Separate neural subsystems within 'Wernicke's area'

Richard J. S. Wise,[1] Sophie K. Scott,[2] S. Catrin Blank,[1] Cath J. Mummery,[3] Kevin Murphy[4] and Elizabeth A. Warburton[5]

[1]*MRC Clinical Sciences Centre, Cyclotron Unit, Hammersmith Hospital,* [2]*Institute of Cognitive Neuroscience, UCL,* [3]*National Hospital for Neurology and Neurosurgery and* [4]*Department of Respiratory Medicine, Charing Cross Hospital, London and* [5]*Department of Medicine of the Elderly, Addenbrooke's Hospital, Cambridge, UK*

Correspondence to: Richard Wise, MRC Clinical Sciences Centre, Cyclotron Unit, Hammersmith Hospital, Du Cane Road, London W12 0NN, UK
E-mail: richard.wise@csc.mrc.ac.uk

Summary

Over time, both the functional and anatomical boundaries of 'Wernicke's area' have become so broad as to be meaningless. We have re-analysed four functional neuroimaging (PET) studies, three previously published and one unpublished, to identify anatomically separable, functional subsystems in the left superior temporal cortex posterior to primary auditory cortex. From the results we identified a posterior stream of auditory processing. One part, directed along the supratemporal cortical plane, responded to both non-speech and speech sounds, including the sound of the speaker's own voice. Activity in its most posterior and medial part, at the junction with the inferior parietal lobe, was linked to speech production rather than perception. The second, more lateral and ventral part lay in the posterior left superior temporal sulcus, a region that responded to an external source of speech. In addition, this region was activated by the recall of lists of words during verbal fluency tasks. The results are compatible with an hypothesis that the posterior superior temporal cortex is specialized for processes involved in the mimicry of sounds, including repetition, the specific role of the posterior left superior temporal sulcus being to transiently represent phonetic sequences, whether heard or internally generated and rehearsed. These processes are central to the acquisition of long-term lexical memories of novel words.

Introduction

In the absence of clear definitions about either its functions or its anatomical boundaries (Williams, 1995), 'Wernicke's area' has become a meaningless concept (Bogen and Bogen, 1976). In the model of single word processing by Lichtheim, >100 years old but still the basis of the bedside assessment of aphasic patients, Wernicke's area, localized to the posterior part of the left superior temporal gyrus (STG), stores the encoded memories of familiar heard words, from which there is access to both meaning and speech production (Lichtheim, 1885). In recent years, and depending on the publication, Wernicke's area may comprise: unimodal auditory association cortex located in the left STG anterior to primary auditory cortex in Heschl's gyrus (HG), and responsible for the phonetic analysis of speech (Demonet *et al.*, 1992), or

heteromodal cortex, comprising three architectonic zones in the left temporal and parietal lobes, where the output from both heard and written word form (lexical) systems converge (Mesulam, 1998). Other studies have made much of the greater size of the left planum temporale (PT), lying between HG and the ascending ramus of the posterior sylvian (lateral) sulcus in the supratemporal cortical plane, compared with the right (for review, see Shapleske *et al.*, 1999) (Fig. 1). Although the anatomical asymmetry has been attributed to the dominance of the left hemisphere for language (Geschwind and Galaburda, 1985) and the entire posterior left supratemporal cortical plane is considered by some to be the core of Wernicke's area (Galaburda *et al.*, 1978), neither the speech-specific function of the left PT is
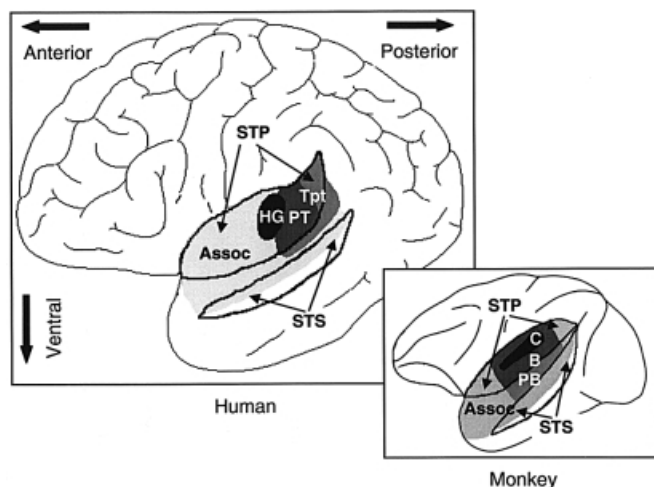
Fig. 1 Depictions of the left superior temporal cortex in the human and the macaque monkey, with the plane of the supratemporal cortex (STP) and inside of the superior temporal sulcus (STS) exposed. Human brain: HG = Heschl's gyrus (including primary auditory cortex); Tpt = supratemporal cortex posterior to HG; PT = planum temporale, part of the supratemporal cortical plane immediately posterior to HG (Shapleske *et al.*, 1999); Assoc = auditory association cortex lateral and anterior to the previous three regions. Monkey brain: C = core (primary auditory cortex); B = belt; PB = parabelt; Assoc = auditory association cortex surrounding the previous three regions.

established (Binder *et al.*, 1996), nor is the claim for anatomical asymmetry universally accepted (Westbury *et al.*, 1999).

In contrast, functional neuroimaging studies of speech perception have drawn attention to the role of lateral auditory projections in speech processing (Binder *et al.*, 1996, 2000; Belin *et al.*, 2000). The authors of these studies concluded that analysis of the complex acoustic features of the human voice is dependent on neurons within the superior temporal sulcus (STS), which separates the STG and middle temporal gyrus (Fig. 1). In addition, they referred to microelectrode studies in the auditory cortex of non-human primates. Core auditory cortex in monkeys is organized cochleotopically, with individual neurons responding maximally to a pure tone of a particular frequency (Kosaki *et al.*, 1997). It is only in non-primary auditory areas, particularly the so-called parabelt region, lateral to primary auditory cortex (Fig. 1) that individual neurons have been shown to respond maximally to complex sounds (Kosaki *et al.*, 1997), including species-specific vocalizations (Rauschecker *et al.*, 1995). The demonstration that voice perception is dependent on auditory projections to the dorsal bank of the human STS fits well with these observations.

However, it is becoming apparent that the anterior–posterior axis of the temporal lobe is an equally important anatomical dimension in auditory function (Rauschecker, 1998; Romanski *et al.*, 1999; Kaas and Hackett, 1999). There appear to be two streams of auditory processing in primates, one directed anteriorly and the other posteriorly. In a human imaging study that looked at the responses to speech and complex non-speech sounds, heard at varying rates, we demonstrated a speech-specific response in left and right lateral STG, anterior to HG; however, in addition there was a similar but asymmetrical response in the posterior left lateral STG/STS (Mummery *et al.*, 1999). In a further study (Scott *et al.*, 2000), which closely matched stimuli for acoustic complexity, it was demonstrated that the anterior left STS responded only to intelligible stimuli, whereas the posterior left STS responded to the presence of auditory phonetic cues, irrespective of the intelligibility of the stimuli. Therefore, this study demonstrated a clear difference in the responses of the anterior and posterior parts of the left STS.

