NeuroImage xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

NeuroImage

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

# Interacting parallel pathways associate sounds with visual identity in auditory cortices

Jyrki Ahveninen [a,*], Samantha Huang [a], Seppo P. Ahlfors [a], Matti Hämäläinen [a,b,c], Stephanie Rossi [a], Mikko Sams [c], Iiro P. Jääskeläinen [c]

[a] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital/Harvard Medical School, Charlestown, MA, USA
[b] Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA
[c] Brain and Mind Laboratory, Department of Neuroscience and Biomedical Engineering, Aalto University School of Science, Espoo, Finland

ARTICLE INFO

*Article history:*
Received 15 June 2015
Accepted 20 September 2015
Available online xxxx

ABSTRACT

Spatial and non-spatial information of sound events is presumably processed in parallel auditory cortex (AC) "what" and "where" streams, which are modulated by inputs from the respective visual-cortex subsystems. How these parallel processes are integrated to perceptual objects that remain stable across time and the source agent's movements is unknown. We recorded magneto- and electroencephalography (MEG/EEG) data while subjects viewed animated video clips featuring two audiovisual objects, a black cat and a gray cat. Adaptor-probe events were either linked to the same object (the black cat meowed twice in a row in the same location) or included a visually conveyed identity change (the black and then the gray cat meowed with identical voices in the same location). In addition to effects in visual (including fusiform, middle temporal or MT areas) and frontoparietal association areas, the visually conveyed object-identity change was associated with a release from adaptation of early (50–150 ms) activity in posterior ACs, spreading to left anterior ACs at 250–450 ms in our combined MEG/EEG source estimates. Repetition of events belonging to the same object resulted in increased theta-band (4–8 Hz) synchronization within the "what" and "where" pathways (e.g., between anterior AC and fusiform areas). In contrast, the visually conveyed identity changes resulted in distributed synchronization at higher frequencies (alpha and beta bands, 8–32 Hz) across different auditory, visual, and association areas. The results suggest that sound events become initially linked to perceptual objects in posterior AC, followed by modulations of representations in anterior AC. Hierarchical what and where pathways seem to operate in parallel after repeating audiovisual associations, whereas the resetting of such associations engages a distributed network across auditory, visual, and multisensory areas.

© 2015 Published by Elsevier Inc.

## Introduction

Perceptual objects refer to constructs that associate individual events with a specific source agent, and integrate them into entities that remain stable across time and the agent's movements in space (Bizley and Cohen, 2013). Indeed, one readily perceives dynamic information from different senses as belonging to a single entity, for example, when a cat moves about and meows as it goes along, rather than seeing and hearing a series of disparate events. In contrast to this everyday experience, a prominent organizational principle of sensory systems is parallel processing where information of different stimulus attributes is processed in separate sensory channels and feature pathways. The human visual pathway is believed to branch into two major streams that process spatial ("where") and object-identity related ("what") information (Ungerleider and Haxby, 1994). Many studies (Ahveninen et al., 2013; Ahveninen et al., 2006; Altmann et al., 2007; Clarke et al., 2002; Lomber and Malhotra, 2008; Rauschecker and Scott, 2009; Rauschecker and Tian, 2000) support the existence of an analogous division between parallel posterior "where" and anterior "what" pathways in auditory cortices (AC) as well (however, for alternative AC models, see also Bizley and Cohen, 2013; Griffiths and Warren, 2002; Recanzone and Cohen, 2010). What is still unknown is where and at which latencies spatial and identity-related feature information of auditory and visual events becomes linked to objects that remain perceptually constant despite their fluctuations across time and space.

The prevailing view has been that integration of information across feature pathways and sensory modalities occurs through hierarchical convergence of parallel pathways at higher-level association areas (Konorski, 1967). Previous studies suggest that the parallel "where" and "what" auditory pathways converge with their visual counterparts in dorsal/superior frontal (frontal eye fields, FEF; dorsolateral prefrontal cortex, DLPFC) vs. inferior frontal cortex (IFC) areas, respectively (Alain et al., 2001; Altmann et al., 2012; Arnott et al., 2005; Barrett and Hall,

* Corresponding author at: MGH/MIT/HMS-Martinos Center, Bldg. 149 13th Street, Charlestown MA 02129, USA. Fax: +1 617 726 7422.
*E-mail address:* jyrki@nmr.mgh.harvard.edu (J. Ahveninen).

http://dx.doi.org/10.1016/j.neuroimage.2015.09.044
1053-8119/© 2015 Published by Elsevier Inc.

