Listen to Yourself: The Medial Prefrontal Cortex Modulates Auditory Alpha Power During Speech Preparation

Nadia Müller1,2, Sabine Leske1, Thomas Hartmann1, Szabolcs Szebényi1,3,4 and Nathan Weisz1

1Center for Mind/Brain Sciences, Università degli Studi di Trento, 38123 Mattarello (TN), Italy, 2Department of Neurology, Epilepsy Center, University Hospital Erlangen, 91054 Erlangen, Germany, 3Department for Neuropsychology, University of Konstanz, 78457 Konstanz, Germany and 4Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, 6500 Nijmegen, The Netherlands

Address correspondence to Dr Nadia Müller, Università degli Studi di Trento, Center for Mind/Brain Sciences, Via delle Regole 101, 38123 Mattarello (TN), Italy. Email: nadia.mueller@gmail.com

How do we process stimuli that stem from the external world and stimuli that are self-generated? In the case of voice perception it has been shown that evoked activity elicited by self-generated sounds is suppressed compared with the same sounds played-back externally. We here wanted to reveal whether neural excitability of the auditory cortex—putatively reflected in local alpha band power—is modulated already prior to speech onset, and which brain regions may mediate such a top-down preparatory response. In the left auditory cortex we show that the typical alpha suppression found when participants prepare to listen disappears when participants expect a self-spoken sound. This suggests an inhibitory adjustment of auditory cortical activity already before sound onset. As a second main finding we demonstrate that the medial prefrontal cortex, a region known for self-referential processes, mediates these condition-specific alpha power modulations. This provides crucial insights into how higher-order regions prepare the auditory cortex for the processing of self-generated sounds. Furthermore, the mechanism outlined could provide further explanations to self-referential phenomena, such as “tickling yourself”. Finally, it has implications for the so far unsolved question of how auditory alpha power is mediated by higher-order regions in a more general sense.

Keywords: effective connectivity, efference copy, MEG, oscillation, top-down

Introduction

Even if our own voice is often intermingled with external voices, the brain can distinguish between speech sounds that are produced by the brain itself and speech sounds that stem from the external world. A vast amount of literature indicates that the auditory cortex is inhibited when we process self generated compared with played back speech sounds. Most of these studies looked at evoked potentials or evoked magnetic fields (Curio et al. 2000; Houde et al. 2002; Ford and Mathalon 2004; Heinks Maldonado et al. 2005; Martikainen et al. 2005; Baess et al. 2011) and showed that evoked activity is reduced in amplitude for self generated speech sounds compared with externally played back speech sounds even if they had the same (or similar) physical characteristics. Most of these results are interpreted in the framework of the so called “efference copies”, meaning that the motor system is sending a copy of the motor command to the respective sensory area, where cortical discharge elicited by this copy is combined with the sensory feedback (Holst and Mittelstaedt 1950; Sperry 1950; Ford and Mathalon 2004). Beyond that, studies on monkeys show that self produced vocalization lead to reduced neuronal firing rates in a majority of auditory cortical neurons (Ploog 1981; Eliades and Wang 2003). In line with that, recordings in epilepsy patients disclosed a suppression of ongoing activity in middle and superior temporal gyrus neurons (Creutzfeld et al. 1989) and a suppression of gamma power in the temporal lobe during speech production (Towle et al. 2008; Flinker et al. 2010). Most interestingly, animal data (Eliades and Wang 2003) and also the data derived from the intracranial recordings (Creutzfeld and colleagues 2003) point to a suppression of brain activity starting already a few hundred milliseconds before sound onset. These findings suggest that the suppression of neuronal activity in the auditory cortex results could, in part, result from internal modulatory mechanisms prior to sound onset.

