ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 32 (2006) 376 – 387

Clustered functional MRI of overt speech production

Peter Sörös,[a,b,*] Lisa Guttman Sokoloff,[c,d] Arpita Bose,[e] Anthony R. McIntosh,[f,g] Simon J. Graham,[a,f,h] and Donald T. Stuss[f,g,i]

[a]*Imaging Research, Sunnybrook and Women's College Health Sciences Centre, Toronto, Ontario, Canada*
[b]*Toronto Rehabilitation Institute, Toronto, Ontario, Canada*
[c]*Communication Disorders, Baycrest, Toronto, Ontario, Canada*
[d]*Department of Speech–Language Pathology, University of Toronto, Ontario, Canada*
[e]*Department of Psychology, University of Windsor, Ontario, Canada*
[f]*The Rotman Research Institute, Baycrest, Toronto, Ontario, Canada*
[g]*Department of Psychology, University of Toronto, Ontario, Canada*
[h]*Department of Medical Biophysics, University of Toronto, Ontario, Canada*
[i]*Department of Medicine (Neurology, Rehabilitation Science), University of Toronto, Ontario, Canada*

Received 27 October 2005; revised 21 February 2006; accepted 24 February 2006
Available online 2 May 2006

To investigate the neural network of overt speech production, event-related fMRI was performed in 9 young healthy adult volunteers. A clustered image acquisition technique was chosen to minimize speech-related movement artifacts. Functional images were acquired during the production of oral movements and of speech of increasing complexity (isolated vowel as well as monosyllabic and trisyllabic utterances). This imaging technique and behavioral task enabled depiction of the articulo-phonologic network of speech production from the supplementary motor area at the cranial end to the red nucleus at the caudal end. Speaking a single vowel and performing simple oral movements involved very similar activation of the cortical and subcortical motor systems. More complex, polysyllabic utterances were associated with additional activation in the bilateral cerebellum, reflecting increased demand on speech motor control, and additional activation in the bilateral temporal cortex, reflecting the stronger involvement of phonologic processing.

Keywords: Speech production; Articulation; Phonologic processing; Cerebellum; Superior temporal gyrus; fMRI; Clustered image acquisition

Introduction

The production of speech is a highly complex motor task that involves approximately 100 orofacial, laryngeal, pharyngeal, and respiratory muscles (Levelt, 1989). Precise and expeditious timing of these muscles is essential for the production of temporally complex speech sounds, which are characterized by transitions as short as 10 ms between frequency bands (Fitch et al., 1997) and an average speaking rate of approximately 15 sounds per second (Levelt, 1989). The neural basis of the exact and rapid coordination of these highly overlearned movements is not yet entirely clear (Munhall, 2001).

For years, the analysis of brain lesions and the correlation between lesion locations and behavioral deficits were the most important sources of knowledge about the mechanisms underlying speech production (Huber et al., 2000; Rorden and Karnath, 2004). The seminal lesion studies of Paul Broca indicated that the production of speech relies on the functional integrity of the left inferior frontal gyrus (Broca, 1861). The investigation of patients with stroke-related apraxia of speech (AOS) added further insights. AOS is conceptualized as a deficit of transforming phonologic plans and articulatory motor programs to appropriate speech movements (Darley et al., 1975; Croot, 2002). Recent research suggests that patients with AOS fail to retrieve the motor patterns essential for speech production (Aichert and Ziegler, 2004). MRI studies of these patients revealed lesions of and around the left inferior frontal gyrus, in particular Brocas area (Hillis et al., 2004), lesions of the left insula (Dronkers, 1996; Nagao et al., 1999), and of the basal ganglia (Peach and Tonkovich, 2004). Nonfluent progressive aphasia, presenting with apraxia of speech and syntactic deficits, is similarly associated with left inferior frontal and insular atrophy (Gorno-Tempini et al., 2004). However, the results of noninvasive neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), provide growing evidence that complex human skills are not located in highly specialized brain areas but are organized in networks connecting several different areas of both hemispheres instead (Sporns et al., 2004). Thus, a widespread network is most likely to underlie the

* Corresponding author. Imaging Research, Sunnybrook and Women's College Health Sciences Centre, 2075, Bayview Avenue Room SG28, Toronto, Ontario, Canada M4N 3M5.
E-mail address: peter.soros@swri.ca (P. Sörös).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.02.046

production of speech (Hickok, 2001), rather than isolated speech centers.

Characterization of speech production by fMRI has been complicated by motion-correlated head movements and by movements of the articulatory organs (Birn et al., 1999). Both motion inside the field of view (head movement) and motion outside the field of view (movement of the oral cavity, the sinuses, or the pharynx) (Yetkin et al., 1996) might cause magnetic field inhomogeneities masking brain activation or generating artifactual intensity changes. Previous studies have used inner speech (e.g., silent repetition of words or syllables) to overcome these speech-related artifacts (Wildgruber et al., 1996, 2001). Behavioral tasks involving inner speech usually do well minimizing task-related motion but face other disadvantages (Munhall, 2001). First, it is very difficult to monitor behavioral performance using covert speech production. This aspect would be less important in the present study of neurologically healthy adults but might be a major issue in experiments involving stroke patients. Second, speakers cannot hear their own voice while generating silent responses. Hearing ones own speech, however, is important for accurate speech motor control (Jones and Munhall, 2000). Third, different activation magnitudes have been observed in the cortical and subcortical portions of the speech motor system when comparing silent and overt word stem completion (Palmer et al., 2001; Rosen et al., 2000) or silent and overt production of monosyllabic or multisyllabic words (Shuster and Lemieux, 2005). These observations emphasize the importance of tasks involving overt utterances and of imaging techniques which are less susceptible to movement artifacts for the study of speech production.

For this study, we investigated overt nonlexical utterances using event-related fMRI with clustered image acquisition. A neuropsychological model of speech production includes at least two major cognitive processes, the assembling and the execution of a motor plan (Levelt, 1989). To separate these processes, subjects were asked to repeat acoustically presented sublexical speech sounds of different complexity and to perform nonverbal oral movements. Most previous studies investigated the production of lexical utterances. To minimize semantic and syntactic processing, sublexical speech was chosen for the present study. In addition, recently developed fMRI methodology, termed clustered volume (or image) acquisition (Edmister et al., 1999; Fu et al., 2002; Liebenthal et al., 2003; Ojanen et al., 2005; Rimol et al., 2005), compressed image acquisition (Abrahams et al., 2003) or sparse temporal sampling (Tanaka et al., 2000), enables improved investigation of movement- and speech-related brain activation. The principle underlying these techniques is that the entire brain volume is scanned in a fraction of the repetition time (TR), leaving an extended silent interval for auditory stimulation and speech production (Gracco et al., 2005). This technique is made possible by the difference between the rapid movements associated with speech production (Szirtes and Vaughan, 1977) and the comparatively slow rise of the hemodynamic response curve (Birn et al., 2004). Furthermore, with judicious timing of behavioral tasks, it is possible to separate the hemodynamic response associated with the auditory component of speech from the response associated with acoustic noise arising from the process of fMRI signal acquisition (Rimol et al., 2005).

This study has three goals: (i) to trace the distributed neural network of overt speaking; (ii) to characterize brain activation specific for speaking as compared to simple oral movements; and (iii) to characterize brain activation specific for speaking of polysyllabic sequences as compared to the production of an isolated vowel. It is hypothesized that speaking a less complex speech sound such as a single vowel activates a distributed motor network, similar to performing simple oral movements (Dresel et al., 2005). In addition, it is hypothesized that speech sounds of increasing complexity (monosyllabic consonant–vowel and trisyllabic consonant–vowel utterances) are associated with an increasing task demand and with the increased recruitment of additional brain regions, such as the left inferior frontal gyrus and the left anterior insula (Wise et al., 1999; Blank et al., 2002).

Methods

Participants

Blood oxygenation-level-dependent (BOLD) fMRI was acquired in 9 healthy volunteers (4 women, 5 men) with an average age of 26 years (range, 22–32). All participants were right-handed and, except one, native speakers of English. One volunteer's first language was German. This volunteer has lived in an English-speaking country for several years and used English as her primary language. The target speech sounds for the present study are common in both English and German. Volunteers were recruited with the help of the Rotman Research Institute volunteer database and by personal communication. The study was approved by Research Ethics Boards at Baycrest and at Sunnybrook and Women's College Health Sciences Centre, Toronto, Ontario, Canada. Informed consent for participation in the project was obtained from all subjects according to the Declaration of Helsinki.

