

# The Constraints Functional Neuroimaging Places on Classical Models of Auditory Word Processing

A. L. Giraud and C. J. Price

## Abstract

■ Several previous functional imaging experiments have demonstrated that auditory presentation of speech, relative to tones or scrambled speech, activate the superior temporal sulci (STS) bilaterally. In this study, we attempted to segregate the neural responses to phonological, lexical, and semantic input by contrasting activation elicited by heard words, meaningless syllables, and environmental sounds. Inevitable differences between the duration and amplitude of each stimulus type were controlled with auditory noise bursts matched to each activation stimulus. Half the subjects were instructed to say “okay” in response to presentation of all stimuli. The other half repeated back the words and syllables, named the source of the sounds, and said “okay” to the control stimuli (noise bursts). We looked for stimulus effects

that were consistent across task. The results revealed that central regions in the STS were equally responsive to speech (words and syllables) and familiar sounds, whereas the posterior and anterior regions of the left superior temporal gyrus were more active for speech. The effect of semantic input was small but revealed more activation in the inferior temporal cortex for words and familiar sounds than syllables and noise. In addition, words (relative to syllables, sounds, and noise) enhanced activation in the temporo-parietal areas that have previously been linked to modality independent semantic processing. Thus, in cognitive terms, we dissociate phonological (speech) and semantic responses and propose that word specificity arises from functional integration among shared phonological and semantic areas. ■

## INTRODUCTION

Models of auditory word processing date back more than a century. Based on behavioral assessments, analyses of lesion sites, neurophysiological recordings, and computational modelling, several distinct functional and anatomical subcomponents have been established. The earliest studies distinguished between speech perception in the left posterior superior temporal cortex (Wernicke, 1874) and speech production in the left posterior inferior frontal cortex (Broca, 1861). Subsequently, however, there have been limited advances in our understanding of functional anatomical relationships due to the well known problems associating lesion sites to cognitive deficits (see Shallice, 1988 for a review). Recent neuropsychological studies of patients have therefore focused on developing the cognitive rather than anatomical components. In the last two decades, such studies have decomposed the normal language system into a series of many interacting subcomponents. For example, speech perception can be grossly subdivided into (i) acoustic analysis that is not specific to speech sounds; (ii) phonological analysis that is specific to speech sounds; (iii) lexical processing of word stimuli; and (iv) semantic associations for mean-

ingful stimuli (see Caplan, 1992; Levelt, 1989; Patterson & Shewell, 1987).

The validity of these subcomponents, however, is still debated. For instance, are there two different stores for the sounds of words (the input and output lexicons) or is there a single lexicon that can be accessed by input and output processes (Shallice, 1988; Allport & Funnell, 1981)? In particular, connectionist models, derived from computation modeling, have emphasized that many of the cognitive properties that traditional cognitive models are based on can emerge from a system that has a limited number of highly interactive components (Gaskell & Marslen-Wilson, 1997; Valdois, Carbonnel, David, Rousset, & Pellat, 1995; Martin & Saffran, 1992). The implication for neuronal implementation is that a specific process might be subserved by discrete anatomical systems or alternatively result from distributed activity over a limited number of connected regions. Deductions about neuronal implementation, however, cannot be drawn from patient behavior and computational modeling, but require physiological validation with the anatomical precision offered by techniques like functional neuroimaging (Howard et al., 1992; Petersen, Fox, Posner, Mintum, & Raichle, 1988).

The systems level approach that functional neuroimaging offers is not limited to the segregation of anatomical regions that are specialized for discrete

---

Institute of Neurology, London, UK

cognitive processes, but allows for functional specialization that is embodied in the interactions among two or more cortical areas. Functional neuroimaging can assess whether there is a specialized neural system for a particular process or whether the implementation of that process is associated with a specific pattern of distributed activity in neural systems that are implicated in other functions (Friston, 1995). For example, is there a neuronal system for lexical processing that is independent of semantic and phonological systems or does word specificity arise from the distributed pattern of activity across shared cortical systems? The first step in answering such questions with functional neuroimaging is to identify the cortical systems for phonological, semantic, and lexical processing and establish the degree to which they are segregated. This is precisely what we aimed to do in the study reported here.