It has been demonstrated that synchronous oscillatory activity in the alpha frequency band (~10 Hz) are inversely related to the excitability of the respective brain regions (Klimesch et al. 2007; Jensen and Mazaheri 2010), an assumption that has recently received strong support from invasive recordings (Hagens et al. 2011). An increase of alpha power in a sensory region is associated with a functional inhibition of that region when sensory stimuli are processed (Jensen and Mazaheri 2010). This has been shown in the visual modality (Worden et al. 2006; Thut 2006; Romei et al. 2008; Siegel et al. 2008; van Dijk et al. 2008; Bahramisharif et al. 2010; Hanslmayr et al. 2011), in the somatosensory modality (Jones et al. 2010; Hagens et al. 2012; Lange et al. 2012) and recently also in the auditory modality (Gomez Ramirez et al. 1111; Muller and Weisz 2012; Weisz et al. 2014; Frey et al. 2014). The aim of the present study was to investigate if the aforementioned inhibition of the auditory cortex prior and during speech production can also be explained by a top down modulation of auditory alpha power, preceding voice onset. Crucially, any differences in neuronal activity due to differences in sound characteristics (own voice vs. played back own voice) can be ruled out, by measuring brain signals generated in the time intervals preceding sound onset. Our prediction being, that the inhibition of the auditory cortex for self spoken versus played back voices becomes evident in a relative increase in auditory alpha power. Such a finding would give evidence on the processes preceding the modulations of evoked activity in the context of voice perception and would, for the first time, provide evidence on a possible internal mechanism modulating auditory cortex excitability when expecting self generated sensory input.

Materials and Methods

Participants

Twenty right handed volunteers reporting normal hearing participated in the current study (9 m/11 f, mean age 22.6). Participants were
recruited via flyers posted at the University of Konstanz and were paid for their written informed consent prior to taking part in the study. Two participants had to be excluded due to an excessive amount of artifacts.

**Experimental Procedure**

Firstly, participants were introduced to the lab facilities and informed about the experimental procedure, which consisted of 2 phases (voice recordings and main magnetoencephalography [MEG] experiment). For the voice recordings participants were asked to repeat the sound “Aah” 50 times, while their voice was recorded by means of a microphone (Zoom H4 USB microphone). Then on and off set of each “Aah” sound was determined and cut out automatically by a Matlab script so that 50 sound files resulted. After verifying manually that the sounds were cut out correctly they were copied to the stimulation computer for the subsequent MEG experiment. The voice recordings were done in order to keep physical characteristics of the self spoken and externally played back sounds as similar as possible. Loudness of the sounds was adjusted later in the MEG so that participants perceived the self spoken and the externally played back sounds as equally loud. For this purpose a random “Aah” sound was selected and presented to the participant in the MEG scanner. Participants had to rate if the played back sound was louder or weaker compared with the self spoken sound, whereupon loudness of the played back sound was adjusted. This procedure was repeated until the participant rated the played back and self spoken sounds as equally loud. After that the root mean square amplitude of the other recorded sounds was matched to the selected reference sound.

Subsequently, the individual headshapes were collected and the main experiment, consisting of 4 blocks, started. In half of the 4 blocks participants were instructed to say the sound “Aah” after a go signal while in the other half of the blocks they were asked to listen to the sound “Aah” (that was randomly taken from the 50 “Aah” sounds generated before the experiment). Each experimental trial started with a baseline period of 500 ms, upon which a red fixation cross was shown for 1.5 s (preparation period). After 1.5 s, the red fixation cross turned green, which was the go signal instructing participants to either say the sound “Aah” (speak condition) or listen to it (listen condition). The next trial started 2.3 s after sound offset. There were a total of 200 trials. The presentation of visual and auditory stimulation material during MEG recordings was controlled using PsychScope X (Cohen et al. 1995), an open source environment for the design and control of behavioral experiments (http://psy.cki.sissa.it/) and R version 2.11.1 for Mac OS X (http://www.R-project.org). The procedure of the experiment is illustrated in Figure 1.