Experimental tasks

Subjects were asked to repeat acoustically presented sublexical 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 "say ah" (for the vowel condition), "say pa", "say ka", or "say ta" (for the CV condition), "say pataka" (for the $C_1VC_2VC_3$ V condition), "open your mouth", and "make a kiss" (for oral movements). Verbal instructions were transmitted through an fMRI compatible audio system with acoustically padded headphones to reduce acoustic fMRI noise by 25 dB (Silent Scan; Avotec, Stuart, FL, USA). All instructions were spoken by a speech–language pathologist in a sound-attenuated room, digitized at 22050 Hz and stored as a digital sound file. To avoid confusion with the English article "a", the long vowel "ah" was presented. Instructions were delivered at a constant onset-to-onset interstimulus interval of 10 s with the stimulation software Eprime 1.1 (Psychology Software Tools, Pittsburgh, PA, USA). Subjects were asked to perform the given task or to produce the required response immediately after the end of the instruction. 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. To minimize task-switching effects, a blocked

presentation of 5 identical cues was chosen. All instructions were delivered, and all responses were made within the silent interval between the acquisition of the fMR images. For additional clarity, a timing diagram is illustrated in Fig. 1.

Magnetic resonance imaging

Imaging was performed on a 3 T MRI system (Signa 3T/94 hardware configuration, VH3/M4 software configuration; GE Healthcare, Waukesha, WI, USA) with the standard quadrature birdcage head coil. The participant's head was padded by foam cushions to restrict major head movements. For blood oxygenation-level-dependent (BOLD) fMRI (Ogawa et al., 1992), T2*-weighted functional images were acquired using a spiral-in/out pulse sequence (Glover and Law, 2001) (TE 30 ms, flip angle 70°, matrix 64 × 64, FoV 20 cm × 20 cm, 26 axial slices 5 mm thick) that decreases signal drop-out in regions with large magnetic susceptibility gradients (Preston et al., 2004). High-order shimming was performed at the beginning of the fMRI sessions for each volunteer. Clustered image acquisition was implemented with a TR of 10,000 ms, and the data from all slices were acquired in 1800 ms of this time interval. The hemodynamic response function peaks approximately 5 s after the presentation of a movement trigger (Handwerker et al., 2004) or of an auditory stimulus (Hulvershorn et al., 2005). Based on this knowledge, the offset of the verbal instructions was set approximately 5 s prior to the midpoint of the data acquisition. High-resolution, T1-weighted images (3D Fast SPGR, TR 7.2 ms, TE 3.1 ms, IR-prepared TI 300 ms, flip angle 15°, matrix 256 × 192, FoV 22 cm × 16.5 cm, 124 axial slices 1.4 mm thick) were acquired for structural reference. For the offline assessment of response accuracy, the participants' vocal responses were recorded via the microphone channel of the Silent Scan Audio System (Avotec, Stuart, FL, USA) and stored on a PC. To measure response latency, the opening of the jaw was monitored using a fMRI-compatible fiber optic sensor (ShapeSensor, Measurand Inc., Fredericton, NB, Canada) attached to the chin. The latency between the onset of the cue and the onset of the speech-related jaw movement was calculated using a custom-written program in the statistical package R for Mac OS X (http://www.r-project.org/).

Data analysis

Analysis of fMRI data was carried out in a multistage process using the software library FSL (http://www.fmrib.ox.ac.uk/fsl/). Linear registration and correction of head motion were performed using MCFLIRT (Jenkinson et al., 2002). The maximum head displacement with respect to the reference image and the relative voxel displacement were calculated. Brain segmentation and removal of nonbrain tissue were achieved by FSL's Brain Extraction Tool (BET) (Smith, 2002). Spatial smoothing using a Gaussian kernel of 5-mm full-width half maximum and a mean-based intensity normalization of all volumes by the same factor were applied before the statistical analysis. The 6 fMRI sessions obtained for each participant were analyzed independently using general linear modeling as implemented in FSL's fMRI Expert Analysis Tool (FEAT). Independent analyses of each session were chosen to avoid artifacts related to motion correction and filtering. 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).

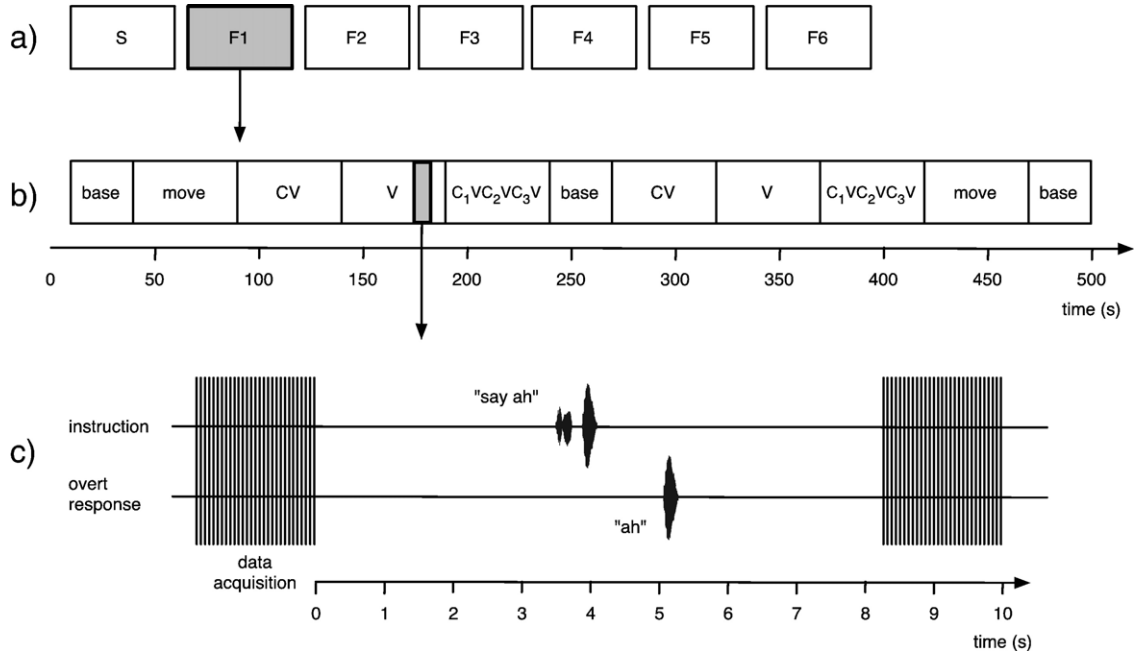

Fig. 1. Timing diagram illustrating behavioral tasks and clustered fMRI data acquisition. (a) The sequence of structural (S) and functional imaging (sessions F1–F6). (b) The sequence of tasks in session F1. This sequence is also used in F3 and F5, while F2, F4, and F6 consist of a different randomization 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.

Because of the long TR (10 s), no temporal autocorrelation between images of one session was assumed. A generalized mixed effects analysis was then carried out to analyze effects across the 6 sessions (Woolrich et al., 2004). Cluster-thresholded activation maps were registered to the high-resolution T1-weighted image. Finally, a mixed effects analysis was performed across all subjects. Because motion-related artefactual signal changes were present in the lateral ventricles in the $C_1VC_2VC_3V$ vs. V contrast, the group analysis was performed after excluding the ventricular region from the statistical calculations. All activation maps were overlaid onto an averaged anatomical template, standardized to Montreal Neurological Institute (MNI) space in radiological convention. To detect local maxima for viewing (Tables 1 and 2), the analysis software Neurolens was used (http://www.neurolens.org). The anatomical location of brain activation was determined by visual inspection and comparison with a detailed neuroanatomical atlas. These results were confirmed using the Talairach daemon (http://ric.uthscsa.edu/projects/tdc/), based on the atlas of Talairach and Tournoux (1988). To compare head movement between conditions, the median relative voxel displacement and the bootstrapped 95% confidence interval were calculated across all participants. For statistical analysis, the Kruskal–Wallis rank sum test was performed using the statistical package R.

