

# Motor cortex maps articulatory features of speech sounds

Friedemann Pulvermüller\*, Martina Huss†, Ferath Kherif‡, Fermin Moscoso del Prado Martin, Olaf Hauk, and Yury Shtyrov

Cognition and Brain Sciences Unit, Medical Research Council, 15 Chaucer Road, Cambridge CB2 2EF, United Kingdom

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved March 17, 2006 (received for review November 17, 2005)

The processing of spoken language has been attributed to areas in the superior temporal lobe, where speech stimuli elicit the greatest activation. However, neurobiological and psycholinguistic models have long postulated that knowledge about the articulatory features of individual phonemes has an important role in their perception and in speech comprehension. To probe the possible involvement of specific motor circuits in the speech-perception process, we used event-related functional MRI and presented experimental subjects with spoken syllables, including [p] and [t] sounds, which are produced by movements of the lips or tongue, respectively. Physically similar nonlinguistic signal-correlated noise patterns were used as control stimuli. In localizer experiments, subjects had to silently articulate the same syllables and, in a second task, move their lips or tongue. Speech perception most strongly activated superior temporal cortex. Crucially, however, distinct motor regions in the precentral gyrus sparked by articulatory movements of the lips and tongue were also differentially activated in a somatotopic manner when subjects listened to the lip- or tongue-related phonemes. This sound-related somatotopic activation in precentral gyrus shows that, during speech perception, specific motor circuits are recruited that reflect phonetic distinctive features of the speech sounds encountered, thus providing direct neuroimaging support for specific links between the phonological mechanisms for speech perception and production.

cell assembly | functional MRI | perception–action cycle | mirror neurons | phonetic distinctive feature

Neurological theories of language have a long-standing tradition of distinguishing specialized modular centers for speech perception and speech production in left superior temporal and inferior frontal lobes, respectively (1–3). Such separate speech-production and -perception modules are consistent with a number of neuroimaging studies, especially the observations that frontal circuits become most strongly active during speech production and that speech input primarily activates the left superior temporal gyrus and sulcus (4–6). Superior temporal speech-perception mechanisms in humans may be situated in areas homologous to the auditory belt and parabelt areas in monkeys (5, 7, 8). In macaca, this region includes neurons specialized for species-specific calls (9, 10). Therefore, it appeared to be reasonable to postulate a speech-perception module confined to temporal cortex specifically processing acoustic information that is immanent to speech.

In contrast to this view, neurobiological models have long claimed that speech perception is connected to production mechanisms (11–16). Similar views have been proposed in psycholinguistics. For example, the direct realist theory of speech perception (17, 18) postulates a link between motor and perceptual representations of speech. According to the motor theory of Liberman *et al.* (19, 20), speech perception requires access to phoneme representations that are conceptualized as both speech-specific and innate. In contrast, neurobiological models put that links between articulatory and perceptual mechanisms are not necessarily speech-specific but may resemble action–perception links that are documented for a range of

nonlinguistic actions (21–23). These links can be explained by neuroscience principles. Because speaking implies motor activity and also auditory stimulation by the self-produced sounds, there is correlated neuronal activity in auditory and articulatory motor systems, which predicts synaptic strengthening and formation of specific articulatory–auditory links (24, 25).

Action theories of speech perception predict activation of motor circuits during speech processing, which is an idea that is supported, to an extent, by recent imaging work. During the presentation of syllables and words, areas in the left inferior frontal and premotor cortex become active along with superior temporal areas (26–30). Also, magnetic-stimulation experiments suggest motor facilitation during speech perception (31–33), and magnetoencephalography revealed parallel near-simultaneous activation of superior temporal areas and frontal action-related circuits (34). This latter finding supports a rapid link from speech-perception circuits in superior temporal lobe to the inferior frontal premotor and motor machinery for speech. The reverse link seems to be effective, too, because during speaking, the superior temporal cortex was activated along with areas in left inferior motor, premotor, and prefrontal cortex, although it was ensured that self-produced sounds could not be perceived through the auditory channel (35). This evidence for mutual activation of left inferior frontal and superior temporal cortex in both speech production and perception suggests access to motor programs and action knowledge in speech comprehension (23).

However, to our knowledge, the critical prediction of an action theory of speech perception has not been addressed. If motor information is crucial for recognizing speech, information about the critical distinctive features of speech sounds must be reflected in the motor system. Because the cortical motor system is organized in a somatotopic fashion, features of articulatory activities are controlled in different parts of motor and premotor cortex. The lip cortex lies superior to the tongue cortex (36–39). Therefore, the critical question is whether these specific motor structures are also differentially activated by incoming speech signals. According to recent theory and data, auditory perception–action circuits with different cortical topographies have an important role in language comprehension, especially at the level of word and sentence meaning (22, 23, 30, 40). However, it remains to be determined whether such specific perception–action links can be demonstrated at the level of speech sounds, in the phonological and phonetic domain, as a perception–action model of speech perception would predict. In the set of func-

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: fMRI, functional MRI; ROI, region of interest; MNI, Montreal Neurological Institute; EMG, electromyogram.

\*To whom correspondence should be addressed. E-mail: friedemann.pulvermuller@mrc-cbu.cam.ac.uk.

†Present address: Faculty of Education, University of Cambridge, Cambridge CB2 2PQ, United Kingdom.

‡Present address: Functional Imaging Laboratory, University College London, London WC1N 3BG, United Kingdom.