The anterior and posterior parts of the superior temporal cortex have very different anatomical connections. Whereas the anterior STS in non-human primates projects widely to high order, amodal association cortex (Jones and Powell, 1970), the posterior superior temporal cortex has reciprocal connections with dorsolateral frontal cortex via the superior longitudinal fasciculus (arcuate fasciculus) (Gloor, 1997). Common functional consequences of lesions around the posterior part of the Sylvian sulcus in humans are disordered repetition and speech production (Benson, 1979). We have re-analysed three of our group's previously published PET studies (Warburton *et al.*, 1996; Murphy *et al.*, 1997; Mummery *et al.*, 1999) and one unpublished study to investigate, first, whether there is a local neural system within the posterior superior temporal cortex that responds to both hearing speech and the recall of words during verbal fluency tasks. A functional conjunction of activations during both the perception and the mental rehearsal of words identifies a system central to language acquisition, whereby the transient representation of sequences of phonemes and their rehearsal, covert or overt, ultimately results in long-term lexical memories. Secondly, we wished to investigate whether there is also a posterior left temporal system that responds to the motor act of speech, identified as a region where the task-dependent activations are related to speech production, independent of the speaker's perception of his own utterances. Such a system must exist to bind speech perception with production during the rehearsal of novel words to acquire lexical memories

Methods
Subjects
Twenty-six right-handed, healthy male volunteers took part in four experiments. Each subject gave informed, written consent. All spoke English as their first language. The studies were approved by the Administration of Radioactive Substances Advisory Committee (Department of Health, UK) and the research ethics committees at the Hammersmith Hospital and the National Hospital for Neurology and Neurosurgery.
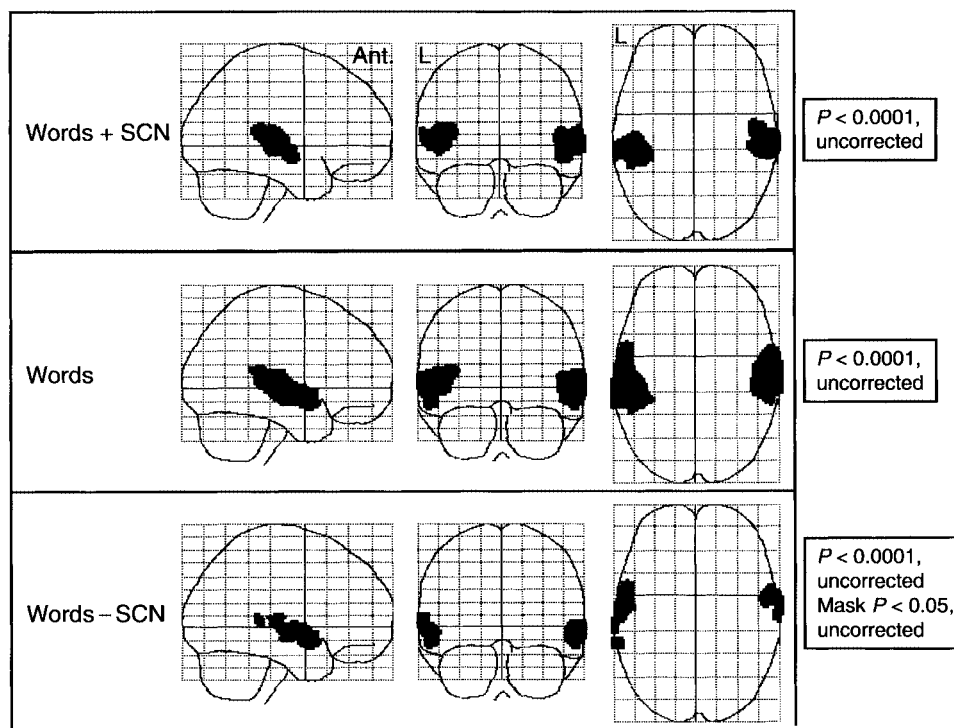
Fig. 2 Experiment 1: statistical parametric maps displayed as sagittal, coronal and axial projections. All voxels significant at $P < 0.0001$, uncorrected, are displayed as black overlays for the three analyses: the conjunction of linear increases in activity with increasing rates of hearing both words and signal correlated noise (Words + SCN); linear increases in activity with increasing rates of hearing words (Words); and linear increases in activity with increasing rates of hearing words once those voxels that also responded to SCN had been masked at a threshold of $P < 0.05$, uncorrected (Words − SCN). Ant. = anterior; L = left.

PET scanning

Brain activation was measured using PET. The dependent variable in functional imaging studies is the haemodynamic response: a local increase in synaptic activity is associated with increased local metabolism, coupled to an increase in regional cerebral blood flow (rCBF). Water labelled with a positron-emitting isotope of oxygen ($H_2^{15}O$) was used as the tracer to demonstrate changes in rCBF, equivalent to changes in tissue concentration of $H_2^{15}O$. The resolution of the technique meant that the activity at the level of neural systems (i.e. local populations of many millions of synapses) was observed. Analysis involved relating changes in local tissue activity (normalized for global changes in activity between scans) to the behavioural task. Each subject had seven to 12 estimations of rCBF, made with a Siemens/CPS ECAT Exact HR+ (962) (Experiment 1) or a Siemens CTI 985B (Experiments 2–4) PET camera, at 8–10 min intervals. The order of stimuli was randomized within and across subjects in each experiment. For each scan, 296–444 MBq of $H_2^{15}O$ (depending on the scanner sensitivity) was administered as a slow intravenous bolus, and the total counts per voxel during the build-up phase of radioactivity served as an estimate of rCBF. Data acquisition was performed in 3D mode, with the lead septa between detector rings removed, with one 90 s acquisition frame beginning at the start of the rise of the head curve. Stimuli were presented to the subjects, or the subjects performed specific tasks, for 75 s, starting 15 s before the arrival of $H_2^{15}O$ in the brain, and covering the critical measurement period of rapid build-up of $H_2^{15}O$ in the brain over 30 s. After measured attenuation correction, images were reconstructed by filtered back projection (Hanning filter, cut-off frequency 0.5 Hz).