Results

Head motion

The median relative voxel displacement was 0.09 mm (CI, 0.07–0.12 mm) in the baseline condition, 0.07 mm (CI, 0.05–0.12 mm) in the movement condition, 0.08 mm (CI, 0.06–0.10 mm) in the V condition, 0.08 mm (CI, 0.05–0.11 mm) in the CV condition, and 0.06 mm (CI, 0.04–0.10 mm) in the $C_1VC_2VC_3V$ condition (Kruskal–Wallis $\chi^2 = 1.81$, $df = 4$, $P = 0.77$).

Behavioral data

The mean response accuracy of all participants was 99.7%. The average latency between stimulus onset and jaw movement onset was 1270 ms (range: 1050–1620 ms) for the movement condition, 1770 ms (range: 1410–2080 ms) for the V condition, 1970 ms (range: 1620–2440 ms) for the CV condition, and 2940 ms (range: 2520–3230 ms) for the $C_1VC_2VC_3V$ condition.

Table 1
Coordinates of group brain activation

Region	Side	V vs. baseline				V vs. movement			
		x	*y*	*z*	*Z* value	*x*	*y*	*z*	*Z* value
Superior frontal gyrus, BA 9	L	–	–	–	–	−2	52	26	3.91
Superior frontal gyrus, BA 6	L	−2	6	60	7.51	–	–	–	–
	R	6	−2	66	9.99	–	–	–	–
Medial frontal gyrus, BA 10	L	–	–	–	–	−4	62	14	4.01
	R	–	–	–	–	2	54	8	3.36
Cingulate gyrus, BA 32	L	−8	12	34	7.01	–	–	–	–
	R	4	12	34	8.12	–	–	–	–
Cingulate gyrus, BA 24	L	−2	0	34	5.84	–	–	–	–
	R	4	8	34	6.95	–	–	–	–
Precentral gyrus, BA 6	R	60	−4	42	6.55	–	–	–	–
Precentral gyrus, BA 4	R	48	−12	44	9.18	–	–	–	–
Insula, BA 13	L	−44	−6	0	5.96	–	–	–	–
	R	44	2	−6	6.1	–	–	–	–
Putamen	L	−24	6	−4	7.07	–	–	–	–
	R	22	−2	2	8.13	–	–	–	–
Lateral globus pallidus	L	−22	−6	−2	7.08	–	–	–	–
	R	20	0	0	7.62	–	–	–	–
Thalamus, ventral lateral nucleus	L	−12	−14	8	7.68	–	–	–	–
	R	12	−14	6	6.99	–	–	–	–
Thalamus, medial dorsal nucleus	L	−10	−20	6	8.29	–	–	–	–
	R	10	−20	8	7.46	–	–	–	–
Superior temporal gyrus	L	−50	−18	−2	9.44	−36	−30	14	4.06
	R	44	4	−14	6.83	48	−32	16	3.78
Transverse temporal gyrus, BA 41	L	−38	−26	10	6.59	−46	−28	10	3.42
	R	–	–	–	–	38	−30	12	4.14
Middle temporal gyrus	L	−58	4	−10	6.16	−60	−34	4	3.45
Posterior cerebellar lobe, pyramis	L	−18	−66	−28	7.33	–	–	–	–
Posterior cerebellar lobe, declive	L	−12	−68	−22	7.15	–	–	–	–
	R	–	–	–	–	–	–	–	–
Posterior cerebellar lobe, uvula	L	−14	−68	−24	7.17	–	–	–	–
	R	14	−68	−24	8.33	–	–	–	–
Red nucleus	L	−8	−18	−6	6.1	–	–	–	–
	R	8	−18	−6	5.88	–	–	–	–

Peak *x*, *y* and *z* coordinates of the *Z*-statistic activation maps in MNI space and the corresponding *Z* values for the V vs. baseline and V vs. movement comparisons. Brodmann areas (BA) are given where appropriate. In the V vs. baseline condition, activated clusters are reported for $Z > 4$ and cluster-corrected $P < 0.01$. In the V vs. movement comparison, activated clusters are reported for $Z > 2.3$ and cluster-corrected $P < 0.01$. V denotes utterance of "ah".

Table 2
Coordinates of group brain activation

Region	Side	CV vs. V				$C_1VC_2VC_3V$ vs. CV				$C_1VC_2VC_3V$ vs. V			
		x	*y*	*z*	*Z* value	*x*	*y*	*z*	*Z* value	*x*	*y*	*z*	*Z* value
Inferior frontal gyrus	L	–	–	–	–	–	–	–	–	−50	28	−12	3.74
Insula, BA 13	L	–	–	–	–	–	–	–	–	−38	−24	4	3.91
	R	–	–	–	–	–	–	–	–	36	−24	14	4.18
Putamen	R	–	–	–	–	28	−10	12	3.87	–	–	–	–
Caudate, tail	L	–	–	–	–	–	–	–	–	−18	−26	20	3.93
Caudate, body	L	–	–	–	–	–	–	–	–	−18	−20	24	3.77
Superior temporal gyrus	L	–	–	–	–	–	–	–	–	−38	−24	8	4.34
	R	70	−24	0	4.12	–	–	–	–	60	−14	2	4.52
Transverse temporal gyrus, BA 41	L	–	–	–	–	–	–	–	–	−38	−26	10	3.19
	R	–	–	–	–	–	–	–	–	36	−26	10	3.72
Middle temporal gyrus	L	–	–	–	–	−52	−16	−6	3.35	−64	−18	−8	4.04
	R	–	–	–	–	–	–	–	–	64	−6	−8	3.98
Inferior temporal gyrus	L	–	–	–	–	–	–	–	–	−58	−8	−18	3.37
Posterior cerebellar lobe, pyramis	L	–	–	–	–	–	–	–	–	−22	−64	−28	4.38
	R	–	–	–	–	–	–	–	–	22	−62	−28	4.04
Posterior cerebellar lobe, declive	L	–	–	–	–	–	–	–	–	−16	−64	−22	3.89
	R	–	–	–	–	–	–	–	–	12	−72	−22	4.49
Posterior cerebellar lobe, uvula	L	–	–	–	–	–	–	–	–	−18	−66	−24	4.49
	R	–	–	–	–	–	–	–	–	16	−68	−24	3.31
Posterior cerebellar lobe, tonsil	L	–	–	–	–	–	–	–	–	−28	−62	−34	3.86
	R	–	–	–	–	–	–	–	–	26	−62	−32	2.96
Anterior cerebellar lobe, culmen	L	–	–	–	–	–	–	–	–	−22	−62	−26	3.79
	R	–	–	–	–	–	–	–	–	26	−62	−26	2.57
Anterior cerebellar lobe, dentate	L	–	–	–	–	–	–	–	–	−18	−56	−24	3.54
	R	–	–	–	–	–	–	–	–	18	−62	−24	3.49

Peak *x*, *y* and *z* coordinates of the *Z*-statistic activation maps in MNI space and the corresponding *Z* values for the CV vs. V, $C_1VC_2VC_3V$ vs. CV and $C_1VC_2VC_3V$ vs. V contrasts. Brodmann areas (BA) are given where appropriate. Activated clusters are reported for $Z > 2.3$ and cluster-corrected $P < 0.01$. V denotes utterance of "ah"; CV, utterance of "pa", "ta" or "ka"; $C_1VC_2VC_3V$, utterance of "pataka".

Vowel utterance

Brain activation associated with speaking the vowel "ah", compared to baseline, is presented in Fig. 2. Speaking involves a widespread, bilateral motor network including the pyramidal and extrapyramidal system. Activation in the frontal and cingulate cortex included the supplementary motor area (SMA), the cingulate motor area (CMA) and the primary motor cortex (M1) in both hemispheres. Subcortical activation was found in the thalamus, the bilateral globus pallidus and the putamen. In addition, activation was present in the bilateral superior and the left middle temporal plane and in the bilateral posterior insula. Lobule VI (Schmahmann, 2000) and the red nucleus were activated bilaterally. Table 1 summarizes the coordinates in MNI space, the corresponding anatomical label of activated brain regions, and their *Z* scores.

Fig. 3 displays brain activation associated with production of the vowel "ah" vs. brain activation related to oral movements. Production of the vowel "ah" is characterized by activation in the left middle and in the bilateral superior temporal gyrus (Table 1).