© 2006 by The National Academy of Sciences of the USA

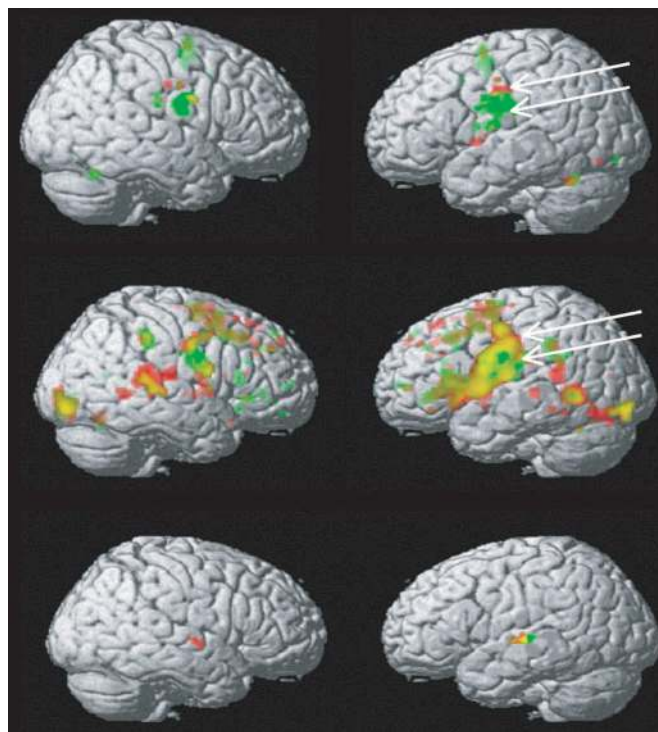
tional MRI (fMRI) experiments presented here, we investigated this critical prediction and sought to determine whether information about distinctive features of speech sounds is present in the sensorimotor cortex of the language-dominant left hemisphere during speech processing.

As language sounds are being produced by different articulators, which are somatotopically mapped on different areas of motor and premotor cortex, the perception of speech sounds produced by a specific articulator movement may activate its motor representation in precentral gyrus (13, 15, 24). To map the somatotomy of articulator movements and articulations, localizer tasks were administered by using fMRI. Subjects had to move their lips or tongue (motor experiment) and silently produce consonant–vowel (CV) syllables, starting with the lip-related (bilabial) and tongue-related (here, alveolar) phonemes [p] and [t] (articulation experiment). These tasks were performed by the same subjects who, in a different experiment, had listened to a stream of CV syllables that also started with [p] and [t] intermixed with matched patterns of signal-correlated noise while sparse-imaging fMRI scanning was administered (speech-perception experiment). Comparison of evoked hemodynamic activity among the motor, articulation, and listening experiments demonstrated for the first time that precentral motor circuits mirror motor information linked to phonetic features of speech input.

## Results

Motor movements strongly activated bilateral sensorimotor cortex, silent articulation led to widespread sensorimotor activation extending into inferior frontal areas, and speech perception produced the strongest activation in superior temporal cortex bilaterally (Fig. 1). An initial set of analyses looked at the laterality of cortical activation in the motor, articulation, and speech-perception experiments in regions of interest (ROIs) placed in these most active areas. In the motor task, maximal activation in sensorimotor cortex did not reveal any laterality. Activation peaks were seen at the Montreal Neurological Institute (MNI) standard coordinates ( $-56, -6, 26$ ; and  $62, 0, 28$ ) in the left and right hemispheres, respectively, and the statistical comparison did not reveal any significant differences (ROIs:  $\pm 59, -3, 27$ ;  $P > 0.2$ , no significant difference). In the silent articulation task, left hemispheric activation appeared more widespread than activity in the right nondominant hemisphere. However, statistical comparison between peak activation regions in each hemisphere, which were located in precentral gyri close to maximal motor activations, failed to reveal a significant difference between the hemispheres (ROI:  $\pm 59, -3, 27$ ;  $P > 0.2$ , no significant difference). Critically, the comparison of an inferior-frontal area overlapping with Broca's area and the homotopic area on the right, where language laterality is commonly found, revealed stronger activation on the left than on the right during silent articulation [ROIs:  $\pm 50, 15, 20$ ;  $F(1,11) = 6.3$ ,  $P < 0.03$ ]. Activation maxima during the speech-perception experiment were seen at MNI coordinates ( $-38, -34, 12$ ) and ( $42, -24, 10$ ) and were greater in the left superior temporal cortex than in the right [ROIs:  $\pm 40, -29, 11$ ;  $F(1,11) = 7.72$ ,  $P < 0.01$ ]. Because language-related activation was left-lateralized and *a priori* predictions on phoneme specificity were on left-hemispheric differential activation, detailed statistical analyses of phoneme specificity focused on the left language-dominant hemisphere.