**Data Acquisition**

The MEG recordings were carried out using a 148 channel whole head magnetometer system (MAGNETM 2500 WH, 4D Neuroimaging, San Diego, USA) installed in a magnetically shielded chamber (Vacuum Hanau). Prior to the recordings, individual head shapes were collected using a digitizer. Participants lay in a comfortable supine position and were asked to keep their eyes open and to fixate a cross displayed by a video projector (JVC TM, DLA G11E) outside of the MEG chamber and projected to the ceiling in the MEG chamber by means of a mirror system. Participants were instructed to hold still and to avoid eye blinks and movements as best as possible. A video camera installed inside the MEG chamber allowed the investigator to monitor participants throughout the experiment. MEG signals were recorded with a sampling rate of 678.17 Hz and a hardwired high pass filter of 0.1 Hz. The recorded and RMS matched “Aah” sounds (see above) were presented through a tube system with a length of 6.1 m and a diameter of 4 mm (Etymotic Research, ER30). Structural images were acquired with a Philips MRI Scanner (Philips Gyroscan ACS T 1.5 T, field of view 256 x 256 x 200 sagittal slices).

**Procedure**

Firstly, participants were introduced to the lab facilities and informed about the experimental procedure, which consisted of 2 phases (voice recordings and main magnetoencephalography [MEG] experiment). For the voice recordings participants were asked to repeat the sound “Aah” 50 times, while their voice was recorded by means of a microphone (Zoom H4 USB microphone). Then on and off set of each “Aah” sound was determined and cut out automatically by a Matlab script so that 50 sound files resulted. After verifying manually that the sounds were cut out correctly they were copied to the stimulation computer for the subsequent MEG experiment. The voice recordings were done in order to keep physical characteristics of the self spoken and externally played back sounds as similar as possible. Loudness of the sounds was adjusted later in the MEG so that participants perceived the self spoken and the externally played back sounds as equally loud. For this purpose a random “Aah” sound was selected and presented to the participant in the MEG scanner. Participants had to rate if the played back sound was louder or weaker compared with the self spoken sound, whereupon loudness of the played back sound was adjusted. This procedure was repeated until the participant rated the played back and self spoken sounds as equally loud. After that the root mean square amplitude of the other recorded sounds was matched to the selected reference sound.

Subsequently, the individual headshapes were collected and the main experiment, consisting of 4 blocks, started. In half of the 4 blocks participants were instructed to say the sound “Aah” after a go signal while in the other half of the blocks they were asked to listen to the sound “Aah” (that was randomly taken from the 50 “Aah” sounds generated before the experiment). Each experimental trial started with a baseline period of 500 ms, upon which a red fixation cross was shown for 1.5 s (preparation period). After 1.5 s, the red fixation cross turned green, which was the go signal instructing participants to either say the sound “Aah” (speak condition) or listen to it (listen condition). The next trial started 2.3 s after sound offset. There were a total of 200 trials. The presentation of visual and auditory stimulation material during MEG recordings was controlled using PsychScope X (Cohen et al. 1995), an open source environment for the design and control of behavioral experiments (http://psy.cki.sissa.it/) and R version 2.11.1 for Mac OS X (http://www.R-project.org). The procedure of the experiment is illustrated in Figure 1.

**Data Analysis**

**Preprocessing**

We analyzed the data sets using Matlab (The MathWorks, Natick, MA, Version 7.5.0 R 2007b) and the Fieldtrip toolbox (Oostenveld et al. 2011). From the raw continuous data, we extracted epochs of 5 s lasting from 2.5 s before onset of the red fixation cross to 2.5 s after onset of the red fixation cross. This was done for the 2 conditions separately (self spoken sound, played back sound) and resulted in 100 trials for each condition. As participants could not avoid blinking sufficiently we decided to perform an independent component analysis (ICA) in order to minimize the influence of the blinks. For ICA correction we first did a coarse visual artifact rejection, removing trials including strong muscle artifacts and dead or very noisy channels. After coarse artifact rejection the data sets (concatenated across conditions) were downsampled to 300 Hz. On a subset of trials an ICA was performed (RUNICA, Delorme and Makeig 2004) and the affected components (eye movements) visually selected. After that ICA was again applied to the data sets of the 2 original conditions and the raw data were reconstructed with the respective components removed. Finally, the resulting data sets were again visually inspected for artifacts and the residual artifactual trials rejected. To ensure a similar signal to noise ratio across conditions, the trial numbers were equalized for the compared conditions (self spoken vs. played back) by random omission (60–90 trials remained). Finally, data were downsampled to 500 Hz.