Functional neuroimaging studies have used two distinct approaches to segregate the anatomical components of speech processing. One involves manipulating the task to selectively weight specific subcomponents. For example, Binder et al. (1997), Démonet et al. (1992), and Démonet, Price, Wise, & Frackowiak (1994a) used tasks that required subjects to make an overt decision on the phonological content of heard nonwords (e.g., detecting the presence of /d/ and /b/ in the nonword “redozaabu”) or the semantic content of real words (e.g., deciding if animal names were “native to the United States and used by people”). These studies have indicated that phonological processing is more reliant on perisylvian regions and semantic processing is more reliant on extrasylvian regions (e.g., the left inferior temporal cortex and the left posterior temporo-parietal cortex). However, they are not representative of natural speech processing and activation differences will inevitably include differential task strategies (Démonet, Price, Wise, & Frackowiak, 1994b). For instance, in order to decide if the target sounds /d/ and /b/ are present in a heard nonword: (i) the target sounds must be kept in memory throughout the task, (ii) attention must be paid to the auditory input, (iii) when the stimuli are presented the nonword must be segmented into a sequence of component parts or syllables (e.g., “red” “oz” “a” “bu”), (iv) this sequence must also be held in memory while (v) a decision is made as to whether the targets are present in the correct order. Finally, (vi) a response is generated to communicate the decision.

The other approach is to manipulate the auditory stimulus during online speech processing. For instance, several neuroimaging experiments have explored the neural correlates of speech processing by contrasting activation elicited by words, pseudowords, reversed words, tones, and noise bursts (Binder et al., 2000; McCrory et al., 2000; Mummery, Ashburner, Scott, & Wise, 1999; Binder et al., 1997; Price, Wise, Warburton, et al., 1996; Zatorre et al., 1996; Zatorre, Meyer, Gjedde, & Evans, 1996; Mazoyer et al., 1993; Howard et al., 1992).

In the most comprehensive study, Binder et al. (2000) demonstrated that dorsal regions of the superior temporal gyri were equally responsive to tones, reversed speech, pseudowords, and words relative to noise, but more ventral regions in the lateral STS responded to speech sounds with equivalent responses for words, pseudowords, and reversed words relative to tones. The lateral STS were therefore associated with acoustic rather than linguistic aspects of speech processing. Another recent study by Belin et al. (2000) provides other relevant data. These authors found that the STS regions identified by Binder et al. (2000) are not specific to speech but equally activated by other nonspeech vocal sounds (e.g., laughs and sighs) relative to nonvocal sounds (finger snaps, footsteps, bells). Moreover, they demonstrated that the voice-selective areas might respond selectively to a combination of both high- and low-frequency components characteristic of voices, since excluding one or the other by filtering significantly reduced neuronal activity.

While central regions in the STS are clearly critical for acoustic processing of speech, areas specialized for phonological, semantic, and lexical processing have not been identified by manipulating stimulus type. For instance, Binder et al. (2000) found no significant differences between words, pseudowords, and reversed words unless the statistical threshold was reduced to  $p < .05$  uncorrected for multiple comparisons. There are several potential explanations for the null results. One is that functional neuroimaging is not sufficiently sensitive to distinguish the different components of speech input. For instance, phonological, lexical, and semantic processing may activate “implicitly” whenever complex auditory stimuli are presented thereby minimizing differences between stimuli that do or do not have semantic associations (see Price, Wise, & Frackowiak, 1996). Another explanation is that specialization for higher cognitive functions is not modular but arises from changes in the pattern of activity across distributed units (see above). Methods for analyzing functional and effective connectivity have been developed but they currently rely on prior anatomical models of cognitive function and therefore have not yet been applied to the study of speech perception. In the current study, we therefore sought to dissociate phonological, lexical, and semantic processing by attempting to manipulate these variables independently in a  $2 \times 2$  factorial design.