Monosyllabic utterance

Contrasting the CV condition with the vowel condition revealed a small activated area within the right middle and superior temporal gyri (Table 2). Comparing the $C_1VC_2VC_3V$ condition with the CV condition displayed activation in the left middle temporal gyrus and in the right basal ganglia (Table 2).

Polysyllabic utterance

Brain activation associated with the production of a $C_1VC_2VC_3V$ nonword ("pataka"), compared to the production of "ah", is presented in Fig. 4. Producing "pataka" vs. "ah" activated the bilateral superior, the bilateral middle and the left inferior temporal gyrus (Table 2). The left caudate and parts of the left anterior and posterior cerebellar lobes were also activated.

To differentiate between auditory stimulation and phonologic processing, brain activation associated with speech production and facial movements was compared to baseline. During facial movements vs. baseline, the bilateral posterior superior temporal cortex was activated similar to the vowel condition vs. baseline (data not shown). The $C_1VC_2VC_3V$ condition, however, was associated with a more extended and stronger activation of the bilateral superior temporal compared to the movement condition (data not shown).

Discussion

The distributed neural network of speaking was investigated using fMRI during the production of oral movements, vowels, and syllables. Speaking was associated with activation in bilateral cortical and subcortical motor centers as well as with activation in the bilateral superior temporal gyrus. Speaking a single vowel and performing simple oral movements involved almost identical activation of the pyramidal and extrapyramidal motor system. More complex, polysyllabic utterances were

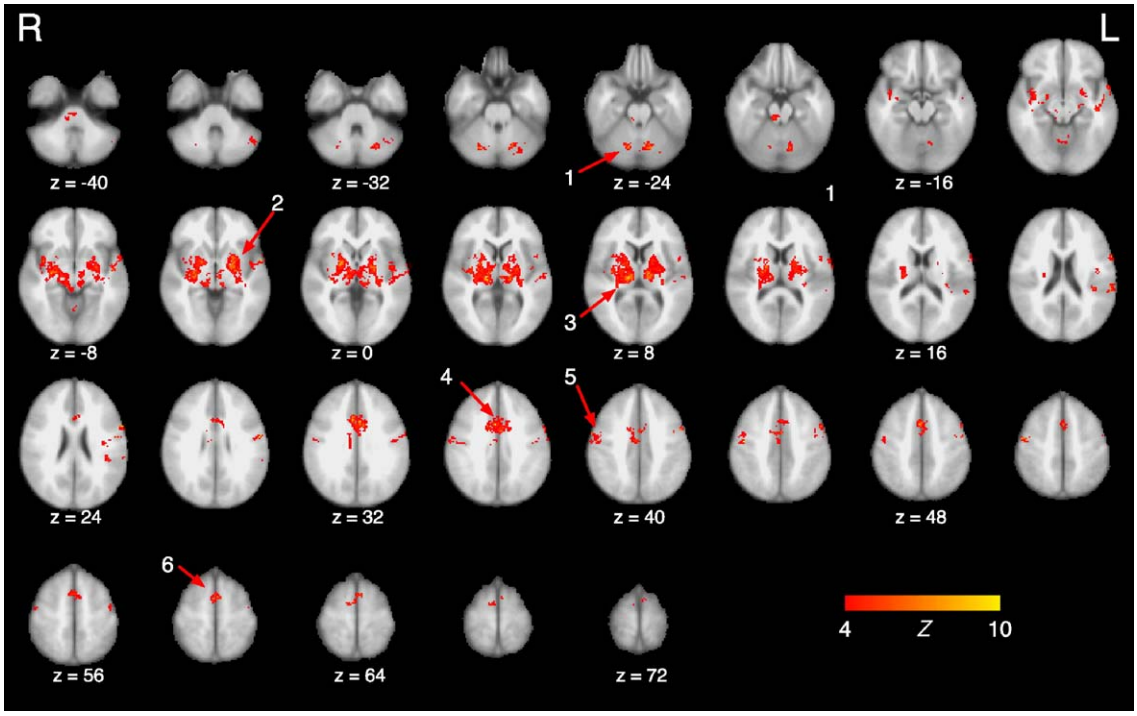

Fig. 2. Group activation map, V vs. baseline condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen primarily in the bilateral posterior cerebellar lobe (1), the basal ganglia (2), the thalamus (3), the cingulate motor area (4), the primary motor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

associated with additional activation in the bilateral cerebellum and the bilateral temporal cortex. With the sensitivity of a 3 T MRI scanner, an optimized imaging sequence and a large field of view it was possible to disclose the neural network of speech production from the SMA at the cranial end to the red nucleus at the caudal end. A schematic illustration of the articulo-phonologic network found in the present study is depicted in Fig. 5. These results will be discussed in relation to the existing functional imaging literature involving speech-related brain function.

The neural network of speaking

The first aim of this study was to identify the articulo-phonologic brain network related to speech production. To avoid confounding variables such as semantic and syntactic processing, the production of the vowel "ah" was investigated. With the exception of sensorimotor activation associated with overt speech production, the involvement of various other brain areas is controversial. In a study on reiterating syllables differing in their articulatory and phonologic demand, brain activation was almost

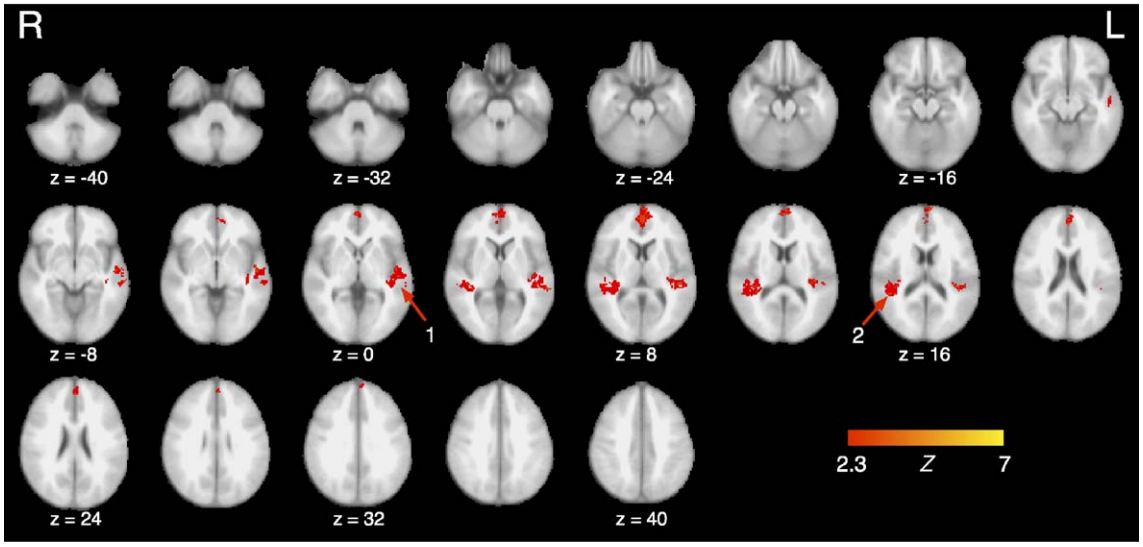

Fig. 3. Group activation map, V vs. movement condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen primarily in the left middle (1) and the bilateral superior temporal gyrus (2).