**Evoked Activity**

In order to replicate the results of previous studies for quality control purposes, we assessed the evoked activity elicited by the sound stimuli. First, data were high pass filtered by 1 Hz and low pass filtered by 45 Hz. Evoked activity was obtained by averaging the single trials. This was done for both conditions separately (self spoken vs. played back, equal trial numbers). Evoked activity was then tested statistically by point wise 2 tailed paired samples t tests.

**Spectral Power Analyses**

Time frequency distributions of the epochs preceding self spoken and externally played back sounds were compared at the sensor and source level. We estimated task related changes in oscillatory power using a multitaper FFT time frequency transformation (Percival 1993)
with frequency dependent Hanning tapers (time window: ∆z = 4/Δf, sliding in 50 ms steps) were calculated power from 3 to 30 Hz in steps of 1 Hz and for both conditions separately. The obtained time frequency representations were then baseline normalized (baseline: 400 to 100 ms before onset of the red fixation cross, relative change).

In order to test if power modulations are significantly different between conditions (expecting self spoken vs. played back own voice) we performed a nonparametric cluster based permutation test of the baseline normalized time frequency representations (Maris and Oostenveld 2007), test based on 2 tailed paired t tests. This test was chosen to correct for multiple comparisons.

As a next step, we estimated the generators of the sensor effects in source space using the frequency domain adaptive spatial filtering algorithm Dynamic Imaging of Coherent Sources (DICS, Gross et al. 2001). For each participant an anatomically realistic headmodel (Nolte 2003) was created and leadfields for a 3 dimensional grid covering the entire brain volume (resolution: 1 cm) calculated. Together, with the sensor level cross spectral density matrix (2 time intervals early 0.5 s and late 1 1.5 s, 13 ± 3 Hz, multitaper analysis, conditions concated nated), we could estimate common spatial filters, optimally passing information for each grid point while attenuating influences from other regions for the frequency and time window of interest (according to the cluster permutation test at sensor level: 0.5 1.5 s, 13 ± 3 Hz).

The common spatial filters were then applied to the Fourier transformed data for both conditions separately (same parameters). After that the resulting activation volumes were interpolated onto the individual MRI. In cases where we could not get a structural scan (5 out of 18), we created “pseudo” individual MRIs that were created based on an affine transformation of the headshape of an Montreal Neurological Institute (MNI) template and the individually gained headshape points. The interpolation activation volumes were then normalized to a template MNI brain provided by the SPM8 toolbox (http://www.fil.ion.ucl.ac.uk/spm/software/spm8). Finally, source solutions for the 2 conditions were compared using a voxel wise dependent samples t statistic. From that analysis the left auditory cortex (Brodmann Areas 21/22 and Brodmman Area 41), the right precentral cortex and the medial prefrontal cortex (BA 8) were derived as main regions showing a significant increase of alpha power for self generated versus externally played back sounds. This is illustrated in the results. To get a better estimate of how alpha power in the auditory cortex is modulated we averaged the power within the left auditory cortex for each participant and for both conditions separately. These values were then tested against baseline values by 2 tailed paired t tests and for both conditions separately. Beyond that, we tested the baseline values of the peak condition against the baseline values of the listen condition again by a 2 tailed paired t test to rule out the possibility that the relative effects were determined due to baseline differences.

Power Power Correlations

After spectral power analysis we aimed at shedding light onto the question of how the condition specific relative alpha increases in the auditory cortex are mediated. We therefore correlated left auditory alpha power with low frequency power (from 2 26 Hz) in all other regions of the brain (for 1 1.5 s). We did this in MNI grid space.