The four conditions were: (1) environmental sounds; (2) heard words corresponding to the names of the sounds (e.g., “dog” and “hammer”); (3) the simplest nonword phonological stimuli, (e.g., “tete”), which we refer to as the syllable condition; and (4) the baseline conditions. The latter were noise bursts with no phonological or semantic content but matched for duration, amplitude, and temporal envelope to each sound (Condition 4a), word (Condition 4b), and syllable (Condition 4c). Semantic input was presented in Conditions 1 and 2

but not 3 and 4. Phonological input was present in Conditions 2 and 3 but not 1 and 4. Lexical input was present in Condition 2 only. To minimize word associations during the syllable condition, we (i) made the syllabic stimuli as unwordlike as possible by excluding syllabic combinations that might have familiar associations (e.g., “dada”); and (ii) used a blocked presentation so that subjects were not expecting to hear familiar words. See Methods and list of stimuli in the Appendix for further details.

Half the subjects were instructed to say “okay” in response to every stimulus (word, syllable, sound, or baseline). The “okay” response was selected to ensure that the subjects attended to the stimuli, made an equivalent response during all conditions, and were not tempted to repeat or name the activation stimuli. In a passive listening condition, one cannot exclude the possibility of subvocal naming in the activation conditions only. The other half of the subjects were instructed to repeat or name the words, syllables, and sounds and say “okay” in response to each acoustic control stimulus. In this instance, the “okay” response partially controls for articulatory mechanisms (Moore & Price, 1999; Price, Moore, Humphreys, Frackowiack, & Friston, 1996). The different tasks allowed us to assess the effect

of stimulus type during different attentional sets. We looked for effects that were consistent across task.

## RESULTS

We report five classes of activation:

1. Common to words, syllables, and sounds relative to their stimulus-specific baselines
2. Phonological input (words and syllables more than sounds or baselines)
3. Semantic input (words and sounds more than syllables and baselines)
4. Lexical input (specific to words)
5. Other (specific to sounds or syllables)

The results are summarized in Table 1, which indicates the main effects over all 12 subjects, and for each group of 6 subjects individually.

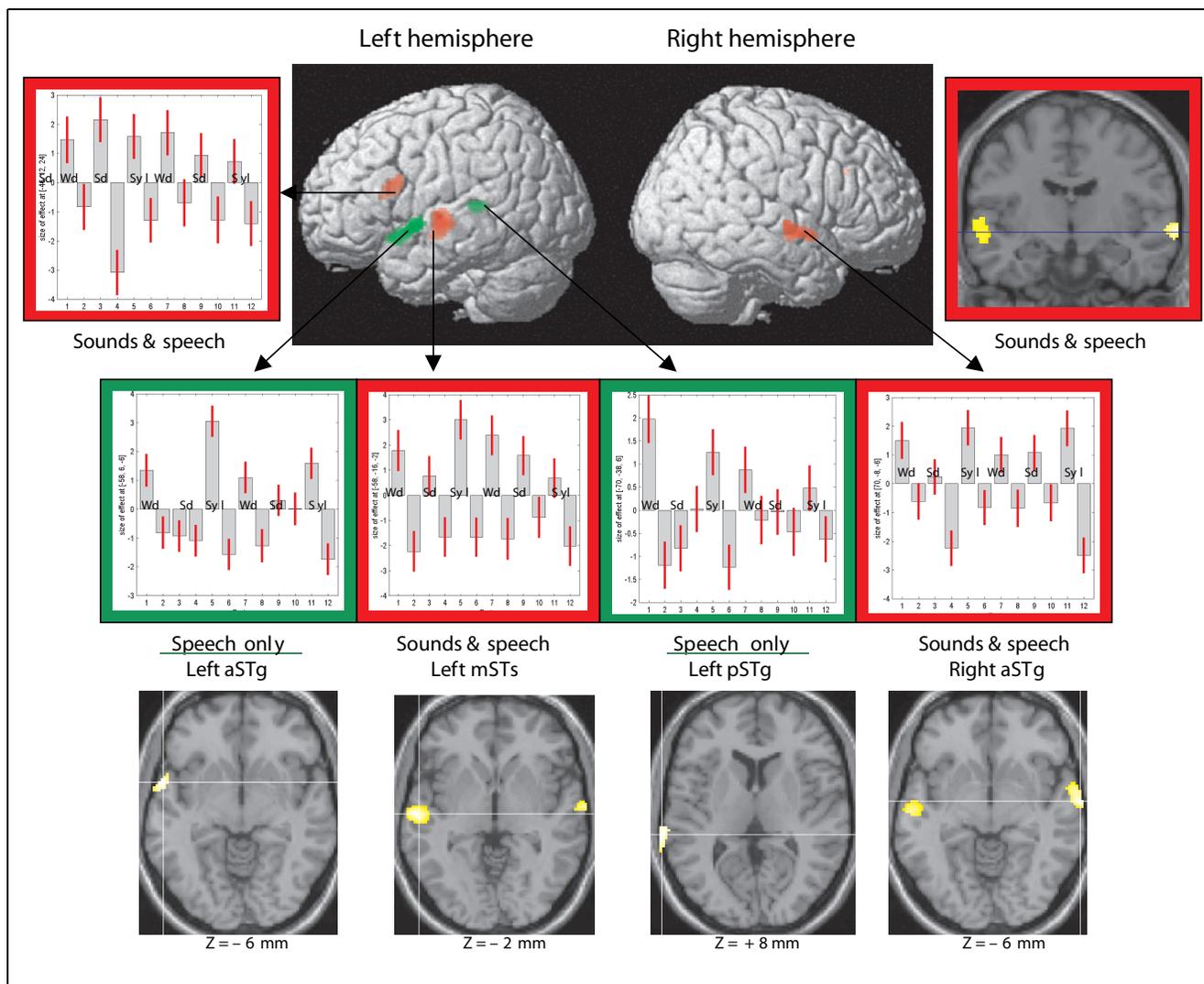
### Common to Words, Syllables, and Sounds Relative to Noise Bursts

