THE ACT OF SPEAKING demands simultaneous control of respiratory, laryngeal, and articulatory systems. The underlying mechanisms have been studied since the late 17th century, and the difference between language formation and the control of speaking has been commented on by various authors [see McFarland and Lund (15) and Whitcombe (30) for reviews].

The principles of speech mechanisms were outlined by Gutzmann (9), but the mechanisms of neural control of breathing for speech remained obscure until Draper et al. (3) described how, during a continuous utterance (counting), coordinated activity of expiratory and inspiratory muscles resulted in the production of a constant expiratory pressure while the thorax and lung were becoming smaller. This analysis has been broadly confirmed in studies of singing (22). All these studies, together with reviews of both respiratory function in speech (11) and phonatory mechanisms (1), testify to the complexity of the required neural control mechanisms.

It is well established that partial or complete inability to speak properly as part of motor linguistic impairment results from lesions centered on the posterior third of the inferior frontal gyrus in the dominant hemisphere (2). More recently, a view has developed in the clinical literature that classical Broca’s aphasia only occurs when damage extends to areas in the frontal and parietal lobes and also to deeper structures (16). However, beyond this, it is difficult to assess the importance of Broca’s area in the central nervous organization of “speaking” without language processing. The significance of the controversies surrounding the function of Broca’s area and the associated disorder of motor linguistics was that it made it necessary to design an experimental paradigm that would permit the determination of the cerebral areas concerned with speaking while at the same time minimizing the higher processes involved in language formation and recall.

We chose to use positron emission tomography (PET) in an experimental design where the combination of different speech-related conditions would allow us to 1) highlight breathing for speech and vocalization while eliminating the effects of articulation and 2) highlight articulation independently of breathing for speech and vocalization. We have utilized repetition of a simple phrase that requires only labial movements (12). Results of these studies have been presented in preliminary form (17).

METHODS

Subjects, Measurements, and Protocol

Six healthy right-handed men (age range, 22–46 yr) were studied; they were naive to the purpose of the study. Local ethical committee approval (Hammersmith Hospitals’ Medical Ethics Committee) and permission to administer radioactivity (Administration of Radioactive Substances Advisory Committee of the Department of Health, UK) were obtained. Each subject gave written informed consent. All studies were performed in the supine position with the subject’s head in the PET scanner while undertaking a series of speech-related tasks (see Experimental Paradigm). Breathing was measured by using a calibrated direct current-coupled respiratory inductance plethysmograph (23), with bands placed around the chest and abdomen from which the following respiratory variables were derived: tidal volume, inspiratory and expiratory time, minute ventilation, and changes in functional residual capacity. Electromyographic (EMG) activity from the submental region and from a masseter muscle was monitored by using appropriately placed surface EMG electrodes (model H82P, ARBO) and a purpose-built isolated amplifier (band width, 10–1,000 Hz). Sound was recorded with a microphone attached to the skin on one side of the thyroid cartilage. Video recordings of the face were made throughout the experiment. Data were recorded onto a magnetic tape recorder (TEAC RD135-T DAT Recorder, Tokyo, Japan) for subsequent analysis. The EMG data, together with the sound and video recordings, were examined to ensure that the subject had
performed the required speech condition correctly. The mean values of the respiratory variables were compared for each condition by use of an analysis of variance (ANOVA) and Fisher's least significant difference (5).