In the motor-localizer experiment, distinct patterns of cortical activation were elicited in the lip and tongue movement conditions (Fig. 1 *Top*). In the left and right inferior precentral and postcentral gyri, the activation related to lip movements (shown in red) was dorsal to that of tongue movements (shown in green). This finding supports somatotopic organization of motor cortices responsible for articulator movements and is in good agreement

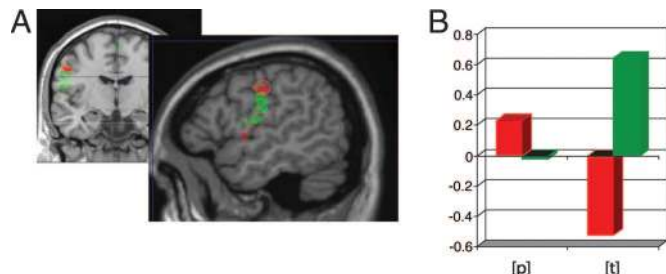


**Fig. 1.** Activation elicited by actions and perceptions related to the lips and tongue. (*Top*) Repetitive lip movements (red) and tongue movements (green) tested against rest as a baseline ( $P < 0.05$ , family-wise-error-corrected). (*Middle*) Repetitive silent articulation of syllables including the lip-related phoneme [p] (red) and the tongue-related phoneme [t] (green) tested against rest ( $P < 0.05$ , false-discovery-rate-corrected). (*Bottom*) Listening to spoken syllables, including the lip-related phoneme [p] (red) and the tongue-related phoneme [t] (green) tested against listening to noise with spectrotemporal characteristics that matched those of the syllables ( $P < 0.05$ , family-wise-error-corrected).

with refs. 36 and 39. There was also specific activation of the foot of the left precentral gyrus (where it touches the pars opercularis) to lip movements, and there was some left dorsolateral prefrontal activation in the vicinity of the frontal eye field to tongue movements. Left-hemispheric activation maxima for lip movements were present at MNI standard coordinates ( $-48, -10, 36$ ) and ( $-50, -10, 46$ ) and for tongue movements at ( $-56, -8, 28$ ) and ( $-56, 0, 26$ ). Note that these activation maxima were in left pre- and postcentral sensorimotor cortex. All activations extended into the precentral gyrus.

The articulation task elicited widespread activation patterns in frontocentral cortex, especially in the left hemisphere, which differed between articulatory gestures (Fig. 1 *Middle*). Consistent with the results from the motor task, the articulations involving the tongue (t phoneme) activated an inferior precentral area (shown in green), whereas the articulations involving the lips (p phoneme) activated motor and premotor cortex dorsal and ventral to this tongue-related spot (shown in red). In addition, there was activation in dorsolateral prefrontal and temporo-parietal cortex and in the cerebellum. In precentral gyrus, activation patches for the tongue-related actions resembled and overlapped between articulatory and motor tasks and similar overlap was also found between lip-related motor and articulatory actions (see arrows in Fig. 1 and see ref. 38).

ROIs for comparing activity related to different speech stimuli were selected on the basis of both localizer studies, the motor and articulatory experiments. Within the areas that were activated in the motor experiments, the voxels in the central sulcus, which were most strongly activated by lip and tongue movements,



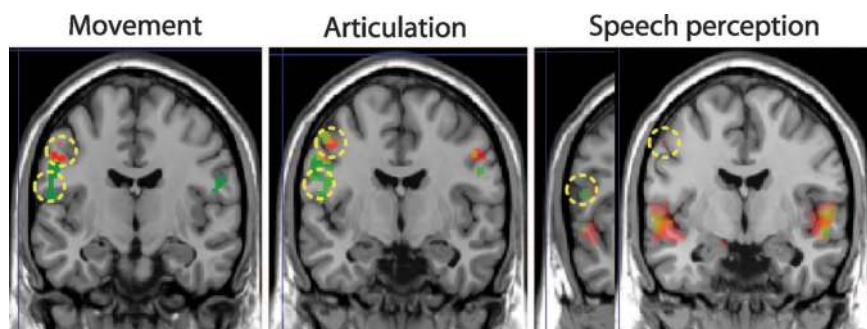
**Fig. 2.** Phoneme-specific activation in precentral cortex during speech perception. (*A*) Frontal and sagittal slices of the standard MNI brain showing activation during lip (red) and tongue (green) movement along with the ROIs centered in precentral gyrus. (*B*) Activation (arbitrary units) during listening to syllables including [p] and [t] in precentral ROIs where pronounced activation for lip (red) and tongue (green) movements was found. The significant interaction demonstrates differential activation of these motor areas related to perception of [p] and [t].

respectively, were selected as the centers of 8-mm radius ROI (lip-central ROI:  $-56, -8, 46$ ; tongue-central ROI:  $-60, -10, 25$ ). Anterior to these postcentral ROIs, a pair of ROIs in precentral cortex were selected (lip precentral ROI:  $-54, -3, 46$ ; tongue premotor ROI:  $-60, 2, 25$ ). The precentral ROIs were also differentially activated by the articulation of the lip-related phoneme [p] and the tongue-related phoneme [t] in the articulation experiment (arrows in Fig. 1). This dissociation was substantiated by ANOVA, revealing a significant interaction of place of articulation and region factors for the precentral ROIs [ $F(1,11) = 6.03, P < 0.03$ ]. The noise- and speech-perception experiment produced strong activation in superior temporal cortex. Consistent with refs. 41 and 42, noise elicited activation in Heschl's gyrus and superior temporal cortex surrounding it. Additional noise-related activation was found in the parietotemporal cortex, especially in the supramarginal gyrus. Speech stimuli activated a greater area in superior temporal cortex, including large parts of superior temporal gyrus and extending into superior temporal sulcus (Fig. 1 *Bottom*), which probably represents the human homologue of the auditory parabelt region (5, 9). The perception of syllables including the lip-related phoneme [p] produced strongest activation at MNI coordinate  $(-60, -12, 0)$ , whereas that to the tongue-related phoneme [t] was maximal at coordinate  $(-60, -20, 2)$ , which is 8 mm posterior-superior to the maximum activation elicited by [p]. This finding is consistent with distinct local cortical processes for features of speech sounds in superior temporal cortex (43, 44).