At a corrected level of significance—irrespective of task—words, syllables, and sounds relative to control

**Table 1.** Details of Activations

1. Common effects ( $W - N$ and $Sd - N$ and $Syl - N$ )			ME	$Wd - N$	$Sd - N$	$Syl - N$	Name	Okay	
L. mid/anterior STS	BA 22	-56, -14, -4	<b>5.6</b>	<b>4.6</b>	3.0	<b>4.5</b>	<b>5.0</b>	<b>5.2</b>	
R. mid/anterior STS	BA 22	68, -10, -4	<b>5.9</b>	3.2	3.5	<b>4.9</b>	<b>5.8</b>	4.0	
R. anterior STS	BA 22	54, 2, -8	<b>5.6</b>	3.7	3.1	<b>4.8</b>	<b>5.0</b>	3.8	
L. inf. frontal (Broca's)	BA 44	-46, 12, 24 -44, 22, 14	<b>5.9</b> <b>4.9</b>	3.2	<b>4.4</b>	3.5	4.3	<b>4.7</b>	
Anterior cingulate		6, 22, 34	<b>5.0</b>	3.2	3.0	3.4	3.9	3.2	
2. Phonological input ( $W + Syl$ ) > ( $Sd + N$ )			ME	$Wd - N$	$Syl - N$	$Wd - Sd$	$Syl - Sd$	Name	Okay
L. posterior STg	BA 42/22	-70, -38, 6	<b>5.1</b>	3.9	4.1	3.3	3.0	<b>5.7</b>	2.8
L. anterior STg		-62, 6, -6	<b>6.5</b>	3.6	<b>6.2</b>	2.7	4.3	<b>5.7</b>	<b>4.8</b>
3. Semantic input ( $W + Sd$ ) > ( $Syl + N$ )			ME	$Wd - N$	$Sd - N$	$Wd - Syl$	$Sd - Syl$	Name	Okay
L. dorsal inferior frontal		-48, 10, 36	4.2	2.9	2.8	3.5	2.7	3.5	3.0
L. inf. temporal	BA 20/21	-54, -8, -16 -68, -42, -14	3.2 2.9	2.7 2.4	3.3 1.9	1.6 <sup>a</sup> 2.2	1.5 1.9	2.3 2.3	2.4 2.2
4. Word specific ( $W - N$ and $W - Sd$ and $W - Syl$ )			ME	$Wd - N$	$W - Syl$	$Wd - Sd$	Name	Okay	
R. post. temporo-parietal	BA 22/39	64, -66, 22	3.5	3.5	3.0	2.8	<b>4.5</b>	<i>ns</i>	
L. post. temporo-parietal	BA 22/39	-52, -56, 18 -34, -74, 42	3.3 3.2	2.3 2.3	3.6 3.7	2.7 2.4	3.6 1.9	1.7 2.9	