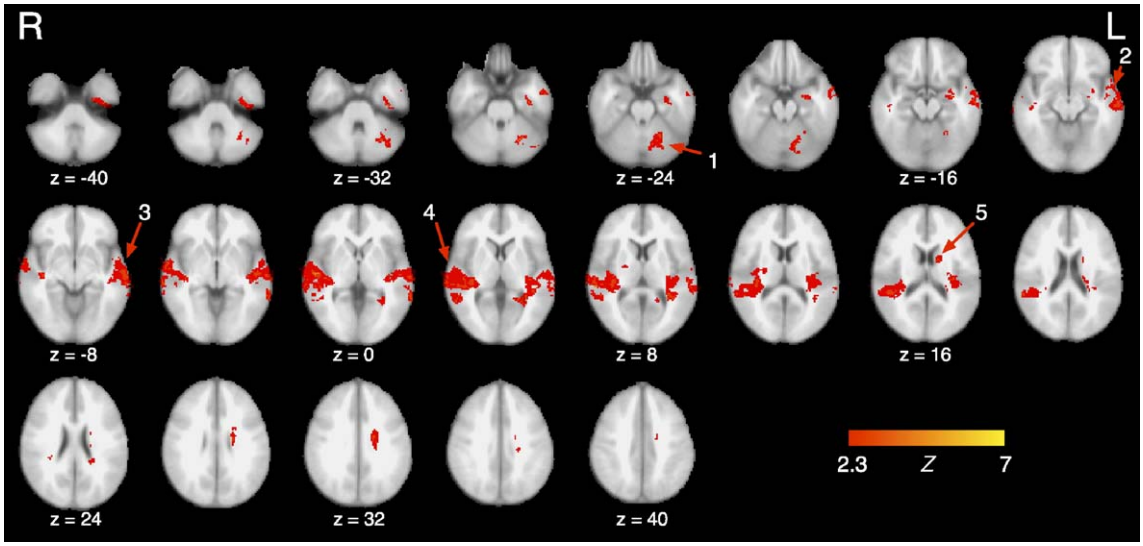

Fig. 4. Group activation map, $C_1VC_2VC_3V$ vs. V condition. Brain activation was averaged across all subjects and registered to MNI space. Activation is seen in the left inferior frontal gyrus (2), in the bilateral middle temporal gyrus (3) and in the bilateral superior temporal gyrus (4). Activation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

limited to the sensorimotor cortex (Riecker et al., 2000). In contrast, repetition of single words (Wise et al., 1999) or syllables (Riecker et al., 2005) was associated with distributed cortical and subcortical activity. In the present study, activation was found in areas related to the planning, execution and control of movements (mainly in the SMA, M1 and cerebellum) and to the auditory and phonologic processing of speech units (middle and superior temporal gyrus).

Supplementary motor areas

Activation of the bilateral SMA is a consistent finding in imaging studies on voluntary movements such as coordinated orofacial movements (Dresel et al., 2005), swallowing (Martin et al., 2004), or speech production (Murphy et al., 1997; Riecker et al., 2005). The SMA proper, at the medial wall of the hemisphere (BA 6), is involved in planning, initiation, and control of movements (Picard and Strick, 2001). In addition to the SMA, cingulate motor areas (CMA), located in the anterior cingulate cortex (ACC) of both hemispheres, were activated here. With its dense connections to the motor cortex, the SMA and the ACC (Paus, 2001) play an important role in the control of movements such as motor speech production (Fig. 5). Cortical stimulation experiments of large portions of the monkey ACC elicited a variety of calls (Paus, 2001). In a human positron emission tomography (PET) study, the production of pronouns or letters (Paus et al., 1993) and the repetition of single words were associated with activation in the ACC (Wise et al., 1999).

Motor and premotor cortex

As expected (Petersen et al., 1988), speaking was associated with bilateral activation in the face area of M1. The location of the motor face area, inferior and lateral of the motor hand area, has been determined by cortical stimulation mapping (Penfield and Boldrey, 1937) and noninvasive brain imaging (Corfield et al., 1999). Imaging studies on articulatory movements consistently found M1 activation in both hemispheres (Wise et al., 1999; Riecker et al., 2000, 2005). These results, including the present study, support clinical observations and neurophysiological data, suggesting a bilateral cortical representation of midline muscles (Muellbacher et al., 1999). Due to the time constraints of clustered image acquisition in the present study, only a portion of the brainstem was scanned, and further research will be required to map activation in the corresponding brainstem nuclei. In addition to M1, activation of the right precentral (BA 6) was found, which is thought to reflect the importance of the BA 6 for the planning and execution of movements.

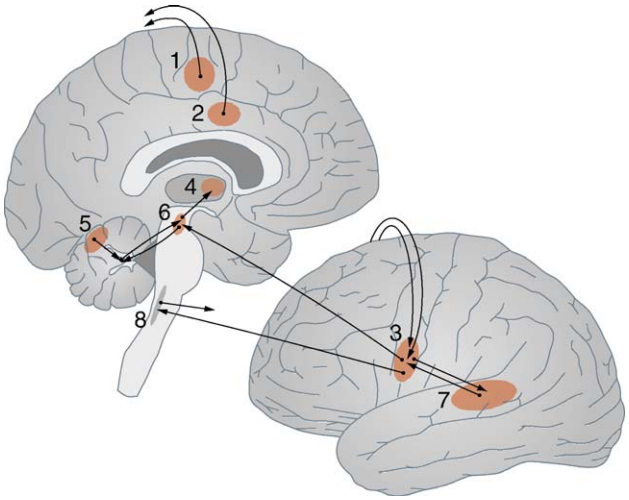

Fig. 5. The neural network of speech production. Areas activated during speaking in the present study are shown in red. Schematic fiber tracts connecting those areas are represented by black arrows. Only main areas of activation and main fiber tracts are shown. The supplementary motor area (1) and the cingulate motor areas (2) are connected with the primary motor cortex (3). Several connections exist between the cortical and the subcortical motor system. 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, such as the nucleus hypoglossus, were outside the field of view (8). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Thalamus and basal ganglia

Speaking a vowel compared to baseline is associated with activation of the bilateral ventro-lateral and medio-dorsal thalamus, the bilateral putamen and the bilateral globus pallidus. The basal ganglia, the thalamus, and the cerebellum are connected by distinct and parallel circuits (Paradiso et al., 2004). The basal ganglia play an important role in the initiation and control of voluntary fine movement and motor sequences as indicated by the examination of neurological deficits in patients with Parkinson syndrome and by functional imaging studies (Boecker et al., 1998). The thalamus is not only a relay for ascending sensory input but is also involved in the preparation of movement (Paradiso et al., 2004). In an experiment of single word repetition, activation of the left posterior pallidum was found (Wise et al., 1999). Overt repetition of a monosyllable was associated with activation of the bilateral putamen, pallidum, and thalamus (Riecker et al., 2005).

Cerebellum

Fig. 2 illustrates left-lateralized paravermal activation in lobules VI, VIII, and IX. The location of this activation comes close to the representation of externally paced vertical tongue movements and of lip movements in a previous fMRI study of cerebellar topography (Grodd et al., 2001). The importance of the intact cerebellum for speech production is emphasized by clinical observations that patients with ataxic dysarthria frequently suffer from lesions in the vermal and paravermal areas (Marien et al., 2001).

Red nucleus

The red nucleus, located in the rostral midbrain, is reciprocally connected with the cerebellar nuclei (Fig. 5). In addition, it receives afferents from the motor cortex and sends efferents to the thalamus (Massion, 1988). The red nucleus is involved in movement coordination as indicated by invasive recordings in cats (Lavoie and Drew, 2002) as well as in somatosensory (Liu et al., 2000) and pain processing demonstrated by fMRI in humans (Dunckley et al., 2005).

Temporal lobe

Activation of the bilateral superior temporal gyrus as well as of the left middle and transverse temporal gyrus was found when comparing the production of a vowel with baseline (Fig. 2). Due to the clustered image acquisition used here, the temporal evolution of the BOLD signal is not available. As speech comprehension and phonologic processing are represented in partly overlapping areas of the temporal cortex, the present experiment cannot determine to what extent these processes contribute to the observed signal changes. Activation of the transverse temporal gyrus (Heschl's gyrus) and of adjacent cortical areas is most likely due to the auditory processing of the spoken instruction and of the volunteer's own response. Volunteers were able to perceive their own voice clearly during fMRI in part through bone conduction, in spite of the sound-attenuating headphones worn. A combined analysis of 4 PET studies found activation in the supratemporal plane associated with the speaker's own voice (Wise et al., 2001). Because a clustered fMRI acquisition technique was used with a TR of 10 s, and because the BOLD hemodynamic response function peaks about 5 s after the onset of the auditory stimulus (Hulvershorn et al., 2005), scanner noise is unlikely to contribute significantly to auditory activation.

The production of a basic speech sound

The second aim of this study was to characterize brain activation specific for speaking as compared to simple oral movements. By comparing the overt production of a vowel with the execution of simple oral movements, it was possible to identify neural processes characteristic for speech production such as phonologic processing and movement coordination. The vowel condition was associated with stronger activation in the bilateral superior and transverse temporal gyri as well as the left middle temporal gyrus compared to the movement condition. Increased activation in the transverse temporal gyrus is primarily due to the overt response and the additional auditory processing in the vowel condition compared to the movement condition. Activation in the posterior superior and middle temporal gyrus is thought to be, at least in part, associated with phonologic processing (Wise et al., 2001).