To test whether the lip and tongue motor areas that are implied in motor and articulatory processing were differentially activated by the perception of lip- and tongue-related phonemes, speech-evoked activation was investigated in the postcentral and precentral ROIs. In postcentral ROIs, the interaction of the factors place of articulation and ROI did not reach significance

( $F < 1, P > 0.2$ , no significant difference). In contrast, precentral ROIs revealed a significant interaction of these factors [ $F(1,11) = 7.25, P < 0.021$ ] because of stronger [t]- than [p]-induced activity in the ventral ROI and the opposite pattern, stronger [p]- than [t]-related blood flow in the dorsal precentral region (Figs. 2 and 3).

These results, obtained from only one *a priori*-defined pair of regions, were complemented by a broader analysis of activation in a larger part of the motor strip, which, to this end, was divided into smaller regions. A chain of seven such regions [spheres with a radius of 5 mm and a center-to-center distance of 5 mm in the vertical ( $z$ ) direction] was aligned along the central sulcus of the standard MNI brain between vertical ( $z$ ) coordinates 20 and 50. Anterior to it, at a distance of 5 mm, an additional line of regions in precentral cortex was defined. This subdivision of precentral and central cortex resulted in an array of  $2 \times 7 = 14$  regions, for each of which activation values were obtained for each experiment (motor, articulation, and speech perception), condition (lip- vs. tongue-related), and subject. As previously done, motor and articulation activity were computed relative to the rest baseline and for the speech-perception condition, activation was calculated relative to averaged noise-evoked activity. Three-way ANOVAs were run with the design articulator (tongue/lips)  $\times$  frontality (precentral/central)  $\times$  dorsality (seven regions, inferior to superior). For all three experiments, significant two-way interactions of the articulator and dorsality factors emerged [motor:  $F(6,66) = 8.13, P < 0.0001$ ; articulation:  $F(6,66) = 8.43, P < 0.0001$ ; speech perception:  $F(6,66) = 3.18, P < 0.008$ ], proving differential activation of frontocentral cortex by the lip- and tongue-related actions, articulations, and speech perceptions. When contrasting activity in the upper vs. lower three regions ( $y$  coordinate,  $20\text{--}30$  vs.  $40\text{--}50$ ) in precentral and central cortex between tongue and lip motor movements, a significant difference emerged again [ $F(1,11) = 6.97, P < 0.02$ ] because of relatively stronger activation for lip movements in the superior regions and stronger tongue movement activation in inferior precentral regions. The same difference was also present for the articulation task, where syllables including the lip-related phoneme [p] led to stronger superior activation, whereas the tongue-related [t] articulations elicited relatively stronger hemodynamic responses in inferior regions [ $F(1,11) = 8.43, P < 0.001$ ]. Crucially, also the listening task showed the same differential activation with relatively stronger dorsal activation to the lip phoneme [p] and relatively stronger inferior activation to the tongue phoneme [t] [ $F(1,11) = 4.84, P < 0.05$ ]. Note that the clearest differential activation was determined in precentral cortex for both articulation and speech perception, whereas for the nonverbal motor task, differential activation was most pronounced at central sites. Because the interactions of the vowel factor with the articulator factor or topographical variables did not reach significance, there was no strong support for differential local vowel-specific motor activation or a change of the



**Fig. 3.** Frontal slices showing differential activation elicited during lip and tongue movements (*Left*), syllable articulation including [p] and [t] (*Center*), and listening to syllables including [p] and [t] (*Right*). Note the superior-central activations for lip-related actions and perceptions (red) and the inferior-central activations for tongue-related ones (green). The significance thresholds are  $P < 0.05$ , family-wise-error-corrected for movements, false-discovery-rate-corrected for articulations, and uncorrected for speech perception.

regionally specific differential activation for [p] and [t] between vowel contexts from this set of analyses.

Analysis of the electromyogram (EMG) data failed to reveal significant differences in neurophysiological activity of the lip muscles when subjects listened to syllables starting with [p] and [t]. The EMG after the syllable presentation did not indicate any neurophysiological distinctions between the different types of stimuli (rms of activity in 200- to 400-ms time interval: 0.54 vs. 0.46 mV;  $F < 2$ ,  $P > 0.2$ , no significant difference). Analysis of longer and later time intervals also failed to suggest differential activation of articulator muscles during syllable perception in a sparse imaging context. This result does not support the possibility that the differential activation in precentral cortex seen during the fMRI study was related to overt or subliminal motor activity in the peripheral motor system elicited by syllable perception.

## Discussion

Differential activation of the cortical motor system, in central and precentral areas, was observed during passive listening to phonemes produced by different articulators (lips and tongue) at different places of articulation. Localizer experiments using nonverbal motor movements and silent articulatory movements involving the lips and tongue showed that the differential activation in precentral cortex to speech sounds was consistent with the somatotopic representation in motor and premotor cortex of the relevant articulatory movements and nonverbal movements of the articulators. The precentral regions consistently activated by nonverbal movements, syllable articulation, and speech perception demonstrate a shared speech-sound-specific neuronal substrate of these sensory and motor processes.

This set of findings is evidence that information about specific articulatory and motor features of speech sounds is accessed in speech perception. Because we defined the areas of differential activation in articulation and motor tasks that involved the tongue and lips, respectively, we can state that the differential activation of these same areas in the speech-perception experiment includes information about articulatory features of speech sound stimuli. This finding supports neurobiological models of language postulating specific functional links in perisylvian cortex between auditory and articulatory circuits processing distinctive features of speech (11, 13, 15, 24, 45). The data also provide support for cognitive models postulating perception-action links in phonological processing (for example, the direct realist theory) (17, 18).

