NeuroImage 147 (2017) 32–42

Contents lists available at ScienceDirect

NeuroImage

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

# Auditory cortical delta-entrainment interacts with oscillatory power in multiple fronto-parietal networks

Anne Keitel*, Robin A.A. Ince, Joachim Gross, Christoph Kayser*

*Institute of Neuroscience and Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, UK*

ARTICLE INFO

*Keywords:*
Auditory entrainment
Speech processing
MEG
Delta band
Prosodic parsing

ABSTRACT

The timing of slow auditory cortical activity aligns to the rhythmic fluctuations in speech. This entrainment is considered to be a marker of the prosodic and syllabic encoding of speech, and has been shown to correlate with intelligibility. Yet, whether and how auditory cortical entrainment is influenced by the activity in other speech–relevant areas remains unknown. Using source-localized MEG data, we quantified the dependency of auditory entrainment on the state of oscillatory activity in fronto-parietal regions. We found that delta band entrainment interacted with the oscillatory activity in three distinct networks. First, entrainment in the left anterior superior temporal gyrus (STG) was modulated by beta power in orbitofrontal areas, possibly reflecting predictive top-down modulations of auditory encoding. Second, entrainment in the left Heschl's Gyrus and anterior STG was dependent on alpha power in central areas, in line with the importance of motor structures for phonological analysis. And third, entrainment in the right posterior STG modulated theta power in parietal areas, consistent with the engagement of semantic memory. These results illustrate the topographical network interactions of auditory delta entrainment and reveal distinct cross-frequency mechanisms by which entrainment can interact with different cognitive processes underlying speech perception.

## 1. Introduction

While listening to speech, rhythmic auditory cortical activity aligns to the quasi-rhythmic regularities arising from stress, syllabic rate, or phonemes. This entrainment of brain activity to speech is particularly prominent in the delta (below 4 Hz) and theta (4–8 Hz) frequency bands (Ahissar et al., 2001; Aiken and Picton, 2008; Di Liberto et al., 2015; Giraud and Poeppel, 2012; Gross et al., 2013; Kayser et al., 2015; Luo and Poeppel, 2007; Ng et al., 2012) and is particularly strong in the auditory cortex (Aiken and Picton, 2008; Gross et al., 2013). Given that the degree of entrainment is predictive of speech intelligibility and comprehension rates (Ahissar et al., 2001; Ding et al., 2014; Ghitza et al., 2012; Ghitza and Greenberg, 2009; Luo and Poeppel, 2007; Peelle and Davis, 2012; Peelle et al., 2013), the alignment of auditory cortical activity to the speech envelope has been proposed to subserve a number of important functions, such as the parsing or encoding of acoustic and phonological features (Cogan and Poeppel, 2011; Ghitza, 2013; Ghitza et al., 2012; Giraud and Poeppel, 2012; Howard and Poeppel, 2010; Peelle and Davis, 2012), or the selection of sensory streams (Schroeder and Lakatos, 2009). While auditory entrainment is frequently used as a marker for speech encoding, the network-level interactions that shape the underlying neural processes remain unknown.

Language processing depends on a large network of interconnected brain areas (Bornkessel-Schlesewsky et al., 2015; Friederici and Gierhan, 2013; Hickok and Poeppel, 2007; Poeppel, 2014; Rauschecker and Scott, 2009), with ventral and dorsal pathways linking auditory cortex with fronto-parietal regions implicated in extracting acoustic, lexical or categorical information (Alho et al., 2014; Bornkessel-Schlesewsky et al., 2015; Friederici and Gierhan, 2013; Hickok and Poeppel, 2007; Schomers et al., 2015; Smalle et al., 2015; Wilson et al., 2004). Interestingly, many studies have shown that auditory perception depends on the state of rhythmic activity, such as theta, alpha or beta activity, in auditory or frontal regions (Henry et al., 2014; Henry and Obleser, 2012; Neuling et al., 2012; Ng et al., 2012; Strauss et al., 2015). Furthermore, oscillatory power in different frequency bands and fronto-parietal brain regions has been quantified in numerous studies investigating cognitive control or attention (de Lange et al., 2008; Hipp et al., 2011; Schroeder and Lakatos, 2009; Stoll et al., 2016; van de Vijver et al., 2011). In the context of speech perception, the state of alpha power has been found to co-vary with experimental manipulations such as attention (Kelly et al., 2006; Klimesch et al., 1998), listening effort (Obleser et al., 2012), and speech intelligibility (Obleser and Weisz, 2012). In addition, changes in theta power have been linked to lexical retrieval processes and