The production of complex utterances

The third aim of the present study was to investigate brain activation associated with utterances of different complexity. Human speech is based on complex sequences of rapidly changing sounds, rather than isolated speech sounds such as the vowel "ah". In contrast to the majority of previously published reports (Murphy et al., 1997; Wise et al., 1999; Riecker et al., 2005) (for an exception, see Riecker et al., 2000), the present study investigated the production of nonword utterances of increasing complexity and increasing demand on articulatory and phonologic processing. CV syllables and the trisyllabic sequence "pataka" ($C_1VC_2VC_3V$) were chosen for the investigation of complex speech. "Pataka" involves different places of articulation, is widely used to assess speech motor deficits, and lacks linguistic content.

Contrasting the complex $C_1VC_2VC_3V$ condition with the basic vowel condition, activation was found in the left inferior frontal gyrus, the left cerebellum, the left caudate nucleus and the bilateral superior and middle temporal gyri (Fig. 4). Comparing the production of a CV syllable with the V condition and the $C_1VC_2VC_3V$ condition reveals the increasing recruitment of neural resources necessary for the production of more demanding utterances. During the production of CV syllables, the right middle and superior temporal gyri were activated compared to the production of an isolated vowel. Production of a polysyllabic $C_1VC_2VC_3V$ utterance, compared to the monosyllabic CV, is associated with increased need for motor control and phonologic processing as indicated by the activation of the left middle temporal gyrus and in the right basal ganglia.

Inferior frontal gyrus

Since the lesion studies of Broca the inferior frontal gyrus have been regarded as essential for speech production (Broca, 1861). Functional imaging studies provided evidence that different subregions of the inferior frontal gyrus subserve, among other functions, phonologic, semantic, and syntactic processing. BA 44, part of Broca's area, is activated by the generation of complex articulatory movements of oral and laryngeal musculature (Horwitz et al., 2003). In addition, a recent investigation using intraoperative stimulation provided evidence for a functional connection between the inferior frontal gyrus and the orofacial motor cortex in humans (Greenlee et al., 2004). In patients with an infarction of the left middle cerebral artery, a lesion of the inferior temporal gyrus was

frequently associated with a speech motor deficit (Hillis et al., 2004). Overt naming of familiar objects (Etard et al., 2000; Sörös et al., 2003), repetition of a simple phrase (Murphy et al., 1997), or repetition of single words (Wise et al., 1999) in contrast were not associated with activation of the inferior frontal gyrus. Taken together, these observations might suggest that fluent speech and the production of unfamiliar syllabic sequences such as "pataka", but not the production of a single vowel or a CV syllable, are characterized by increased demand for phonologic processing and are associated with activation in Broca's area.

Cerebellum and basal ganglia

Although the vowel condition was associated with posterior paravermal activation, additional activation in the anterior paravermal region was found in the $C_1VC_2VC_3V$ condition. Clinical observations suggested that damage to the anterior vermal and paravermal regions is frequently found in patients with cerebellar dysarthria (Marien et al., 2001). Additional activation in the left anterior cerebellar lobe and in the left caudate nucleus probably reflects the increased demand on motor control in the production of a $C_1VC_2VC_3V$ sequence compared to the production of a vowel. The present results support the long-held concept that the cerebellum is engaged in motor timing and coordination (Eccles et al., 1967), in particular in the control of sequential movements (Catalan et al., 1998; Haaland et al., 2004). In addition, activation of the right putamen was found when contrasting $C_1VC_2VC_3V$ and CV production (Table 2). Activation of the bilateral putamen is part of the neural network of speaking, as indicated by the production of "ah" vs. baseline (Table 1). In a human fMRI study, the timing of sequential movements was associated with activity in the right putamen (Garraux et al., 2005). Thus, increased activation of the right putamen in the $C_1VC_2VC_3V$ condition probably reflects the necessity of accurate timing when producing a multisyllable utterance.

Insula

Speaking a vowel as compared to baseline activated the bilateral posterior insula (Table 1). The posterior insula, prominently involved in pain processing, is also thought to mediate self-awareness of movements as indicated by a recently published study on stroke patients with anosognosia (Karnath et al., 2005). It is hypothesized that the posterior insula, as a multimodal area with strong connections to motor, somatosensory and auditory centers, integrates auditory and proprioceptive features of speech production and contributes to one's own consciousness of speaking.

No significant activation, however, was found in the left anterior insula. An influential study in chronic stroke patients linked apraxia of speech (AOS) with a lesion in the left anterior insula (Dronkers, 1996). This result was corroborated by subsequent lesions studies in patients with AOS (Nagao et al., 1999) and primary progressive aphasia (Nestor et al., 2003). A recently published investigation on anatomical, diffusion-weighted and perfusion-weighted imaging in stroke patients with AOS, in contrast, did not find an association between AOS and an insular lesions. In this study, lesions of the left posterior inferior frontal gyrus (Brocas area, BA 44) were associated with AOS (Hillis et al., 2004). In addition to the aforementioned lesion studies (Dronkers, 1996; Nagao et al., 1999; Nestor et al., 2003), several neuroimaging studies supported the notion that the left insula is crucial for the formulation of an articulatory plan. These studies used tasks with varying lexical items, repetition of heard nouns (Wise et al., 1999), naming the months of the year (Riecker et al., 2000), or pronouncing different words (Kuriki et al., 1999). Other studies, including the present work, used repetitions of identical items, repetitive saying of a simple phrase (Murphy et al., 1997) or repetitive saying of mono- and polysyllabic items and of a noun (Riecker et al., 2000) and did not find insular activation on the group level. In contrast, an fMRI study on externally paced repetition of the syllable "pa" reported activation of the left anterior insula (Riecker et al., 2005). These observations might indicate that the anterior insula is predominantly involved in varying rather than repetitive speech production. Activation of the left insula appears to be especially sensitive to the distinctive features of the experimental design chosen (Indefrey and Levelt, 2000).

Temporal lobe

As discussed earlier, the interpretation of superior and middle temporal cortex activation is not trivial for the present paradigm. The vowel, the CV and the $C_1VC_2VC_3V$ condition involve a spoken instruction and an overt response. Instruction and response are longer, however, in the $C_1VC_2VC_3V$ condition which is presumably correlated with increased auditory processing compared to the vowel condition. Recent research gave rise to the hypothesis that the posterior superior temporal cortex is not only crucial for auditory processing (Lütkenhöner and Steinsträter, 1998) but also for phonologic processing (Hickok et al., 2003; Rimol et al., 2005). It is hypothesized that increased activation of the bilateral middle and superior temporal gyrus in the $C_1VC_2VC_3V$ condition compared to the vowel condition reflects increased demand on phonologic processing associated with the more complex trisyllabic response. This interpretation is supported by a fMRI study on covert naming of objects with names consisting of one to four syllables (Okada et al., 2003). The posterior superior temporal cortex was activated, in relation to the word length, in all subjects during covert naming (Okada et al., 2003). The contrasts CV vs. V and $C_1VC_2VC_3V$ vs. CV indicate a remarkable lateralization of temporal lobe activity in the present experiment. In auditory processing, temporal variations of stimuli activated primarily the left temporal cortex. Spectral variations, in contrast, primarily activated the right temporal cortex (Zatorre and Belin, 2001). As a CV syllable contains more complex spectral features than a single vowel, the stronger activation of the right superior temporal gyrus in the CV condition compared to the V condition might reflect increased spectral processing in CV production. In the $C_1VC_2VC_3V$ condition, relative to CV, stronger activation of the left temporal lobe probably reflects increased temporal processing during production of a multisyllable utterance.

Methodological considerations

Our data indicated that high-field fMRI with clustered image acquisition is a valuable tool for the characterization of neural networks involved in overt speech production. Head motion, a major concern in previous fMRI studies on speech production, was similar during baseline and during overt speech. The study design, however, has limitations.

Clustered image acquisition requires a meticulous coordination between stimulation and data acquisition as functional scans of the brain are acquired for only one time point per event. In the present study, the offset of the verbal instructions was set 5 s prior to the midpoint of the data acquisition. The timing of the data acquisition

is difficult as the evolution of the hemodynamic response differs between individuals and between brain areas (Handwerker et al., 2004) and might be altered by the different lengths of the required responses. The authors do not believe that these factors considerably affect the observed pattern of brain activation because the hemodynamic response lasts about 7–10 s (Handwerker et al., 2004). Based on previous research, a 5-s interval between response onset and data acquisition is appropriate to detect BOLD signal changes at the peak or close to the peak of the hemodynamic response curve.

