham stimulation condition (Figure 4.2 b), but not in the real stimulation condition (Figure 4.2 c) in auditory sensors. Interestingly, the two sensors (Figure 4.2 a, highlighted with ‘o’) showing the strongest modulation effect in sham stimulation condition were not among the sensors where significant M100 suppression effects were found in pre-TMS session. Selecting those two sensors, the same within-subject ANOVA analysis performed at each time point within the first 400 ms of stimulus onset showed that the modulation effect peaked at the time point (86 ms) within the M100 component time window (Figure 4.3 a).

Since we are interested in the effects of TMS on M100 suppression effect and no significant difference (paired t-test, t(9) = 0.13, *p* = 0.90) was found in the pre-TMS M100 suppression effect between sham (mean = 3.53e-14; SD = 5.42e-14) and real stimulation condition (mean = 3.34e-14; SD = 5.81e-14) (Figure 4.2 a & d), the post-TMS M100 suppression effects were referenced to the pre-TMS M100 suppression effect (through subtraction) before being subjected to a 2 (sham vs. real stimulation) by 2 (post-TMS1 vs. post-TMS2) within-subject ANOVA analysis. This analysis showed an increase of pre-TMS referenced M100 suppression effect (F(1,9) = 7.21, *p* = 0.03) from post-TMS1 (mean = -1.11e-14; SD = 2.81e-14) to post-TMS2 (mean = 2.36e-14; SD = 3.80e-14). Importantly, there was a significant interaction effect between stimulation condition and testing session (F(1,9) = 7.71, *p* = 0.02). Post-hoc analysis showed that there was a significant increase of M100 suppression effect (t(9) = 3.31, *p* = 0.01) from post-TMS1 (mean = -2.31e-14; SD = 5.35e-14) to post-TMS2 (mean = 3.11e-14; SD = 5.25e-14) in sham stimulation condition, but not in real stimulation condition (post-TMS1: mean = 8.64e-16; SD = 5.69e-14; post-TMS2: mean = 1.61e-14; SD = 5.85e-14; t(9) = -1.19, *p* = 0.27). The M100 suppression effect in post-TMS2 session was higher than the average of M100 suppression effects in pre-TMS and post-TMS1 sessions in 8 out of the 10 tested participants (Figure 4.3 b). None of the pre-TMS referenced post-TMS M100 suppression effects differed from 0 (post-TMS1 in sham condition: t(9) = -1.36, *p* = 0.21; post-TMS2 in sham condition: t(9) = 1.88, *p* = 0.09; post-TMS1 in real condition: t(9) = 0.05, *p* = 0.96; post-TMS2 in real condition: t(9) = 0.87, *p* = 0.41) and no other tested effects reached statistical significance (main effect of stimulation condition: F(1,9) = 0.03, *p* = 0.87).

To find the source of the increased M100 suppression effect in post-TMS2 session of sham stimulation condition (an amplitude decrease for self-generated tones vs. an amplitude increase for external tones), another 2 (self-generated vs. external) by 3 (testing sessions) within-subject ANOVA analysis was performed on the amplitude of M100 components in sham stimulation condition and it revealed a significant interaction effect (F(2,18) = 5.35, *p* = 0.02). Post-hoc analysis showed two significant simple effects, with the amplitude of M100 component for self-generated tones in post-TMS2 session being smaller than the amplitudes of M100 component for external tones in the same testing session (t(9) = 2.57, *p* = 0.03) and for self-generated tones in post-TMS1 session (t(9) = 2.44, *p* =0.04). The first simple effect indicated the significant M100 suppression effect in post-TMS2 session. The latter simple effect revealed that the increased M100 suppression effect in post-TMS2 session was due to a decrease of M100 amplitudes for self-generated tones. No other effects from the ANOVA analysis reached statistical significance (main effect of tones:

F(1,9) = 3.77, p = 0.08; main effect of testing sessions: F(2,18) = 0.26, p = 0.78). See Figure 4.4 for an illustration of M100 evoked responses in all testing sessions.

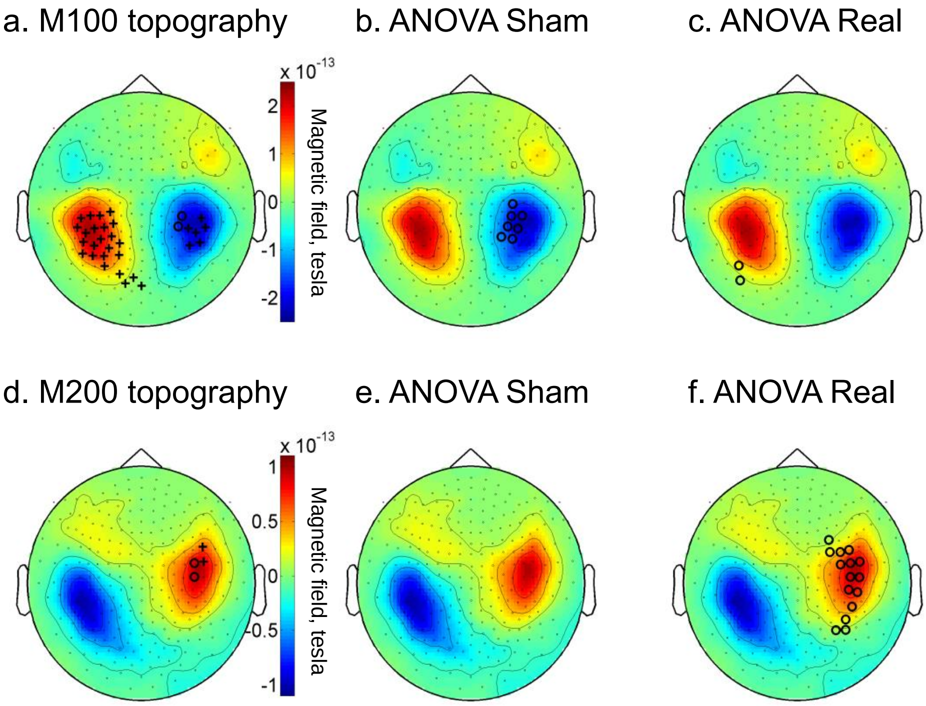
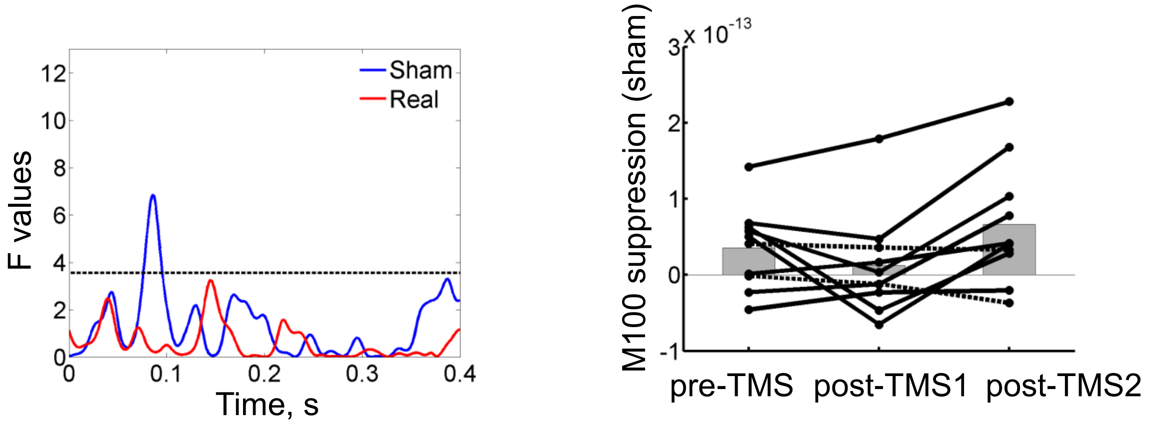
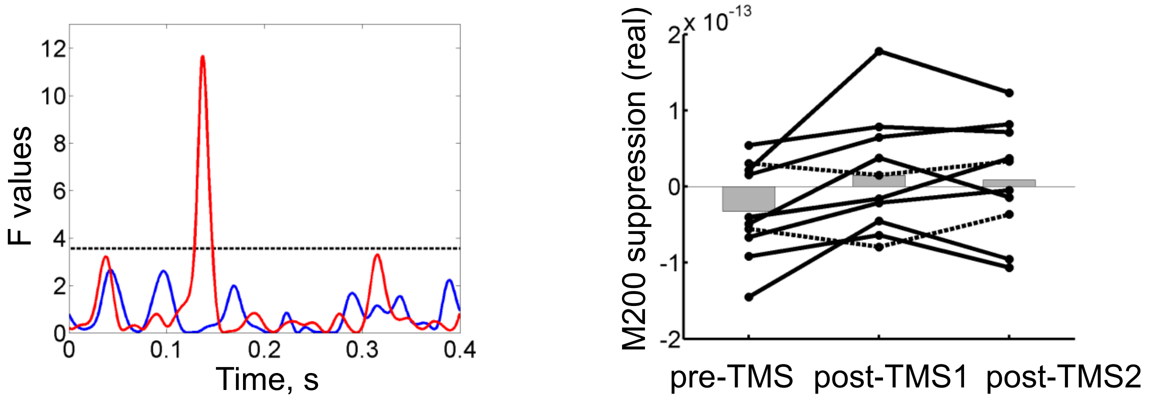


Figure 4.2 Topographies of evoked responses and results of ANOVA analyses. (a) shows the topography of the M100 component activation (70-100 ms; grand average of the external tone evoked responses in the pre-TMS session in real stimulation condition). Sensors marked with '+' show significant M100 suppression effects ($p < 0.05$ with paired t-tests; no correction for multiple comparisons). Sensors marked with 'o' show largest F values in the ANOVA analysis in (b) and are selected for the analysis shown in Figure 4.3 a. (b) and (c) show the significant sensors (marked with 'o'; without correction for multiple comparisons) from the within-subject ANOVA analysis of M100 suppression over testing sessions, in the sham and real stimulation condition, respectively. Significant sensors are overlaid on the M100 topography. (d) shows the topography of the M200 activation (125-155 ms; grand average of the external tone evoked responses in the pre-TMS session in real stimulation condition). Sensors marked with '+' show significant M200 effects ($p <0.05$ with paired t-tests; no correction for multiple comparisons). Sensors marked with 'o' show largest F values in the ANOVA analysis in (e) and are selected for the analysis shown in Figure 4.3 c. (e) and (f) show the significant sensors (marked with 'o'; without correction for multiple comparisons) from the within-subject ANOVA analysis of M200 suppression over testing sessions, in the sham and real stimulation condition, respectively. Significant sensors are overlaid on the M200 topography. In all figures, only the significant sensors that are part of a cluster with at least two sensors are shown.



a. F values of M100 suppression

b. M100 suppression in sham condition



c. F values of M200 suppression

d. M200 suppression in real condition

Figure 4.3 Results of the within-subject ANOVA analysis over time for selected sensors (a & c) and individual effects (b & d). (a) shows the F values from the within-subject ANOVA analysis of M100 suppression effect in sham (blue) and real (red) stimulation condition. There is a clear peak at 86 ms in the sham stimulation condition. The dashed line (also in c) shows the cut-off F value for statistical significance of the ANOVA analysis (α = 0.05; no correction for multiple comparisons). (b) shows the changes of M100 suppression effects over testing sessions in the sham stimulation condition for all participants. Each line represents one participant. In 8 out of 10 participants (solid lines), the M100 suppression effect in the post-TMS2 session is higher than the average M100 suppression effect of the first two testing sessions. (c) and (d) show the same analysis done with sensors showing strong M200 suppression modulation in real stimulation condition. There is a clear peak of F values at 137 ms in the real stimulation condition and in 8 out of 10 participants (solid lines), the M200 suppression effect in the pre-TMS session is higher than the average of the two post-TMS testing sessions.