First, a template grid was created (using a template head model based on a segmented template MNI brain provided by the SPM8 toolbox). Using this template grid an individual grid was generated by warping the template grid to the individual MRI for each participant separately. Importantly, the warped individual grids have an equal number of points with equal positions in MNI space, so that the individual grids of different participants can be compared directly (grid points Subject 1 correspond to grid points of Subject 2). These individual MNI grids were then used for source analysis. Source analysis was done for the single trials and using the DICS beam former algorithm (MNI grid, 1 1.5 s after red trigger onset, 13 Hz ± 3, same settings as for alpha power source analysis despite the use of the individual MNI grids). We calculated source solutions for frequencies from 2 to 26 Hz in increments of 2 Hz. Thereby, power values for each participant, each condition, each trial, each frequency and each grid point were obtained. We then calculated correlations between alpha power at the reference voxel, which was defined as the grid point being closest to the main alpha power effect as derived from source analysis (MNI coordinates: 55 28 2, left auditory cortex), and all other grid points. We repeated this for all frequencies (2 26 Hz) and Fisher z transformed the correlation values afterwards. We thereby obtained a 2 D matrix for both conditions (grid points × frequencies).

Finally, the frequency × grid point maps were tested for significant differences between conditions across subjects using a nonparametric cluster based permutation test (Maris and Oostenveld 2007); neighbors were defined as grid points that had a distance of <3 cm resulting in average 75 neighbors per grid point, which reflects ~3% of all grid points). This analysis yielded that alpha power in the left auditory cortex is strongly correlated with low frequency power (6 14 Hz) in the medial prefrontal cortex when participants expect a self generated sound. To get a better estimate of how connectivity between the medial prefrontal cortex and the auditory cortex is mediated in both conditions separately we averaged the correlation values within the significant region for each participant and for both conditions separately. These values were then tested against correlation values (within the same region) that were obtained during the baseline period by 2 tailed paired t tests, and for both conditions separately.

Partial Directed Coherence Between Auditory Cortex and Medial Frontal

As a final step we wanted to elucidate the direction of information flow between the auditory cortex (MNI coordinates: 55 28 2) and the medial prefrontal cortex (peak voxel of correlation effect, MNI coordinates: 4 4 6), assessed via partial directed coherence (PDC, Baccala and Sameshima 2001). PDC is a measure of effective coupling that is based on multivariate autoregressive (MVAR) modeling. For a pair of voxels the information flow can be assessed in both directions. We first projected the raw time series into source space by multiplying the raw time series for both conditions separately with a common spatial filter. The spatial filter was created using the LCVM beamformer (Van Veen et al. 1997) and the concatenated data of both conditions (2 26 Hz, time window including baseline and activation 0.5 to 1.5 s). We thereby obtained time series for both conditions and both sources (auditory cortex, medial prefrontal cortex) separately. For these time series a MVAR model was fitted (“bsmat”). The model order was set to 15, according to previous analysis approaches (Supp et al. 2007; Weisz et al. 2014). Then a Fourier transform was performed on the resulting coefficients of the MVAR model. These Fourier transformed coefficients were then used to calculate partial directed coherence between the auditory and medial prefrontal cortex. The PDC values were base line normalized using the baseline interval ( 0.5 to 0) by first subtracting and then dividing the values by the values of the baseline interval. Finally, the PDC values were tested for differences between conditions (speak vs. listen) using paired t tests.

Results

The current study aimed at disentangling brain activity preceding the processing of participants’ own voice that was either self spoken or played back externally. We investigated brain activity on a local and on a network level in the time interval before voice onset and with a focus on low frequency oscillatory power.

Evoked Responses

The event related response was significantly stronger for the externally played back speech sound compared with the self generated one between 150 and 200 ms after sound onset (un corrected). This is comparable to the previous literature (Curio et al. 2000; Houde et al. 2002; Flinker et al. 2010). Results are shown in Figure 2 (upper panel).