Analyses

The data were analysed using statistical parametric mapping, version SPM99 (Wellcome Department of Cognitive Neurology). Each individual's data were realigned to remove head movements between scans, normalized into a standard stereotactic space, and smoothed using an isotropic 10 mm, full width, half-maximum Gaussian kernel to account for individual variation in gyral anatomy and to improve the signal-to-noise ratio (Friston *et al.*, 1995*a*). Individual studies were rejected if there were incomplete axial slices between 40 mm below and 50 mm above the plane of the anterior and posterior commissures of the normalized images, to ensure that there had been inclusion of all the temporal and inferior parietal lobes, with the exception of the ventral surface of the temporal poles. In practice, incomplete volumes were only encountered in two out of nine subjects in one
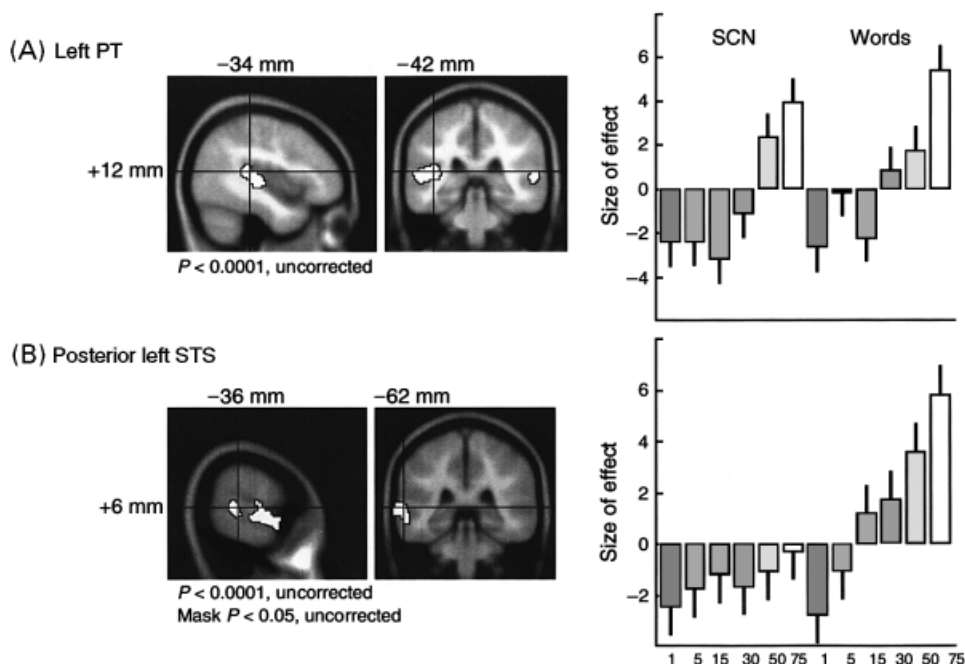
Fig. 3 Experiment 1: the results for the left PT (**A**) and posterior left STS (**B**). The peak voxels (cross hairs) are shown on sagittal and coronal slices of the MRI T_1-weighted template (the averaged image from 125 scans of normal subjects) available in the SPM99 software. All voxels significant at $P < 0.0001$, uncorrected, are displayed as white overlays on the images. The coordinates for the peaks are given for MNI space, the stereotactic space employed by SPM99. On the right of the figure, for both peak voxels, each condition (*x*-axis), coded on a grey-scale from low to high rates of presentation of the stimuli, is plotted against the size of its effect (*y*-axis) in the weighted contrast (i.e. –6, –5, –3, 1, 4, 9) across conditions.

study (Experiment 2). Specific effects were investigated using appropriate contrasts to create statistical parametric maps of the *t*-statistic (Friston *et al.*, 1995*b*). We used an analysis of covariance with global counts as confound to remove the effect of global changes in perfusion across each individual's scans (Friston *et al.*, 1990). The thresholds for significance are described under the presentation of the results of the individual studies. SPM99 displays a list of the peaks (>4 mm apart) within an activated region. We identified and reported in detail only those peaks located within superior temporal cortex. Peaks located within HG and the PT were identified by using published probability maps (Penhune *et al.*, 1996; Westbury *et al.*, 1999), following a correction for the differences in the coordinate systems between the Talairach and Tournoux atlas (1988) (used in the probability maps) and the stereotactic space employed by SPM99, created at the Montreal Neurological Institute (Evans *et al.*, 1993) (http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html). In practice, the corrections for coordinates in superior temporal cortex were never >3 mm in any one axis. The location of the other peaks were made with reference to the Talairach and Tournoux atlas (1988). In the figures, the PET activations are displayed on the average template of 125 T_1-weighted MRI normal scans available in SPM99, using the Montreal Neurological Institute's coordinate system.

Individual experimental designs, analyses, results and comments
Experiment 1
Design
Six subjects heard either bisyllabic nouns or signal correlated noise (SCN) (Mummery *et al.*, 1999). SCN was prepared by taking the time–amplitude envelopes of a selection of the bisyllabic nouns and multiplying these envelopes with white noise. The resulting sounds contained no phonetic cues, but retained the rhythm and syllabic segmentation of words (Rosen, 1992). The rates of the stimuli were varied across scans (1, 5, 15, 30, 50 and 75 per min), so that each subject heard each type of stimulus six times, once each for the six different rates.

Analyses
Each scan was entered as a separate condition. Appropriate contrasts, centred around zero (i.e. –6, –5, –3, 1, 4, 9), were used to show voxels where activity increased approximately linearly with the rates of hearing SCN alone and words alone. The threshold was set at $P < 0.05$, corrected for analysis across the whole brain volume.

Table 1 *The peak activations in posterior temporal cortex observed in Experiments 1–4: their coordinates in Talairach and Tournoux stereotactic space (x, y and z, relative to the anterior commissures), their Z-scores and their significance, corrected for analyses across the whole brain volume, are shown*

	Left hemisphere					Right hemisphere				
	x	*y*	*z*	*Z*	*P*	*x*	*y*	*z*	*Z*	*P*
Experiment 1										
Linear response to increasing rates of hearing both SCN and words										
Anterior STS						+51	–08	–06	7.3	<0.001
HG	–46	–23	+04	7.8	<0.001 (50–75%)					
Lateral STG	–53	–27	+05	7.8	<0.001	+65	–21	+07	>8.0	<0.001
PT	–42	–32	+13	7.0	<0.001 (26–45%)					
	–49	–36	+13	6.7	<0.001 (46–65%)	+49	–25	+09	>8.0	<0.001 (46–65%)
Linear response to increasing rates of hearing words without response to SCN										
Anterior STS	–57	+04	–08	5.0	0.02	+59	–02	–03	6.5	<0.001
Mid-STS	–59	–16	–01	5.3	0.005					
Posterior STS	–61	–35	+06	5.0	0.02					
Experiment 2										
Noun and verb generation contrasted with rest state										
Posterior STS	–63	–37	+06	6.6	<0.001					
PT	–57	–42	+22	5.3	0.004 (26–45%)					
Posterior MTG	–57	–36	–07	4.5	>0.1					
Experiment 3										
Perception of own utterances										
Anterior lateral STG						+63	+05	–07	7.2	<0.001
Mid-STS	–51	–15	+03	7.4	<0.001	+61	–17	+03	7.7	<0.001
PT	–44	–34	+11	7.7	<0.001 (26–45%)	+43	–29	+11	7.5	<0.001 (26–45%)
Response to voicing, speaking and mouthing, each contrasted with silent rehearsal										
Medial temporo-parietal junction	–42	–40	+20	5.7	0.001					
Experiment 4										
Correlation of activity with the rate of hearing stimuli + the rate of retrieving words										
Posterior STS	–63	–34	+02	3.9	>0.1					
					0.002 for spatial extent significance					

For HG and PT, the probability that the peak voxel lay within the designated cortical region is also shown: the maximum probability for any one voxel from the published maps is 100% for HG (Penhune *et al.*, 1996) and 65% for PT (Westbury *et al.*, 1999). The location of the other peaks, in the superior and middle temporal gyri (STG and MTG), the superior temporal sulcus (STS) and the temporoparietal junction, were made with reference to the Talairach and Tournoux atlas (1988).