---

* Corresponding authors.
*E-mail addresses:* anne.keitel@glasgow.ac.uk (A. Keitel), christoph.kayser@glasgow.ac.uk (C. Kayser).

http://dx.doi.org/10.1016/j.neuroimage.2016.11.062
Received 13 July 2016; Accepted 25 November 2016

semantic working memory (Bastiaansen and Hagoort, 2006; Bastiaansen et al., 2005), while changes in beta and gamma power have been associated with a predictive coding framework, representing top-down predictions and bottom-up prediction errors during language processing (Arnal and Giraud, 2012; Arnal et al., 2011; Lewis and Bastiaansen, 2015). However, the functional interpretation of the processes indexed by the oscillatory activity extracted from fronto-parietal regions usually rests on the implicit assumption that changes in this oscillatory activity correlate with, or even causally relate to, changes in the auditory cortical speech representations provided by the entrainment of neural activity to the speech envelope (Arnal and Giraud, 2012; Kayser et al., 2015; Park et al., 2015). For example, changes in frontal alpha during enhanced listening effort have been suggested to relate to changes in the precision of auditory cortical entrainment (Kayser et al., 2015). However, no study to date has directly tested the hypothesis that changes in fronto-parietal oscillatory processes directly correlate with the fidelity of dynamic speech representations in temporally entrained auditory cortical activity.

To address the functional dependence between auditory speech entrainment and the state of oscillatory activity in fronto-parietal regions, we used source-localised MEG data obtained while participants listened to a continuous story (Gross et al., 2013). We chose an unbiased approach, in that we did not limit our analyses to specific and pre-defined anatomical regions of interest within frontal or parietal lobes. The reason for this was that previous studies have implied a vast number of regions in language processes, or provided only very coarse localizations of these. Likewise, we included a wide range of frequency bands when assessing the state of oscillatory activity in fronto-parietal regions (from delta to gamma), as previous studies have implied a wide range of rhythmic processes in speech perception. As a result, we systematically quantified the relation between auditory cortical speech entrainment and states of oscillatory power in fronto-parietal regions across a wide range of frequency bands.

## 2. Materials and methods

### 2.1. Participants and data acquisition

MEG data were acquired from 23 healthy, right-handed participants (12 female, mean age 26.9 ± 7.9 years [$M \pm SD$]) as part of a previous study (Gross et al., 2013). All participants provided written informed consent prior to testing. The experiment was approved by a local ethics committee (University of Glasgow, Faculty of Information and Mathematical Sciences), and conducted in compliance with the Declaration of Helsinki.

MEG-recordings were obtained with a 248-magnetometers whole-head MEG system (MAGNES 3600 WH, 4-D Neuroimaging; sampling rate: 1017 Hz). The participants' head positions were measured at the beginning and end of each run via 5 coils placed on the forehead and behind the ears. Head position was co-digitised with head-shape (FASTRAK®, Polhemus Inc., VT, USA). Participants sat upright and fixated a cross projected centrally on screen with a DLP projector while listening to two auditory presentations in a pseudo-randomised order. Sounds were presented binaurally via plastic earpieces and 5-m long plastic tubes connected to a sound pressure transducer. Stimulus presentation was controlled with Psychtoolbox (Brainard, 1997) for MATLAB (The MathWorks, Inc.).