Pre voice Power Differences Sensor Level

In a first step, we assessed differences in low frequency power for self spoken versus externally played back voices on sensor level. We found a significant power increase (cluster P < 0.05) peaking between 10 and 16 Hz and encompassing frontal and
Source Localization of Alpha Power Differences Before Voice Onset

In order to get a better estimate of where in the brain the low frequency power modulations take place, we performed a source analysis of the power modulations derived from sensor level (10-16 Hz, 0.5-1 s and 1-1.5 s after onset of the red fixation cross). Source results indicate that, besides extra auditory areas (right precentral, medial dorsolateral prefrontal cortex), the left auditory cortex shows a strong relative increase of alpha power, becoming most evident in the last part of the preparation period (1-1.5 s, P<0.01, including Brodmann areas 21/22/41). There were no significant differences in auditory activity during baseline (P=0.19). Interestingly, when extracting the power modulations for the 2 conditions separately, it turned out that the relative increase of alpha power is due to a decrease of alpha power compared with baseline when participants prepare to process their externally played back voice (P<0.01), while this alpha power decrease seems to be abolished (no statistical difference compared with baseline) when participants expect to process their self spoken voice. For comparison see Figure 2 (lower panel).

Beside the auditory power modulations, alpha power was increased in the medial dorsolateral prefrontal cortex (P<0.01, BA 9 and 32) and the right precentral cortex (P<0.01, BA 4). Within these extra auditory regions alpha power was significantly increased compared with baseline when participants prepared to speak (P<0.01), while we found no significant differences in alpha power compared with baseline when participants prepared to listen (Fig. 3).

Note, that alpha power in the right precentral region was already modulated during baseline (probably due to the blocked design). During baseline alpha power was significantly decreased in the "speak blocks" compared with the "listen blocks". Baseline differences of the entire brain are shown in Supplementary Fig. 1.

Modulations of Connectivity with the Left Auditory Cortex Before Voice Onset

The second main analysis tackled the question of how the condition specific alpha power modulations in the auditory cortex are mediated. Therefore we looked at power power correlations between alpha power in the left auditory cortex and low frequency power in the other regions of the brain. This was conducted on a single trial level and for the time interval preceding voice onset when auditory alpha power modulations were strongest (1-1.5 s after onset of the red fixation cross). We took strong power power correlations as an indicator of a possible communication between the accordant brain regions (Park et al. 2011). A cluster based permutation test revealed the medial prefrontal cortex as the main region differentially communicating with the left auditory cortex when participants expected to listen to their self produced versus externally played back voice (cluster P<0.05, BA 11). The effect was strongest for power correlations between 6 and 14 Hz. When looking at the effect more precisely, it turned out that communication between the left auditory cortex and the medial prefrontal cortex was significantly enhanced compared with baseline when participants expected their self spoken voice (P<0.05) and significantly reduced compared with baseline when they expected their played back voice (P<0.05). See Figure 4 (upper panel) for comparison.
Finally we wanted to assess the direction of the information flow between the left auditory and the medial prefrontal cortex. In order to do that, we calculated Partial Directed Coherence between the 2 regions. Information flow from the auditory cortex to the medial prefrontal cortex showed no significant differences between conditions (all $P > 0.05$). In contrast, information flow from the medial prefrontal cortex to the left auditory cortex was significantly enhanced when participants prepared for speaking ($P < 0.05$). The effect was strongest for the time interval slightly preceding the main auditory alpha power modulations (0.7–1 s after onset of the red fixation cross). For comparison see Figure 4 (lower panel).

**Discussion**

In the present study we investigated if and how pre speech brain activity is modulated when participants expect a self spoken sound. We concentrated on modulations in low frequency power with a focus on the auditory cortex and its communication with non auditory regions. Results show that the alpha power suppression, typically present when participants expect sounds, is absent in the left auditory cortex when participants expect their own voice. They further show that the medial prefrontal cortex mediates this effect. The absence of the auditory alpha power suppression can be interpreted as inhibition of the auditory cortex when participants expect self spoken sounds. This is in line with the previous literature postulating an inhibition of the auditory cortex when processing self spoken sounds and extends the previous literature by showing that brain activity in the auditory cortex is inhibited already before sound onset and on a macroscopic scale. So far a suppression of brain activity in the auditory cortex before speech onset has only been shown for ongoing activity in single neurons (Creutzfeldt et al. 1989; Eliades and Wang 2003) and not for local field potentials. In addition, the mediation of the auditory cortex's increase in alpha power via medial prefrontal cortex suggests a mechanism of how auditory cortex excitability is adjusted. This gives new insights into the processes going on within the tested paradigm and, crucially, provides first evidence of how auditory alpha power could be causally modulated by higher order regions.