As a spoken instruction was used, it was impossible to discriminate between superior temporal gyrus activation due to the instruction and activation due to the overt response. One possible alternative is to eliminate this interference through the use of abstract visual cues in future experiments. This study was designed to investigate differences in brain function between oral movements and utterances of different complexity, resulting in 4 experimental conditions and a total measurement time of approximately 60 min. In future studies, the number of conditions should be reduced, based on the results of the present experiment. A shorter measurement time is especially desirable in studies on patients with speech motor deficits.

Conclusion

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 (Fig. 5). 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 the 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.

Acknowledgments

The authors acknowledge Fred Tam for his expert help in setting up the MRI sequences and Gianfranco Pellicori for the technical assistance, in particular for writing the Eprime stimulation script. The authors wish to thank Dr. Gary Glover for providing his spiral-in/out pulse sequence. This study was supported by the Heart and Stroke Foundation of Ontario Centre for Stroke Recovery and by Baycrest's Posluns Centre for Stroke and Cognition. In addition, P.S. received personal funding from the Heart and Stroke Foundation of Ontario Centre for Stroke Recovery.

References

Abrahams, S., Goldstein, L., Simmons, A., Brammer, M., Williams, S., Giampietro, V., Andrew, C., Leigh, P., 2003. Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. Hum. Brain Mapp. 20 (1), 29–40.

Aichert, I., Ziegler, W., 2004. Syllable frequency and syllable structure in apraxia of speech. Brain Lang. 88 (1), 148–159.

Birn, R., Bandettini, P., Cox, R., Shaker, R., 1999. Event-related fMRI of tasks involving brief motion. Hum. Brain Mapp. 7 (2), 106–114.

Birn, R., Cox, R., Bandettini, P., 2004. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. NeuroImage 23 (3), 1046–1058.

Blank, S., Scott, S., Murphy, K., Warburton, E., Wise, R., 2002. Speech production: Wernicke, broca and beyond. Brain 125 (Pt. 8), 1829–1838.

Boecker, H., Dagher, A., Ceballos-Baumann, A., Passingham, R., Samuel, M., Friston, K., Poline, J., Dettmers, C., Conrad, B., Brooks, D., 1998. Role of the human rostral supplementary motor area and the basal ganglia in motor sequence control: investigations with H2 15O PET. J. Neurophysiol. 79 (2), 1070–1080.

Broca, P., 1861. Remarques sur le siége de la faculté du langage articulé suivies d'une observation d'aphémie (perte de la parole). Bull. Soc. Anat. 6, 330–357.

Catalan, M., Honda, M., Weeks, R., Cohen, L., Hallett, M., 1998. The functional neuroanatomy of simple and complex sequential finger movements: a PET study. Brain 121 (Pt. 2), 253–264.

Corfield, D., Murphy, K., Josephs, O., Fink, G., Frackowiak, R., Guz, A., Adams, L., Turner, R., 1999. Cortical and subcortical control of tongue movement in humans: a functional neuroimaging study using fMRI. J. Appl. Physiol. 86 (5), 1468–1477.

Croot, K., 2002. Diagnosis of AOS: definition and criteria. Semin. Speech Lang. 23 (4), 267–280.

Darley, F., Aronson, A., Brown, J., 1975. Motor Speech Disorders. Saunders.

Dresel, C., Castrop, F., Haslinger, B., Wohlschlaeger, A., Hennenlotter, A., Ceballos-Baumann, A., 2005. The functional neuroanatomy of coordinated orofacial movements: sparse sampling fMRI of whistling. NeuroImage 28 (3), 588–597.

Dronkers, N., 1996. A new brain region for coordinating speech articulation. Nature 384 (6605), 159–161.

Dunckley, P., Wise, R., Fairhurst, M., Hobden, P., Aziz, Q., Chang, L., Tracey, I., 2005. A comparison of visceral and somatic pain processing in the human brainstem using functional magnetic resonance imaging. J. Neurosci. 25 (32), 7333–7341.

Eccles, J., Ito, M., Szentágothai, J., 1967. The Cerebellum as a Neuronal Machine. Springer-Verlag, Verlag.

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

Etard, O., Mellet, E., Papathanassiou, D., Benali, K., Houdé, O., Mazoyer, B., Tzourio-Mazoyer, N., 2000. Picture naming without Broca's and Wernicke's area. NeuroReport 11 (3), 617–622.

Fitch, R., Miller, S., Tallal, P., 1997. Neurobiology of speech perception. Annu. Rev. Neurosci. 20, 331–353.

Fu, C., Morgan, K., Suckling, J., Williams, S., Andrew, C., Vythelingum, G., McGuire, P., 2002. A functional magnetic resonance imaging study of overt letter verbal fluency using a clustered acquisition sequence: greater anterior cingulate activation with increased task demand. NeuroImage 17 (2), 871–879.

Garraux, G., McKinney, C., Wu, T., Kansaku, K., Nolte, G., Hallett, M., 2005. Shared brain areas but not functional connections controlling movement timing and order. J. Neurosci. 25 (22), 5290–5297.

Glover, G., Law, C., 2001. Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. Magn. Reson. Med. 46 (3), 515–522.

Gorno-Tempini, M., Dronkers, N., Rankin, K., Ogar, J., Phengrasamy, L., Rosen, H., Johnson, J., Weiner, M., Miller, B., 2004. Cognition and anatomy in three variants of primary progressive aphasia. Ann. Neurol. 55 (3), 335–346.

Gracco, V., Tremblay, P., Pike, B., 2005. Imaging speech production using fMRI. NeuroImage 26 (1), 294–301.