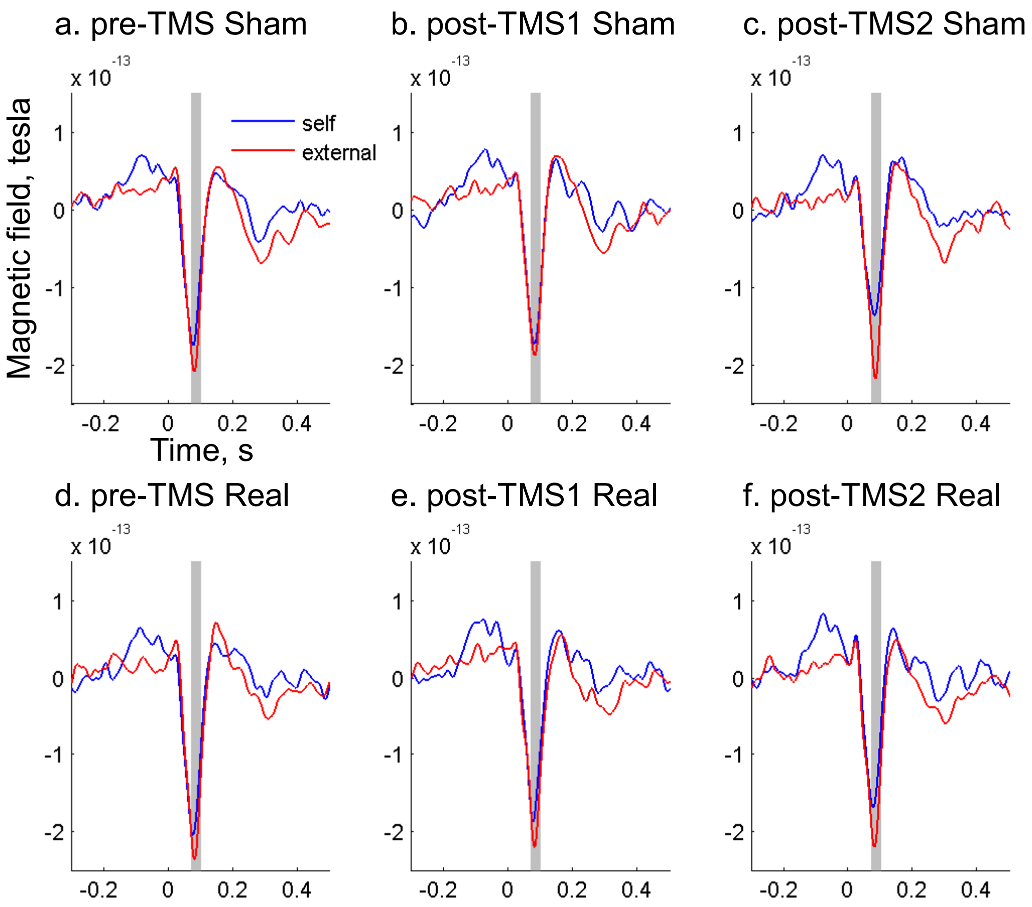


Figure 4.4 Illustration of M100 components in all testing sessions. In sham stimulation condition, there is an increase of M100 suppression effect in post-TMS2 session (c) from pre-TMS session (a) and post-TMS1 session (b). In the real stimulation condition, the M100 suppression effect stays comparable over testing sessions (d-f). The shaded area shows the post-stimulus time window (70-100 ms) selected for analysis. Evoked components are computed as the average of highlighted sensors (‘o’) in Figure 4.2 a.

4.3.3 M200 suppression pattern reversed after TMS

The within-subject ANOVA analysis of M200 suppression effects over testing sessions revealed a significant effect in real stimulation condition (Figure 4.2 f), but not in sham stimulation condition (Figure 4.2 e) in auditory sensors. Again, the two sensors showing the largest F values of the ANOVA analysis were not among the sensors (Figure 4.2 d, highlighted with ‘o’) showing significant M200 suppression effects in pre-TMS session of the real-stimulation condition. The within-subject ANOVA analysis over time with these two sensors showed a clear peak of F values at the time point (137 ms) within the M200 component time window (Figure 4.3 c) only in the real stimulation condition.

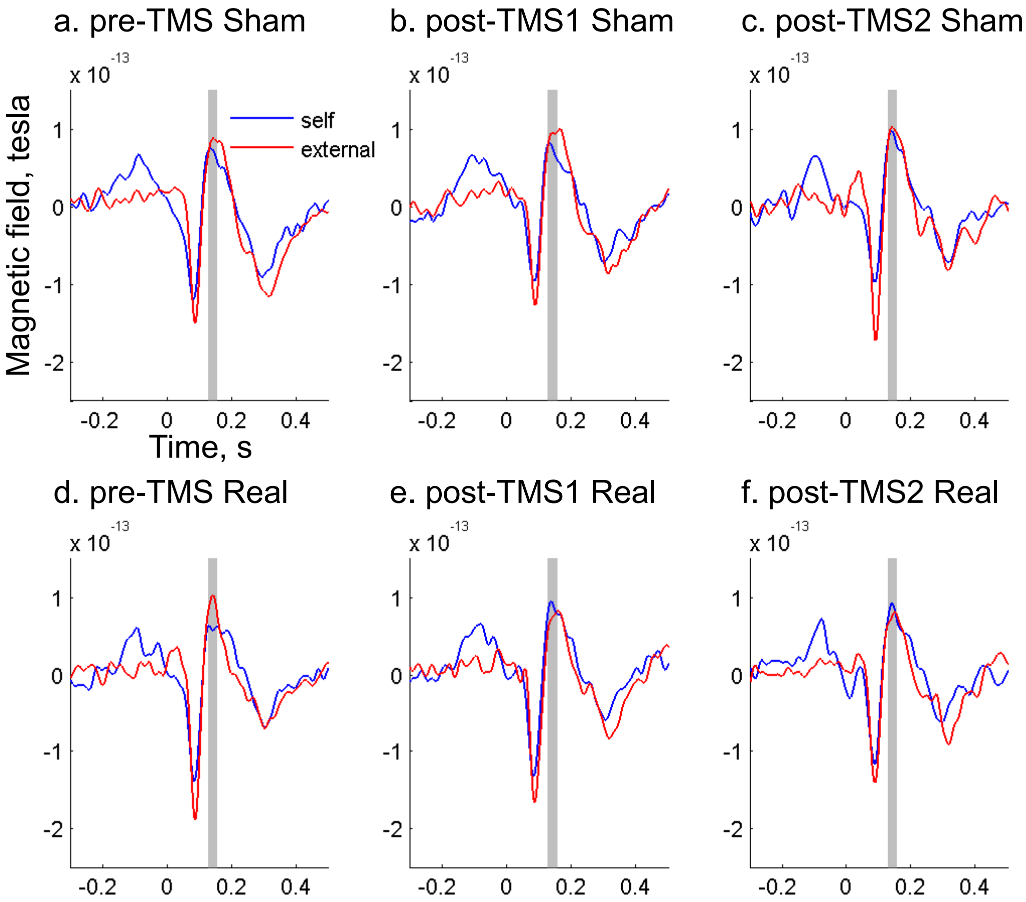


Figure 4.5 Illustration of M200 components in all testing sessions. In sham stimulation condition, M100 suppression effect stays comparable over testing sessions (a-b). In real stimulation condition, the M200 suppression effect pattern is observable in pre-TMS session (d), but it reverses in both post-TMS sessions (e & f), with the amplitude of M200 component being larger for self-generated tones than for external tones. The shaded area shows the post-stimulus time window (125-155 ms) selected for analysis. Evoked components are computed as the average of highlighted sensors (‘o’) in Figure 4.2 d.

Post-TMS M200 suppression effects were subjected to a 2 (sham vs. real stimulation) by 2 (post-TMS1 vs. post-TMS2) within-subject ANOVA analysis after being referenced to the pre-TMS M200 suppression effect (through subtraction), for which no significant differences were found between sham stimulation condition (mean = -1.17e-14; SD = 5.93e-14) and real stimulation condition (mean = -3.27e-14; SD = 6.23e-14; t(9) = 1.58, *p* = 0.15; Figure 4.5 a & b). This analysis revealed a significant main effect of stimulation condition (F(1,9) = 6.69, *p* = 0.03), with the pre-TMS referenced post-TMS M200 suppression effects being significantly smaller in real stimulation condition (mean = 4.45e-14; SD = 4.14e-14) than in sham stimulation condition (mean = -3.53e-16; SD = 3.05e-14). Actually, the post-TMS M200 suppression effects in real stimulation condition reversed signs and should now rather be called ‘M200 enhancement’, with both pre-TMS referenced post-TMS1 M200 suppression (t(9) = 2.76, *p* = 0.02) and post-TMS2 M200 suppression (t(9) = 3.63, *p* = 0.01) being significantly larger than 0, i.e., M200 component was higher for self-generated tones than external tones. The pre-TMS referenced post-TMS1 (t(9) = -0.35, *p* = 0.73) and post-TMS2 (t(9) = 0.48, *p* = 0.64) in sham condition did not differ from 0. No other effects from the ANOVA analysis reached statistical significance (main effect of sessions: F(1,9) = 0.07, *p* = 0.79; interaction effect: F(1,9) = 0.34, *p* = 0.58). 8 out of the 10 tested participants showed the M200 suppression effect changes after TMS (i.e., the mean of both post-TMS M200 suppression effects was smaller than the pre-TMS M200 suppression effect, Figure 4.3 d).

A 2 (self-generated vs. external) by 3 (testing sessions) within-subject ANOVA analysis of the amplitude of M200 components in real stimulation condition and it revealed a significant interaction effect (F(2,18) = 6.74, *p* = 0.01). However, no simple effects reached statistical significance in post-hoc analysis. No other effects from the ANVOA analysis reached statistical significance (main effect for tones: F(1,9) = 0.02, *p* = 0.89; main effect for testing sessions: F(2,18) = 0.02, *p* = 0.98). See Figure 4.5 for an illustration of M200 evoked responses in all testing sessions.

4.3.4 Correlation between M100 and M200 suppression effects

Having observed changes of M100 suppression effect in sham stimulation condition and changes of M200 suppression effect in real stimulation condition, here the relationship between the two was investigated with a correlation analysis. Across participants, the average M100 suppression effect across all testing sessions was significantly correlated with the average M200 suppression effect across all testing sessions (Figure 4.6; Spearman’s rho = 0.70, *p* = 0.03; Spearman skipped correlation 95% CI = [0.37 1]). This correlation cannot be simply explained by the autocorrelation of MEG data as the magnitude of M100 component and the magnitude of M200 component were not significantly correlated across participants (Spearman’s rho = 0.02, *p* = 0.97; Spearman skipped correlation 95% CI = [-1 0.88]).

4.4 Discussion

In this study, the role of the cerebellum in adapting to action-effect delay was investigated

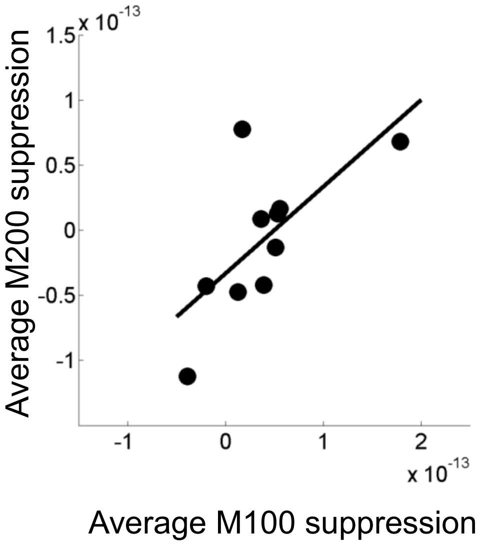


Figure 4.6 M100 suppression effect and M200 suppression effect are correlated across participants. Each point represents the average magnitude of M100/M200 suppression effect across all testing sessions of one participant.

with TMS and MEG. The effect of actions on auditory processing was measured in three separate testing sessions, with the tone onset always being delayed for around 100 ms following the tone-generating action. The classic M100 suppression and M200 suppression effects for self-generated tones were replicated in the presence of the delay. In sham stimulation condition, there was a significant increase of M100 suppression effect in the last testing session and no changes were found in M200 suppression effect. Interestingly, in the real stimulation condition (suppressive stimulation over the right cerebellum), the increase of M100 suppression effect disappeared, but M200 suppression effect changed right after the stimulation. There was also a significant correlation between M100 suppression effect and M200 suppression effect across participants.

In sham stimulation condition, a significant increase of M00 suppression effect in the last testing session (post-TMS2) was found. Interestingly, the strongest increase of M100 suppression effect over testing sessions was found at sensors where no significant M100 suppression effect was found in the first testing session (pre-TMS). Most sensors that showed significant M100 suppression effect in the first testing session (pre-TMS) did not have significant changes of M100 suppression effect over testing sessions. The fact that there was a significant M100 suppression effect in the presence of stimuli onset delay is consistent with the idea that actions have a general suppressive effect on auditory processing regardless of prediction (Cao & Gross, 2015a; Schröger et al., 2015). Actually, SA effect can be obtained under a range of stimulus onset delays up to at least 1000 ms (Bäß et al., 2008; van Elk, Salomon, Kannape, & Blanke, 2014). Or, it is also possible that a prediction mechanism that is not temporally well-tuned to the stimulus onset exists. The increase of M100 suppression effect in the last testing session suggests that at least 200 trials of learning are needed to adapt to a short stimulus onset delay following actions, which is consistent with the results from Aliu et al. (2009). Importantly, the increased M100 suppression effect supports the idea that there is a specific prediction effect of actions on auditory processing that takes into account the temporal delay (Blakemore, Frith, et al., 1999; Wolpert & Ghahramani, 2000). This suggests an update in the forward model over testing sessions that tunes to the temporal delay. In our study, the forward model updating effect is only present in the hemisphere ipsilateral to the action finger, but Aliu et al. (2009) showed the effect predominantly in the opposite hemisphere. The source of this discrepancy is unclear.