Please cite this article as: Ahveninen, J., et al., Interacting parallel pathways associate sounds with visual identity in auditory cortices, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.09.044

UNCORRECTED PROOF

2006; Clarke et al., 2002; Leung and Alain, 2011; Maeder et al., 2001; Romanski et al., 1999; Weeks et al., 1999). Furthermore, in humans there is abundant evidence of super-additive (or sub-additive) multisensory interactions in areas such as the posterior superior temporal sulcus (pSTS) and middle temporal gyrus (MTG) (Beauchamp et al., 2004; Bischoff et al., 2007; Calvert et al., 2000; Raij et al., 2000; Werner and Noppeney, 2010), which may also include neurons contributing to face and voice identity integration (von Kriegstein et al., 2005). The existence of a slow higher-order feature-attribute and crossmodal binding mechanism is also supported by psychophysical evidence of temporal judgments (Fujisaki and Nishida, 2010). The hierarchical convergence model is confronted by increasing evidence of interactions across sensory areas during multisensory stimulus processing (Senkowski et al., 2008). Indeed, even primary AC neurons can be crossmodally modulated (Bizley et al., 2007; Budinger et al., 2006; Pekkola et al., 2005; Raij et al., 2010), and these influences become progressively stronger in non-primary areas (Calvert et al., 1997; Kayser et al., 2005; Romanski et al., 1999; Smiley et al., 2007).

In contrast to multisensory areas such as pSTS and MTG that are activated by a variety of object-related events (Beauchamp et al., 2004), modulations of "unimodal" AC areas have been mostly studied with visual cues related to sound production (Calvert et al., 1997; Jääskeläinen et al., 2004b; Pekkola et al., 2005; van Wassenhove et al., 2005). These cues have, thus, predicted spectrotemporal (McGurk and MacDonald, 1976) or spatial (Bonath et al., 2007) properties of isolated events, rather than tapping into perceptual objects that remain constant despite their fluctuations across time and space. Interestingly, indices of a specific role of AC areas in object processing were found in a single-unit study in rhesus monkeys, which documented a small number of neurons that responded to specific voice types coming from a particular direction in posterior non-primary ACs (Tian et al., 2001). Although most of the neurons sampled in this study were predominantly spatially selective, the linking of feature representations to more complex perceptual entities could be initiated in this area. Since response latencies were not examined in this pioneering monkey study, it remains to be determined whether such an effect occurs at early latencies driven by bottom-up processing, or whether it results from later feedback signals from higher-order cortical regions. Furthermore, whether these neurons are sensitive to corresponding inputs from other sensory modalities was not tested.

The goal of the present study is to examine non-invasively in humans in which anatomical areas and at which latencies information from parallel auditory pathways becomes associated with the perceptual object's stable visual identity, beyond cues related to voice/sound production itself. Functional properties of AC neurons have been often studied by measuring neuronal adaptation, or repetition suppression, of neuronal responses to recurring stimulus attributes (Jääskeläinen et al., 2011). In populations sensitive to the feature of interest, a release from adaptation is expected when the consecutive "adaptor" and "probe" stimuli differ from one another with respective to this attribute (Ahveninen et al., 2006; Altmann et al., 2007). In ACs, such effects can be measured by using the MEG/EEG response N1, which shows adaptation effects with, for example, speech-related visual stimuli (Jääskeläinen et al., 2004b). Analogously, one might also test whether two consecutive stimuli are related to the same multisensory object or not: a release from adaptation is expected after two events related to different objects (e.g., a cat of different color but similar voice).

