Clustered fMRI of speech production

Peter Sörös[1,2], Lisa Guttman Sokoloff[3,4], Arpita Bose[5], Anthony R. McIntosh[6,7], Simon J. Graham[1,6,8] & Donald T. Stuss[6,7,9]

[1]Imaging Research, Sunnybrook Health Sciences Centre, Toronto, [2]Toronto Rehabilitation Institute, Toronto, [3]Communication Disorders, Baycrest, Toronto, [4]Department of Speech-Language Pathology, University of Toronto, [5]Department of Psychology, University of Windsor, [6]The Rotman Research Institute, Baycrest, Toronto, [7]Department of Psychology, University of Toronto, [8]Department of Medical Biophysics, University of Toronto, [9]Department of Medicine (Neurology, Rehabilitation Science), University of Toronto, Ontario, Canada

peter.soros@sri.utoronto.ca

Introduction

The production of speech is a **highly complex motor task** that involves approximately 100 orofacial, laryngeal, pharyngeal, and respiratory muscles. The neural basis of the exact and rapid coordination of these highly overlearned speech movements is not yet entirely clear. The investigation of the neural correlates of overt speech production by fMRI has been complicated by **artefacts** due to motion-correlated head movements and movements of the articulatory organs.

The **aim** of this study [1] was to trace the **distributed neural network** involved in overt speaking of non-lexical speech sounds of different complexity. To minimize speech-related artefacts, **clustered image acquisition** was used.

Methods

Participants
BOLD fMRI was acquired in 9 healthy volunteers (4 women, 5 men; average age = 26 years; age range = 22 - 32 years).

Experimental tasks
Subjects were asked to repeat acoustically presented **sub-lexical speech sounds of different complexity** and to perform oral movements without vocalization. The required responses were the vowel (**V**) "ah", a consonant-vowel (**CV**) syllable (either "pa", "ka", or "ta"), a **$C_1VC_2VC_3V$** utterance ("pataka"), and **oral movements** (opening the mouth or protruding the lips). Instructions were transmitted through an fMRI compatible audio system (Avotec, Stuart, FL, USA) at a constant onset-to-onset interstimulus interval of 10 s. Six experimental sessions were performed. Each session comprised 6 separate blocks of speech, 2 blocks of oral movement (50 s each) and 3 blocks of baseline (30 s). During the baseline, no verbal instructions were given and no responses were performed (Fig. 1).

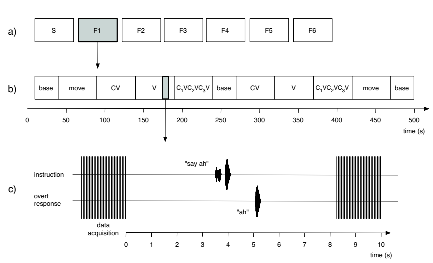

FIGURE 1: Timing diagram illustrating behavioral tasks and clustered fMRI data acquisition. a) The sequence of structural (S) and functional imaging (sessions F1-F6). b) The sequence of tasks in session F1. This sequence is also used in F3 and F5, while F2, F4, and F6 consist of a different randomization of movement and articulatory tasks. Base denotes baseline; move, oral movement; V, CV, $C_1VC_2VC_3V$, speech production tasks. c) Event-related clustered fMRI acquisition. Both the auditory cue and the verbal response fall within the silent interval between multislice data acquisition. The speech waveforms represent the instruction (upper trace) and the overt response (lower trace), both recorded by an fMRI-compatible microphone.

Magnetic resonance imaging
Imaging was performed on a 3 Tesla GE MRI system with a quadrature birdcage head coil. For BOLD fMRI, a **spiral-in/out pulse sequence** [2] and **clustered image acquisition** [3] were used (TR = 10 s). Based on the timecourse of the hemodynamic response function, the offset of the verbal instructions was set approximately 5 s prior to the midpoint of the data acquisition. High-resolution, T1-weighted images were acquired for structural reference.

Data analysis
Analysis of fMRI data was carried out using the software library **FSL** [4]. Preprocessing included linear registration, brain segmentation, spatial smoothing and intensity normalization. The 6 fMRI sessions obtained for each participant were analyzed independently using **general linear modeling**. Statistic parametric (Z score) images were thresholded using clusters determined by $Z > 4$ and a (corrected) cluster significance threshold of $P = 0.01$ (for the task vs. baseline conditions) or by $Z > 2.3$, $P < 0.01$ (for the comparison of task conditions). A **mixed-effects analysis** was then carried out to analyze effects across the 6 sessions of single participants. Finally, a **mixed-effects analysis** was performed across all subjects.