**Experimental Paradigm**

The phrase “Buy Bobby a poppy” was utilized in an experimental paradigm with four conditions to highlight the various cerebral activations required to enunciate this phrase continuously (Table 1). Condition A consisted of saying the phrase repetitively at a normal conversational level. Condition B consisted of the subject's mouthing the phrase repetitively without making any sound, whereas condition C required the subject to vocalize the phrase without moving the mouth or tongue (i.e., no articulation of the phrase occurred). Condition D consisted of the subjects saying the phrase to themselves silently, without any muscle contraction (i.e., no articulation or vocalization). Examination of Table 1 shows that the phrase is articulated in conditions A and B but not in C and D, whereas the phrase is vocalized in conditions A and C but not in B and D. We expected, on the basis of pilot studies, that breathing for speaking would be present in A and C but not in B and D (Fig. 1). For the purposes of this study, we have separated the neural components of "speech" into five elements: language, breathing for speech, vocalization, articulation, and hearing. Table 1 shows those elements to be present in each of the four experimental conditions. We reasoned that contrasting the pattern of neural activation in the brain between conditions A and B would highlight areas associated with breathing for speech and vocalization and hearing oneself, with articulation present, while the contrast between conditions C and D would highlight the same areas, with articulation absent. Similarly, contrasting conditions A and C would highlight areas associated with articulation when control of breathing for speech and vocalization are present, whereas contrasting conditions B and D would highlight the areas associated with articulation when control of breathing for speech and vocalization are absent.

**Training**

Before the day of the study, subjects were instrumented as described above. They were familiarized with the speech conditions that they would be expected to perform by giving them a fixed set of instructions for each condition (Table 2) together with a demonstration by one of the investigators. Particular care was taken not to give any guidance on the breathing pattern during the performance of any of these conditions.

**PET Scanning**

PET scanning was performed, measuring relative regional cerebral blood flow (rCBF) by recording the regional distribution of cerebral radioactivity after the iv slow bolus injection of radiolabeled water (H215O). 15O is a positron emitter with a half-life of 2.1 min. Scans were performed in random order during the four conditions described above (Table 1; Fig. 1); each was repeated three times. Before each scan, the appropriate instruction was given, as used during the familiarization visit. Each scan was separated by a period of 10 min to allow for sufficient decay of radioactivity.

---

**Table 1. Elements present in 4 speech-related conditions**

<table>
<thead>
<tr>
<th>Breathing for Speech and Vocalization</th>
<th>No Breathing for Speech and No Vocalization</th>
<th>Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articulation</td>
<td>Condition A: Speech</td>
<td>Condition B: Mouthing speech</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>Language Breathing for speech</td>
</tr>
<tr>
<td>+</td>
<td>0</td>
<td>Vocalization</td>
</tr>
<tr>
<td>+</td>
<td>0</td>
<td>Articulation</td>
</tr>
<tr>
<td>+</td>
<td>0</td>
<td>Hearing</td>
</tr>
<tr>
<td>Condition C: Unarticulated speech</td>
<td>Condition D: Internal speech</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>Language Breathing for speech</td>
</tr>
<tr>
<td>+</td>
<td>0</td>
<td>Vocalization</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>Articulation</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>Hearing</td>
</tr>
</tbody>
</table>

The 4 speech-related conditions studied have been placed into appropriate rows (depending on whether articulation was present (+) or absent (0)) and appropriate columns (depending on whether or not there was breathing for speech and vocalization). Under each condition presence or absence of elements within that condition is indicated. Column on right lists these elements. Subtraction of condition B from A and of condition D from C leaves breathing for speech, vocalization, and hearing. Subtraction of condition C from A and of condition D from B leaves articulation.
Vision scanner (2T), utilizing relative T1 weighting to give good gray-white matter contrast.

