Primary auditory cortex activation by visual speech: an fMRI study at 3 T

Johanna Pekkola,1,2,3 CA Ville Ojanen,1,3 Taina Autti,2 Iiro P. Jääskeläinen,1,3,4 Riikka Möttönen,1,3 Antti Tarkiainen3 and Mikko Sams1,3

1 Laboratory of Computational Engineering; 2 Helsinki Medical Imaging Center, Helsinki University Central Hospital, Helsinki, Finland; 3 Advanced Magnetic Imaging Centre, Helsinki University of Technology, P.O. Box 3000, FIN-02015 HUT, Espoo; 4 Massachusetts General Hospital, Massachusetts Institute of Technology, Harvard Medical School Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA, USA

CA, 1,3 Corresponding Author and Address: pekkojo@lce.hut.c.t
Received 5 October 2004; accepted 12 November 2004

Recent studies have yielded contradictory evidence on whether visual speech perception (watching articulatory gestures) can activate the human primary auditory cortex. To circumvent confounds due to inter-individual anatomical variation, we defined our subjects’ Heschl’s gyri and assessed blood oxygenation-dependent signal changes at 3 T within this confined region during visual speech perception and observation of moving circles. Visual speech perception activated Heschl’s gyri in nine subjects, with activation in seven of them extending to the area of primary auditory cortex. Activation was significantly stronger during visual speech perception than during observation of the moving circles. Further, a significant hemisphere by stimulus interaction occurred, suggesting left Heschl’s gyrus specialization for visual speech processing.

Key words: Auditory cortex; fMRI; Human; Visual speech perception

INTRODUCTION

Observing lip movements of a speaker facilitates perception of heard speech [1] and can even alter the auditory percept [2]. While these phenomena have been behaviourally well described, the underlying neural mechanisms have remained unclear, one of the most pertinent questions being the stage at which visual information influences the acoustic information processing stream.

Several EEG and MEG studies have demonstrated modulation of auditory cortex responses to speech sounds by simultaneously presented visual articulations [4–8], and recent functional magnetic resonance imaging (fMRI) studies have suggested that the adjacent multisensory superior temporal sulcus acts as an integration site [9], mediating enhanced activity of the sensory-specific auditory cortex [10] during audiovisual speech perception.

Concurrently, fMRI and positron emission tomography studies have consistently reported secondary auditory cortex and superior temporal sulcus activations by visible articulatory gestures [11–15], in the absence of heard speech. Whether this effect is limited to the hierarchically higher areas, or whether primary auditory cortex also contributes to the processing of visual speech information has, however, remained an open question [11,14]. Anatomically, primary auditory cortex (auditory koniocortex, BA41) is located within the transverse Heschl’s gyrus on the superior temporal plane. It comprises approximately the medial half of Heschl’s gyrus [3], its lateral part and the surrounding temporal plane being covered by lesser granular secondary auditory cortex (BA42/22).

The first group-level fMRI study showing human auditory cortex activation during visual perception of articulatory gestures [11] also reported significant signal changes within primary auditory cortex. However, a subsequent per-participant fMRI study contradicted these results [14]. The imaged area in that study consisted of four 10 mm-thick slices centered on Heschl’s gyri. Although several clusters exhibited activation by both auditory stimulation and visual speech, after standard-space alignment their peaks fell outside the range of Heschl’s gyri, according to the probabilistic map of Penhune et al. [16].

The location, form, and extent of Heschl’s gyri vary substantially between individuals [3,16], rendering group-level analyses wherein individual brains are spatially averaged ineffective in assessing signal changes within primary auditory cortex. The probabilistic map of Penhune et al. [16] is based on manually defining Heschl’s gyri in high-resolution MR images of a limited population (20 subjects) and transferring this information to a standard template. Thus, verification of primary auditory cortex activity can still be complicated by individually variable gyral morphology and errors in standard-space alignment.

Here, we studied primary auditory cortex activation during visual perception of articulatory gestures using an individual definition of primary auditory cortex based on cortical anatomy. We hypothesized that the possibly weak and spatially limited primary auditory cortex activations could be more readily observed through individual-level analysis that circumvents standard-space alignment. Further, to enhance sensitivity, we used higher magnetic
field strength (3 T) and smaller voxel size (3.4 × 3.4 × 3.4 mm) than previously [11,14].