In real stimulation condition, the increase of M100 suppression effect disappeared after the right cerebellar stimulation with an inhibitory TMS protocol. This suggests that the stimulation specifically impaired the process of forward model updating. It is also interesting to note that no changes were found in any of the auditory sensors in real stimulation condition, suggesting that the general suppressive effect of actions on auditory processing is not affected by cerebellar stimulation. There might be different mechanisms for the prediction based (forward model) and general suppressive effects of actions. Interestingly, the M200 suppression pattern changed to M200 enhancement pattern after the stimulation, which is also right hemisphere lateralized. Note that no changes in M200 suppression were found in sham stimulation condition. So the change of M200 suppression is a direct effect of the right cerebellar stimulation.

It is clear that the M100 suppression changes reflect the forward model updating. But the interpretation of the M200 suppression changes is a bit ambiguous. One possibility is that the M200 suppression changes are related to changes of high-level cognitive aspects of the sound. For example, suppression of EEG measured P2 component has been related to the sense of agency, i.e., knowing oneself as the generator of the stimulus (Timm et al., 2016). However, we did not measure any cognitive aspect of the sound from participants. This possibility remains highly speculative. A more likely possibility is that the M200 suppression effect may be related to the process of forward model updating. The significant correlation between M100 suppression effect and M200 suppression effect supports this idea. The positive correlation possibly means that little updating (small M200 suppression) is needed when the forward model is suitable for the situation (large M100 suppression; note the negative sign of M200 suppression effect). Since our results suggest that the cerebellum is a likely place where the forward model updating occurs, the M200 suppression effect may relate to a process of feeding information (prediction errors) back to the cerebellum to make the forward model updating possible. This hypothesis also receives support from studies showing that the cerebellum encodes prediction errors and is sensible to the temporal information of self-generated stimuli (Blakemore et al., 2001; Imamizu et al., 2000; Schlerf, Ivry, & Diedrichsen, 2012; Watkins, Patel, Davis, & Howell, 2005). For example, in an fMRI study, participants were asked to read sentences aloud. Significant increases of cerebellar activation were found when the auditory feedback was delayed for 200 ms as compared to normal feedback (Watkins et al., 2005).

5 General conclusions

There is a profound influence of actions on auditory processing. The influence is evidenced in reduced neural responses (amplitudes of auditory evoked components, firing rates of auditory neurons) and behavioural measurements (e.g., perceived loudness) for self-generated sounds, known as sensory attenation effect. At least two separate mechanisms are underlying the influence. One is based on the rationale that self-generated sounds are predicted, which is represented by the concepts of efference copy, corollary discharge and the forward model. The other emphasizes the general sensory gating effect of actions regardless of prediction. A series of experiments in this thesis provided support for both accounts. Study 4 revealed that pre-stimulus alpha power is a likely candidate for the neural implementation of prediction in auditory cortex. It also suggested that a sequence of oscillatory events in different frequencies represent the information carried by self-generated sounds, which is consistent with the predictive coding framework. This is the first comprehensive study investigating the neural responses (neural oscillations) around the time of self-generated stimuli onset using MEG. Study 5 provides the first causal evidence, using TMS, for the role of the cerebellum in learning to predict self-generated sounds when they are delayed for presentation following actions. This supports a predictive nature of self-generated sounds. Another novel finding is that the late-latency auditory evoked responses (M200 component) may be involved in the process of learning. The presence of SA in the beginning of the learning provided support for the sensory gating account. Study 1 and 3 provided further support for the sensory gating account by demonstrating SA when the sound is unpredictable (study 1) and possibly when motor activation is indirectly induced by reading action-related words (study 3). Study 2 showed a cultural difference in perceiving others generated sounds, which revolves the discrepant findings from different cultural regions and supports the idea that auditory perception is under higher cognitive control. These studies contribute to our understanding of the role of the motor system in auditory processing by demonstrating two different mechanisms in the process and elucidating the underlying neural implementation.

Although there is a general sensory gating of auditory processing from actions, it is not clear whether this is functionally meaningfully or just a byproduct from the architecture of neural systems. The sensory prediction account of SA is best studied in the literature. It indicates that the influence from actions forms an integral part of sensory processing. Attention may modulate SA effect due to the pervasive influence of attention on sensory processing, but the attention account cannot fully explain SA. For example, the increse of M100 suppression effect in study 5 can hardly be explained by attention or sensroy gating, whereas it makes perfect sense from the perspective of prediction. In this sense, the preactivation account may be rather trivial as it claims the activation of sensory representation from actions, which can be viewed as a model of the realization of prediction. Therefore, the sensory prediction account should be given the most weight in explaining SA and maybe also in understanding action's influence on perception.

Appendix 1

Word stimuli (German) used in study 3. English translations are included in the parentheses.

Real words			Pseudowords		
Sound verb	Sound + Action verb	Abstract verb	Sound verb	Sound + Action verb	Abstract verb
brodeln (to seethe)	husten (to cough)	bessern (to improve)	sposeln	huseln	bekkern
hallen (to resound)	keuchen (to pant)	folgern (to conclude)	rellen	kenchen	lelgern
prasseln (to patter)	kreischen (to screech)	irren (to err)	knosseln	flieschen	ürben
rauschen (to whoosh)	quengeln (to whine)	mogeln (to cheat)	fieschen	drengeln	soseln
ticken (to tick)	röcheln (to wheeze)	schummeln (to cheat)	nokken	lücheln	schunneln
tosen (to roar)	schluchzen (to sob)	sehnen (to yearn)	soden	schmucknen	särnen
tuckern (to chug)	schwatzen (to gabble)	täuschen (to fool)	huhsern	schletzen	telschen
zirpen (to chrip)	stöhnen (to groan)	trotzen (to defy)	zaspen	stölzen	kretzen
zwitschern (to twitter)	stottern (to stutter)	wundern (to marvel)	zwarschern	stürtern	fursern
donnern (to thunder)	schreien (to cry)	büffeln (to swot)	tennern	schwieen	güpfeln
klingen (to chink)	grölen (to bawl)	grübeln (to brood)	spinken	frülen	fröbeln
klirren (to clank)	johlen (to yell)	hadern (to quarrel)	spörren	järlen	wasern
krachen (to crack)	mampfen (to chomp)	hassen (to hate)	flasten	lempfen	wossen
plätschern (to dabble)	prusten (to snort)	schulden (to owe)	blotschern	ruschen	schunsen
schallen (to echo)	schnalzen (to chirrup)	zaudern (to tarry)	schakken	schralben	ziesern
surren (to buzz)	schnaufen (to wheeze)	zweifeln (to doubt)	huhben	schraumen	preimeln

Appendix 2

Robust correlation results in source space using Spearman skipped correlations (study 4):
Between SA (source) and pre-stimulus alpha power increase: Spearman’s rho = -0.65, 95% CI = [-0.93 -0.12];
Between SA (source) and post-stimulus alpha phase locking decrease: Spearman’s rho = 0.89, 95% CI = [0.78 0.98];
Between SA (source) and post-stimulus high gamma power decrease: Spearman’s rho = 0.67, 95% CI = [0.17 0.92];
Between SA (source) and post-stimulus low gamma power decrease: Spearman’s rho = 0.65, 95% CI = [0.18 0.93];
Between pre-stimulus alpha power increase and post-stimulus high gamma power decrease: Spearman’s rho = -0.80, 95% CI = [-0.95 -0.41];
Between post-stimulus alpha phase locking decrease and post-stimulus high gamma power decrease: Spearman’s rho = 0.73, 95% CI = [0.31 0.91].

Supplementary figures

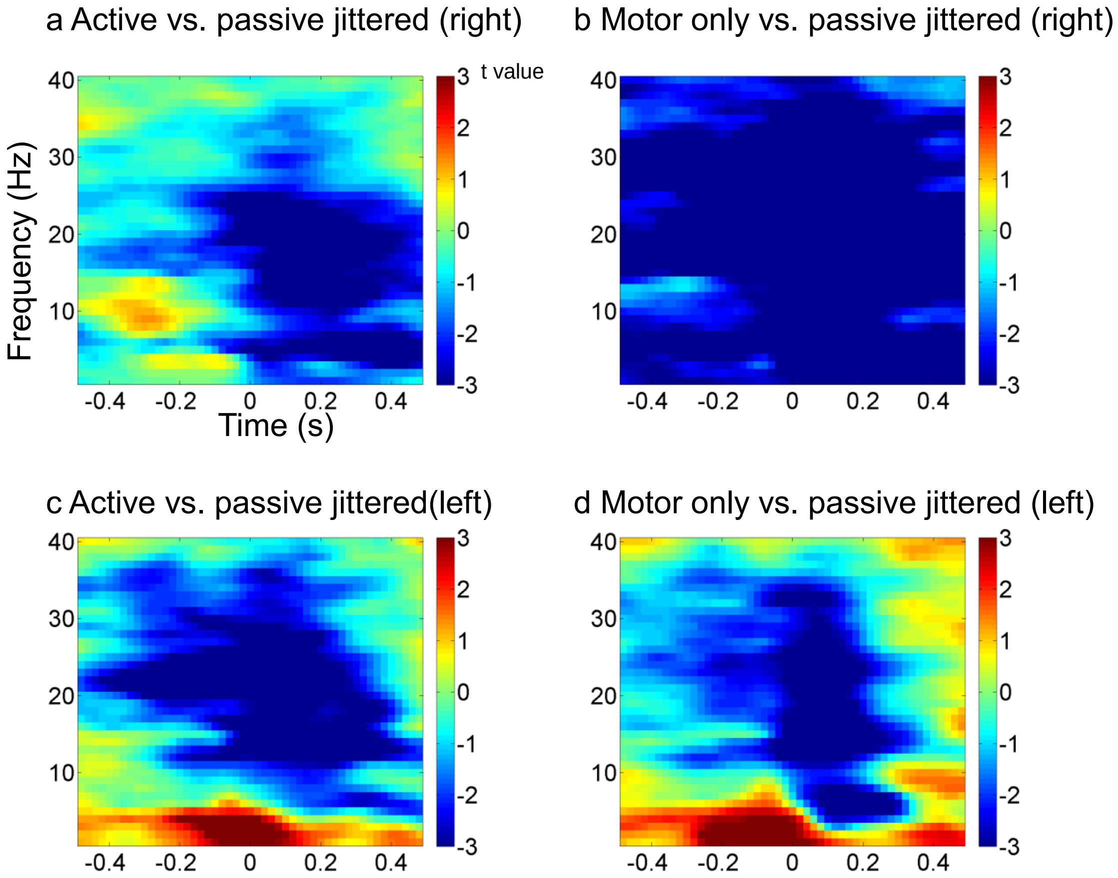


Figure S3.1 Power comparisons between active and passive jittered condition in the right (a) and left (c) auditory sensors, between motor only and passive jittered condition in the right (b) and left (d) auditory sensors.