The anatomical names, coordinates (Talairach & Tournoux, 1988), and Z scores of all activations. Results are from all 12 subjects except in last two columns where “name” refers to main effect from naming/repetition group and “okay” refers to main effect from group that said “okay” to all stimuli. Z scores in bold reached a corrected level of significance ( $Z > 4.4$ ).  $Wd$  = words;  $Sd$  = sounds;  $Syl$  = syllables;  $N$  = noise; ME = main effect; L = left; R = right; sup. = superior; inf. = inferior; ant. = anterior; post. = posterior; ST = superior temporal; STS = superior temporal sulcus.  
<sup>a</sup>Effect of words relative to syllables was only detected in group of subjects that named/repeated stimuli.



**Figure 1.** Activations common to sounds and speech relative to noise (red) and specific to speech (green) rendered onto left and right hemispheres of a template brain with a coronal cut (right) and axial cuts (bottom row) showing precise anatomical localization. Abbreviations as in Table 1. Histograms illustrate relative blood flow response in each activated area for each of the 12 conditions (1–6 from Study 1 and 7–12 from Study 2). Bar 1 = repeating words; Bar 2 = saying “okay” to noises matched to words; Bar 3 = naming sounds; Bar 4 = saying “okay” to noises matched to sounds; Bar 5 = repeating syllables; Bar 6 = saying “okay” to noises matched to syllables; Bar 7 = saying “okay” to words; Bar 8 = saying “okay” to noises matched to words; Bar 9 = saying “okay” to sounds; Bar 10 = saying “okay” to noises matched to sounds; Bar 11 = saying “okay” to syllables; Bar 12 = saying “okay” to noises matched to syllables.

conditions activated: (i) central regions of bilateral STS (BA 22/21), lateral, anterior, and ventral to the primary auditory cortices; (ii) the left posterior inferior frontal gyrus (Broca’s area); and (iii) the anterior cingulate. Figure 1 illustrates the anatomical localization of these activations (in red) and plots of the relative activations across all 12 conditions. Table 1 also illustrates the consistency of these effects over stimulus type (words, sounds, and syllables) and subject group (name/repeat or say “okay”).

### Specific to Phonological Input

At a corrected level of significance, irrespective of task, words and syllables relative to noise and familiar sounds

increased activation in the left superior temporal gyrus, both posterior and anterior to the areas activated by familiar sounds. Figure 1 illustrates the anatomical localization of these activations (in green) and plots of the relative activations across all 12 conditions. Activation in the posterior area was highly significant for the 6 subjects who were required to repeat/name the stimuli ( $Z$  score = 5.6) but was much weaker ( $Z$  score = 2.8) in the six subjects who said “okay” to each stimulus (see Table 1 and Figure 1).

### Semantic Input

Words and sounds relative to noise and syllables did not reveal activation at a corrected level of significance.

**Table 2.** Semantic Regions Identified in this and Previous Studies

<i>Study</i>	<i>Task</i>	<i>Stimuli</i>	<i>Left</i>	<i>Right</i>
<i>Left anterior inferior temporal cortex</i>				
Giraud & Price, this study	Auditory repetition and listening	Sounds and words	-54, -8, -16	
Gorno-Tempini et al., 1998	Viewing faces	Familiar–unfamiliar	-54, -8, -26	
Gorno-Tempini et al., 1998	Reading words	Familiar–unfamiliar	-48, -16, -28	
Mummary et al., 1998	Semantic > phonological decisions	Object names	-44, -22, -16	
Mummary et al., 1999	Semantic > perceptual decisions	Pictures of objects	-68, -20, -28	
Vandenberghe et al., 1996	Semantic > perceptual decisions	Pictures of objects	-44, -10, -28	
<i>Bilateral temporo-parietal junction</i>				
Giraud & Price, this study	Auditory repetition	Words–syllables	-50, -54, 20	56, -66, 20
McCrorry et al., 1999	Auditory repetition	Words–pseudowords	-52, -48, 26	<i>ns</i>
Démonet et al., 1994a, 1994b	Semantic > phonological decisions	Auditory words	-44, -60, 24	40, -58, 12
Gorno-Tempini et al., 1998	Viewing faces	Familiar–unfamiliar	-52, -56, 24	<i>ns</i>
Gorno-Tempini et al., 1998	Reading words	Familiar–unfamiliar	-60, -60, 20	<i>ns</i>
Mummary et al., 1998	Semantic > phonological decisions	Object names	-50, -68, 20	48, -64, 22
Mummary et al., 1999	Semantic > perceptual decisions	Pictures of objects	-66, -60, 16	62, -60, 24
Price, Moore, & Frackowiak, 1996	Silent reading > rest	Written words	-52 -54, 16	53 -56 16