Results (Table 1 and Figs 2 and 3)

Peaks of activity common to SCN and words were present in left HG, the left and right lateral STG and the left and right PT. A response to words alone was observed in the left and right STS, anterior to the coronal plane of HG, but also in the posterior left STS.

Comment

Speech-specific responses in this contrast with SCN were confined to the lateral STG and STS. Anterior to HG there was no apparent asymmetry. As SCN lacks both phonetic information and the periodicity (voicing) that gives speech its pitch and intonation, it cannot be inferred from the symmetry of the left and right anterior responses that the two hemispheres responded to the same acoustic features in the speech signal (Belin *et al.*, 2000; Scott *et al.*, 2000). It was evident that the response in the posterior left STS was speech-specific, whereas in the posterior right STS it was not.

Experiment 2

Design

Seven subjects were scanned during the following three conditions, with four scans per condition (Warburton *et al.*, 1996).

A. Rest, when the subjects were told to 'empty your mind'.
B. Verb generation, when the subjects had to think of as many verbs as they could in the time available (15 s), without vocalization, in response to basic level, concrete nouns (e.g. shirt: wash, iron, mend, etc.).
C. As B, but the subjects had to think of basic level nouns in response to hearing a superordinate noun (e.g. fish: cod, salmon, perch, etc.).

Sagittal   Coronal   Axial

−64 mm   −58 mm

+6 mm   +22 mm
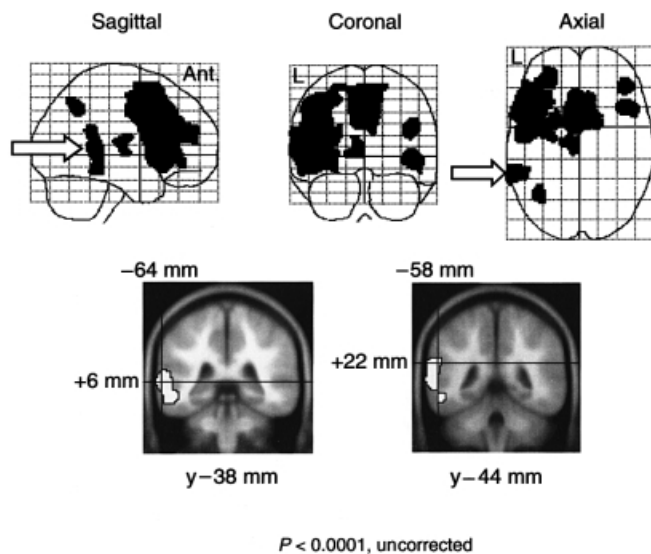
y−38 mm   y−44 mm

P < 0.0001, uncorrected

Fig. 4 Experiment 2: statistical parametric maps displayed as sagittal, coronal and axial projections in the upper half of the figure. All voxels significant at *P* < 0.0001, uncorrected, are displayed as black overlays for the one analysis. Extensive activations, described in the text, include a peak in the caudal left STS (white arrow). In the lower half of the figure, the two significant peaks in the caudal left temporal cortex are displayed on averaged MRI templates, using the same method described in Fig. 2, with the posterior left STS in the left coronal image and the left PT in the right coronal image.

Analysis
One contrast, (B + C) – A, was analysed. The threshold was set at *P* < 0.05, corrected.

Results (Table 1, Fig. 4)
There were extensive, predominantly left-lateralized activations in premotor and dorsolateral prefrontal cortex, with additional activations in left and right frontal opercular cortex and right dorsolateral prefrontal cortex. There were also activations in the left temporal lobe, comprising a main peak in the posterior left STS, within 5 mm of the coordinates of the peak in the posterior STS observed in Experiment 1. There were additional, smaller peaks in the lateral aspect of the PT and in the middle temporal gyrus (the latter being below the threshold set for significance).

Comment
Although cued word retrieval is a complex task, involving many psychological processes and widely distributed neural systems, posterior left temporal lobe subsystems were identified that included a peak in the posterior STS. Therefore, in the posterior left STS there was a conjunction of activity for perceiving words, observed in Experiment 1, and for retrieving words from long-term lexical semantic memory.

Experiment 3
Design
Six subjects were taught the phrase 'buy Bobby a poppy' (Murphy *et al.*, 1997). The place of articulation for the consonants (i.e. the location of the supralaryngeal restriction to air flow) was at the lips. There were four conditions, as follows.
A. Repeatedly saying the phrase out loud.
B. Mouthing the phrase, with lip movements but no voicing or adduction of the false vocal cords (as occurs during whispering).
C. Using a single, voiced vowel sound ('uh') to repeatedly sound out the phrase without movement of the articulators.
D. Thinking of the phrase repeatedly.

Analyses
Conditions A and C were associated with breathing patterns typically observed during normal speech (Murphy *et al.*, 1997). A contrast of (A – B) + (C – D) was used in the original publication to investigate the cortical control of breathing during speech (with, additionally, motor control of vocal cord adduction). This contrast also included the auditory cortical responses to the subjects' own utterances, which were not discussed in the original study but are now presented below. We also performed a new analysis, investigating the conjunction of activity in the contrasts of A with D, B with D and C with D. This identified only those voxels activated by all three conditions, A, B and C, relative to condition D. The three contrasts were entered in the order, C–D, A–D and B–D. This specifies the order of orthogonalization. Orthogonalization ensures that any effect modelled by one contrast cannot be explained by another, enabling a test for the conjunction of independent effects. Because we used a common baseline (condition D) the original contrasts were not orthogonal, but were rendered so after appropriate rotation in SPM99. The voxels revealed by this conjunction analysis were associated with speech production, independent of the perception of own utterances, which was not present during silent mouthing of the phrase (condition B). The threshold was set at *P* < 0.05, corrected.

