NeuroImage 147 (2017) 895–903

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

The role of brain oscillations in predicting self-generated sounds

Liyu Cao*, Gregor Thut, Joachim Gross

Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, UK

ARTICLE INFO

Keywords:
Auditory perception
Neural oscillations
Prediction
Sensory attenuation

ABSTRACT

Being able to predict self-generated sensory consequences is an important feature of normal brain functioning. In the auditory domain, self-generated sounds lead to smaller brain responses (e.g., auditory evoked responses) compared to externally generated sounds, which is usually referred to as the sensory attenuation effect. Here we investigated the role of brain oscillations underlying this effect. With magnetoencephalography, we show that self-generated sounds are associated with increased pre-stimulus alpha power and decreased post-stimulus gamma power and alpha/beta phase locking in auditory cortex. All these oscillatory changes are correlated with changes in evoked responses, suggesting a tight link between these oscillatory events and sensory attenuation. Furthermore, the pre- and post- oscillatory changes correlate with each other across participants, supporting the idea that they constitute a neural information processing sequence for self-generated sounds. In line with findings of alpha oscillations reflecting feedback and gamma oscillations feedforward processes and models of predictive coding, we suggest that pre-stimulus alpha power represent prediction and post-stimulus gamma power represent prediction error, which is further processed with post-stimulus alpha/beta phase resetting. The correlation between these oscillatory events is further validated with cross-trial analysis, which provides additional support for the proposed information processing sequence that might reflect a general mechanism for the prediction of self-generated sensory input.

1. Introduction

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 (Blakemore et al., 1998; Martikainen et al., 2005; Schafer and Marcus, 1973) – a phenomenon known as sensory attenuation. For example, a 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 (Martikainen et al., 2005).

A forward model has been proposed to account for this effect (Blakemore et al., 1999; Ramnani, 2006; Wolpert and Ghahramani, 2000). The model posits that along with a motor command, an efference copy (von Holst and 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 and results in a modulation of the brain responses depending on the match between the real and the predicted sensory signal (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 is reflected in a decreased evoked response (note that the similar idea was already put forward by von Holst and Mittelstaedt (1950)). 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 rhythms and prediction errors are communicated along feedforward connections via gamma rhythms (Bastos et al., 2015; Michalareas et al., 2016; Wang, 2010).

Our study addresses the following three questions: 1) How is the pre-stimulus prediction of expected sensory consequences of an action reflected in the oscillatory activity of sensory brain areas? Neural oscillations in low frequency bands (below 20 Hz) are likely candidates for the implementation of sensory attenuation for several reasons. First, these oscillations are tightly linked to excitability changes in neural populations (Jensen and Mazaheri, 2010; Thut et al., 2012; Weisz et al., 2011), and therefore may mediate gain control for the processing of incoming sensory information. Second, a number of

---

* Correspondence to: School of Psychology, University of Glasgow, 58 Hillhead Street, G12 8QB Glasgow, UK.
E-mail address: liyuc@psy.gla.ac.uk (L. Cao).

http://dx.doi.org/10.1016/j.neuroimage.2016.11.001
Received 2 September 2016; Accepted 1 November 2016

studies provide 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 et al., 2008). Finally, 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 and Giraud, 2012; Busch et al., 2009; Lakatos et al., 2007; Mathewson et al., 2009). We therefore hypothesized that pre-stimulus changes in low frequency oscillations may reflect a prediction process, which is generated by the forward model to implement the suppression of post-stimulus responses for sensory attenuation. Indeed, some studies already provided evidence that pre-stimulus alpha power is higher in the sensory cortex when speech or visual stimuli are self-induced by movement (Müller et al., 2014; Stenner et al., 2014).