Image Processing and Data Analysis

Image manipulations and calculations were performed on SPARC computers (Sun Computers, London, UK) with the use of ANALYZE (BRU, Mayo Foundation, Rochester, MN) and PROMATLAB (Mathworks) software. Statistical parametric modeling software (Wellcome Dept. of Cognitive Neurology) was used to create statistical maps of significant relative rCBF changes (7, 8). First, all scans of each individual were realigned to the first scan of the series, thus ensuring spatial congruency. All PET scans and the MR image were transformed to standard stereotactic space (27) to account for individual differences in brain size and shape (7, 8). PET images were filtered in three dimensions by use of a 10-mm FWHM isotropic Gaussian kernel filter for the following reasons: 1) to improve signal-to-noise ratio, 2) to reduce intersubject variability caused by differences in local anatomy, and 3) to condition the data so that they conform more closely to the Gaussian field theory (8) that forms the basis for the model from which the statistical inferences are drawn when correcting for multiple comparisons. Subsequently, a pixel-based analysis of covariance, treating “global” activity (reflecting global blood flow) as the covariate, controlled for state-dependent differences in global blood flow associated with the different conditions (7, 8). A simple general linear model (8) was used to model relative rCBF as a linear combination of condition, subject, and global effects. In this analysis, each scan was treated as an independent observation. Thereafter, contrasts of condition effects were assessed by using t-statistics (at each voxel), and a transformation into normal-distribution Z statistics was performed. The resulting sets of spatially distributed Z values constitute statistical parametric maps showing regions of significant condition-associated relative rCBF changes (7, 8). These regions were displayed with a statistical threshold at either P < 0.001 and Z > 3.10 or P < 0.01 and Z > 2.33 (not corrected for multiple comparisons). Within each region, the local maxima of significant relative rCBF changes were then derived in terms of x, y, and z coordinates (see definitions in Fig. 4 legend). The anatomic locations of local maxima were defined both with reference to the stereotactic atlas of Talairach and Tournoux (27) and also by superimposition on the group mean MR image that had been transformed into standard stereotactic space (as described above).

The orthogonal design (Table 1) of the experimental paradigm outlined above allowed two principal contrasts to be defined, focusing on areas associated with control of breathing for speech and vocalization (A–B and C–D), on the one hand, and articulation (A–C and B–D), on the other hand. In each case, the two contrasts could be combined to increase the power of the analysis. The elements of (A–B) + (C–D) double the contribution of breathing for speech and vocalization (together with hearing) but eliminate the contribution from articulation. Similarly, the elements of (A–C) + (B–D) double the contribution from articulation but eliminate the contributions from breathing for speech, vocalization, and hearing. Areas activated during breathing for speech and vocalization cannot be distinguished from each other in the experimental paradigm used in the present study.

RESULTS

Performance of the Task

All subjects reported that they were able to perform the tasks as instructed. During condition A (speaking

<table>
<thead>
<tr>
<th>Table 2. Exact instructions given to subjects both during training phase and before each scan during experimental runs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condition A: Speech</strong></td>
</tr>
<tr>
<td>I want you to say the phrase “Buy Bobby a poppy” repeatedly in a normal speaking voice and at a normal speaking rate until I tell you to stop.</td>
</tr>
<tr>
<td><strong>Condition B: Moutned speech</strong></td>
</tr>
<tr>
<td>I want you to mouth the phrase “Buy Bobby a poppy” repeatedly without making any sound and at a normal speaking rate until I tell you to stop.</td>
</tr>
<tr>
<td><strong>Condition C: Unarticulated speech</strong></td>
</tr>
<tr>
<td>I want you to say to yourself the phrase “Buy Bobby a poppy” repeatedly and at a normal speaking rate until I tell you to stop. In time with this, I want you to make an “aah” sound; you should do this with your mouth slightly open but without moving your tongue, lips, or jaw. Please ensure that you keep repeating the phrase to yourself in your head in time with the sound.</td>
</tr>
<tr>
<td><strong>Condition D: Internal speech</strong></td>
</tr>
<tr>
<td>I want you to say to yourself the phrase “Buy Bobby a poppy” repeatedly and at a normal speaking rate until I tell you to stop. There should be no sound and no movement. Please ensure that you keep repeating the phrase to yourself in your head.</td>
</tr>
</tbody>
</table>

Scanning Protocol and Data Acquisition

Subjects lay supine, with eyes closed, in a dimmed room with low ambient noise. They were positioned on an adjustable scanner table attached to the PET camera (model 953B, Siemens CTI). The head was positioned to minimize movement during the tasks and was supported in a stable position by use of an individually fitted head mold. This ensured that during execution of the various tasks, the head movement would be <2 mm. Emission data were acquired in three-dimensional mode with interdetector collimating septa removed (28) and corrected for effects of tissue attenuation by use of measurements from a transmission scan (68Ge/68Ga ring sources) carried out before the first emission scan.