The motor theory of speech perception (MTSP; refs. 19 and 20) also postulated reference to motor gestural information in speech processing. Therefore, the data presented here are compliant with this postulate immanent to this theory. However, it is an *a priori* that neuroimaging evidence cannot, *per se*, prove the necessity of motor information access for the cognitive process under investigation. Therefore, the data set presented here cannot prove that motor representation access is necessary for speech perception, as the MTSP claims. Crucially, however, these data suggest a shared neural substrate (in precentral cortex) of feature-specific perceptual and articulatory phonological processes and of body-part-specific nonlinguistic motor processes. This finding may be interpreted as evidence against one of the major claims of the MTSP (namely, the idea that “speech is special,” being processed in a speech-specific module not accessed when nonlinguistic actions are performed). More generally, the idea that motor links are required for discriminating between speech sounds needs to be considered in the light of the evidence for categorical speech perception in early infancy (46) and in animals (47, 48).

The results indicate that the precentral cortex is the site where the clearest and most consistent differential somatotopic activation was produced by motor movements of the articulators, by

articulatory movements performed with these same articulators, and, last, by the perception of the speech sounds elicited by moving these very articulators. Because the precentral gyrus includes both primary motor and premotor cortex, it may be that both of these parts of the motor system participate not only in articulatory planning and execution but also in speech perception. Because primary and premotor cortex cannot be separated with certainty on the basis of the methods used in this study, the question of distinguishable functional roles of these areas in the speech-perception process remains to be determined. A related interesting issue addresses the role of attention and task contexts. Previous research indicated that inferior-frontal and frontocentral activation to speech can occur in a range of experimental tasks and even under attentional distraction (29, 34), but there is also evidence that inferior-frontal activation is modulated by phonological processes (49, 50). The influence of attention on the dynamics of precentral phoneme-feature-specific networks remains to be determined.

Our results might be interpreted in the sense that perceptual representations of phonemes, or perceptual phonemic processes, are localized in precentral cortex. However, we stress that such a view presuming phonemes as a cognitive and concrete neural entity, which is immanent to some models, is not supported by the data presented here. Instead, these data are compatible as well with models postulating phonological systems in which phonemic distinctive features are the relevant units of representation, independent of the existence of an embodiment of phonemes *per se* (44, 51, 52). We believe that the least-biased linguistic interpretation of our results is as follows: articulatory phonetic features [+alveolar] (or, perhaps more accurately, [+tongue-tip]) and [+bilabial] are accessed in both speech perception and listening, and this access process involves regionally specific areas in precentral gyrus. Remarkably, the regionally specific differential activation for [p] and [t] did not significantly change between the vowel contexts of [ɪ] or [æ] in which the consonants were presented. This observation is consistent with the suggestion that the different perceptual patterns corresponding to the same phoneme or phonemic feature are mapped onto the same gestural and motor representation.

These results indicate that perception of speech sounds in a listening task activates specific motor circuits in precentral cortex, and that these are the same motor circuits that also contribute to articulatory and motor processes invoked when the muscles generating the speech sounds are being moved. This outcome supports a general theory of language and action circuits realized by distributed sensorimotor cell assemblies that include multimodal neurons in premotor and primary motor cortex and in other parts of the perisylvian language areas as well (13, 22–24). Generally, the results argue against modular theories, which state that language is stored away from other perceptual and motor processes in encapsulated modules that are devoted to either production or comprehension of language and speech. Rather, it appears that speech production and comprehension share a relevant part of their neural substrate, which also overlaps with that of nonlinguistic activities involving the body parts that express language (21–23, 53).

## Materials and Methods

**Subjects.** The participants of the study were 12 (seven females and five males) right-handed monolingual native speakers of English without any left-handed first-order relatives. They had normal hearing and no history of neurological or psychiatric illness or drug abuse. The mean age of the volunteers was 27.3 years (SD, 4.9). Ethical approval was obtained from the Cambridge Local Research Ethics Committee.

**Stimuli and Experimental Design.** Volunteers lay in the scanner and were instructed to avoid any movement (except for the subtle

movements required in the articulatory and motor tasks). The study consisted of three experiments that were structured into four blocks and administered in the order of (i) a speech-perception experiment, which was divided into two blocks (24 min each); (ii) an articulation-localizer experiment (5.4 min); and (iii) a motor-localizer experiment (3.2 min). The motor and articulation tasks were administered by using block designs with continuous MRI scanning, whereas sparse imaging with sound presentation in silent gaps was used in the speech-perception experiment. Here, we describe these experiments in the reverse order because the motor and articulatory experiments were used to localize the critical areas for the analysis of the speech-perception experiment.

**Motor-Localizer Experiment.** Subjects had to perform repetitive minimal movements of the lips and tip of the tongue upon written instruction. The word “lips” or “tongue” appeared on a computer screen for 16 s, and subjects had to perform the respective movement during the entire period. In an additional condition, a fixation cross appeared on the screen and the subjects had to rest. Each task occurred four times, thus resulting in 12 blocks, which were administered in pseudorandom order (excluding direct succession of two blocks of the same type). Minimal alternating up-and-down movements of the tongue tip and the lips were practiced briefly outside the scanner before the experiment. Subjects were instructed to minimize movements, and they avoided somatosensory self-stimulation by lip closure or tongue-palate contact as much as possible.