Accumulating evidence further suggests that crossmodal influences on AC function are mediated by inter-regional synchronization of neuronal oscillations (for a review, see Senkowski et al., 2008). Whereas local integration effects may occur at the high-frequency gamma band (~30–100 Hz) (for a review, see Jensen et al., 2007), longer-range coupling might be supported by lower frequencies at which neurons have more robust spike timing delays (Engel et al., 2001; Ermentrout and Kopell, 1998; Roelfsema et al., 1997). Non-invasive evidence of crossmodal oscillatory mechanisms has been obtained using both EEG (Doesburg et al., 2008; Hipp et al., 2011; van Driel et al., 2014; von Stein et al., 1999) and MEG (Alho et al., 2014) in humans. Long-range synchronization at the 2–4 Hz delta and 4–8 Hz theta bands could help provide low-level timing information of the speech rhythm from visual system to AC neurons (Arnal et al., 2011; Luo et al., 2010; Schroeder et al., 2008). Oscillatory phase synchronization at the alpha (8–16 Hz) (van Driel et al., 2014) and gamma (30–100 Hz) (Doesburg et al., 2008) frequencies may also play a role in temporal integration of audiovisual inputs during multisensory object formation. Therefore, it is conceivable that coordinated oscillatory interactions across ACs and visual areas also contribute to the formation of perceptual objects.

Here, we specifically hypothesized that regions of AC will show release from adaptation when a visual cue changes the object identity, indicative of a presence of neuron groups within AC that are associated with perceptual objects. We further hypothesized that the activity in such AC regions will exhibit oscillatory functional connectivity with other brain areas processing auditory and visual object-feature and spatial information, and that the connectivity patterns will be different for repeating vs. changing audiovisual object associations. To test these hypotheses, we used an adaptation paradigm that taps into the formation of non-speech related audio-visual perceptual objects. Stimulus-related and interregional oscillatory processes during audiovisual processing were estimated using a multimodal cortically constrained MEG/EEG source modeling approach (Huang et al., 2014; Lin et al., 2006; Sharon et al., 2007). In this approach, the MEG and EEG source locations are restricted to the cortical mantle derived from anatomical MRI to reduce the potential solution space (Dale and Sereno, 1993). Additional improvements are achieved by combining the complementary information provided by simultaneously measured MEG and EEG, which helps provide better accuracy and smaller point spread of the source estimates than either modality alone (Ding and Yuan, 2013; Henson et al., 2009; Liu et al., 2002; Sharon et al., 2007).

## Materials and methods

### *Subjects, task and stimuli*

Eleven subjects (age 21–50 years, five females) were studied. One subject of the initial $N = 12$ was excluded due to excessive blink artifacts. The subjects had normal hearing and normal or corrected-to-normal vision. Human subjects' approval was obtained and voluntary consents were signed before each measurement, in accordance with the experimental protocol approved by the Massachusetts General Hospital Institutional Review Board.

During MEG/EEG acquisition, the subjects watched 3.1-s animated movie clips consisting of audiovisual Adaptor/Probe stimulus pairs (Fig. 1a). In each clip, cartoon faces of a gray cat and a black cat first appeared at the screen 5° to the right and left from a central fixation mark. Each cat subtended ~1.7° x 1.7°. At $t_1 = 1$ s, the black cat meowed (adaptor stimulus, meow sound and a visual presentation of the opening of the mouth, duration 450 ms). After that both cats moved to the center and returned to either their original places or switched sides. At $t_2 = 2.4$ s, the cat that had ended up in the black cat's initial position uttered a second meow (Probe stimulus, duration 450 ms), which sounded exactly the same as the Adaptor meow. Thus, 50% of the trials included a *repetition of audiovisual events linked to the same object* ($AV_{Repeat}$): the black cat meowed twice in a row in the same place. The other 50% of the trials included a *visual identity change* ($A_{Repeat}V_{Change}$), that is, the black and gray cats meowed in a row with exactly the same voice in the same location. The meow sound stimulus (Fig. 1b), modified from a public domain recording, was presented at a clearly audible and comfortable level, being always simulated from the same direction as the corresponding visual stimulus using generic head-related transfer functions (Algazi et al., 2001). The position of the meowing cat was counterbalanced across the subjects ($N_{Right} = 6$, $N_{Left} = 5$) and controlled as a between-subject nuisance factor in our statistical analyses.

UNCORRECTED PROOF

Please cite this article as: Ahveninen, J., et al., Interacting parallel pathways associate sounds with visual identity in auditory cortices, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.09.044

Download English Version:

https://daneshyari.com/en/article/6024398

Download Persian Version:

https://daneshyari.com/article/6024398

Daneshyari.com