For each measurement of relative rCBF, 10 mCi of H215O were given via an indwelling cannula placed in an antecubital vein. Emission data were collected sequentially over 150 s, comprising a 30-s background scan (scan A), a 30-s delay, and a 90-s scan (during which the different speech conditions were performed) after tracer arrival in the brain (scan B). H215O was given as a slow bolus 15-s before scan B (25). At the same time, the subject was told to start the appropriate speech condition. The integrated counts during scan B were corrected for background activity (from scan A). This gave an estimate of rCBF. Because no arterial blood samples were taken (14), no calibration was possible, and the term “relative” rCBF implies that there was no absolute quantification of rCBF. The corrected emission scan data were then reconstructed as 31 axial planes (plane thickness, 3.44 mm) with the use of filtered backprojection and a Hanning filter (cut-off frequency, 0.5 cycles/pixel). The resolution of the resulting raw images (determined by the physical characteristics of PET and the camera used) was 8.5 (x) × 8.5 (y) × 4.3 (z) mm at full-width half-maximum (FWHM) filtration (26). For further processing, image planes were displayed in a 128 × 128 pixel format, with a pixel size of 2 × 2 mm.

Magnetic Resonance Imaging

On a separate occasion, a magnetic resonance (MR) image of the subject’s brain was obtained using a 2-Tesla Siemens Vision scanner (2T), utilizing relative T1 weighting to give good gray-white matter contrast.
aloud), normal conversational loudness and a normal speaking rate were easily maintained. During condition B (mouthing silently), the throat microphone confirmed the absence of sound production, and the video camera, together with the EMG recordings, confirmed that the subjects were mouthing appropriately. During condition C (vocalized without articulation, ah-ahah-ah-ahah), the cadence was similar to condition A, and the video confirmed that the mouth was not moving. During condition D (internal speech), there was no evidence of sound production, facial movement, or EMG activity; the subjects reported that they were saying the phrase to themselves as requested.

Respiratory Data

An example of the breathing patterns from one individual throughout one scan for each of the four conditions is shown in Fig. 2, and the average values for the individual respiratory variables during these scans and also before scanning in the absence of any speech (Rest) are shown for each condition in Fig. 3. Speech, whether articulated or unarticulated, showed a controlled, near constant, expiratory airflow lasting 8–10 s, interrupted by rapid inspirations of ~0.75-s duration. The similarity between the breathing variables in these two conditions (A and C) was confirmed for the group as a whole (P > 0.05 by ANOVA; Fig. 3) for all the variables shown. Breathing variables in the two conditions without vocalization (B and D) were similar (P > 0.05 by ANOVA; Fig. 3) to each other and to the breathing in the "Rest" condition. However, the breathing traces showed a variable degree of "control" of expiratory airflow, as judged by periods of constancy of slope of the expired volume trace (together with minor slope disturbances), in condition B in four of the six subjects. This is not represented in Fig. 2, nor were these disturbances enough to affect the mean respiratory variables (Fig. 3).

MR Imaging

MR structural images of the brain were obtained in five of the six subjects.

Relative rCBF Measurements

rCBF associated with control of articulation. The mean results of the main contrast (A−C) + (B−D) are presented in Fig. 4. This showed a significant bilateral activation in the sensorimotor cortex (smc1), together with activation in the cerebellum bilaterally (Z = 3.5) and on the right side in the thalamus/caudate nucleus (Z = 3.6); there was no activation in the supplementary motor area (SMA). The bar graph shows the relative rCBF (arbitrarily adjusted to a global value of 50 ml·min⁻¹·dl⁻¹) at the pixel of maximum significance for each condition for the right-sided smc1 activation alone. There were significant differences in the adjusted rCBF values between conditions A or B with C or D, with D being significantly lower than C. Thus adjusted rCBF was greater in these conditions in which articulation was present.