SUBJECTS AND METHODS

Ten healthy, right-handed, native Finnish-speaking (aged 21–30 years, 7 males) volunteers signed an informed consent prior to participation. The study was approved by the ethics committee of the Hospital District of Helsinki and Uusimaa.

The stimuli (vowels and circles) were prepared from video clips of a female speaker who silently articulated four Finnish vowels (/a/, /o/, /i/, /y/). Vowels showed the articulations per se (Fig. 1). To design a moving control stimulus with analogous time-varying characteristics (circles), a semi-transparent regular-shaped blue oval was added to the mouth region of a still image of the talker’s face and designed to expand and constrict in four alternative directions (vertical, horizontal, left and right oblique; Fig. 1), following a time-scale and range of movement similar to that of the lip movements. A fixation mark was constantly visible on the mouth region of the face.

We used a block-design with intermittent 30-s active and baseline blocks. The experiment consisted of two 6-min runs (vowels and circles), the order of which was randomized across subjects. The runs shared the same baseline condition (watching a still face image). During the active blocks, the subjects were shown either the visual articulations (vowels run) or the moving circles (circles run).

All subjects were told that the articulations represented vowels (/a/, /o/, /i/, /y/) and were allowed to view the stimuli in silence before scanning. They were instructed to observe what the person was saying and press a button when any articulation occurred twice in succession. During the circles run, they were to detect occurrence of two circles successively moving in the same direction. To balance out the motor responses, target sequence occurrence was set to 10/run. Excluding this, stimuli were presented in randomized order (duration 560 ms, stimulus-onset asynchrony 620 ms).
A 3 T GE Signa system and a quadrature birdcage head coil was used to obtain 120 gradient echo-planar volumes (TE 32 ms, TR 2500 ms, flip angle 90°, 28 axial oblique slices, voxel size 3.4 × 3.4 × 3.4 mm) per run, depicting blood oxygenation-dependent (BOLD) contrast [17]. A high-resolution T1-weighted sagittal whole-head volume (voxel size 1 × 1 × 1.4 mm) was acquired for anatomical co-alignment. Data were analyzed using fMRI Expert Analysis Tool (FEAT) software, version 3.1. Pre-processing included discarding the first two volumes, non-brain tissue extraction, motion correction [18], Gaussian spatial smoothing (FWHM 5 mm), intensity normalization, and high-pass temporal filtering. Time-series analyses were done using general linear model with local autocorrelation correction [19]. The statistical maps were thresholded at Z > 2.3 for subsequent clustering and a cluster-wise significance threshold set at p < 0.01 (corrected for multiple comparisons across the acquisition volume).

A region depicting Heschl’s gyrus was defined separately for both hemispheres in each subject’s high-resolution MRIs. Maximum percentage signal changes within each Heschl’s gyrus were determined for each run and entered into a 2 × 2 ANOVA with stimulus type and hemisphere as the factors. We used a definition for Heschl’s gyri similar to that of Penhune et al. [16]: its anterior boundary defined as the anterior edge of the first Heschl’s gyrus, posterior boundary as Heschl’s sulcus, medial boundary as a straight line between the medial ends of the transverse sulcus and Heschl’s sulcus, and lateral boundary as the tip of Heschl’s gyrus or, if Heschl’s gyrus merged to the temporal plane without a clear end-point, as continuation of the anterior and posterior boundaries to the lateral border of the temporal plane. Although these landmarks are distinguishable in high-resolution MRIs, differentiating primary from secondary auditory cortex within Heschl’s gyri is non-trivial, since no transition between cytoarchitectonic areas can be seen in MRIs. Defining primary auditory cortex functionally would require a stimulus that exclusively and reliably activates primary auditory cortex. In our pilot studies, pure tone stimuli also activated the lateral parts of Heschl’s gyri, making them ineffective in depicting our region of interest.

A more specific determination for primary auditory cortex might be achieved by identifying tonotopic-responsive auditory areas. The tonotopic organization may, however, approach the limits of the current spatial resolution of fMRI, and its complexity is still unclear [20]. Therefore, we relied on neuroanatomical data localizing the primary auditory cortex to the medial half of each Heschl’s gyrus [3]; the extension of the activations to this area was assessed visually by two experienced radiologists (T.A. and J.P.).

RESULTS

The subjects detected over 90% of the target sequences during both runs (vowels 94 ± 7% and circles 91 ± 2.3%, mean ± s.e.m.).