**Auditory Alpha Power Modulations**

As described above we found a decrease of auditory alpha power compared with baseline in the “listen” condition and an abolishment of that effect in the “speak” condition. There were no significant differences of auditory alpha activity during the baseline. This points to a relative inhibition of that brain region (Klimesch et al. 2007; Jensen and Mazaheri 2010), in this case a relative inhibition of the auditory cortex (Gomez Ramirez et al. 2011; Muller and Weisz 2012; Weisz et al. 2014) and is therefore consistent with previous literature postulating an inhibition of the auditory cortex when processing self generated sounds compared with externally played back ones (Creutzfeldt et al. 1989; Curio et al. 2000; Houde et al. 2002; Eliades and Wang 2003; Ford and Mathalon 2004). A growing number of studies in the visual and somatosensory system (Worden et al. 2000; Thut 2006; Romei et al. 2008; Siegel et al. 2008; van Dijk et al. 2008;
Jones et al. 2010; Händel et al. 2011; Haegens et al. 2012; Lange et al. 2012) and also in the auditory system (Gomez Ramirez et al. 2011; Muller and Weisz 2012; Müller et al. 2013; Weisz et al. 2014; Frey et al. 2014) have convincingly shown that alpha power modulations have the potential to dynamically adjust the excitability of brain regions with respect to the according task demands (e.g., attention, near threshold detection, memory), and thereby make neuronal processing adaptive and most effective (Jensen and Mazaheri 2010). Based on that the current results can be interpreted as follows: the auditory system is by default in an inhibited state to filter out the vast amount of auditory information it is exposed to. This is in accord with the observation that alpha in sensory cortices is high when subjects are awake and not engaged in any task (Basar et al. 1997; Klimesch et al. 2007; Jensen and Mazaheri 2010). If participants expect to process an external sound they reduce auditory alpha power in order to enhance processing capacities for the incoming auditory stimulus, as it is the case in “normal” auditory perception (Gomez Ramirez et al. 2011; Hartmann et al. 2012; Muller and Weisz 2012; Weisz et al. 2014). In contrast, if participants expect a self generated sound, auditory alpha power is kept high meaning that, in that case, processing capacities are not enhanced compared with baseline. We thus propose that even if alpha power is not increased beyond baseline levels, the “relative increase” (i.e., the one arising from the condition contrast) goes in line with the hypothesis of an inhibition of processing when participants prepare to listen to their self spoken voice. Such a mechanism could explain why we process and perceive self spoken sounds differentially from played back ones. Interestingly, it has been shown that an increase in alpha power has an impact on neuronal firing (Haegens et al. 2011) and also on event related responses (Basar and Stampsper 1985; Ergenoglu et al. 2004; Klimesch et al. 2007). The present findings could thus be the prerequisite of the reported inhibition of the auditory cortex during the processing of self generated sounds as reported in literature (Creutzfeldt et al. 1989; Curio et al. 2000; Houde et al. 2002; Eliades and Wang 2003; Ford and Mathalon 2004; Heinks Maldonado et al. 2005), however such a direct relation would have to be tested within further studies. Also findings postulating that the suppression in the auditory cortex is very specific for the self generated sounds and does not block auditory processing in general thereby will have to be taken into account (McGuire et al. 1996; Heinks Maldonado et al. 2005; Fu et al. 2006). We here suggest that the abolishment of the usual alpha power reduction when expecting self generated sounds is an active and top down modulated process helping to differentiate between self spoken and externally played back sounds. According to the connectivity results, which are explained in more detail below, this seems to be indeed the case.

**Left Auditory Alpha Power Modulations**

We found that the condition specific alpha power modulations are lateralized to the left auditory cortex. This is in line with...
the literature on the processing of self-generated speech sounds showing that the suppression effects are dominant in the left auditory cortex (Curio et al. 2000; Houde et al. 2002; Heinks Maldonado et al. 2005; Kauramaki et al. 2010).