The pattern of activation of smc1 for one representative subject, co-registered on his own structural MR image, is shown in Fig. 5A. The coronal, sagittal, and transverse planes are displayed at the group mean coordinates of the right-sided focus of activation applied onto that subject's MR images. This clearly shows the cortical sites of these activations.

In view of the evidence that the breathing in condition B may have been "contaminated" by some "breathing for speech" (see above), we examined the component (A−C) of the main contrast (Table 1; Fig. 6A). The foci of activation, smc1, were still present and had the same coordinates. The only difference was that the areas of
activation were more circumscribed, with a highly significant Z score in spite of the reduced statistical power of the contrast.

rCBF associated with control of breathing for speech and vocalization and hearing. The mean results of the main contrast, \((A - B) + (C - D)\), are presented in Fig. 7, which highlights bilateral areas within the sensorimotor cortex (smc2) together with a unilateral activation in a strip on the surface of the precentral gyrus (mc), most probably in the motor cortex (Brodmann area 4/6).

Fig. 4. Group results (averaged across all runs for all subjects) of relative regional cerebral blood flow (rCBF) increases associated with articulation obtained from orthogonal combination of both contrasts \((A - C)\) and \((B - D)\) are shown as through projections onto representations of stereotactic space (27). Areas reaching statistical significance \((P < 0.001; \ Z > 3.10)\) are imaged. Sagittal image views brain from side; coronal image is from back; transverse image is from top. R, right; VPC, vertical plane through posterior commissure; VAC, vertical plane through the anterior commissure; nos. are coordinates in stereotactic space. x, Distance (in mm) to right (+) or left (-) of midsagittal (interhemispheric) line; y, distance anterior (+) or posterior (-) to vertical plane through anterior commissure (VAC); z, distance above (+) or below (-) intercommissural line through anterior and posterior commissures. Areas of activation of increasing significance are shown as arbitrary gray scale ranging from light gray to black. smc1, Sensorimotor cortex; t/c, thalamus/caudate nucleus; c, cerebellum. Coordinates of peak activation in smc1 (left and right) are tabulated beneath coronal image, together with Z statistic (far right). Lower right, adjusted rCBF for right side smc1 for all subjects by condition. Bars show increased activity in conditions A and B where articulation is present. Solid marks represent Fisher’s least significant difference between any mean value at \(P < 0.05\). Other significant activations are seen in coronal and transverse images within c and t/c areas on right.

Fig. 3. Indicated respiratory variables were averaged across period of scan for each of the speech-related conditions and then averaged across subjects. Mean values are shown (error bars are SE of mean) for each of 4 conditions. Resting period of 1 min (rest) was recorded when subject had been prepared and was in scanner but before any experimental runs had taken place. Height of bars represents Fisher’s least significant difference between any mean value at \(P < 0.05\). These are only calculated for those parameters that have a significant effect across condition. Overall level of ventilation \((V\dot{I})\) is not different across conditions. Expected breathing pattern (see Fig. 1) can be seen with a shorter inspiratory duration \((TI)\), larger expiratory duration \((TE)\), and larger total volume \((VT)\) in those conditions that have “breathing for speech and vocalization” in them. Those conditions predicted to have little modulation of breathing (Rest, Mouthing, and Internal speech) are similar.
junction). Reduction of the level of significance to \( P < 0.01 (Z > 2.33) \) demonstrates a similar area of activation on the other side (mc, shown for a single subject in Fig. 5B). The coordinates of mc shown in Fig. 7 are the site of the focus of activation, but the area of activation extends above and below this level by at least 4 mm in both directions. The coordinates of the region with no activation (“hole”) lying between mc and smc2 are similar to that of smc1 in the other main contrast (described above; Fig. 4). Activations were also present in the SMA in the midline \( (Z = 3.8) \), in the thalamus bilaterally \( (Z = 4.1) \), and in the cerebellum in the midline \( (Z = 3.94) \). The bilateral activations seen in the superior part of the temporal lobes, presumably associated with “hearing” oneself speaking, were expected because “hearing” could not be removed by this contrast (Table 1; METHODS).