Greenlee, J., Oya, H., Kawasaki, H., Volkov, I., Kaufman, O., Kovach, C.,

Howard, M., Brugge, J., 2004. A functional connection between inferior frontal gyrus and orofacial motor cortex in human. J. Neurophysiol. 92 (2), 1153–1164.
Grodd, W., Hülsmann, E., Lotze, M., Wildgruber, D., Erb, M., 2001. Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. Hum. Brain Mapp. 13 (2), 55–73.
Haaland, K., Elsinger, C., Mayer, A., Durgerian, S., Rao, S., 2004. Motor sequence complexity and performing hand produce differential patterns of hemispheric lateralization. J. Cogn. Neurosci. 16 (4), 621–636.
Handwerker, D., Ollinger, J., D'Esposito, M., 2004. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. NeuroImage 21 (4), 1639–1651.
Hickok, G., 2001. Functional anatomy of speech perception and speech production: psycholinguistic implications. J. Psycholinguist. Res. 30 (3), 225–235.
Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J. Cogn. Neurosci. 15 (5), 673–682.
Hillis, A., Work, M., Barker, P., Jacobs, M., Breese, E., Maurer, K., 2004. Re-examining the brain regions crucial for orchestrating speech articulation. Brain 127 (Pt. 7), 1479–1487.
Horwitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, K., Zilles, K., Braun, A., 2003. Activation of Broca's area during the production of spoken and signed language: a combined cytoarchitectonic mapping and PET analysis. Neuropsychologia 41 (14), 1868–1876.
Huber, P., Gutbrod, K., Ozdoba, C., Nirkko, A., Lövblad, K., Schroth, G., 2000. Zur Geschichte der Aphasiologie und Sprachlokalisation im Gehirn. Schweiz. Med. Wochenschr. 130 (3), 49–59.
Hulvershorn, J., Bloy, L., Gualtieri, E., Redmann, C., Leigh, J., Elliott, M., 2005. Temporal resolving power of spin echo and gradient echo fMRI at 3T with apparent diffusion coefficient compartmentalization. Hum. Brain Mapp. 25 (2), 247–258.
Indefrey, P., Levelt, J., 2000. The neural correlates of language production. In: Gazzaniga, M.S. (Ed.), The New Cognitive Neurosciences, 2nd ed. MIT Press, pp. 845–866.
Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. NeuroImage 17 (2), 825–841.
Jones, J., Munhall, K., 2000. Perceptual calibration of F0 production: evidence from feedback perturbation. J. Acoust. Soc. Am. 108 (3 Pt. 1), 1246–1251.
Karnath, H., Baier, B., Nägele, T., 2005. Awareness of the functioning of one's own limbs mediated by the insular cortex? J. Neurosci. 25 (31), 7134–7138.
Kuriki, S., Mori, T., Hirata, Y., 1999. Motor planning center for speech articulation in the normal human brain. NeuroReport 10 (4), 765–769.
Lavoie, S., Drew, T., 2002. Discharge characteristics of neurons in the red nucleus during voluntary gait modifications: a comparison with the motor cortex. J. Neurophysiol. 88 (4), 1791–1814.
Levelt, W., 1989. Speaking: From Intention to Articulation. MIT Press.
Liebenthal, E., Ellingson, M., Spanaki, M., Prieto, T., Ropella, K., Binder, J., 2003. Simultaneous ERP and fMRI of the auditory cortex in a passive oddball paradigm. NeuroImage 19 (4), 1395–1404.
Liu, Y., Pu, Y., Gao, J., Parsons, L., Xiong, J., Liotti, M., Bower, J., Fo, P., 2000. The human red nucleus and lateral cerebellum in supporting roles for sensory information processing. Hum. Brain Mapp. 10 (4), 147–159.
Lütkenhöner, B., Steinsträter, O., 1998. High-precision neuromagnetic study of the functional organization of the human auditory cortex. Audiol. Neuro-otol. 3 (2–3), 191–213.
Marien, P., Engelborghs, S., Fabbro, F., De Deyn, P., 2001. The lateralized linguistic cerebellum: a review and a new hypothesis. Brain Lang. 79 (3), 580–600.
Martin, R., MacIntosh, B., Smith, R., Barr, A., Stevens, T., Gati, J., Menon, R., 2004. Cerebral areas processing swallowing and tongue movement are overlapping but distinct: a functional magnetic resonance imaging study. J. Neurophysiol. 92 (4), 2428–2443.
Massion, J., 1988. Red nucleus: past and future. Behav. Brain Res. 28 (1–2), 1–8.
Muellbacher, W., Artner, C., Mamoli, B., 1999. The role of the intact hemisphere in recovery of midline muscles after recent monohemispheric stroke. J. Neurol. 246 (4), 250–256.
Munhall, K., 2001. Functional imaging during speech production. Acta. Psychol. (Amst) 107 (1–3), 95–117.
Murphy, K., Corfield, D., Guz, A., Fink, G., Wise, R., Harrison, J., Adams, L., 1997. Cerebral areas associated with motor control of speech in humans. J. Appl. Physiol. 83 (5), 1438–1447.
Nagao, M., Takeda, K., Komori, T., Isozaki, E., Hirai, S., 1999. Apraxia of speech associated with an infarct in the precentral gyrus of the insula. Neuroradiology 41 (5), 356–357.
Nestor, P., Graham, N., Fryer, T., Williams, G., Patterson, K., Hodges, J., 2003. Progressive non-fluent aphasia is associated with hypometabolism centred on the left anterior insula. Brain 126 (Pt. 11), 2406–2418.
Ogawa, S., Tank, D., Menon, R., Ellermann, J., Kim, S., Merkle, H., Ugurbil, K., 1992. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc. Natl. Acad. Sci. U. S. A. 89 (13), 5951–5955.
Ojanen, V., Möttönen, R., Pekkola, J., Jääskeläinen, I., Joensuu, R., Autti, T., Sams, M., 2005. Processing of audiovisual speech in Broca's area. NeuroImage 25 (2), 333–338.
Okada, K., Smith, K., Humphries, C., Hickok, G., 2003. Word length modulates neural activity in auditory cortex during covert object naming. NeuroReport 14 (18), 2323–2326.
Palmer, E., Rosen, H., Ojemann, J., Buckner, R., Kelley, W., Petersen, S., 2001. An event-related fMRI study of overt and covert word stem completion. NeuroImage 14 (1 Pt. 1), 182–193.
Paradiso, G., Cunic, D., Saint-Cyr, J., Hoque, T., Lozano, A., Lang, A., Chen, R., 2004. Involvement of human thalamus in the preparation of self-paced movement. Brain 127 (Pt. 12), 2717–2731.
Paus, T., 2001. Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat. Rev., Neurosci. 2 (6), 417–424.
Paus, T., Petrides, M., Evans, A., Meyer, E., 1993. Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J. Neurophysiol. 70 (2), 453–469.
Peach, R., Tonkovich, J., 2004. Phonemic characteristics of apraxia of speech resulting from subcortical hemorrhage. J. Commun. Disord. 37 (1), 77–90.
Penfield, W., Boldrey, E., 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain 60 (Pt. 4), 389–443.
Petersen, S., Fox, P., Posner, M., Mintun, M., Raichle, M., 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331 (6157), 585–589.
Picard, N., Strick, P., 2001. Imaging the premotor areas. Curr. Opin. Neurobiol. 11 (6), 663–672.
Preston, A., Thomason, M., Ochsner, K., Cooper, J., Glover, G., 2004. Comparison of spiral-in/out and spiral-out BOLD fMRI at 1.5 and 3 T. NeuroImage 21 (1), 291–301.
Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G., Haider, H., Grodd, W., 2000. Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. Brain Lang. 75 (2), 259–276.
Riecker, A., Mathiak, K., Wildgruber, D., Erb, M., Hertrich, I., Grodd, W., Ackermann, H., 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology 64 (4), 700–706.
Rimol, L., Specht, K., Weis, S., Savoy, R., Hugdahl, K., 2005. Processing of sub-syllabic speech units in the posterior temporal lobe: an fMRI study. NeuroImage 26 (4), 1059–1067.
Rorden, C., Karnath, H., 2004. Using human brain lesions to infer function: a relic from a past era in the fMRI age? Nat. Rev., Neurosci. 5 (10), 813–819.

Rosen, H., Ojemann, J., Ollinger, J., Petersen, S., 2000. Comparison of brain activation during word retrieval done silently and aloud using fMRI. Brain Cogn. 42 (2), 201–217.

Schmahmann,, 2000. MRI Atlas of the Human Cerebellum. Academic Press.

Shuster, L., Lemieux, S., 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang. 93 (1), 20–31.

Smith, S., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155.

Sörös, P., Cornelissen, K., Laine, M., Salmelin, R., 2003. Naming actions and objects: cortical dynamics in healthy adults and in an anomic patient with a dissociation in action/object naming. NeuroImage 19 (4), 1787–1801.

Sporns, O., Chialvo, D., Kaiser, M., Hilgetag, C., 2004. Organization, development and function of complex brain networks. Trends Cogn. Sci. 8 (9), 418–425.

Szirtes, J., Vaughan, H., 1977. Characteristics of cranial and facial potentials associated with speech production. Electroencephalogr. Clin. Neurophysiol. 43 (3), 386–396.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme.

Tanaka, H., Fujita, N., Watanabe, Y., Hirabuki, N., Takanashi, M., Oshiro, Y., Nakamura, H., 2000. Effects of stimulus rate on the auditory cortex using fMRI with 'sparse' temporal sampling. NeuroReport 11 (9), 2045–2049.

Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B., Grodd, W., 1996. Functional lateralization of speech production at primary motor cortex: a fMRI study. NeuroReport 7 (15–17), 2791–2795.

Wildgruber, D., Ackermann, H., Grodd, W., 2001. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. NeuroImage 13 (1), 101–109.

Wise, R., Greene, J., Büchel, C., Scott, S., 1999. Brain regions involved in articulation. Lancet 353 (9158), 1057–1061.

Wise, R., Scott, S., Blank, S., Mummery, C., Murphy, K., Warburton, E., 2001. Separate neural subsystems within 'Wernicke's area'. Brain 124 (Pt. 1), 83–95.

Woolrich, M., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S., 2004. Multilevel linear modelling for fMRI group analysis using Bayesian inference. NeuroImage 21 (4), 1732–1747.

Yetkin, F., Haughton, V., Cox, R., Hyde, J., Birn, R., Wong, E., Prost, R., 1996. Effect of motion outside the field of view on functional MR. AJNR Am. J. Neuroradiol. 17 (6), 1005–1009.

Zatorre, R., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. Cereb. Cortex 11 (10), 946–953.