Results

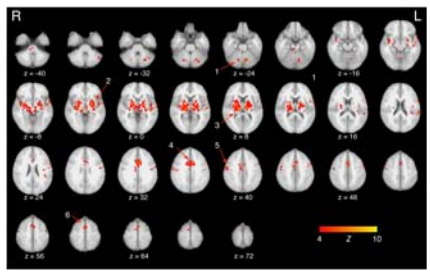

FIGURE 2: Group activation map, V vs. baseline condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen primarily in the bilateral posterior cerebellar lobe (1), the basal ganglia (2), the thalamus (3), the cingulate motor area (4), the primary motor cortex (5), and the supplementary motor area (6).

Vowel vs. orofacial movement

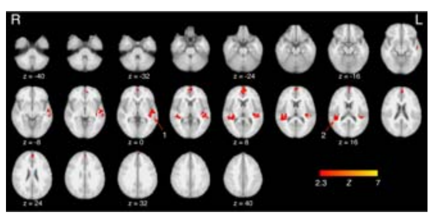

FIGURE 3: Group activation map, V vs. movement condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen primarily in the left middle (1) and the bilateral superior temporal gyrus (2).

Polysyllabic utterance vs. vowel

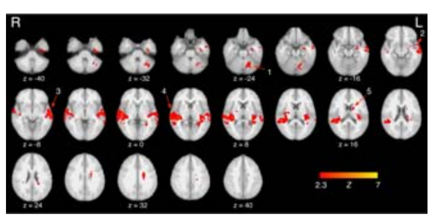

FIGURE 4: Group activation map, $C_1VC_2VC_3V$ vs. V condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen in the left cerebellum (1), the left inferior frontal gyrus (2), the bilateral middle temporal gyrus (3), the bilateral superior temporal gyrus (4) and in the left caudate nucleus (5).

Right middle temporal and right superior temporal activation was seen in the CV condition relative to the vowel condition (data not shown). Activation in the left middle temporal gyrus and in the right basal ganglia was present in the $C_1VC_2VC_3V$ vs. CV comparison (data not shown).

Discussion

The results of this study identified a **distributed articulo-phonologic network** that consisted of cortical and subcortical motor areas as well as bilateral temporal regions (Figs. 2 and 5).

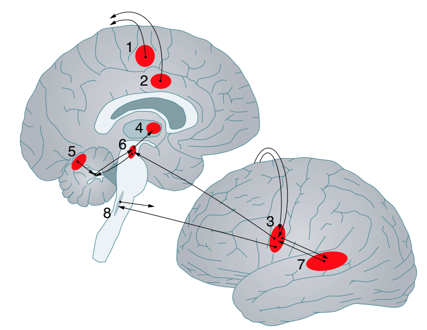

FIGURE 5: The neural network of speech production. Main areas of activation (red) and main fiber tracts are shown. The supplementary motor area (1) and the cingulate motor areas (2) are connected with the primary motor cortex (3). Subcortical activation was found in the thalamus (4), the basal ganglia (not shown), the red nucleus (6) and in the vermal and paravermal cerebellum (5). In addition, the bilateral posterior superior temporal gyrus (7) was activated. The brain stem nuclei innervating the articulatory organs were outside the field of view (8).

Compared to the production of an isolated vowel ("ah"), the production of a $C_1VC_2VC_3V$ sequence ("pataka") poses a higher demand of phonologic processing and articulatory sequencing. Increased task demand was represented by **increased activation in a left-lateralized caudate nucleus-cerebellum circuit**, presumably involved in speech motor control, particularly in the production of rapid movement sequences.

In addition, production of a $C_1VC_2VC_3V$ sequence was associated with increased activation in subregions of the **bilateral temporal lobe**, especially the bilateral posterior superior temporal gyrus. Our data support the notion that the **posterior temporal gyrus** is, dependent on the phonologic demand, involved in the phonologic processing of speech production [5].

References

[1] P. Sörös, L. Sokoloff, A. Bose, A. McIntosh, S. Graham, D. Stuss, Clustered functional MRI of overt speech production, Neuroimage 2006 Apr 20; [Epub ahead of print].

[2] G. Glover, C. Law, Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts, Magn Reson Med 46 (3) (2001) 515–22.

[3] W. Edmister, T. Talavage, P. Ledden, R. Weisskoff, Improved auditory cortex imaging using clustered volume acquisitions, Hum Brain Mapp 7 (2) (1999) 89–97.

[4] S. Smith, M. Jenkinson, M. Woolrich, C. Beckmann, T. Behrens, H. Johansen-Berg, P. Bannister, M. De Luca, I. Drobnjak, D. Flitney, R. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. Brady, P. Matthews, Advances in functional and structural MR image analysis and implementation as FSL, Neuroimage 23 Suppl 1 (2004) S208–19.

[5] K. Okada, G. Hickok, Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI, Brain Lang 98 (1) (2006) 112–7.

Acknowledgments

This study was supported by the Heart and Stroke Foundation of Ontario Centre for Stroke Recovery, the Ontario Research and Development Challenge Fund, GE Healthcare Canada, Baycrest's Poslun Centre for Stroke and Cognition and the J.F. McDonnell Foundation.