The bar graphs in Fig. 7 show the relative rCBF (arbitrarily adjusted to a global value of 50 ml·min\(^{-1}\)·dl\(^{-1}\)) at the pixel of maximum significance, for each condition for the right-sided smc2 area and for the right-sided mc area. For the smc2 activation, there were significant differences in the adjusted rCBF values between conditions A and B; between A, B, or C with D; but not between B and C or A and C. Thus adjusted rCBF in smc2 tended to be greater in this area when there was breathing for speech together with vocalization and to be lowest in the conditions with least modulation of breathing. For the mc activation,
the bar graph shows that there were no significant
differences between the adjusted rCBF values in condi-
tions A, B, and C but they were all significantly greater
than those in condition D.

The pattern of activation for one representative
subject, co-registered on his own structural MR image,
is shown in Fig. 5B. The coronal, sagittal, and trans-
verse planes are at the group mean coordinates of the
right-sided focus of activation (smc2) applied to the
individual subject’s images and show the cortical sites
of the main activations, including that of mc (bilater-
ally in this individual).

The evidence of some degree of control of expiratory
airflow in condition B compared with condition D in
some of the subjects led to a suspicion that the breath-
ing pattern seen might have been contaminated by
some breathing for speech as in conditions A and C. If
this were true, then any contrast involving scans taken
during condition B might be similarly contaminated.
We therefore examined the contrast (C−D) alone (Table
1; Fig. 6B). In spite of the reduced statistical power of
this contrast, a similar pattern of activation was found
as in the main contrast (A−B) + (C−D), except that
areas smc2, and especially mc, were not so clearly
defined.

Other Areas Activated

Neither of the above contrasts highlighted any unilat-
eral activations classically associated with speaking
a language. In particular, Broca’s area centered on the
triangular part of the left inferior frontal cortex (approx-
imate Talairach coordinates: x, −40 to −60; y, +10 to
+12; and z, +8 to +12) showed no foci of activation.

DISCUSSION

We have demonstrated bilateral areas in the sensori-
motor cortex that are activated in the control of breath-
ing for speech and vocalization. In addition, we have
identified separate bilateral areas in the sensorimotor
cortex that are activated in the control of articulation.

Experimental Paradigm

The phrase. The cadence and repetitive nature of the
task resembled that utilized by Draper and co-workers
(3) in counting, in which evidence of sophisticated
neural control of the respiratory musculature had already
been found. A major advantage of the phrase (12) was
that it consisted predominantly of labial consonants.
Thus all the mouthing activity was visible to the investigator and could be monitored. The importance of this was that the control conditions C and D, in which an absence of mouth movement was required, could be shown to have been carried out correctly. In addition, inspection of mouth movements could be made qualitatively between conditions A and B to assess their similarity.

Language content. The task used a phrase which was overlearned, repetitive, and required no particular attention. There was therefore minimal language content, with little probability of finding activation in the classic Broca’s area. Those aspects of speaking and language that are associated with activation in the left dorsolateral prefrontal cortex have been recently studied by Warburton et al. (29) using PET; noun and verb generation were the optimal stimulus to such activation. Herholz et al. (10), using MR imaging-guided PET, showed that the most common anatomic areas of activation with verb generation were in the left inferior frontal cortex centered around the classic Broca’s site in the pars triangularis. The subtraction paradigm used in the present study (Table 1) was designed to remove the pars triangularis. The subtraction paradigm used in the present study (Table 1) was as expected (Fig. 1A) was as expected (Fig. 1A) was realized (Fig. 2A). This was also true for unarticulated speech (Fig. 1C; compare Fig. 2C). The expected breathing pattern during the mouthing of speech, i.e., no obvious change from a normal resting level (Fig. 1B), was not completely realized because of the evidence that a degree of respiratory modulation that resembled control of breathing for speech was present in some subjects. The breathing pattern during internal speech (Fig. 2D) was as expected (Fig. 1D). The normal mean resting level of ventilation (Fig. 3) maintained throughout the various tasks testifies that the subjects, when vocalizing, were doing so at a normal conversational level and when not vocalizing were not hyperventilating. This suggests that they were relaxed and had had sufficient training.