Results (Table 1, Figs 5 and 6)
Peaks of activity in response to own utterances were observed in the anterior right lateral STG, left and right mid STS and left and right PT. There was no separate peak in the posterior left STS. Associated with the motor gestures of speech, there were, as previously reported, activations in posterior frontal cortex; however, in addition, there was an activation in the depth of the most posterior part of the left sylvian sulcus, at the most medial part of the junction of the STG with the inferior parietal lobe.
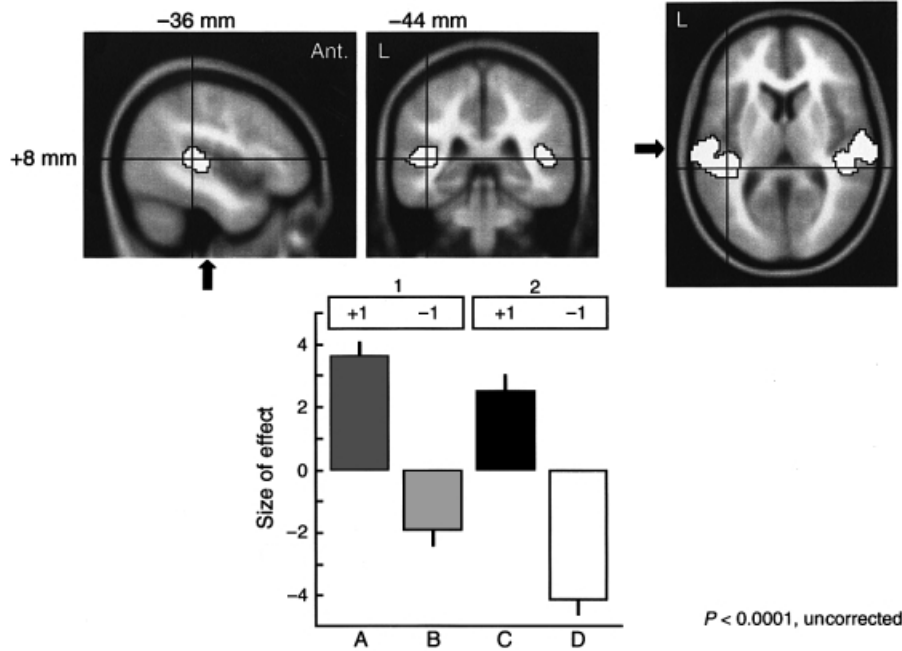
Fig. 5 Experiment 3: sagittal, coronal and axial slices on the MRI T_1-weighted template to show the peak in the left PT in the conjunction of the contrasts of speech (condition A) with mouthing (condition B) and voicing (condition C) with silent rehearsal (condition D), orthogonalized in that order. The method of display is the same as that employed in Fig. 2. It is apparent that activations are also present in the right PT and in the left and right HG (the plane of the left HG is depicted by a black arrow in the sagittal and axial projections). In the lower half of the figure, each condition (*x*-axis) is plotted against the size of its effect (*y*-axis) in the left PT, with the contrasts and their orthogonalization order shown above the plot.
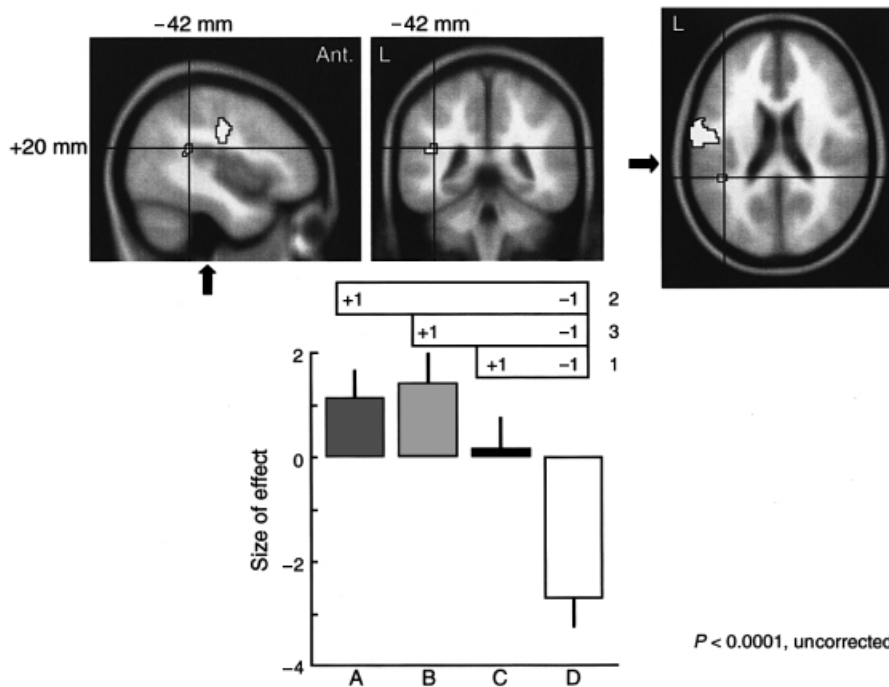

Fig. 6 Experiment 3: sagittal, coronal and axial slices on the MRI T_1-weighted template to show the peak in the medial left temporoparietal junction for the conjunction of the separate contrasts of voicing (condition C), speech (condition A) and mouthing (condition B) with silent rehearsal (condition D), orthogonalized in that order. The method of display is the same as employed in Fig. 2. In the lower half of the figure, each condition (*x*-axis) is plotted against the size of its effect (*y*-axis) in the medial left temporoparietal junction, with the contrasts and their orthogonalization order shown above the plot.

Comment

When contrasted with mentally rehearsing the phrase, a task involving no auditory input or auditory attention, the perception of own utterances produced bilateral supra-temporal activations that did not, unlike the response to hearing words observed in Experiment 1, extend ventrally into the posterior left STS. In addition, a posterior left temporal/inferior parietal system was identified that responded to the motor act of speech, independent of the speaker's perception of his own utterances.

Experiment 4

Design

Seven subjects took part in a study of noun generation (as in Experiment 2) and counting. The seven conditions, one scan per condition, were as follows.

A. Rest, when the subjects were asked to 'empty your mind'.
B. Noun retrieval, when the subjects had to think of basic level nouns after hearing a superordinate noun cue, one stimulus every 30 s, without speaking. Immediately following the scan, the subjects performed the task again out loud, with their responses recorded, to give an estimate of the number of basic level nouns generated per minute.
C. As B, but the stimuli were heard every 10 s.
D. As B, but the stimuli were heard every 2 s and the subject was told to only think of one response per stimulus.
E. As B, with one stimulus every 30 s, but the subjects were asked to speak their responses, which were recorded. One of the subjects did not complete this condition because of scan failure.
F. The subjects were asked to count silently from 1000 (1001 . . . 1002 . . . 1003 . . ., etc.). A root of one thousand was used to slow up the rate of counting, to approximate it to the rate of retrieving nouns in conditions B–E. At the end of the scan the subject was asked to name the number he had reached.
G. As F, but the subjects counted aloud. One of the subjects did not complete this condition because of scan failure.
Thus, the subjects only spoke their responses during scanning in conditions E and G.

Analyses

The contrast of (E + G) – (A + B + C + D + F) was used to determine the temporal lobe response to the sound of own utterances. Experiments 1–3 demonstrated that the posterior left STS responded to both hearing words (but not self-monitoring of speech output) and to retrieving words from memory. To investigate this further, the rate of hearing the stimuli plus the rate of generating responses for each scan for each subject was used as covariate, with the rate of hearing own utterances in conditions E and G excluded. The correlation in any voxel was not significant at the level of $P < 0.05$, corrected for analysis across the whole brain
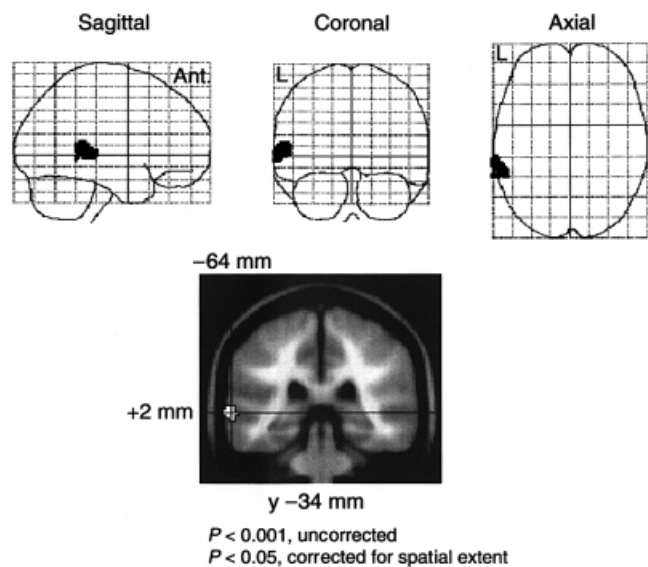

Fig. 7 Experiment 4: statistical parametric maps displayed as sagittal, coronal and axial projections in the upper half of the figure, demonstrating the posterior left STS where activity for word perception and retrieval was additive ($P < 0.001$, uncorrected for voxel-level significance; $P < 0.05$, corrected for spatial extent significance). In the lower half of the figure, this region is displayed on a coronal slice of the averaged MRI template, using the same method described in Fig. 2.

volume. As there was only one scan per condition for each subject, the power of this analysis was relatively low. Therefore, the voxel-level significance (peak intensity) was set at $P < 0.001$, uncorrected, but to avoid false positives the significance for the spatial extent of each activated region (i.e. the number of voxels in a cluster) was set at $P < 0.05$, corrected (Poline *et al.*, 1997).