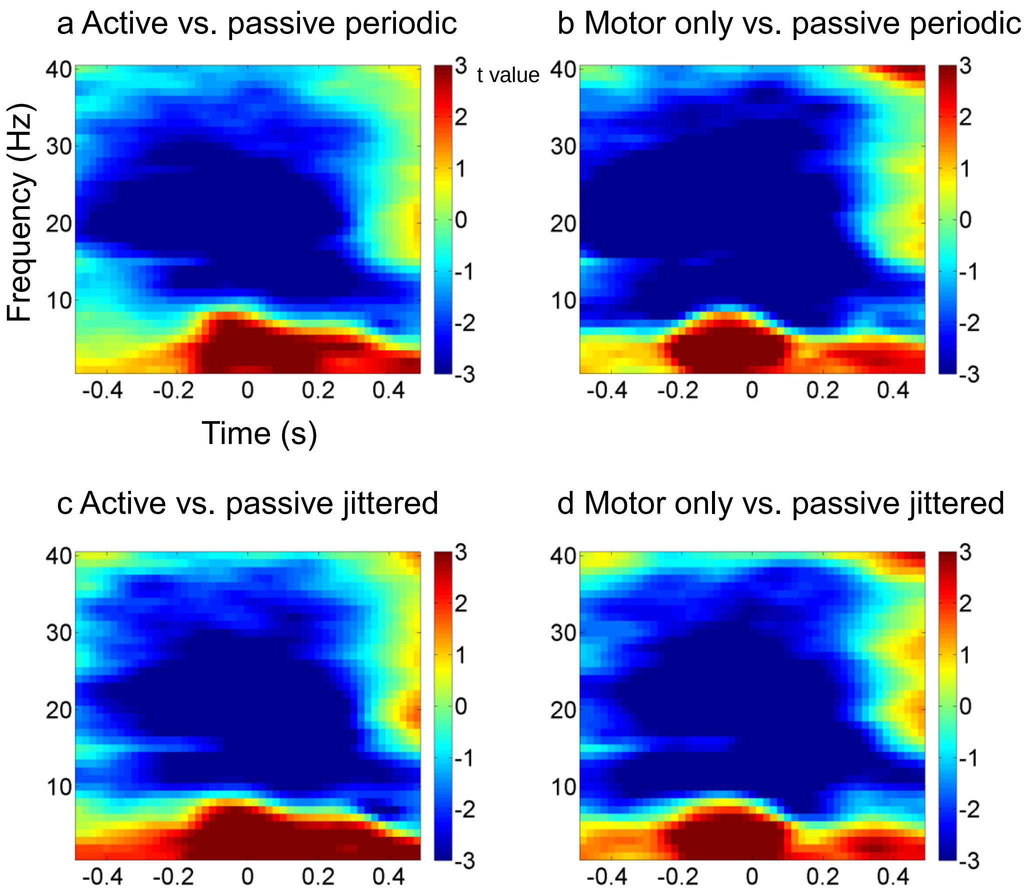
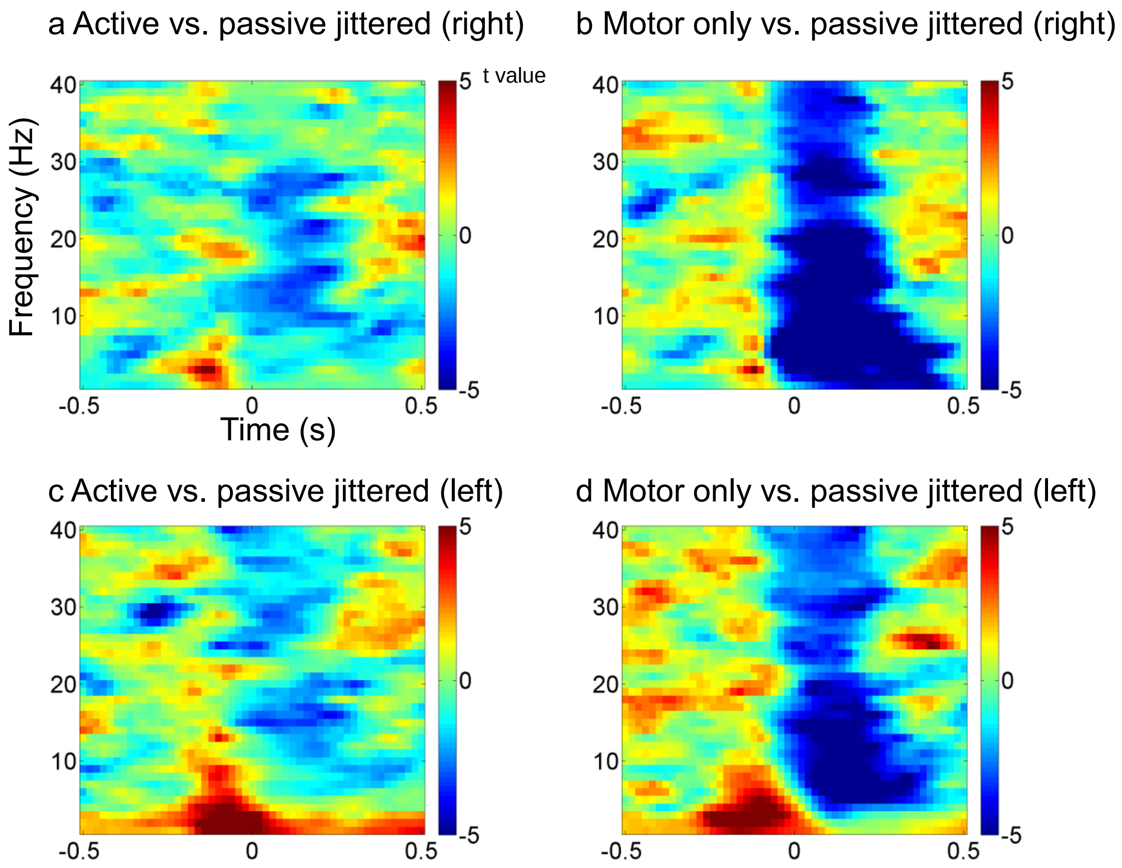


Figure S3.2 Power comparisons between different conditions made with data from motor sensors. See the heading of the figure for the information of conditions being compared.



a Active vs. passive jittered (right)

b Motor only vs. passive jittered (right)

c Active vs. passive jittered (left)

d Motor only vs. passive jittered (left)

Figure S3.3 Phase locking value (PLV) comparisons between active and passive jittered condition in the right (a) and left (c) auditory sensors, between motor only and passive jittered condition in the right (b) and left (d) auditory sensors.

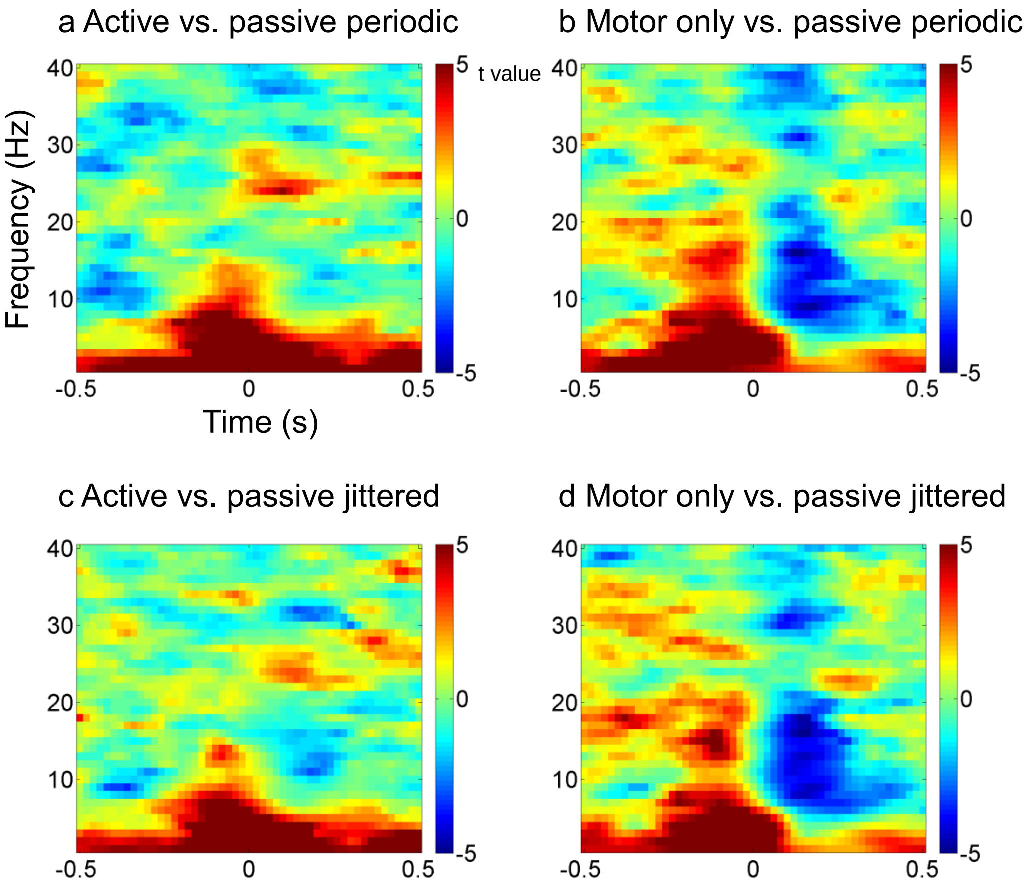
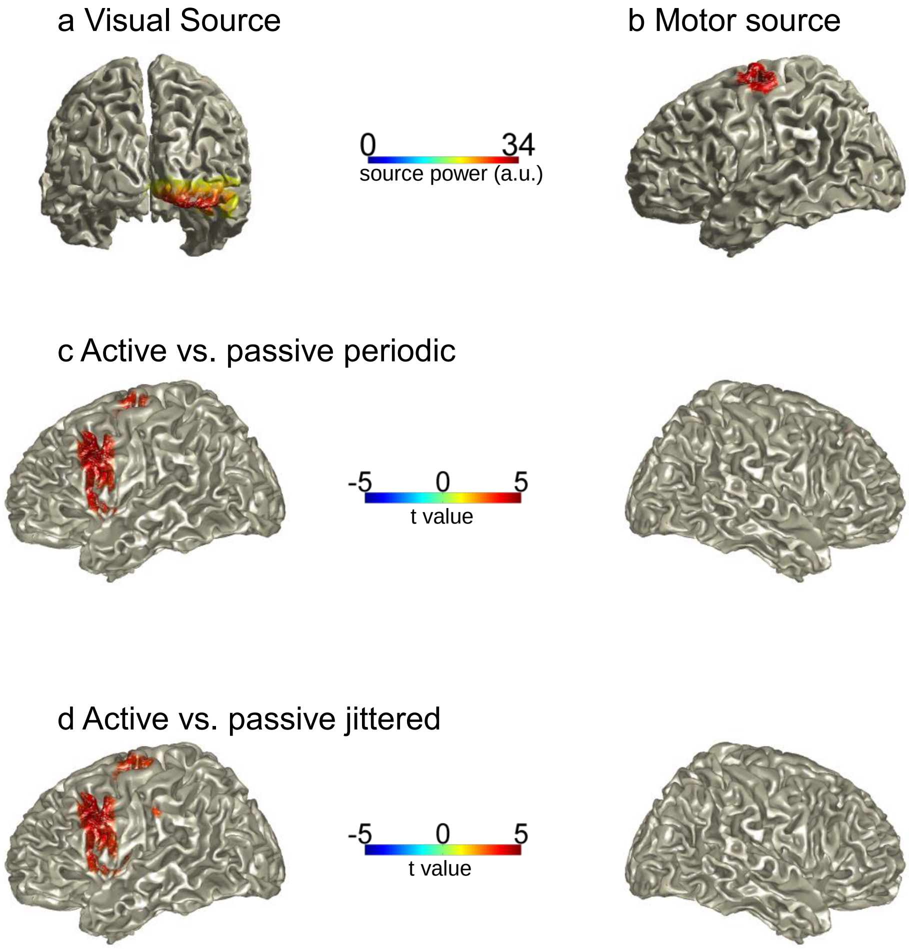


Figure S3.4 PLV comparisons between different conditions made data from motor sensors. See the heading of the figure for the information of conditions being compared.



a Visual Source

b Motor source

c Active vs. passive periodic

d Active vs. passive jittered

Figure S3.5 Source results for visual components. Visual evoked responses (90 to 160 ms) are localized in visual cortex (predominantly in the right hemisphere) (a) and motor evoked responses (-20 to 40 ms) are localized in motor cortex (b). In the post-stimulus time window (90 to 160 ms), there is stronger response in left motor cortex for active condition than for both passive conditions (c & d).