The anatomical names and coordinates (Talairach & Tournoux, 1988) of activations elicited by: (A) listening to words relative to tools in Binder et al. (2000); (B) listening to words relative to listening to “signal correlated noise equivalents” (SCN) in Mummary et al. (1999); and (C) listening to vocal sounds (speech and nonspeech) relative to nonvocal sounds and SCN in Belin et al. (2000); and (D) activations in the same regions from the study reported in this paper. The latter is accompanied by either “sounds” (activated by sounds and speech) or “speech” (activated by speech only). Abbreviations: as in Table 1. Anterior is “ant.,” posterior is “post.,” middle is “mid.”

When the threshold was lowered to  $p < .001$  uncorrected, activation for words and sounds but not syllables was detected in two regions of the inferior temporal cortex. The inferior frontal activation (see above) was also found to be more extensive (in the dorsal direction) for words and familiar sounds than syllables. Although these effects are weak and did not survive a correction for multiple comparisons, it is interesting that the inferior temporal areas have previously been associated with semantic processing using attention-demanding task manipulations (see Table 2).

The consistency with previous studies suggests that our findings are not false-positives, rather they indicate greater semantic activation for words and familiar sounds than syllables or noise.

### Word Specificity

For word repetition only, a region in the right posterior temporo-parietal cortex was more active than any other condition and this effect reached a corrected level of significance. When the threshold was lowered to  $p < .001$  uncorrected, the left posterior temporo-parietal junction was also revealed. Although the  $Z$  score on the left did not reach a corrected level of significance,

McCrorry et al. (2000) have reported the same left posterior temporo-parietal area for repetition of words relative to pseudowords (see Table 2). Furthermore, other studies have associated both left and right temporo-parietal areas with semantic tasks on faces, objects, and visual words even when there is no auditory input and auditory output is controlled (Gorno-Tempini et al., 1998; Mummary, Patterson, Hodges, & Price, 1998; Binder et al., 1997; Démonet et al., 1994a, 1994b, see Table 2b). Therefore, the function of our word-specific activations cannot be equated with the auditory input lexicon but suggests more semantic activation for words than syllables and sounds.

### Sound and Syllable Specificity

There were no sounds- or syllable-specific activations at a corrected level of significance.

### Summary of Results

Our experimental design allowed us to dissociate (i) central regions of bilateral STS, which respond to words, syllables, and familiar sounds relative to noise bursts; (ii) left anterior and posterior STg, which were more active

for words and syllables than familiar sounds; and (iii) left inferior temporal regions that were more responsive to stimuli with semantic associations (familiar words and sounds relative to syllables and noise).

The effects of semantic input and word specificity did not survive the correction for multiple comparisons. Nevertheless, the areas identified at  $p < .001$  uncorrected (left inferior temporal and bilateral posterior temporo-parietal cortices) correspond to activations that have previously been linked to semantic processing on the basis of task manipulations that specifically weight semantic retrieval (see Table 2).

## DISCUSSION

In this study, we used functional neuroimaging to identify the distributed neuronal systems activated by words, syllables, and environmental sounds. By comparing the different patterns of activation elicited by the different stimuli, we attempt to segregate neural activity (a) common to speech and sound perception; (b) specific to speech/phonological input; (c) specific to semantic input; and (d) specific to lexical input (words).