**Articulation-Localizer Experiment.** Subjects had to “pronounce silently, without making any noise” the syllables [pI], [pæ], [tI], and [tæ] upon written instruction. The letter string “PIH,” “PAH,” “TIH,” or “TAH” appeared on a computer screen for 16 s, and subjects had to perform the respective articulatory movement repeatedly during the entire period. In an additional condition, a fixation cross appeared on screen and subjects had to rest. Each instruction occurred four times, thus resulting in 20 blocks, which were administered in pseudorandom order.

**Speech-Perception Experiment.** Speech and matched noise stimuli were presented through nonmagnetic headphones. Volunteers were asked “to listen and attend to the sounds and avoid any movement or motor activity during the experiment.” To reduce the interfering effect of noise from the scanner, which would acoustically mask the stimuli and lead to saturation of auditory cortical area activation (54), the so-called “sparse imaging” paradigm was applied. Stimuli appeared in 1.6-s gaps between MRI scans, each also lasting 1.6 s. All speech and noise stimuli were 200 ms; their onset followed the preceding scanner noise offset by 0.4–1.2 s (randomly varied). We presented 100 stimuli of each of eight stimulus types (see next paragraph) and 100 null events with silence in the gap between scans in a pseudorandom order, which was newly created for each subject.

The speech stimuli included different tokens of the following four syllable types: [pI], [pæ], [tI], and [tæ]. A large number of syllables of each type were spoken by a female native speaker of English and recorded on computer disk. Items that were maximally similar to each other were selected. These similar syllables from the four types were then adjusted so that they were exactly matched for length, overall sound energy,  $F_0$  frequency and voice onset time. Matched noise stimuli were generated for each recorded syllable by performing a fast-Fourier transform analysis of the spoken stimuli and subsequently modulating white noise according to the frequency profile and the envelope of the syllable (55). These spectrally adjusted signal-correlated noise stimuli were also matched for length, overall sound energy, and  $F_0$  frequency to their corresponding syllables and among each other. From each recorded syllable and each noise pattern, four

variants were generated by increasing and decreasing  $F_0$  frequencies by 2% and 4%, thus resulting in 20 different syllables and 20 different noise patterns (5 from each type; 40 different stimuli).

**Imaging Methods.** Subjects were scanned in a 3-T magnetic resonance system (Bruker, Billerica, MA) using a head coil. Echo-planar imaging (EPI) was applied with a repetition time of 3.2 s in the speech-perception task (1.6 s of scanning and 1.6 s of silent gap) and 1.6 s in the motor and articulation-localizer tasks. The acquisition time (TA) was 1.6 s, and the flip angle was 90°. The functional images consisted of 91 slices of  $64 \times 64$  voxels covering the whole brain. Voxel size was  $3 \times 3 \times 2.25$  mm, with an interslice distance of 0.75 mm, thus resulting in an isotropic scanning volume. Isotropic scanning was applied because the major prediction was about activation differences along the  $z$  axis. Imaging data were processed by using the Automatic-analysis for neuroimaging (aa) package (Cognition and Brain Sciences Unit, Medical Research Council) with SPM2 software (Wellcome Department of Cognitive Neurology, London). Images were corrected for slice timing and then realigned to the first image by using sinc interpolation. Field maps were used to correct for image distortions resulting from inhomogeneities in the magnetic field (56, 57). Any nonbrain parts were removed from the T1-weighted structural images by using a surface-model approach (“skull-stripping;” see ref. 58). The EPI images were coregistered to these skull-stripped structural T1 images by using a mutual information coregistration procedure (59). The structural MRI was normalized to the 152-subject T1 template of the MNI. The resulting transformation parameters were applied to the coregistered EPI images. During the spatial normalization process, images were resampled with a spatial resolution of  $2 \times 2 \times 2$  mm. Last, all normalized images were spatially smoothed with a 6-mm full-width half-maximum Gaussian kernel, and single-subject statistical contrasts were computed by using the general linear model (60). Low-frequency noise was removed with a high-pass filter (motor localizer: time constant, 160 s; articulation task, 224 s; listening task, 288 s). The predicted activation time course was modeled as a box-car function (blocked motor and articulatory experiments) and as a box-car function convolved with the canonical hemodynamic response function that is included in SPM2 (sparse-imaging speech-perception experiment).

Group data were analyzed with a random-effects analysis. A brain locus was considered to be activated in a particular condition if 27 or more adjacent voxels all passed the threshold of  $P = 0.05$  corrected. For correction for multiple comparisons, the family-wise-error and false-discovery-rate methods were used (61). Stereotaxic coordinates for voxels with maximal  $t$  values within activation clusters are reported in the MNI standard space, which very closely resembles the standardized Talairach (62) space (63).

We computed the average parameter estimates over voxels for each individual subject in the speech-perception task for spherical ROIs in the left hemisphere (radius, 8 mm). The ROIs were defined on the basis of the random-effects analysis as differentially activated by movement and articulation types. This definition was done by using MARSBAR software utility (63). Values for each of the four syllable types that were subtracted by the values obtained for the average of all noise stimuli were subjected to ANOVA, including the following factors: ROI (lip or tongue movement-related activity foci), place of articulation of the speech sound (bilabial/lip-related [p] vs. alveolar/tongue tip-related [t]), and vowel included in the speech sound ([I] vs. [æ]). Additional statistical analyses were performed on ROIs aligned along the motor strip, as described in *Results*. For the analysis of cortical laterality, ROIs were defined as spheres with a radius of 1 cm around the points of maximal activation in the

left and right hemispheres (comparisons: movements vs. rest, articulations vs. rest, and speech/sound perception vs. silence). Because coordinates of activation maxima were similar in the two hemispheres, the average of the left- and right-hemispheric activation maxima was used to define the origins of symmetric ROI pairs used for statistical testing. Additional ROIs were defined based on refs. 30 and 64 so that they comprised a large portion of Broca's area and its right-hemispheric counterpart.