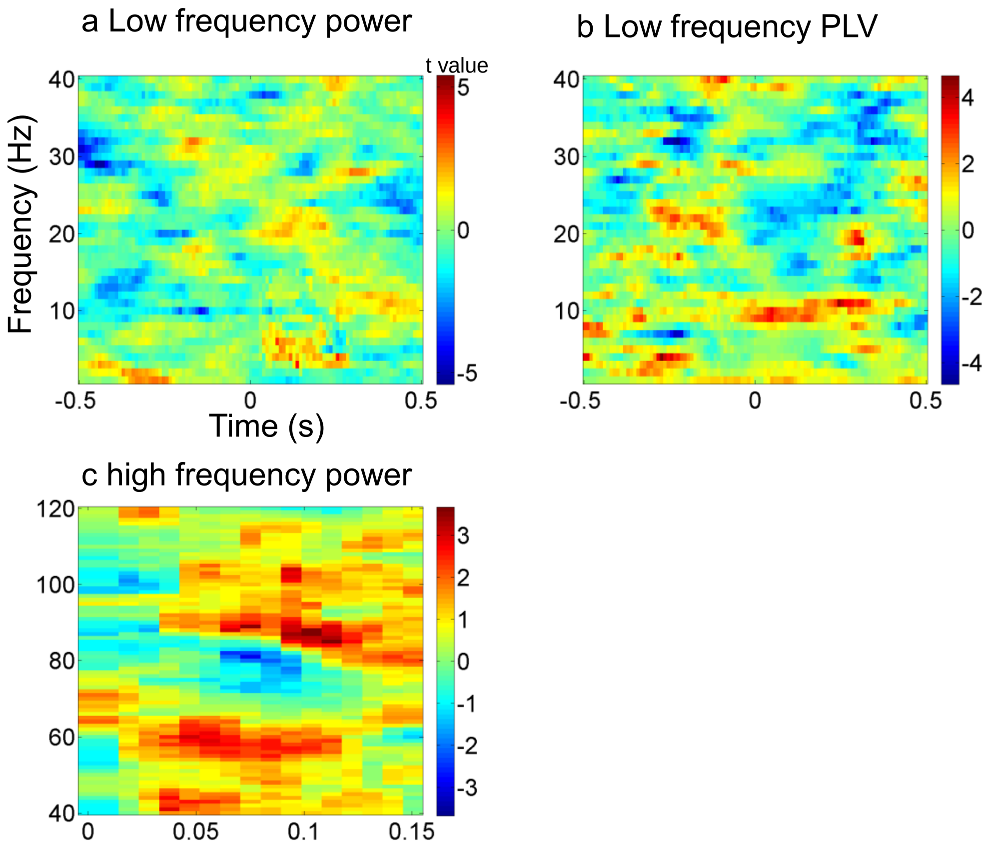


Figure S3.6 Correlation maps between SA (source) and oscillatory differences between active and passive periodic conditions. (a) each point in the map is a t value (converted from Spearman’s rho value) for the correlation between SA (source) and the power difference at the time-frequency represented by the point. (b) and (c) are the same correlation but with low frequency phase locking value and high frequency power, respectively.

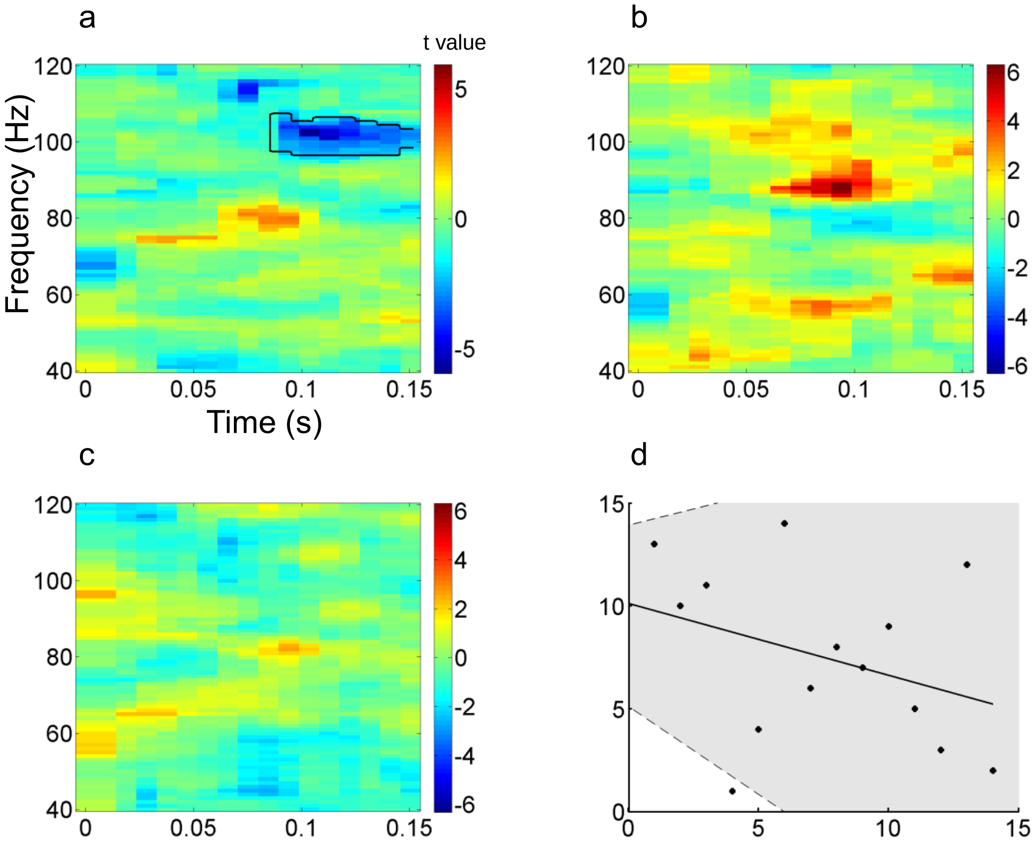


Figure S3.7 Correlation maps between pre-stimulus alpha power increase (averaged over 9 to 12 Hz and -400 to -60 ms) and post-stimulus gamma power changes over time and frequency (a), between post-stimulus alpha phase locking value decrease (averaged over 9 to 11 Hz and 0 to 150 ms) and post-stimulus gamma power changes across time and frequency (b) and between post-stimulus alpha power decrease (averaged over 10 to 13 Hz and 0 to 150 ms) and post-stimulus gamma power changes across time and frequency. (d) shows the scatter plot between post-stimulus alpha power decrease and post-stimulus high gamma power decrease (Spearman’s rho = -0.35, *p* = 0.22, CI = [-0.95 0.30]).

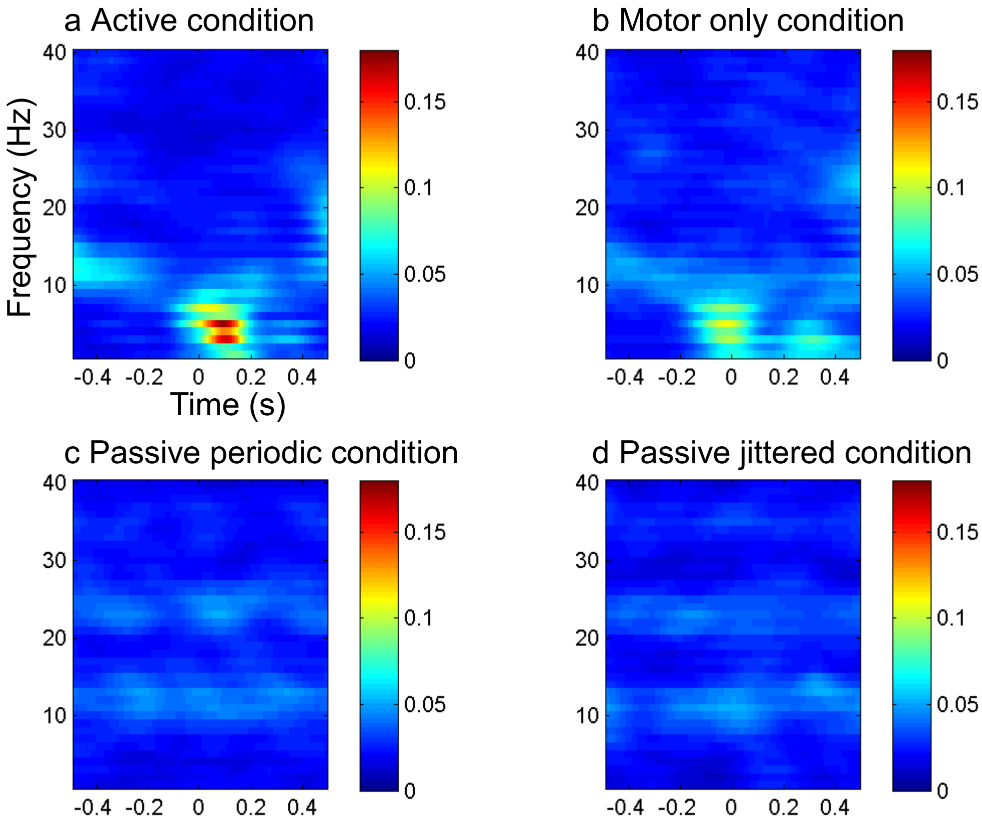


Figure S3.8 Granger causality results in the time frequency domain showing the causal influence of motor cortex over left auditory cortex in all the 4 conditions.

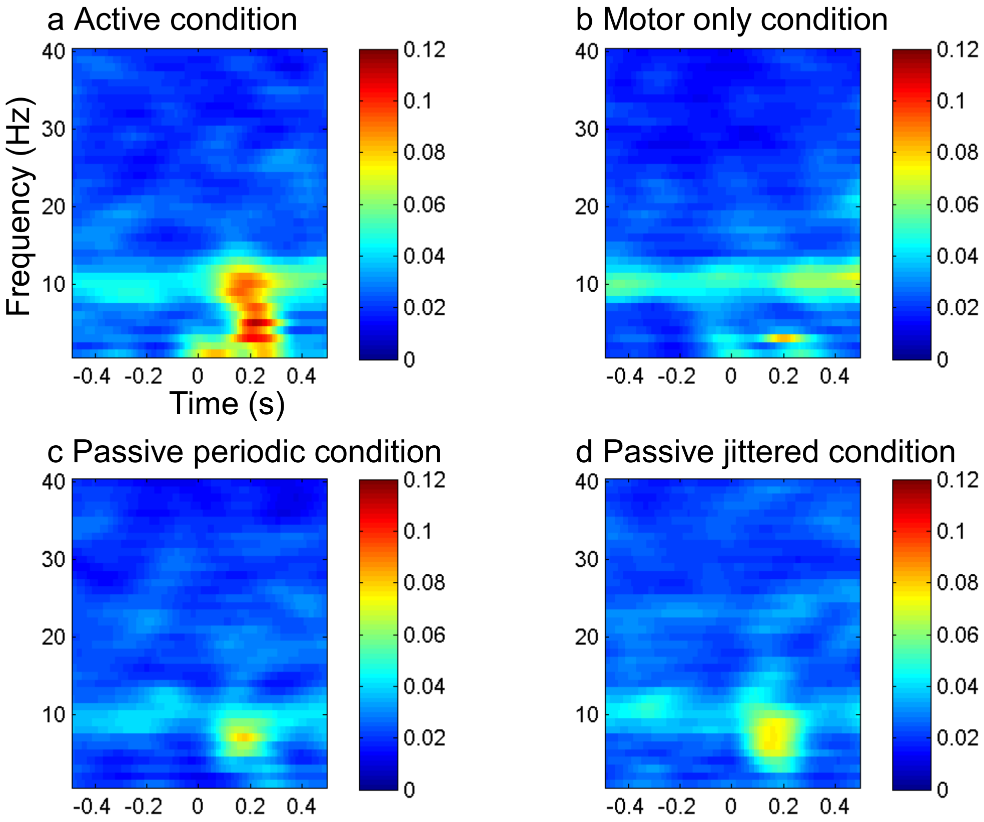


Figure S3.9 Granger causality results in the time frequency domain showing the causal influence of left auditory cortex over motor cortex in all the 4 conditions.

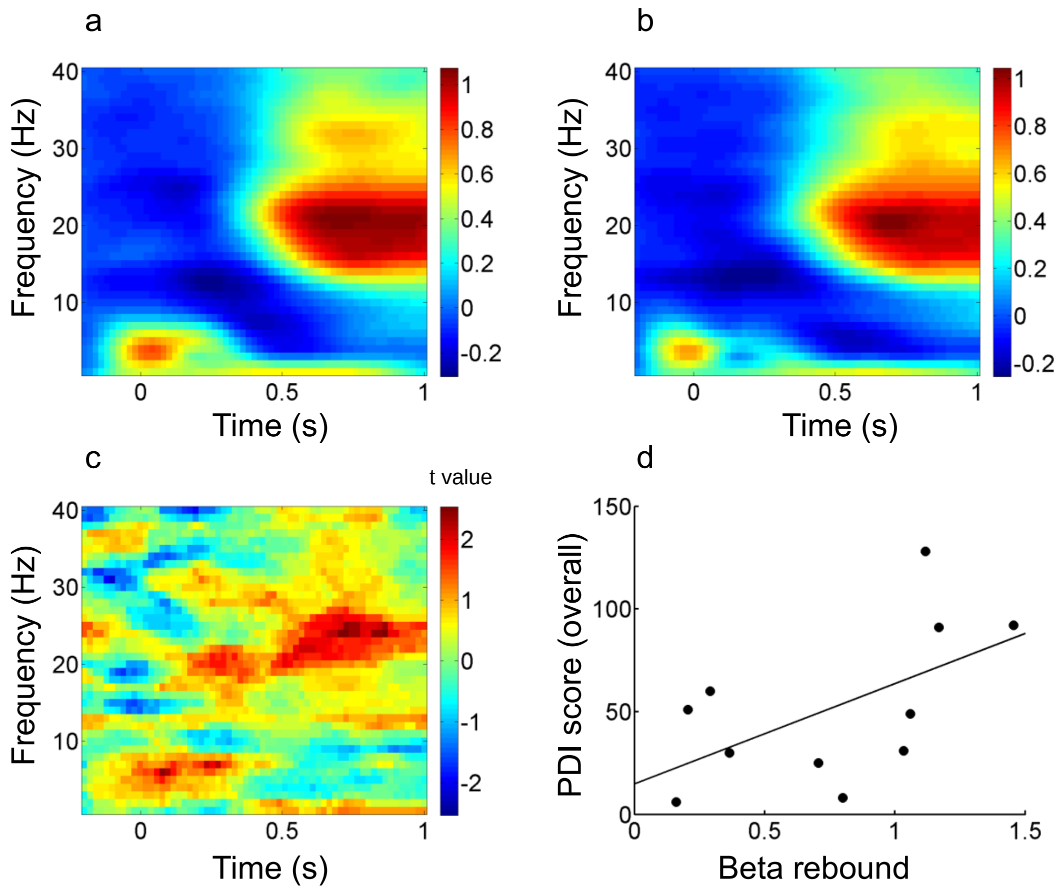


Figure S3.10 (a) and (b) show the beta rebound in active and motor only condition, respectively. The power data are baseline corrected (-500 to -100 ms) and averaged over three sensors that show strongest motor activation in motor only condition. (c) is the correlation map between PDI overall score and induced power in motor only condition. Note the high t values in the period of beta rebound. (d) is the scatter plot between PDI overall score and beta rebound (24-26 Hz, 650 to 900 ms) in motor only condition (Spearman’s rho = 0.62, *p* = 0.048).