Statistical Inference

Because we had a priori hypotheses for the regions of interest that we planned to test with our experimental paradigm, i.e., sensorimotor cortex, SMA, cerebellum, and Broca’s area, we made no correction in our statistical analysis for multiple comparisons (8). Of the other brain regions for which we had no strong a priori hypotheses, only the activation in the superior temporal gyrus survived correction for multiple comparisons (P < 0.001).

Validity of Using Elements of Speech as Independent Variables

The validity of manipulating the elements of a task that has both cognitive and sensorimotor components as if they were simple arithmetical symbols entirely depends on the absence of an interaction between these elements. This is clearly highlighted in the present studies (Table 1) in which, with the objective of increasing the power of the analysis, the results from simple contrasts between two conditions have been combined, utilizing an orthogonal design. This makes the subtraction complex. The problems that this may cause have been extensively reviewed by Friston (6). Formal investigation by Sidtis et al. (24) of these problems during speaking with minimum language content concluded that often, but not always, statistical maps produced by compound subtractions (as in the present study) diverge more from the results of nonimaging studies of the functional anatomy of speech than do those produced by simple subtractions. Examination of a linguistic task combined with finger movements (13) demonstrated that cognitive and behavioral processes do not combine in a simple additive manner. Nevertheless, in the present study, if images produced by complex subtractions are compared with those from simple subtractions (Fig. 4 compared with Fig. 6A; Fig. 7 compared with Fig. 6B), they are found to be surprisingly alike. Indeed, the focus of activation for control of breathing for speech and vocalization with articulation present (A–B) were similar to the foci when articulation was present (A–C). There is further support for the use of complex subtractions in the fact that, in the simple subtraction (A–C; Table 1), the superior temporal lobe activation associated with hearing disappears (even though the nature of the sound being produced is very different) as would be predicted. Addition of the second simple subtraction (B–D), which does not contain hearing, does not interfere with the disappearance of hearing in the complex contrast (Fig. 4).

General Motor Control

The cerebellum was activated in both main contrasts defining areas concerned with articulation (Fig. 4) and breathing for speech with vocalization (Fig. 7). By contrast, the SMA was activated only in the contrast defining control of breathing for speech and vocalization (Fig. 7). These activations are entirely typical of those found with any volitional action (18), presumably reflecting planning and control mechanisms. The apparent lack of SMA activation in the contrast defining areas concerned with the control of articulation is likely to have resulted from an effective subtraction of SMA activations in the components of the contrast (Table 1). It is of interest that these general motor control areas are activated in speaking, even when speech is overlearned, repetitive, and requires no attention.

Articulation

The main contrast designed to show areas in the cerebrum activated with articulation uncontaminated
by other activities has shown clearly defined areas in the sensorimotor cortex (smc1) with very high statistical probability (Fig. 4). The areas are symmetrical and large. They are entirely compatible with the map of Penfield and Rasmussen (see Fig. 22 in Ref. 20), with the homunculus showing the lip representation on the motor homunculus dorsal to the jaw representation and ventral to the face representation but occupying an area larger than both. The tongue representation on the map is more ventral and close to the Sylvian fissure. In the present study, the tongue should not have been tensed or moved, but there is no objective evidence of this. Some tongue movement may be associated with the ventral extensions of the activations, especially seen on the left side in the main contrast (Fig. 4) but not seen in the simple subtraction (Fig. 6A) in which condition C is contrasted with condition A alone.