**EMG Movement-Control Experiment.** Although subjects were instructed explicitly to refrain from any motor movement during the speech-perception experiment, one may argue that syllable perception might evoke involuntary movements that may affect the blood oxygenation level-dependent response. Because EMG provides accurate information about muscle activity even if it does not result in an overt movement, this method was applied in a control experiment. Because temporal gradients of strong magnetic fields produce huge electrical artifacts and thus noise in the EMG recording, this experiment was performed outside the scanner, in the EEG laboratory of the Cognition and Brain Sciences Unit of the Medical Research Council. To avoid repetition effects, a different population of eight subjects (mean age, 25.3; SD, 6.8; all satisfying the criteria for subject selection described above) was tested. The

experiment was an exact replication of the fMRI speech-perception experiment, including also the scanner noise produced during sparse imaging. One pair of EMG electrodes were placed to the left and right above and below the mouth, respectively (total of eight electrodes) at standard EMG recording sites targeting the orbicularis oris muscle (65). This muscle would show strong EMG activity during the articulation of a lip-related consonant and during lip movements but would do so much less during tongue-related phoneme production or tongue movements. During articulations of syllables starting with [p] and [t], EMG responses at these electrodes immediately differentiated between these phonemes, with maximal differences occurring at 200–400 ms. EMG responses to speech stimuli were evaluated by calculating averages of rms values for all four bipolar recordings and averaging data across electrodes in the time window of 200–400 ms after syllable onset. Additional longer and later time intervals were analyzed to probe the consistency of the results.

We thank Rhodri Cusack, Matt Davis, Ingrid Johnsrude, Christian Schwarzbauer, William Marslen-Wilson, and two anonymous referees for their help at different stages of this work. This research was supported by the Medical Research Council of the United Kingdom and European Community Information Society Technologies Program Grant IST-2001-35282.