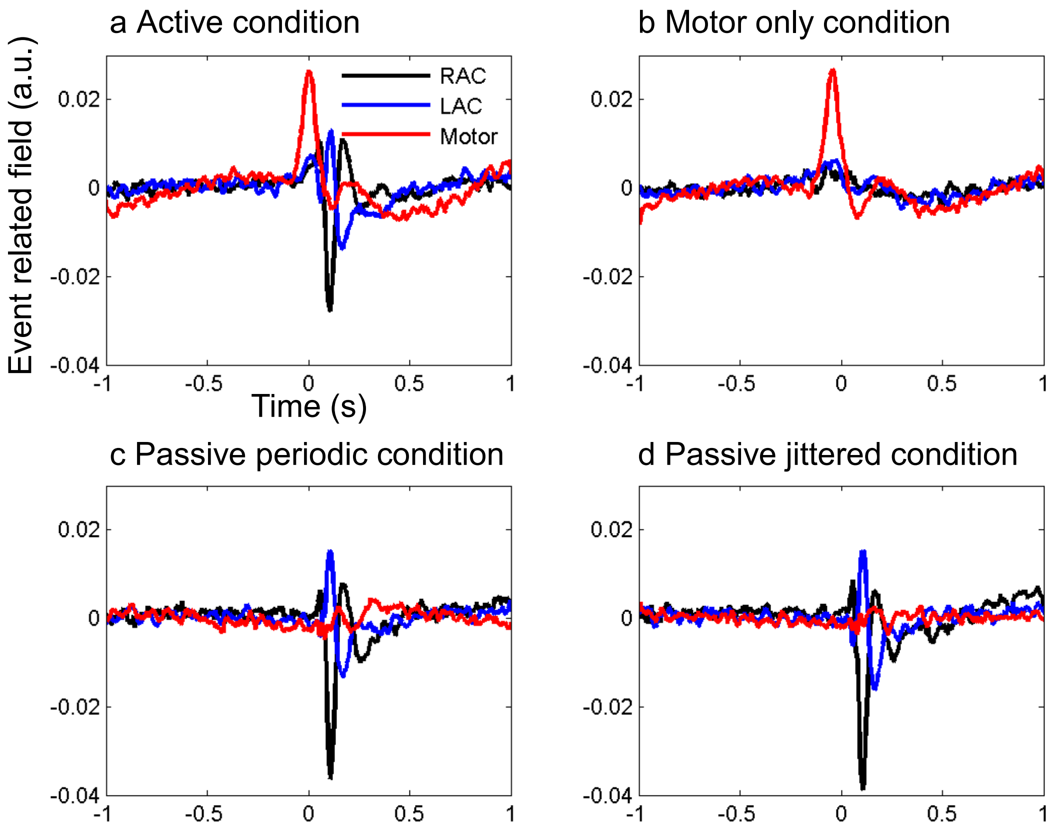


Figure S3.11 Source reconstructed time series data in different cortical regions and different conditions that are used for Granger causality analysis. The same right auditory voxel is used for all other source space analyses. The signal for motor cortex is average over 8 voxels.

References

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. doi:10.1162/jocn.2009.21055

Andres, M., Finocchiaro, C., Buiatti, M., & Piazza, M. (2015). Contribution of motor representations to action verb processing. *Cognition, 134*, 174-184. doi:10.1016/j.cognition.2014.10.004

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences, 16*(7), 390-398. doi:10.1016/j.tics.2012.05.003

Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci, 14*(6), 797-801. doi:10.1038/nn.2810

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412. doi:10.1016/j.jml.2007.12.005

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (CD□ROM): Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

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., & Prinz, W. (2015). My partner is also on my mind: social context modulates the N1 response. *Experimental Brain Research, 233*(1), 105-113. doi:10.1007/s00221-014-4092-9

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*(6), 248-254.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in psychology, 4*. doi:10.3389/fpsyg.2013.00328

Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J. C., Costumero, V., . . . Ávila, C. (2002). Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. *Cerebral Cortex, 22*(11), 2554-2563.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol., 59*, 617-645.

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

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., . . . Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron, 85*(2), 390-401. doi:10.1016/j.neuron.2014.12.018

Bauer, M., Stenner, M. P., Friston, K. J., & Dolan, R. J. (2014). Attentional modulation of alpha/beta and gamma oscillations reflect functionally distinct processes. *J Neurosci, 34*(48), 16117-16125. doi:10.1523/JNEUROSCI.3474-13.2014

Baumgarten, T. J., Schnitzler, A., & Lange, J. (2014). Prestimulus Alpha Power Influences Tactile Temporal Perceptual Discrimination and Confidence in Decisions. *Cereb Cortex*, bhu247.

Behroozmand, R., Oya, H., Nourski, K. V., Kawasaki, H., Larson, C. R., Brugge, J. F., . . . Greenlee, J. D. (2016). Neural Correlates of Vocal Production and Motor Control in Human Heschl's Gyrus. *J Neurosci, 36*(7), 2302-2315. doi:10.1523/JNEUROSCI.3305-14.2016

Bennett, C., Arroyo, S., & Hestrin, S. (2013). Subthreshold mechanisms underlying state-dependent modulation of visual responses. *Neuron, 80*(2), 350-357. doi:10.1016/j.neuron.2013.08.007

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software, 31*(10), 1-21.

Besl, P. J., & McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*(2), 239 - 256. doi:10.1109/34.121791

Biemann, C., Heyer, G., Quasthoff, U., & Richter, M. (2007). *The Leipzig Corpora Collection – Monolingual corpora of standard size*. Paper presented at the Proceedings of Corpus Linguistics 2007. Available from: http://www.birmingham.ac.uk/research/activity/corpus/publications/conference-archives/2007-birmingham.aspx.

Billington, J., Baron-Cohen, S., & Bor, D. (2008). Systemizing influences attentional processes during the Navon task: An fMRI study. *Neuropsychologia, 46*(2), 511-520.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of cognitive neuroscience, 11*(5), 551-559.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport, 12*(9), 1879-1884. doi:Doi 10.1097/00001756-200107030-00023

Blakemore, S. J., Smith, J., Steel, R., Johnstone, C. E., & Frith, C. D. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol Med, 30*(5), 1131-1139.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nat Neurosci, 1*(7), 635-640. doi:10.1038/2870

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1999). The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. *Neuroimage, 10*(4), 448-459.

Borra, T., Versnel, H., Kemner, C., van Opstal, A. J., & van Ee, R. (2013). Octave effect in auditory attention. *Proc Natl Acad Sci U S A, 110*(38), 15225-15230. doi:10.1073/pnas.1213756110

Boulenger, V., Silber, B. Y., Roy, A. C., Paulignan, Y., Jeannerod, M., & Nazir, T. A. (2008). Subliminal display of action words interferes with motor planning: a combined EEG and kinematic study. *J Physiol Paris, 102*(1-3), 130-136.

Brown, H., Adams, R. A., Parees, I., Edwards, M., & Friston, K. J. (2013). Active inference, sensory attenuation and illusions. *Cognitive Processing, 14*(4), 411-427. doi:10.1007/s10339-013-0571-3

Busch, N. A., Dubois, J., & VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci, 29*(24), 7869-7876. doi:10.1523/JNEUROSCI.0113-09.2009

Cao, L., & Gross, J. (2015a). Attention Wins over Sensory Attenuation in a Sound Detection Task. *Plos One, 10*(8), e0136585. doi:10.1371/journal.pone.0136585

Cao, L., & Gross, J. (2015b). Cultural Differences in Perceiving Sounds Generated by Others: Self Matters. *Front Psychol, 6*, 1865. doi:10.3389/fpsyg.2015.01865

Cao, L., & Hu, Y. M. (2016). Beta Rebound in Visuomotor Adaptation: Still the Status Quo? *J Neurosci, 36*(24), 6365-6367. doi:10.1523/JNEUROSCI.1007-16.2016

Cao, L., Klepp, A., Schnitzler, A., Gross, J., & Biermann-Ruben, K. (2016). Auditory perception modulated by word reading. *Experimental Brain Research*, 1-9. doi:10.1007/s00221-016-4706-5

Cao, L., Thut, G., & Gross, J. (2016). The role of brain oscillations in predicting self-generated sounds. *Neuroimage*. doi:http://dx.doi.org/10.1016/j.neuroimage.2016.11.001

Capilla, A., Schoffelen, J. M., Paterson, G., Thut, G., & Gross, J. (2012). Dissociated α-band modulations in the dorsal and ventral visual pathways in visuospatial attention and perception. *Cereb Cortex*, bhs343.

Cardoso-Leite, P., Mamassian, P., Schuetz-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

Chapman, C. E., Bushnell, M. C., Miron, D., Duncan, G. H., & Lund, J. P. (1987). Sensory perception during movement in man. *Experimental Brain Research, 68*(3), 516-524.

Chapman, C. E., Jiang, W., & Lamarre, Y. (1988). Modulation of lemniscal input during conditioned arm movements in the monkey. *Experimental Brain Research, 72*(2), 316-334.

Christensen, M. S., Lundbye-Jensen, J., Geertsen, S. S., Petersen, T. H., Paulson, O. B., & Nielsen, J. B. (2007). Premotor cortex modulates somatosensory cortex during voluntary movements without proprioceptive feedback. *Nature Neuroscience, 10*(4), 417-419. doi:10.1038/nn1873

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: From origin to function. *Behavioral and Brain Sciences, 37*(02), 177-192.

Craddock, N., & Owen, M. J. (2010). The Kraepelinian dichotomy - going, going... but still not gone. *Br J Psychiatry, 196*(2), 92-95. doi:10.1192/bjp.bp.109.073429

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

Cross, S. E., Hardin, E. E., & Gercek-Swing, B. (2011). The what, how, why, and where of self-construal. *Pers Soc Psychol Rev, 15*(2), 142-179. doi:10.1177/1088868310373752

Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp, 9*(4), 183-191.

Daliri, A., & Max, L. (2015). Electrophysiological evidence for a general auditory prediction deficit in adults who stutter. *Brain Lang, 150*, 37-44. doi:10.1016/j.bandl.2015.08.008

de Vega, M., Moreno, V., & Castillo, D. (2013). The comprehension of action-related sentences may cause interference rather than facilitation on matching actions. *Psychol Res, 77*(1), 20-30. doi:10.1007/s00426-011-0356-1

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

Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage,, 41*(2), 354-362.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*(1), 176-180.

Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology, 89*(4), 2194-2207. doi:10.1152/jn.00627.2002

Eliades, S. J., & Wang, X. (2013). Comparison of auditory-vocal interactions across multiple types of vocalizations in marmoset auditory cortex. *Journal of Neurophysiology, 109*(6), 1638-1657. doi:10.1152/jn.00698.2012

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nat Rev Neurosci, 10*(1), 48-58. doi:10.1038/nrn2536

Flinker, A., Chang, E. F., Kirsch, H. E., Barbaro, N. M., Crone, N. E., & Knight, R. T. (2010). Single-trial speech suppression of auditory cortex activity in humans. *J Neurosci, 30*(49), 16643-16650. doi:10.1523/JNEUROSCI.1809-10.2010

Fontolan, L., Morillon, B., Liegeois-Chauvel, C., & Giraud, A. L. (2014). The contribution of frequency-specific activity to hierarchical information processing in the human auditory cortex. *Nat Commun, 5*, 4694. doi:10.1038/ncomms5694

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

Ford, J. M., Palzes, V. A., Roach, B. J., & Mathalon, D. H. (2013). 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

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in psychology, 2*. doi:10.3389/fpsyg.2011.00154

Frey, J. N., Mainy, N., Lachaux, J. P., Muller, N., Bertrand, O., & Weisz, N. (2014). Selective modulation of auditory cortical alpha activity in an audiovisual spatial attention task. *J Neurosci, 34*(19), 6634-6639. doi:10.1523/JNEUROSCI.4813-13.2014

Friston, K. J. (2005). A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci, 360*(1456), 815-836. doi:10.1098/rstb.2005.1622

Friston, K. J., Bastos, A. M., Pinotsis, D., & Litvak, V. (2015). LFP and oscillations-what do they tell us? *Curr Opin Neurobiol, 31*, 1-6. doi:10.1016/j.conb.2014.05.004

Gentsch, A., Kathmann, N., & Schutz-Bosbach, S. (2012). Reliability of sensory predictions determines the experience of self-agency. *Behavioural Brain Research, 228*(2), 415-422. doi:10.1016/j.bbr.2011.12.029

Gentsch, A., & Schutz-Bosbach, S. (2011). I Did It: Unconscious Expectation of Sensory Consequences Modulates the Experience of Self-agency and Its Functional Signature. *Journal of cognitive neuroscience, 23*(12), 3817-3828.

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., & Ávila, C. (2006). Reading cinnamon activates olfactory brain regions. *Neuroimage, 32*(2), 906-912.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica, 37*(3), 414-&.

Green, D. M., McKey, M. J., & Licklider, J. C. R. (1959). Detection of a pulsed sinusoid in noise as a function of frequency. *The Journal of the Acoustical Society of America, 31*(11), 1446-1452.

Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-Response Characteristic of Auditory Observers Detecting Signals of a Single Frequency in Noise: The Probe-Signal Method. *The Journal of the Acoustical Society of America, 44*(6), 1513-1523.

Greenlee, J. D., Jackson, A. W., Chen, F., Larson, C. R., Oya, H., Kawasaki, H., . . . Howard, M. A., 3rd. (2011). Human auditory cortical activation during self-vocalization. *Plos One, 6*(3), e14744. doi:10.1371/journal.pone.0014744

Gross, J., Baillet, S., Barnes, G. R., Henson, R. N., Hillebrand, A., Jensen, O., . . . Schoffelen, J. M. (2013). Good practice for conducting and reporting MEG research. *Neuroimage, 65*, 349-363. doi:10.1016/j.neuroimage.2012.10.001

Gross, J., Schnitzler, A., Timmermann, L., & Ploner, M. (2007). Gamma oscillations in human primary somatosensory cortex reflect pain perception. *PLoS Biol, 5*(5), e133. doi:10.1371/journal.pbio.0050133

Grüsser, O. J. (1995). On the history of the ideas of efference copy and reafference. *Clio Med, 33*, 35-55.

Haegens, S., Händel, B. F., & Jensen, O. (2011). Top-down controlled alpha band activity in somatosensory areas determines behavioral performance in a discrimination task. *The Journal of Neuroscience, 31*(14), 5197-5204.

Haegens, S., Nacher, V., Luna, R., Romo, R., & Jensen, O. (2011). alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. *Proc Natl Acad Sci U S A, 108*(48), 19377-19382. doi:10.1073/pnas.1117190108

Hall, P., & Welsh, A. H. (1985). Limit-Theorems for the Median Deviation. *Annals of the Institute of Statistical Mathematics, 37*(1), 27-36. doi:Doi 10.1007/Bf02481078

Han, S., & Northoff, G. (2008). Culture-sensitive neural substrates of human cognition: a transcultural neuroimaging approach. *Nature Reviews Neuroscience, 9*(8), 646-654.

Hardwick, R. M., Lesage, E., & Miall, R. C. (2014). Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain Stimul, 7*(5), 643-649. doi:10.1016/j.brs.2014.04.009

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc Natl Acad Sci U S A, 95*(25), 15061-15065.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*(2), 301-307.

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values ofd′. *Behavior Research Methods, Instruments, & Computers, 27*(1), 46-51.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., & Ford, J. M. (2007). Relationship of Imprecise Corollary Discharge in Schizophrenia to Auditory Hallucinations. *Archives of General Psychiatry, 64*(3), 286-296.

Hesse, M. D., Nishitani, N., Fink, G. R., Jousmaki, V., & Hari, R. (2009). Attenuation of Somatosensory Responses to Self-Produced Tactile Stimulation. *Cerebral Cortex, 20*(2), 425-432. doi:10.1093/cercor/bhp110

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.

Horvath, J. (2014). The role of mechanical impact in action-related auditory attenuation. *Cogn Affect Behav Neurosci, 14*(4), 1392-1406. doi:10.3758/s13415-014-0283-x

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research, 1626*, 54-65. doi:10.1016/j.brainres.2015.03.038

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.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J Cogn Neurosci, 14*(8), 1125-1138. doi:10.1162/089892902760807140

Hughes, G. (2015). ERP and behavioral evidence of increased sensory attenuation for fear-related action outcomes. *Biol Psychol, 111*, 8-13. doi:10.1016/j.biopsycho.2015.08.002

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

Hughes, G., & Waszak, F. (2011). ERP correlates of action effect prediction and visual sensory attenuation in voluntary action. *Neuroimage, 56*(3), 1632-1640. doi:10.1016/j.neuroimage.2011.02.057

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., PuÈtz, B., . . . Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature, 403*(6766), 192-195.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Ann N Y Acad Sci, 1169*, 58-73. doi:10.1111/j.1749-6632.2009.04579.x

Izawa, J., Criscimagna-Hemminger, S. E., & Shadmehr, R. (2012). Cerebellar contributions to reach adaptation and learning sensory consequences of action. *J Neurosci, 32*(12), 4230-4239. doi:10.1523/JNEUROSCI.6353-11.2012

Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Frontiers in Human Neuroscience, 4*, 186. doi:10.3389/fnhum.2010.00186

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

Kauramaki, J., Jaaskelainen, I. P., Hari, R., Mottonen, R., Rauschecker, J. P., & Sams, M. (2010). Lipreading and covert speech production similarly modulate human auditory-cortex responses to pure tones. *J Neurosci, 30*(4), 1314-1321. doi:10.1523/JNEUROSCI.1950-09.2010

Kessler, K., Cao, L., O'Shea, K. J., & Wang, H. (2014). A cross-culture, cross-gender comparison of perspective taking mechanisms. *Proc Biol Sci, 281*(1785), 20140388. doi:10.1098/rspb.2014.0388

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627-633.

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: four markers for a link between auditory and conceptual brain systems. *J Neurosci, 28*(47), 12224-12230. doi:10.1523/JNEUROSCI.3579-08.2008

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cogn Process, 8*(3), 159-166. doi:10.1007/s10339-007-0170-2

Kitayama, S., & Park, J. (2014). Error-Related Brain Activity Reveals Self-Centric Motivation: Culture Matters. *Journal of Experimental Psychology: General, 143*(1), 62-70. doi:10.1037/a0031696.supp

Klepp, A., Niccolai, V., Buccino, G., Schnitzler, A., & Biermann-Ruben, K. (2015). Language-motor interference reflected in MEG beta oscillations. *Neuroimage, 109*, 438-448. doi:10.1016/j.neuroimage.2014.12.077

Klepp, A., Weissler, H., Niccolai, V., Terhalle, A., Geisler, H., Schnitzler, A., & Biermann-Ruben, K. (2014). Neuromagnetic hand and foot motor sources recruited during action verb processing. *Brain Lang, 128*(1), 41-52. doi:10.1016/j.bandl.2013.12.001

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *J Cogn Neurosci, 24*(3), 698-706. doi:10.1162/jocn_a_00167

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex, 49*(9), 2449-2461. doi:10.1016/j.cortex.2012.12.012

Kok, P., Rahnev, D., Jehee, J. F., Lau, H. C., & de Lange, F. P. (2011). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex, bhr310.*

Kuhn, S., Seurinck, R., Fias, W., & Waszak, F. (2010). The internal anticipation of sensory action effects: when action induces FFA and PPA activity. *Frontiers in Human Neuroscience, 4*. doi:10.3389/fnhum.2010.00054

Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron, 53*(2), 279-292. doi:10.1016/j.neuron.2006.12.011

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*, 263. doi:10.3389/fnhum.2013.00263

Macerollo, A., Chen, J. C., Korlipara, P., Foltynie, T., Rothwell, J., Edwards, M. J., & Kilner, J. M. (2016). Dopaminergic treatment modulates sensory attenuation at the onset of the movement in Parkinson's disease: A test of a new framework for bradykinesia. *Movement Disorders, 31*(1), 143-146. doi:10.1002/mds.26493

Malassis, R., Del Cul, A., & Collins, T. (2015). Corollary Discharge Failure in an Oculomotor Task Is Related to Delusional Ideation in Healthy Individuals. *Plos One, 10*(8). doi:10.1371/journal.pone.0134483

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods, 164*(1), 177-190.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253. doi:10.1037/0033-295X.98.2.224

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

Mast, F. W., & Ellis, A. W. (2015). Internal Models, Vestibular Cognition, and Mental Imagery: Conceptual Considerations. *Multisens Res, 28*(5-6), 443-460.

Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., & Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J Neurosci, 29*(9), 2725-2732. doi:10.1523/JNEUROSCI.3963-08.2009

Miall, R. C., & Wolpert, D. M. (1996). Forward Models for Physiological Motor Control. *Neural Netw, 9*(8), 1265-1279.

Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-Beta and Gamma Rhythms Subserve Feedback and Feedforward Influences among Human Visual Cortical Areas. *Neuron*. doi:10.1016/j.neuron.2015.12.018

Mifsud, N. G., Oestreich, L. K., Jack, B. N., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2016). Self-initiated actions result in suppressed auditory but amplified visual evoked components in healthy participants. *Psychophysiology*. doi:10.1111/psyp.12605

Mirabella, G., Iaconelli, S., Spadacenta, S., Federico, P., & Gallese, V. (2012). Processing of hand-related verbs specifically affects the planning and execution of arm reaching movements. *Plos One, 7*(4), e35403. doi:10.1371/journal.pone.0035403

Mock, J. R., Foundas, A. L., & Golob, E. J. (2011). Modulation of sensory and motor cortex activity during speech preparation. *Eur J Neurosci, 33*(5), 1001-1011. doi:10.1111/j.1460-9568.2010.07585.x

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nat Commun, 5*, 5255. doi:10.1038/ncomms6255

Mukamel, R., Ekstrom, A. D., Kaplan, J., Iacoboni, M., & Fried, I. (2010). Single-neuron responses in humans during execution and observation of actions. *Curr Biol, 20*(8), 750-756. doi:10.1016/j.cub.2010.02.045

Müller, N., Leske, S., Hartmann, T., Szebenyi, S., & Weisz, N. (2014). Listen to Yourself: The Medial Prefrontal Cortex Modulates Auditory Alpha Power During Speech Preparation. *Cerebral Cortex*. doi:10.1093/cercor/bhu117

Niccolai, V., Klepp, A., Weissler, H., Hoogenboom, N., Schnitzler, A., & Biermann-Ruben, K. (2014). Grasping hand verbs: oscillatory beta and alpha correlates of action-word processing. *Plos One, 9*(9), e108059. doi:10.1371/journal.pone.0108059

Niell, C. M., & Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron, 65*(4), 472-479. doi:10.1016/j.neuron.2010.01.033

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in medicine and biology, 48*(22), 3637. doi:10.1088/0031-9155/48/22/002

Oestreich, L. K. L., Mifsud, N. G., Ford, J. M., Roach, B. J., Mathalon, D. H., & Whitford, T. J. (2015). Subnormal sensory attenuation to self-generated speech in schizotypy: Electrophysiological evidence for a 'continuum of psychosis'. *International Journal of Psychophysiology, 97*(2), 131-138. doi:10.1016/j.ijpsycho.2015.05.014

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience, 2011.*

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences, 16*(10), 504-510. doi:10.1016/j.tics.2012.08.009

Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2012). Robust correlation analyses: false positive and power validation using a new open source matlab toolbox. *Front Psychol, 3*, 606. doi:10.3389/fpsyg.2012.00606

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring delusional ideation: the 21-item Peters et al. Delusions Inventory (PDI). *Schizophrenia Bulletin, 30*(4), 1005-1022.