During visual speech perception, nine subjects showed significant (Z > 2.3, p < 0.01, corrected) BOLD signal changes within the left Heschl’s gyrus. Notably, in seven of those (subjects 1–7, Fig. 2) the activation extended to the medial half of the gyrus, the anatomical area of the left primary auditory cortex. Five of these subjects also exhibited activation in the right Heschl’s gyrus, and the signal changes extended to its medial half in three.

Six subjects also showed significant activation within the left Heschl’s gyrus when watching the moving circles, and in three of these activation extended to the medial half of the gyrus. This activation was bilateral in one subject. All subjects showing Heschl’s gyrus activation while watching the moving circles also displayed activation while watching articulatory gestures. One subject showed no significant activation during either condition.

The 2 × 2 ANOVA demonstrated significant main effects for both stimulus type and hemisphere (Fig. 3): visual speech perception activated Heschl’s gyri more strongly than did watching the moving circles (F(1,9) = 7.9, p < 0.05), and maximum percent BOLD signal changes were greater in the left than in the right Heschl’s gyrus (F(1,9) = 10.3, p = 0.01). There was also a significant hemisphere × stimulus interaction (F(1,9) = 14.8, p < 0.01) due to greater signal strength during visual speech perception than during observation of the moving circles in the left Heschl’s gyrus (p < 0.01).

DISCUSSION

Our results support the findings of Calvert et al. [11], demonstrating human primary auditory cortex activation during visual observation of articulatory gestures, and suggest as a new finding that the left Heschl’s gyrus is specifically involved in the processing of visual speech signals.

The discrepancy between our findings and the results of a previous per-participant study [14] can in part be explained by the relative weakness of the activations: signal changes observed were somewhat weaker than those typically elicited by auditory stimulation at 3 T [21]. Since BOLD signal change amplitude increases with increasing magnetic field strength, conducting the experiment at a higher field strength than before [11,14] might have enhanced detection of auditory cortex activations during visual perception.

Several subjects exhibited activation within Heschl’s gyri while watching the moving circles. Although auditory cortex activation by relatively simple visual stimuli has been documented in animals [22], the activation was significantly stronger during visual speech perception than when viewing circles. This, combined with left-hemisphere dominance for visual speech, suggests that visually perceived articulations possess features that specifically activate the left Heschl’s gyrus (for corroborating results from a study with biological motion vs silent articulation stimuli, see [12]).
We suggest two possible underlying mechanisms for these findings. First, invasive animal studies suggest that auditory cortex receives converging input from the visual modality, possibly via superior temporal sulcus [22], and may thus be recruited in visual information processing. Interestingly, even direct neuronal projections from primary auditory areas to visual striate cortex have been revealed in monkeys [23], although the existence of reciprocal connections is tentative. Audiovisual interactions within the time course of sensory processing have also recently been detected in humans [24]. In future, combining diffusion tensor imaging and audiovisual stimulation designs, for instance, would provide the potential of examining the connectivity of the human auditory cortex.

Alternatively, viewing articulations may enhance primary auditory cortex activity caused by the acoustic noise from the MR scanner analogous to the modulation of auditory cortex function during audiovisual speech perception [4–10]. Further, based on the over-learned association between auditory and visual components of speech, multisensory attention may modulate primary auditory cortex activity by tuning the auditory system to anticipate acoustic features of speech while watching articulatory gestures. The recent observation of activation in the lateral part of Heschl’s gyrus in a silent event-related fMRI experiment [12] makes it, however, doubtful that effects related to scanner noise can fully explain the finding. An experiment comparing blocked and silent fMRI designs might reveal whether concurrent acoustic input is a prerequisite for visual speech-induced primary auditory cortex activation.

Subvocalization during visual speech perception offers yet another possible mediating mechanism. We cannot rule this out, since our subjects received no instructions to avoid subvocalizing during the task, but activity related to auditory imagery is suggested to be limited to secondary auditory areas not involving Heschl’s gyri [25].

CONCLUSION

Our results support the view that visual speech perception either activates the hierarchically early human primary auditory cortex or modulates its activity. Whether this effect occurs early or late in the time domain is beyond the temporal resolution of fMRI and needs to be clarified in future studies. The specific aspects of visual speech contributing to primary auditory cortex activation also need elucidation, as well as the mechanisms mediating this phenomenon.

ACKNOWLEDGEMENTS

This work was supported by grants from the Academy of Finland (206368), the US National Institute of Health (NIH-ROI HD040712), and the Kordelin Foundation. We warmly thank Carol Norris for her help in preparing the manuscript.

REFERENCES