Non auditory Alpha Power Modulations
Despite the absence of alpha power suppression in the auditory cortex we derived an increase of alpha power in a prefrontal region (BA 8), encompassing the medial dorsolateral prefrontal cortex and a power increase in the right precentral cortex. Brodmann area 8 is involved in planning, cognitive control, and maintaining attention (MacDonald 2000; Seamans et al. 2008) and also in guiding decisions (Seamans et al. 2008). A modulation of that region during speech preparation could point to an active disengagement from the expected and to be inhibited auditory input. However, the role of alpha power in higher order regions is not understood so far so that possible modulations within these prefrontal regions during speech preparation will have to be tested by further studies.

Concerning the precentral cortex, it is important to clarify that the effect is due to differences in the baseline (see Supplementary Fig. 1 for comparison). During baseline alpha power is reduced in the left and right precentral cortex for the “speak” compared with the “listen” condition. The left precentral alpha power decrease is still present in the presentation interval what is in line with the modulations we would expect in the precentral cortex during motor preparation (Jasper and Penfield 1949; Pfurtscheller et al. 1996; Sauseng et al. 2009). Interestingly, however, the right precentral power decrease for the “speak” versus “listen” conditions disappears, leading to the impression of an increase of alpha power in the precentral cortex when subjects prepare to speak. These hemispheric differences could be due to the dominance of the left hemisphere for speech production (Llorens et al. 2011; Price et al. 2011).

Auditory Alpha Power Modulation Mediated by the Medial Prefrontal Cortex
Another crucial question to answer was how the condition specific auditory alpha power modulations are mediated. We could elucidate the medial prefrontal cortex (BA 11) as the main region showing increased communication in this process. Crucially, this increase in communication was driven by an increase of unilateral communication from the medial prefrontal cortex to the left auditory cortex. This provides clear evidence for a condition specific modulation of auditory alpha power communicated by the medial prefrontal cortex. The medial prefrontal cortex is involved in self referential thinking (meta analyses, Johnson et al. 2002; Heatherton et al. 2006; Northoff et al. 2006; van der Meer et al. 2010), in comparing the self with others (Moore et al. 2013) and self reflective judgments (Macrae et al. 2004). It is thus also from a theoretical point of view very likely that the medial prefrontal cortex has a crucial role in mediating the excitability of the auditory cortex when we process our own voice. Interestingly, the increase of information flow from the medial prefrontal cortex to the auditory cortex was strongest shortly before the relative increase of auditory alpha power. All in all, we suggest that the medial prefrontal cortex triggers alpha power in the auditory cortex so that self generated sounds are processed less intensely and we can easily distinguish between self generated and external sounds.

Conclusion
With the present study our aims were to disentangling brain activity associated with the expectation of a self spoken sound. We concentrated on alpha power modulations in the auditory cortex and on how these modulations are mediated by non auditory brain regions. We can show that the typical alpha power suppression when participants expect external sounds is absent in the left auditory cortex when participants expect self spoken sounds. This points to a relative inhibition of the auditory cortex that is already present before speech onset and is in line with the previous literature showing a suppression of brain activity mainly during sound production. Importantly, the current findings complement the existing evidence on modulations of evoked activity and rather local modulations in auditory activity (as derived by ECoG and animal studies) by elucidating that also the state/excitability of the auditory cortex is modulated when processing self generated sounds, which became evident in the auditory alpha power modulations. As second main finding we demonstrate that the medial prefrontal cortex, a region known for self referential processes, mediates these condition specific alpha power modulations. This provides crucial insights into how higher order regions prepare the auditory cortex for the processing of self generated sounds and seems interesting as a mechanism itself having the potential to explain similar phenomena related to self referential processing like for instance “tickling yourself”. Beyond that, the findings also have implications for the so far unsolved question of how auditory alpha power is mediated by higher order regions in a more general sense.

Supplementary Material
Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Funding
This work was supported by the European Research Council (grant number: 283404) and the Deutsche Forschungsgemeinschaft (grant number: 4156/2 1).

Notes
We thank Nick Peatfield for proofreading the manuscript. Conflict of Interest: None declared.

References