Results (Table 1, Fig. 7)

There was a bilateral supratemporal response in response to hearing own articulations (not illustrated), closely similar to that observed in Experiment 3. The sum of the rates of hearing stimuli and generating responses correlated with activity within the posterior left STS. The number of voxels in this cluster was significant ($P = 0.002$; $P > 0.1$ in all other clusters).

Comment

This study demonstrated directly a conjunction of activity for single word perception and word retrieval in the posterior left STS. The activity in response to word retrieval was not specific for the recall of exemplars from semantic memory, as the retrieval of numbers also activated this region.

Discussion

The four studies have identified two regions in the left superior temporal cortex, posterior to HG, associated with

Experiment 1

Words – SCN

Experiment 2

Word retrieval from
semantic memory

Experiment 4

Word perception +
word retrieval

Experiment 3
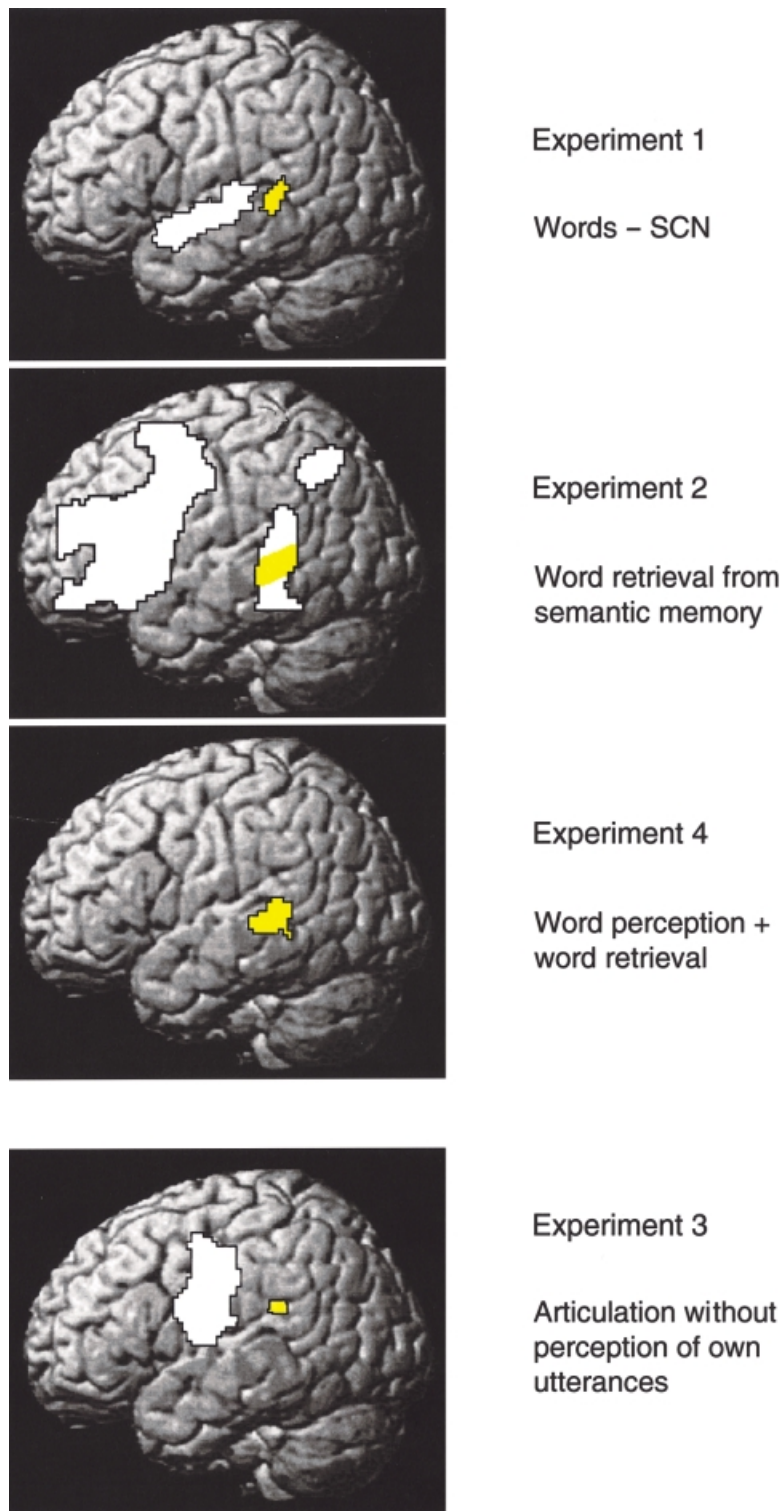
Articulation without
perception of own
utterances

Fig. 8 Summary figure from the four experiments, showing projections of activated voxels (thresholded at $P < 0.001$, uncorrected) on to the MRI template of the lateral surface of the left cerebral hemisphere available in SPM99. For Experiment 1 the left temporal regions are shown where there was a correlation between activity and the rate of hearing words, but not SCN, with the voxels located within the posterior left STS highlighted in yellow and all other voxels shown in white. Using a similar method of display, the left hemisphere regions activated by semantically cued word retrieval (Experiment 2) and the sum of activity for word perception and word retrieval (Experiment 4) are shown. The peak voxels for the posterior left STS in Experiments 1, 2 and 4 were within 4 mm of each other in the x, y and z planes. The voxels activated by articulation but not those responding to hearing own utterances (Experiment 3) include those at the medial left temporoparietal junction, highlighted in yellow and displayed, for illustrative purposes, on the lateral surface of the hemisphere.

the processing of single words (Fig. 8): the posterior STS, which is involved both in the perception of single words (but not of own utterances) and the retrieval of words from memory (both words in response to a semantic cue and words for numbers), and the junction of the STG with the inferior parietal lobe, which was engaged by the motor act of speech, independent of the speaker's perception of his own utterances. The response of the left PT was not selective, responding to complex non-speech sounds (SCN) and the sound of the speaker's own utterances. The lack of speech-specificity of the PT has already been observed in other functional neuroimaging studies. In one study, word perception and perception of tone sequences were contrasted, both directly and indirectly by comparing each condition separately with a silent control condition (Binder *et al.*, 1996). Others observed no signal in PT when contrasting listening to speech with listening to non-speech sounds (Demonet *et al.*, 1992; Zatorre *et al.*, 1992; Belin *et al.*, 2000).