- Wernicke, C. (1874) *Der Aphasische Symptomencomplex. Eine Psychologische Studie auf Anatomischer Basis* (Kohn and Weigert, Breslau, Germany).
- Lichtheim, L. (1885) *Brain* **7**, 433–484.
- Damasio, A. R. & Geschwind, N. (1984) *Annu. Rev. Neurosci.* **7**, 127–147.
- Price, C. J. (2000) *J. Anat.* **197**, 335–359.
- Scott, S. K. & Johnsrude, I. S. (2003) *Trends Neurosci.* **26**, 100–107.
- Zatorre, R. J., Belin, P. & Penhune, V. B. (2002) *Trends Cognit. Sci.* **6**, 37–46.
- Patterson, R. D., Uppenkamp, S., Norris, D., Marslen-Wilson, W., Johnsrude, I. & Williams, E. (2000) in *Proceedings of the 6th International Conference of Spoken Language Processing (ICSLP, Beijing)*, Vol. II, pp. 1–4.
- Davis, M. H. & Johnsrude, I. S. (2003) *J. Neurosci.* **23**, 3423–3431.
- Rauschecker, J. P. & Tian, B. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 11800–11806.
- Tian, B., Reser, D., Durham, A., Kustov, A. & Rauschecker, J. P. (2001) *Science* **292**, 290–293.
- Fry, D. B. (1966) in *The Genesis of Language*, eds. Smith, F. & Miller, G. A. (MIT Press, Cambridge, MA), pp. 187–206.
- Braitenberg, V. (1980) in *L'Accostamento Interdisciplinare allo Studio del Linguaggio*, eds. Braga, G., Braitenberg, V., Cipolli, C., Coseriu, E., Crespi-Reghizzi, S., Mehler, J. & Titone, R. (Franco Angeli, Milan), pp. 96–108.
- Braitenberg, V. & Schüz, A. (1992) in *Language Origin: A Multidisciplinary Approach*, eds. Wind, J., Chiarelli, B., Bichakjian, B. H., Nocentini, A. & Jonker, A. (Kluwer, Dordrecht, The Netherlands), pp. 89–102.
- Fuster, J. M. (1998) *Neuron* **21**, 1223–1224.
- Pulvermüller, F. (1999) *Behav. Brain Sci.* **22**, 253–336.
- Fuster, J. M. (2003) *Cortex and Mind: Unifying Cognition* (Oxford Univ. Press, Oxford).
- Fowler, C. A. (1986) *J. Phonetics* **14**, 3–28.
- Fowler, C. A., Brown, J. M., Sabadini, L. & Wehling, J. (2003) *J. Mem. Lang.* **49**, 396–413.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) *Psychol. Rev.* **74**, 431–461.
- Lieberman, A. M. & Whalen, D. H. (2000) *Trends Cognit. Sci.* **4**, 187–196.
- Jeannerod, M., Arbib, M. A., Rizzolatti, G. & Sakata, H. (1995) *Trends Neurosci.* **18**, 314–320.
- Rizzolatti, G., Fogassi, L. & Gallese, V. (2001) *Nat. Rev. Neurosci.* **2**, 661–670.
- Pulvermüller, F. (2005) *Nat. Rev. Neurosci.* **6**, 576–582.
- Braitenberg, V. & Pulvermüller, F. (1992) *Naturwissenschaften* **79**, 103–117.
- Pulvermüller, F. & Preissl, H. (1991) *Network Comput. Neural Syst.* **2**, 455–468.
- Zatorre, R. J., Evans, A. C., Meyer, E. & Gjedde, A. (1992) *Science* **256**, 846–849.
- Siok, W. T., Jin, Z., Fletcher, P. & Tan, L. H. (2003) *Hum. Brain Mapp.* **18**, 201–207.
- Joanisse, M. F. & Gati, J. S. (2003) *NeuroImage* **19**, 64–79.
- Wilson, S. M., Saygin, A. P., Sereno, M. I. & Iacoboni, M. (2004) *Nat. Neurosci.* **7**, 701–702.
- Hauk, O., Johnsrude, I. & Pulvermüller, F. (2004) *Neuron* **41**, 301–307.
- Fadiga, L., Craighero, L., Buccino, G. & Rizzolatti, G. (2002) *Eur. J. Neurosci.* **15**, 399–402.
- Watkins, K. E., Strafella, A. P. & Paus, T. (2003) *Neuropsychologia* **41**, 989–994.
- Watkins, K. & Paus, T. (2004) *J. Cognit. Neurosci.* **16**, 978–987.
- Pulvermüller, F., Shtyrov, Y. & Ilmoniemi, R. J. (2003) *NeuroImage* **20**, 1020–1025.
- Paus, T., Perry, D. W., Zatorre, R. J., Worsley, K. J. & Evans, A. C. (1996) *Eur. J. Neurosci.* **8**, 2236–2246.
- Penfield, W. & Boldrey, E. (1937) *Brain* **60**, 389–443.
- McCarthy, G., Allison, T. & Spencer, D. D. (1993) *J. Neurosurg.* **79**, 874–884.
- Lotze, M., Seggewies, G., Erb, M., Grodd, W. & Birbaumer, N. (2000) *NeuroReport* **11**, 2985–2989.
- Hesselmann, V., Sorger, B., Lasek, K., Guntinas-Lichius, O., Krug, B., Sturm, V., Goebel, R. & Lackner, K. (2004) *Brain Topogr.* **16**, 159–167.
- Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F. & Perani, D. (2005) *J. Cognit. Neurosci.* **17**, 273–281.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C. & Lutkenhoner, B. (2003) *Cereb. Cortex* **13**, 765–772.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S. & Griffiths, T. D. (2002) *Neuron* **36**, 767–776.
- Obleser, J., Lahiri, A. & Eulitz, C. (2003) *NeuroImage* **20**, 1839–1847.
- Obleser, J., Lahiri, A. & Eulitz, C. (2004) *J. Cognit. Neurosci.* **16**, 31–39.
- Pulvermüller, F. (1992) *Concepts Neurosci.* **3**, 157–200.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P. W. & Vigorito, J. (1971) *Science* **171**, 303–306.
- Kuhl, P. K. & Miller, J. D. (1978) *J. Acoust. Soc. Am.* **63**, 905–917.
- Kluender, K. R., Diehl, R. L. & Killeen, P. R. (1987) *Science* **237**, 1195–1197.
- Burton, M. W., Small, S. L. & Blumstein, S. E. (2000) *J. Cognit. Neurosci.* **12**, 679–90.
- Zatorre, R. J., Meyer, E., Gjedde, A. & Evans, A. C. (1996) *Cereb. Cortex* **6**, 21–30.
- Lahiri, A. & Marslen-Wilson, W. (1991) *Cognition* **38**, 245–294.
- Marslen-Wilson, W. & Warren, P. (1994) *Psychol. Rev.* **101**, 653–675.
- Hauk, O., Shtyrov, Y. & Pulvermüller, F. (2006) *Eur. J. Neurosci.* **23**, 811–821.
- Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., Gurney, E. M. & Bowtell, R. W. (1999) *Hum. Brain Mapp.* **7**, 213–23.
- Shtyrov, Y., Pihko, E. & Pulvermüller, F. (2005) *NeuroImage* **27**, 37–47.
- Jezzard, P. & Balaban, R. S. (1995) *Magn. Reson. Med.* **34**, 65–73.
- Cusack, R., Brett, M. & Osswald, K. (2003) *NeuroImage* **18**, 127–142.
- Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P. M., Federico, A. & De Stefano, N. (2002) *NeuroImage* **17**, 479–489.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. & Suetens, P. (1997) *IEEE Trans. Med. Imaging* **16**, 187–198.
- Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D. & Turner, R. (1998) *NeuroImage* **7**, 30–40.
- Genovese, C. R., Lazar, N. A. & Nichols, T. (2002) *NeuroImage* **15**, 870–878.
- Talairach, J. & Tournoux, P. (1988) *Coplanar Stereotaxic Atlas of the Human Brain* (Thieme, New York).
- Brett, M., Johnsrude, I. S. & Owen, A. M. (2002) *Nat. Rev. Neurosci.* **3**, 243–249.
- Bookheimer, S. (2002) *Annu. Rev. Neurosci.* **25**, 151–188.
- Cacioppo, J. T., Tassinari, L. G. & Berntson, G. G. (2000) *Handbook of Psychophysiology* (Cambridge Univ. Press, Cambridge, U.K.).