### Common to Speech and Sound Perception

Areas associated with perception of speech and environmental sounds were those that were equally responsive to speech and sounds relative to baseline, irrespective of task. This revealed central portions of the STS extending more anteriorly into the superior temporal gyrus on the right. Table 3 shows the similarity between our results and those of Belin et al. (2000), Binder et al. (2000), and Mummery et al. (1999). The Binder et al. and Mummery

et al. data indicated that the STS were specific to speech stimuli. The Belin et al. data indicated that the same regions were specific to vocal stimuli. In contrast, we demonstrate equivalent responses for speech and familiar sounds and suggest that differences between our findings and those of Belin et al. may relate to the category of sounds presented. In the Belin et al. study, there were multiple sources of nonvocal sounds including 14% nature (wind, streams), 29% animals (cries and gallops), 37% man-made (telephone, cars), and 20% musical instruments and the authors did not compare relative activation in the different categories. In our study, there were 50% animal cries and 50% man-made items and both evoked equivalent activation in STS (Animal sounds: right coordinates = 70, -14, -2,  $Z = 3.1$  and 68, -4, -4,  $Z = 2.5$ ; left coordinates = -62, -16, 2,  $Z = 2.5$ ; Man-made sounds: right coordinates = 62, 4, -8,  $Z = 3.1$  and 68, -12, -6,  $Z = 2.1$ ; left coordinates = -54, -10, -8,  $Z = 2.8$ ) with no significant difference in either hemisphere even when the threshold was reduced to  $p < .08$  uncorrected.

In addition, we observed highly significant activation for speech and familiar sounds, irrespective of task, in the left posterior inferior frontal cortex and the anterior cingulate. These areas have not been associated with passive word listening paradigms (Belin et al. 2000; Binder et al., 2000; Mummery et al., 1999; Price, Wise, Warburton, et al., 1996) but Zatorre et al. (1996) found a frontal region with almost the same coordinates (-44, +8, +27) to this study (-44, +14, +22) when a phoneme monitoring task (press a key if a heard word contains a "b") was contrasted to passive listening and Price, Wise, Warburton, et al. (1996) found a slightly more ventral region (-58, +8, +12) for repetition

**Table 3.** Results from Previous Studies

		(A) Words-tones (Binder et al., 2000)	(B) Words-SCN (Mummery et al., 1999)	(C) Vocal-nonvocal (Belin, 2000, Experiment 2)	(D) This study
Mid. STS	Right lateral	+58, -16, -2	+50, -20, 0	+52, -19, -1	+68, -10, -4 (sounds)
		+52, -14, 4	+63, -13, -1		
	Left lateral	-53, -14, -1	-54, -14, 0	-62, -14, 0	-56, -14, -4 (sounds)
Ant. ST	Right lateral	+60, -7, -4	+54, 2, -4	+60, -1, -4	+54, 2, -8 (sounds)
	Left lateral	-58, -7, -1	-48, 2, -12	-	62, 6, -6 (speech)
Post. ST	Left lateral	-52, -42, 6	-54, -38, 8	-	70 -38 6 (speech)
		-59, -33, 3	-54, -28, 4	-	
	Right lateral	-	-	+56, -30, 6	-
Planum temporale	Left medial	-	-	-40, -37, 13	-

A selection of previous studies that have shown activation in the same areas as found for words and familiar sounds.

relative to listening. The Price, Wise, Warburton, et al. (1996) and Zatorre et al. (1996) results indicate that activation in the posterior inferior frontal area increases with attention to phonological input or output. However, we found frontal activation for sounds as well as words and syllables (see Table 1) indicating a role that is not specific to speech/phonological input. We also found equivalent frontal activation for the six subjects who said “okay” to each stimulus and the six subjects who named/repeated the stimulus, which does not appear to be consistent with the demands placed on speech output. Nevertheless, we cannot exclude the possibility that when subjects had to say “okay” to familiar stimuli, the demands on speech output might be as high as when subjects had to name or repeat the stimuli. This would occur if speech output corresponding to the stimulus (the actual name of the stimulus) was activated even when the subjects were instructed to say “okay.”

We suggest that when subjects are required to make a stimulus-specific response (name/repeat or say “okay”), attention to the auditory input is increased relative to passive word listening paradigms (Belin et al., 2000; Binder et al., 2000; Mummery et al., 1999). Furthermore, when the stimuli are familiar (speech and environmental sounds rather than noise), increased attention may increase higher-level processing at both perceptual and production levels. The frontal activation may reflect this higher-level processing but our imaging experiment cannot distinguish whether it is involved at a perceptual or production level. Lesion data have classically associated Broca’s area with speech production, but (i) Dronkers (1996) has demonstrated that the critical site for articulation is the anterior insula not the posterior inferior frontal gyrus; and (ii) lesions to Broca’s area also impair speech perception (Blumstein, 1995). Thus, the inferior frontal area may be part of an input–output circuit where output feeds back to modulate processing at an input level.