### 2.2. Auditory stimuli

The 'forward' story condition consisted of an approximately 7-min long narration ('Pie-man' told by Jim O'Grady, recorded at 'The Moth' storytelling event in New York, 2012). This narration comprises approximately 950 words, with syllables spoken at an average rate of 6.8 Hz. This is in line with studies that establish the syllabic rate typically in the theta-range (4–8 Hz) (Cotton, 1936; Hyafil et al., 2015; Poeppel, 2003). In natural speech, the rate of prominent, or stressed, syllables is approximately one third of all syllables (e.g., Kochanski et al., 2005), essentially placing it in the delta-range (1–4 Hz) (Goswami and Leong, 2013; Greenberg et al., 2003). Consequently, entrainment in the delta band has been proposed to reflect prosodic fluctuations in speech (Ghitza, 2013; Ghitza et al., 2012). The 'backward' condition consisted of the same story played backwards and served as an unintelligible control condition.

### 2.3. Extraction of speech envelope

From the waveform of the acoustic stimulus we computed the wideband speech envelope by band-pass filtering (3rd order Butterworth filter, forward and reverse) into eight bands in the range of 100–10,000 Hz, equidistant on the cochlear frequency map (Smith et al., 2002). Individual band-limited envelopes were obtained using the magnitude of the Hilbert transform and were subsequently averaged to obtain the wide band speech envelope (Chandrasekaran et al., 2009; Drullman, 1995; Gross et al., 2013; Kayser et al., 2015; Smith et al., 2002). The envelope was resampled to 150 Hz for subsequent analysis.

### 2.4. MEG data processing

Data were analysed using MATLAB (The MathWorks, Inc.), including external toolboxes, such as FieldTrip (Oostenveld et al., 2011), and custom-written routines. The MEG signal was detrended and resampled to 150 Hz. Data were band-limited to seven frequency bands (*delta* 1–4 Hz, *theta* 4–8 Hz, *alpha* 8–12 Hz, *low beta* 12–18 Hz, *beta* 18–24 Hz, *high beta* 24–36 Hz, *gamma* 30–48 Hz), using FIR filters (forward and reverse, with 60 dB stop-band attenuation, 1-Hz transition bandwidth, and 0.01 dB pass-band ripple).

### 2.5. MEG source localisation

Individual, T1-weighted structural magnetic resonance images (MRIs) were manually co-registered to the MEG coordinate system by using participants' digitised head shapes. MRIs were further realigned with individual head shapes through an iterative closest point (ICP) algorithm (Besl and Mckay, 1992). MRIs were then segmented to obtain a representation of the brain, including grey and white matter, and cerebrospinal fluid. A single-shell model was used to construct a volume conduction model (Nolte, 2003). Individual anatomical MRIs were linearly transformed to a template (MNI) brain using Fieldtrip/SPM5. Sensor level data were transformed into source space using the linear constraint minimum variance (LCMV) beamformer (Van Veen et al., 1997) on a 4-mm regular grid covering the entire brain (7% regularisation). The optimal orientation for each dipole was computed using the SVD approach. We used the AAL atlas (Automated Anatomical Labelling atlas) to parcellate the template brain into 116 anatomical areas (Tzourio-Mazoyer et al., 2002). As the superior temporal gyrus (STG) comprises a very large and functionally differentiated area in the AAL atlas, we divided it further into an anterior and posterior section (e.g., Friederici, 2002; Hickok and Poeppel, 2007). Specifically, the median of voxel positions along the horizontal plane (i.e., y-coordinates in MNI space) was used as threshold for the division between anterior and posterior parts to obtain equally-sized anterior and posterior STG partitions.

### 2.6. Region-specific analyses

We quantified the entrainment of rhythmic activity to speech within three auditory regions in each hemisphere (Heschl's Gyrus, anterior/posterior superior temporal gyrus [aSTG/pSTG]). For each auditory region, the bandpass-filtered MEG source-space data were Hilbert-transformed to derive the instantaneous phase for each time and grid