Using microelectrode recordings and tracer injections in non-human primates, it has been shown that there are anterior and posterior auditory projections to, respectively, rostral (anterior) prefrontal cortex and dorsolateral prefrontal and premotor cortex (Romanski *et al.*, 1999). In addition to the direct projections from lateral belt regions, which is immediately adjacent to core auditory cortex, to frontal cortex, there are parallel routes with the same frontal lobe terminations: via adjacent anterior temporal regions and through the posterior STG and STS and the parietal lobe (Kaas and Hackett, 1999). It has been proposed that the anterior projections encode information about the object source of a sound, and the posterior projections encode auditory spatial information, analogous to the 'what' and 'where' visual pathways (Rauschecker, 1998; Kaas and Hackett, 1999). Although the anatomical evidence about the local connectivity of the human superior temporal cortex is limited, recent evidence clearly distinguishes between cortex anterior and posterior to HG (Galuske *et al.*, 1999). The former is reciprocally connected via monosynaptic pathways with HG, whereas the latter has no direct connections with HG; however, whether its main afferent input is from cortical or subcortical structures is not known. This difference in connectivity between anterior and posterior human auditory association cortex suggests a difference in function and supports the possibility of dual auditory streams in man.

Knowledge about 'where' directs attention, and the orientation of the eyes and body, towards a sound source. However, visual information also directs other motor responses, such as the arm reaching and finger movements required to grasp a small object (Goodale and Milner, 1992). In audition, sounds cannot be used to direct manipulation of the objects from which they originated but, particularly in humans, they can be used to direct the articulatory muscles, i.e. they can be mimicked. This is most evident in repeating back the utterances of a speaker, but humans can also mimic the vocalizations of other species and make approximations to the sounds made by inanimate objects. Mimicking both

words and non-speech sounds requires that an analysis of the sound structure of the percept is used to direct the muscles of respiration, the larynx, the pharynx, tongue and lips to reproduce the sound and an ability to relate articulatory gestures to the actual sound produced in the self-monitoring of one's own utterances. Of particular importance is the ability to transiently represent the temporal order of the elements, so as not to perceive and repeat, for example, 'tap' as 'pat'. Repeated rehearsal of the temporally ordered elements of words is central to the acquisition of long-term lexical representations of familiar words (Hartley and Houghton, 1996).

We have used the responsiveness of neural systems during word perception, retrieval and production to investigate whether the posterior auditory processing stream observed in non-human primates has developed a role in the human brain to support word rehearsal and lexical acquisition. We propose that the posterior left STS, which is equally responsive to hearing single words and retrieving single words from memory, acts as an interface between word perception and the long-term representations of familiar words held in memory. It may perform this role by transiently representing the temporally ordered sound structure of words, both heard words (the external source) and words retrieved from lexical memory (the internal source). Although silent verbal fluency is a complex task, involving a number of psychological processes, it includes the retrieval of the sound structures of appropriate lexical items and their mental rehearsal in preparation to speak (Warburton *et al.*, 1996). This is inferred from the distribution of activated regions, which include bilateral frontal opercular cortex and left lateral premotor cortex, lesions of which are known to impair severely speech production (Lecours and Lhermitte, 1976; Mohr *et al.*, 1978; Mao *et al.*, 1989; Broussolle *et al.*, 1996). Converging evidence for the importance of the posterior left superior temporal cortex in transiently representing sequences of phonemes in repetition and during word retrieval comes from two single case studies, which used cortical stimulation during epilepsy surgery. Stimulation at electrode pairs over the posterior left STG, close to or overlying the STS, resulted in phonetic errors during repetition and during naming pictures and naming from description (Anderson *et al.*, 1999; Quigg and Fountain, 1999).

A previous study (Fiez *et al.*, 1996) also re-analysed previous studies of hearing words and word retrieval, the latter in response to visually presented word cues. The re-analysis distinguished two posterior regions on the left. The dorsal region, located several millimetres posterior to the PT (Westbury *et al.*, 1999), responded most strongly to hearing words and the ventral region, located close to or within the posterior STS, was activated by word generation. Based on our observations, their ventral region should have been equally activated by hearing words and word generation. Inspection of their data shows that the difference in the magnitude of activation between the dorsal and ventral regions for word generation was four times greater than that

for hearing words and there was little difference in the response of the posterior left STS to hearing words and word retrieval. Therefore, there is consistency between the earlier retrospective analysis of Fiez and colleagues and ours.

In the medial posterior supratemporal cortical plane, at its junction with the inferior parietal lobe, we identified a neural subsystem activated by overt articulation. The results are consistent with the hypothesis that this region acts as an interface between speech perception or lexical recall and speech production. Silent verbal fluency was also associated with activation of the lateral aspect of the left PT, which demonstrated that lexical retrieval is associated with activation spreading from the STS towards the medial temporoparietal junction, with the latter only activated during overt articulation. Although the loci are not identical, a functional MRI study of lexical retrieval without articulation during picture naming has also been associated with several peaks of activity in the posterior left STG (Hickok *et al.*, 2000).

There is an alternative explanation. A previous PET study has been interpreted as indicating that regions encoding articulation modulate the left superior temporal cortex as motor-to-sensory discharges (Paus *et al.*, 1996). This raises the possibility that there may be an effect of proprioceptive feedback from articulatory structures on posterior temporal cortex. The temporal resolution of PET is incapable of settling whether the activation of posterior left superior temporal cortex in our study was pre- or post-articulatory, but a study of picture naming using magneto-encephalography demonstrated that activity in this region occurs prior to articulation (Levelt *et al.*, 1998; see also Hickock and Poeppel, 2000).

The demonstration that neurons in the inferior parietal lobe instruct motor actions has a precedent in a study of patients with right cerebral hemisphere lesions centred on the inferior parietal lobe, close to the temporoparietal junction (Mattingley *et al.*, 1998). It was demonstrated that delay in initiating a right hand movement towards the left in response to a visual cue in the left hemifield was as much due to slowness of motor initiation as to impaired attention to the visual stimulus. The authors concluded that neurons in the inferior parietal lobe act as an interface between a sensory percept and its associated motor response. Our results go further in demonstrating that a motor (speech) response can be associated with temporoparietal activation in response to the retrieval of an internal (lexical) cue, in the absence of a sensory (auditory) percept.

Observing the operation of the locally distributed system in the posterior temporal cortex in response to the word tasks our subjects were asked to perform does not allow us to speculate about its role in everyday speech production. This would require evidence that cued lexical retrieval uses the same system to retrieve lexical memories as that operating during word retrieval associated with propositional speech. Furthermore, we have not established whether the response

of the posterior left STS is only to speech. It remains to be seen whether it is engaged by the overt or covert rehearsal of non-speech sounds with complex temporal sequences, such as bird song, which can be successfully mimicked and learnt by humans.