Pickering, M. J., & Clark, A. (2014). Getting ahead: forward models and their place in cognitive architecture. *Trends in Cognitive Sciences, 18*(9), 451-456. doi:10.1016/j.tics.2014.05.006

Ploner, M., Gross, J., Timmermann, L., Pollok, B., & Schnitzler, A. (2006). Oscillatory activity reflects the excitability of the human somatosensory system. *Neuroimage, 32*(3), 1231-1236. doi:10.1016/j.neuroimage.2006.06.004

Poonian, S. K., McFadyen, J., Ogden, J., & Cunnington, R. (2015). Implicit Agency in Observed Actions: Evidence for N1 Suppression of Tones Caused by Self-made and Observed Actions. *J Cogn Neurosci, 27*(4), 752-764. doi:10.1162/jocn_a_00745

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences, 22*(02), 253-279.

Pulvermüller, F. (2013). Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. *Brain Lang, 127*(1), 86-103. doi:10.1016/j.bandl.2013.05.015

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nat Rev Neurosci, 7*(7), 511-522. doi:10.1038/nrn1953

Reznik, D., Henkin, Y., Levy, O., & Mukamel, R. (2015). Perceived loudness of self-generated sounds is differentially modified by expected sound intensity. *Plos One, 10*(5), e0127651. doi:10.1371/journal.pone.0127651

Reznik, D., Henkin, Y., Schadel, N., & Mukamel, R. (2014). Lateralized enhancement of auditory cortex activity and increased sensitivity to self-generated sounds. *Nat Commun, 5*, 4059. doi:10.1038/ncomms5059

Reznik, D., Ossmy, O., & Mukamel, R. (2015). Enhanced Auditory Evoked Activity to Self-Generated Sounds Is Mediated by Primary and Supplementary Motor Cortices. *Journal of Neuroscience, 35*(5), 2173-2180. doi:10.1523/Jneurosci.3723-14.2015

Robson, S. E., Brookes, M. J., Hall, E. L., Palaniyappan, L., Kumar, J., Skelton, M., . . . Morris, P. G. (2015). Abnormal visuomotor processing in schizophrenia. *NeuroImage: Clinical*.

Romei, V., Brodbeck, V., Michel, C., Amedi, A., Pascual-Leone, A., & Thut, G. (2008). Spontaneous fluctuations in posterior alpha-band EEG activity reflect variability in excitability of human visual areas. *Cerebral Cortex, 18*(9), 2010-2018. doi:10.1093/cercor/bhm229

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

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*(4), 334-343. doi:10.1111/psyp.12024

Sato, A. (2008). Action observation modulates auditory perception of the consequence of others' actions. *Conscious Cogn, 17*(4), 1219-1227. doi:10.1016/j.concog.2008.01.003

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition, 94*(3), 241-255. doi:10.1016/j.cognition.2004.04.003

Sato, M., Mengarelli, M., Riggio, L., Gallese, V., & Buccino, G. (2008). Task related modulation of the motor system during language processing. *Brain Lang, 105*(2), 83-90. doi:10.1016/j.bandl.2007.10.001

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., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science, 181*(4095), 175-177.

Schlerf, J., Ivry, R. B., & Diedrichsen, J. (2012). Encoding of sensory prediction errors in the human cerebellum. *J Neurosci, 32*(14), 4913-4922. doi:10.1523/JNEUROSCI.4504-11.2012

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

Schröger, E., Marzecová, A., & SanMiguel, I. (2015). Attention and prediction in human audition: a lesson from cognitive psychophysiology. *European Journal of Neuroscience, 41*(5), 641-664. doi:10.1111/ejn.12816

Shergill, S. S., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2003). Two eyes for an eye: the neuroscience of force escalation. *Science, 301*(5630), 187. doi:10.1126/science.1085327

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *Am J Psychiatry, 162*(12), 2384-2386. doi:10.1176/appi.ajp.162.12.2384

Shergill, S. S., White, T. P., Joyce, D. W., Bays, P. M., Wolpert, D. M., & Frith, C. D. (2014). Functional magnetic resonance imaging of impaired sensory prediction in schizophrenia. *JAMA Psychiatry, 71*(1), 28-35. doi:10.1001/jamapsychiatry.2013.2974

Shtyrov, Y., Butorina, A., Nikolaeva, A., & Stroganova, T. (2014). Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. *Proceedings of the National Academy of Sciences, 111*(18), E1918-E1923.

Singelis, T. M. (1994). The Measurement of Independent and Interdependent Self-Construals. *Personality and Social Psychology Bulletin, 20*(5), 580-591. doi:Doi 10.1177/0146167294205014

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., . . . Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences of the United States of America, 111*(42), 15220-15225. doi:10.1073/pnas.1416797111

Sirigu, A., Daprati, E., Ciancia, S., Giraux, P., Nighoghossian, N., Posada, A., & Haggard, P. (2004). Altered awareness of voluntary action after damage to the parietal cortex. *Nat Neurosci, 7*(1), 80-84. doi:10.1038/nn1160

Song, W., & Francis, J. T. (2013). Tactile information processing in primate hand somatosensory cortex (S1) during passive arm movement. *Journal of Neurophysiology, 110*(9), 2061-2070. doi:10.1152/jn.00893.2012

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43*(6), 482-489. doi:doi: 10.1037/h0055479

Stenner, M.-P., Bauer, M., Haggard, P., Heinze, H.-J., & Dolan, R. (2014). Enhanced Alpha-oscillations in Visual Cortex during Anticipation of Self-generated Visual

Stimulation. *Journal of cognitive neuroscience, 26*(11), 2540-2551. doi:10.1162/jocn_a_00658

Stenner, M.-P., Bauer, M., Sidarus, N., Heinze, H.-J., Haggard, P., & Dolan, R. J. (2014). Subliminal action priming modulates the perceived intensity of sensory action consequences. *Cognition, 130*(2), 227-235. doi:10.1016/j.cognition.2013.11.008

Sui, J., Hong, Y. Y., Liu, C. H., Humphreys, G. W., & Han, S. (2013). Dynamic cultural modulation of neural responses to one's own and friend's faces. *Soc Cogn Affect Neurosci, 8*(3), 326-332. doi:10.1093/scan/nss001

Sui, J., Liu, C. H., & Han, S. (2009). Cultural difference in neural mechanisms of self-recognition. *Soc Neurosci, 4*(5), 402-411. doi:10.1080/17470910802674825

Synofzik, M., Lindner, A., & Thier, P. (2008). The cerebellum updates predictions about the visual consequences of one's behavior. *Current Biology, 18*(11), 814-818. doi:10.1016/j.cub.2008.04.071

Tan, H., Wade, C., & Brown, P. (2016). Post-Movement Beta Activity in Sensorimotor Cortex Indexes Confidence in the Estimations from Internal Models. *J Neurosci, 36*(5), 1516-1528. doi:10.1523/JNEUROSCI.3204-15.2016

Teufel, C., Kingdon, A., Ingram, J. N., Wolpert, D. M., & Fletcher, P. C. (2010). Deficits in sensory prediction are related to delusional ideation in healthy individuals. *Neuropsychologia, 48*(14), 4169-4172. doi:10.1016/j.neuropsychologia.2010.10.024

Thut, G., Miniussi, C., & Gross, J. (2012). The Functional Importance of Rhythmic Activity in the Brain. *Current Biology, 22*(16), R658-R663. doi:10.1016/j.cub.2012.06.061

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci, 26*(37), 9494-9502. doi:10.1523/JNEUROSCI.0875-06.2006

Tian, X., & Poeppel, D. (2014). Dynamics of Self-monitoring and Error Detection in Speech Production: Evidence from Mental Imagery and MEG. *J Cogn Neurosci*, 1-13. doi:10.1162/jocn_a_00692

Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *Bmc Neuroscience, 14*. doi:10.1186/1471-2202-14-2

Timm, J., Schönwiesner, M., Schröger, E., & SanMiguel, I. (2016). Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. *Cortex, 80*, 5-20.

Trumpp, N. M., Kliese, D., Hoenig, K., Haarmeier, T., & Kiefer, M. (2013). Losing the sound of concepts: damage to auditory association cortex impairs the processing of sound-related concepts. *Cortex, 49*(2), 474-486. doi:10.1016/j.cortex.2012.02.002

Trumpp, N. M., Traub, F., & Kiefer, M. (2013). Masked priming of conceptual features reveals differential brain activation during unconscious access to conceptual action and sound information. *Plos One, 8*(5), e65910. doi:10.1371/journal.pone.0065910

van Baaren, R. B., Maddux, W. W., Chartrand, T. L., de Bouter, C., & van Knippenberg, A. (2003). It takes two to mimic: Behavioral consequences of self-construals. *Journal of Personality and Social Psychology, 84*(5), 1093-1102. doi:10.1037/0022-3514.84.5.1093

van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci, 28*(8), 1816-1823. doi:10.1523/JNEUROSCI.1853-07.2008

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. *Biol Psychol, 102*, 108-117. doi:10.1016/j.biopsycho.2014.06.007

Van Hulle, L., Juravle, G., Spence, C., Crombez, G., & Van Damme, S. (2013). Attention modulates sensory suppression during back movements. *Consciousness and Cognition, 22*(2), 420-429. doi:10.1016/j.concog.2013.01.011

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M. A., Poort, J., van der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci U S A, 111*(40), 14332-14341. doi:10.1073/pnas.1402773111

Varnum, M. E., Grossmann, I., Kitayama, S., & Nisbett, R. E. (2010). The Origin of Cultural Differences in Cognition: Evidence for the Social Orientation Hypothesis. *Curr Dir Psychol Sci, 19*(1), 9-13. doi:10.1177/0963721409359301

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

Voss, M., Bays, P. M., Rothwell, J. C., & Wolpert, D. M. (2007). An improvement in perception of self-generated tactile stimuli following theta-burst stimulation of primary motor cortex. *Neuropsychologia, 45*(12), 2712-2717. doi:10.1016/j.neuropsychologia.2007.04.008

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., . . . Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences, 41*(5), 929-940. doi:10.1016/j.paid.2006.03.017

Walsh, E., & Haggard, P. (2007). The internal structure of stopping as revealed by a sensory detection task. *Experimental Brain Research, 183*(3), 405-410. doi:10.1007/s00221-007-1128-4

Walter, E., Dassonville, P., & Bochsler, T. M. (2009). A specific autistic trait that modulates visuospatial illusion susceptibility. *J Autism Dev Disord, 39*(2), 339-349. doi:10.1007/s10803-008-0630-2

Wang, J., Mathalon, D. H., Roach, B. J., Reilly, J., Keedy, S. K., Sweeney, J. A., & Ford, J. M. (2014). Action planning and predictive coding when speaking. *Neuroimage, 91*, 91-98. doi:10.1016/j.neuroimage.2014.01.003

Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev, 90*(3), 1195-1268. doi:10.1152/physrev.00035.2008

Wang, Y., Yuan, Q., & Xu, Q. (2008). A Preliminary Study on Self-constructionals Scales (SCS) of Chinese-Version. *Chinese Journal of Clinical Psychology, 16*(6), 602-604.

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

Watkins, K., Patel, N., Davis, S., & Howell, P. (2005). Brain activity during altered auditory feedback: an FMRI study in healthy adolescents. *Neuroimage, 26*(Supp 1), 304.

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary observations on tickling oneself. *Nature, 230*(5296), 598-599.

Weiss, C., Herwig, A., & Schuetz-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

Weisz, N., Hartmann, T., Muller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Frontiers in psychology, 2*. doi:10.3389/fpsyg.2011.00073

Wichmann, F. A., & Hill, J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293-1313.

Williamson, R. S., Hancock, K. E., Shinn-Cunningham, B. G., & Polley, D. B. (2015). Locomotion and Task Demands Differentially Modulate Thalamic Audiovisual

Processing during Active Search. *Curr Biol, 25*(14), 1885-1891. doi:10.1016/j.cub.2015.05.045

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat Neurosci, 3 Suppl*, 1212-1217. doi:10.1038/81497

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science, 18*(7), 600-606. doi:10.1111/j.1467-9280.2007.01946.x

Zhang, S., Xu, M., Chang, W. C., Ma, C., Hoang Do, J. P., Jeong, D., . . . Dan, Y. (2016). Organization of long-range inputs and outputs of frontal cortex for top-down control. *Nat Neurosci*. doi:10.1038/nn.4417

Zhou, M., Liang, F., Xiong, X. R., Li, L., Li, H., Xiao, Z., . . . Zhang, L. I. (2014). Scaling down of balanced excitation and inhibition by active behavioral states in auditory cortex. *Nat Neurosci, 17*(6), 841-850. doi:10.1038/nn.3701