2) How is the prediction error reflected in post-stimulus oscillatory activities? We hypothesised that processes related to prediction error are reflected in gamma oscillations (Bauer et al., 2014; Behroozmand et al., 2016), in line with findings showing that gamma oscillations relay feedforward information (e.g., Michalareas et al., 2016). In the context of sensory attenuation, intracranial recordings from neurosurgical participants showed that gamma power (70–150 Hz) was suppressed in response to speech stimuli during speaking as compared to listening (Flinker et al., 2010). Thus reduced gamma power may indicate decreased prediction errors when the stimulus is better predicted through the forward model during speaking as compared to listening. Furthermore, we planned to use correlation analysis to test if there is a link between the prediction related pre-stimulus oscillations and prediction error related post-stimulus oscillations.

3) How does the post-stimulus attenuation of evoked field responses (reflecting sensory attenuation) relate to post-stimulus changes in the frequency domain (decreases of oscillatory power, phase locking or both)? While a decrease in post-stimulus gamma power has recently been reported (Flinker et al., 2010) and a reduction in evoked field responses is a frequent finding in sensory attenuation paradigms, our understanding of how these processes interrelate is still incomplete. Notably, the decrease in post-stimulus gamma power does not seem to contribute to sensory attenuation as reflected in trial-averaged evoked responses (e.g., attenuation of M100 component), given that a low pass filter at around 40 Hz was applied in many studies on evoked responses (e.g., Baess et al., 2011; Martikainen et al., 2005; Müller et al., 2014). For a better understanding of the post-stimulus neural processes underlying sensory attenuation, we hence conducted analysis at the level of single trials. A reduced amplitude of evoked responses after averaging across trials during sensory attenuation could result from an amplitude reduction in single trials, an increased single trial phase jitter or a combination of both. Moreover, since sensory evoked responses are primarily reflected in an increase of power and/or phase locking in the theta frequency band, one may expect that a reduction of power and/or phase locking in the same frequency band contributes to sensory attenuation. Finally, we used correlation analysis to establish possible links between the neuronal processes in the post-stimulus window across the different, relevant frequency bands (e.g., gamma and alpha).

To answer these questions, we conducted a MEG experiment using a well-established sensory attenuation paradigm in the auditory domain, in which neural responses from self-generated and passive stimuli were compared (Baess et al., 2011; Schafer and Marcus, 1973). After confirming the existence of sensory attenuation in auditory cortex, we performed time-frequency analysis for neural activations in auditory cortex to answer these questions.

2. Methods

2.1. Participants, procedure and recording

14 healthy, right-handed volunteers (6 males, mean age=22.6, SD=1.8; all reported normal hearing) 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 1017 Hz.

The stimulus was a pure tone (1000 Hz, 50 ms in duration, 90 dB sound pressure level) delivered through a plastic tube. There were four conditions (100 trials each). In the passive periodic condition, the auditory stimulus was controlled by the computer and was presented once every three seconds. The passive jittered condition was the same with the passive periodic condition except that the stimulus was presented with a jittered inter-stimulus interval between 2000 and 4000 ms (uniform distribution). In the 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. The motor only condition was the same with the active condition except that no stimulus was presented after each movement. We used a light sensor (instead of a response box) to record the movements without noise associated with the finger movement. Every movement unblocked the beam from the light sensor (placed next to participant's right index finger), which then generated a sound stimulus. Participants were asked to close their eyes during 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. The four conditions were presented in a random order and participants were encouraged to take a break in between. The condition with jittered stimulus presentation served to analyze spontaneous fluctuations in preparedness to sounds (after having identified the oscillatory correlates in the active vs passive periodic comparisons). The motor only condition was not further analyzed here.

2.2. Data analysis

Data analysis was performed with Matlab using FieldTrip toolbox (Oostenveld et al., 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 (less than 1.3% in the active condition). 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. Eye movement and heart artefacts were rejected using ICA. On average, 93.6 (SD: 4.1, minimum: 85), 94.0 (SD: 4.4, minimum: 86) and 94.1 (SD: 3.6, minimum: 88) trials remained after this step for the active, passive periodic and passive jittered condition, respectively.

2.3. Evoked responses

In sensor space analysis, MEG signals were low-pass filtered with 30 Hz cut-off frequency. Original magnetometer signals were converted to planar gradient representation. Three sensors from each hemisphere that were predominantly responding at the latency of the M100