In summary, the results from three PET studies have demonstrated a conjunction of activity in the posterior left STS in response to hearing single words and during cued word retrieval. We postulate that this local system transiently represents the temporally ordered sequence of sounds that comprise a heard (external) or retrieved (internal) word, and that it acts as an interface between the perception and long-term mental representations of familiar words. A fourth PET study demonstrated an adjacent local system, at the medial left temporoparietal junction, that acts as an interface between posterior temporal cortex and motor cortex for speech. These two anatomically and functionally separable regions are candidates for systems that must exist to allow us to perceive and rehearse novel words until they are acquired as retrievable lexical memories.

Acknowledgements

References

Anderson JM, Gilmore R, Roper S, Crosson B, Bauer RM, Nadeau S, et al. Conduction aphasia and the arcuate fasciculus: a reexamination of the Wernicke–Geschwind model. Brain Lang 1999; 70: 1–12.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature 2000; 403: 309–12.

Benson DF. Aphasia, alexia and agraphia. New York: Churchill Livingstone; 1979.

Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW. Function of the left planum temporale in auditory and linguistic processing. Brain 1996; 119: 1239–47.

Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, et al. Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 2000; 10: 512–28.

Bogen JE, Bogen GM. Wernicke's region—where is it? Ann NY Acad Sci 1976; 280: 834–43.

Broussolle E, Bakchine S, Tommasi M, Laurent B, Bazin B, Cinotti L, et al. Slowly progressive anarthria with late anterior opercular syndrome: a variant form of frontal cortical atrophy syndromes. J Neurol Sci 1996; 144: 44–58.

Demonet J-F, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise R, et al. The anatomy of phonological and semantic processing in normal subjects. Brain 1992; 115: 1753–68.

Evans AC, Collins DL, Mills SR, Brown RD, Kelly RL, Peters TM. 3D statistical neuroanatomical models from 305 MRI volumes. IEEE Nucl Sci Symp Med Imag Conf 1993: 1813–7.

Fiez JA, Raichle ME, Balota DA, Tallal P, Petersen SE. PET activation of posterior temporal regions during auditory word presentation and verb generation. Cereb Cortex 1996; 6: 1–10.

Friston KJ, Frith CD, Liddle PF, Dolan RJ, Lammertsma AA, Frackowiak RS. The relationship between global and local changes in PET scans. J Cereb Blood Flow Metab 1990; 10: 458–66.

Friston KJ, Ashburner J, Frith CD, Poline J-B, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. Hum Brain Mapp 1995a; 3: 165–89.

Friston KJ, Holmes AP, Worsley KJ, Poline J-B, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 1995b; 2: 189–210.

Galaburda AM, Sanides F, Geschwind N. Human brain. Cytoarchitectonic left-right asymmetries in the temporal speech region. Arch Neurol 1978; 35: 812–7.

Galuske RAW, Schuhmann A, Schlote W, Bratzke H, Singer W. Interareal connections in the human auditory cortex [abstract]. Neuroimage 1999; 9 (6 Pt 2): S994.

Geschwind N, Galaburda AM. Cerebral lateralization. Biological mechanisms, associations, and pathology: I. A hypothesis and a program for research. Arch Neurol 1985; 42: 428–59.

Gloor P. The temporal lobe and limbic system. New York: Oxford University Press; 1997.

Goodale MA, Milner AD. Separate visual pathways for perception and action. [Review]. Trends Neurosci 1992; 15: 20–5.

Hartley T, Houghton G. A linguistically constrained model of short-term memory for nonwords. J Mem Lang 1996; 35: 1–31.

Hickok G, Poeppel D. Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 2000; 4: 131–8.

Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, et al. A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. Neurosci Lett 2000; 287: 156–60.

Jones EG, Powell TP. An anatomical study of converging sensory pathways within the cerebral cortex of the monkey. Brain 1970; 93: 793–820.

Kaas JH, Hackett TA. 'What' and 'where' processing in auditory cortex [news]. Nat Neurosci 1999; 2: 1045–7.

Kosaki H, Hashikawa T, He J, Jones EG. Tonotopic organization of auditory cortical fields delineated by parvalbumin immunoreactivity in macaque monkeys. J Comp Neurol 1997; 386: 304–16.

Lecours AR, Lhermitte F. The 'pure form' of the phonetic disintegration syndrome (pure anarthria): anatamico-clinical report of a historical case. Brain Lang 1976; 3: 88–113.

Levelt WJ, Praamstra P, Meyer AS, Helenius PI, Salmelin R. An MEG study of picture naming. J Cogn Neurosci 1998; 10: 553–67.

Lichtheim L. On aphasia. Brain 1885; 7: 433–84.

Mao C-C, Coull BM, Golper LA, Rau MT. Anterior operculum syndrome. Neurology 1989; 39: 1169–72.

Mattingley JB, Husain M, Rorden C, Kennard C, Driver J. Motor role of human inferior parietal lobe revealed in unilateral neglect patients. Nature 1998; 392: 179–82.

Mesulam MM. From sensation to cognition. [Review]. Brain 1998; 121: 1013–52.

Mohr JP, Pessin MS, Finkelstein S, Funkenstein HH, Duncan GW, Davis KR. Broca aphasia: pathologic and clinical. Neurology 1978; 28: 311–24.

Mummery CJ, Ashburner J, Scott SK, Wise RJ. Functional neuroimaging of speech perception in six normal and two aphasic subjects. J Acoust Soc Am 1999; 106: 449–57.

Murphy K, Corfield DR, Guz A, Fink GR, Wise RJ, Harrison J, et al. Cerebral areas associated with motor control of speech in humans. J Appl Physiol 1997; 83: 1438–47.

Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. Eur J Neurosci 1996; 8: 2236–46.

Penhune VB, Zatorre RJ, MacDonald JD, Evans AC. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. Cereb Cortex 1996; 6: 661–72.

Poline J-B, Worsley KJ, Evans AC, Friston KJ. Combining spatial extent and peak intensity to test for activations in functional imaging. Neuroimage 1997; 5: 83–96.

Quigg M, Fountain NB. Conduction aphasia elicited by stimulation of the left posterior superior temporal gyrus. J Neurol Neurosurg Psychiatry 1999; 66: 393–6.

Rauschecker JP. Cortical processing of complex sounds. [Review]. Curr Opin Neurobiol 1998; 8: 516–21.

Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. Science 1995; 268: 111–4.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nature Neurosci 1999; 2: 1131–6.

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. [Review]. Philos Trans R Soc Lond B Biol Sci 1992; 336: 367–73.

Scott SK, Blank C, Rosen S, Wise RJS. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 2000; 123: 2400–6.

Shapleske J, Rossell SL, Woodruff PW, David AS. The planum temporale: a systematic, quantitative review of its structural, functional and clinical significance. [Review]. Brain Res Rev 1999; 29: 26–49.

Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. Stuttgart: Thieme-Verlag; 1988.

Warburton E, Wise RJ, Price CJ, Weiller C, Hadar U, Ramsay S, et al. Noun and verb retrieval by normal subjects: studies with PET. [Review]. Brain 1996; 119: 159–79.

Westbury CF, Zatorre RJ, Evans AC. Quantifying variability in the planum temporale: a probability map. Cereb Cortex 1999; 9: 392–405.

Williams PL, editor. Gray's anatomy. 38th ed. New York: Churchill Livingstone; 1995.

Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. Science 1992; 256: 846–9.