We found no definite evidence for the articulatory premotor area, concerned with the planning of articulatory movements and recently described from studies of infants in patients with apraxia of speech, to be in the precentral gyrus of the insula on the left side (4). This small area with left-sided coordinates (using the same Talairach system as in the present study) of x = 41, y = -2, z = +10 may possibly have been activated during articulation in the present study (see Fig. 4). However, the nearest area of significant (Z = 2.65) activation in the present study is at coordinates x = -46, y = -6, z = +16. There also appears to be an area significantly (Z = 2.5) activated at similar coordinates on the opposite side. These are close to the coordinates given by Dronkers (4).

Breathing for Speech and Vocalization

Penfield and Bouldrey (19) made the seminal observation that vocalization, usually without lip or tongue movements, may result from electrical stimulation over a defined area of the surface of the exposed lateral motor cortex of man during neurosurgery over a range of sites in either hemisphere. The continuous involuntary sounds produced were terminated at maximum expiratory volume and resumed after an inspiration. The areas we have described as being associated with the control of breathing for speech and vocalization, namely smc2 and mc, are in similar regions of the cortex. The significance of the activations seen in the mc area was weaker than in the smc2 area. The level of significance had to be reduced to P < 0.01 to show that these areas were bilaterally activated. The likelihood that mc is indeed an area concerned with the control of breathing for speech and vocalization is increased by its preservation as an activated area when data from condition B (mouthing speech) was removed in the simple subtraction shown in Fig. 6B. The bar graph for mc in Fig. 7 showing that the mean adjusted rCBF in condition B is at the same level as in conditions A and C (although significantly higher than for condition D) also suggests that breathing for speech (but not vocalization) is contaminating condition B. The higher levels of significance of the activation in area smc2, both in the complex and simple subtractions, give greater confidence that these areas are concerned with breathing for speech and vocalization. The bar graph in Fig. 7 for smc2 also shows a reduction in mean adjusted rCBF for condition B that is below that for conditions A and C but not to the level of condition D. We argue that the effect of contamination of condition B with breathing for speech is of less importance when activations are highly significant (as in smc2) than when such activations are less significant (as in mc).

Ramsay and colleagues (21) reported on activations in the motor cortex bilaterally that occur with a controlled willful expiration without vocalization (see Fig. 3C in Ref. 21). The y and z coordinates of the pixels of maximum significance are similar to those found for the mc area in the present study, but the x coordinate is 8 mm more medial. When the coordinates in the study of Ramsay et al. are compared with those of smc2 in the present study, they are 14–15 mm more anterior, but this is still within the overall areas of activation of both studies. We cannot therefore be certain that areas of the sensorimotor cortex activated by willful expiration are similar to those areas activated by behavioral expiration (as in speech), but there is certainly an overlap. This question cannot be resolved with the present data because of the difficulty of geometrically defining the limits of activations which have length and depth.

Conclusion

These studies have used the simplest possible paradigm to focus on speech of a very limited and stereotyped form so that we can separate the elements of its complex motor control. The challenge now is to extend these studies to more natural speech so that the mechanisms underlying the interaction between language and speaking can be revealed.

We thank Professor R. S. J. Frackowiak, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, and Professor D. Brooks, Medical Research Council Cyclotron Unit, Hammersmith Hospital, London, UK, for helpful preliminary discussions. We also thank Dr. K. Friston, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, for statistical advice during the planning of these experiments. We are grateful to Elizabeth Whitcombe (Spinal Injuries Unit, Stoke Mandeville Hospital) for persistence in urging that this type of study be done.

This work was supported by a Wellcome Trust Retired Investigators Grant to A. Guz. The Wellcome Trust also supported D. R. Corfield and G. R. Fink (Wellcome Clinical Research Fellow).

Address for reprint requests: K. Murphy, Dept. of Respiratory Medicine, Imperial College of Medicine, Charing Cross Campus, St. Dunstan's Rd., London W6 8RP, UK (E-mail: K.Murphy@cxwms.ac.uk).

Received 5 March 1997; accepted in final form 8 July 1997.

REFERENCES