The anterior cingulate may also be part of this input–output circuit. Indeed, it is well established that the anterior cingulate is part of an attentional circuit that serves to regulate cognitive and emotional processing (Bush, Luu, & Posner, 2000) particularly when there is response conflict (George et al., 1994; Taylor, Kornblum, Minoshima, Oliver, & Koeppe, 1994; Bench et al., 1993; Pardo, Pardo, Janer, & Raichle, 1990).

In summary, we have shown that during auditory word processing, activation in central regions of bilateral STS, right anterior STS, and left posterior inferior frontal cortex is not specific to speech or vocal input. We suggest that the central STS regions are involved in the perception of familiar sounds and speech and that activation is modulated by attention to the auditory input. Concurrent activation in the anterior cingulate is consistent with this theory. The left posterior inferior frontal cortex is classically associated with speech output and the demands on speech output may be greater when subjects

say “okay” to familiar relative to unfamiliar stimuli. However, we cannot exclude the possibility that the frontal activation was involved in semantic processing or regulation of acoustic processing.

### Specific to Speech/Phonological Input

For words and syllables relative to familiar and unfamiliar sounds, activation in the left lateral superior temporal cortex extended (a) posteriorly and dorsally and (b) anteriorly around the anterior superior temporal gyrus. Both areas have previously been associated with speech relative to nonspeech (Binder et al., 2000; Mummery et al., 1999) but not for words relative to reversed words (Binder et al., 2000; Price, Wise, Warburton, et al., 1996; Howard et al., 1992) and not for vocal sounds relative to amplitude-modulated noise and scrambled voice stimuli (Belin et al., 2000, Experiment 2). The left anterior and posterior temporal areas therefore appear to be specific to phonological content. They also appear to have very different functions from one another. For instance, anatomical studies in monkeys suggest that anterior projections support recognition and posterior projections may be involved in localization (Romanski et al., 1999). In human subjects, functional neuroimaging studies have shown the anterior temporal area to be specific for phonological input, whereas the posterior area is also involved in phonological output in the absence of phonological input (e.g., during reading, see Price, 2000; Moore & Price, 1999). Indeed, the posterior superior temporal region corresponds to Wernicke’s area where damage clearly results in speech output difficulties with less impact on speech perception.

In the study we report here, activation in the posterior area was more significant when subjects repeated the words and syllables ( $Z = 5.6$ ) than when subjects said “okay” ( $Z = 2.8$ ), whereas responses in the anterior area were equivalent irrespective of task or subject group. One possibility is that the posterior area (Wernicke’s area) is involved in recoding phonological input into phonological output and activation during passive listening conditions (Binder et al., 2000; Mummery et al., 1999) occurs automatically. However, this does not explain why the posterior area responds to stimuli that cannot be articulated such as (i) reversed words (Binder et al., 2000); (ii) noise presented at a frequency of 4 Hz (the average syllabic rate) relative to noise presented at higher rates (less important for speech segmentation, see Giraud et al., 2000); and (iii) unexpected (deviant) tones (Celsis et al., 1999). Wernicke’s area has also been shown to respond to decreased temporal predictability in the spatial sequence of visually presented colored squares (Bischoff-Grethe, Proper, Mao, Daniels, & Berns, 2000). To account for the involvement in both speech input and speech output, Wise et al. (2001) have proposed that Wernicke’s area transiently represents “the temporally ordered sound structure of words, both heard words

(the external source) and words retrieved from lexical memory (the internal source).” This explanation may also parsimoniously account for responses to deviant or unpredicted events (Bischoff-Grethe et al., 2000; Celsis et al., 1999). However, it is not clear how it accounts for the lack of response during sound naming (this study) or picture naming (Moore & Price, 1999). An alternative possibility is that the posterior STS may have many different functions depending on the areas it interacts with (see Price et al., 2001 for similar hypotheses regarding the posterior inferior temporal lobe). We therefore conclude that the function(s) of the left posterior and anterior temporal cortex require further investigation. Our study contributes by segregating the responses in the anterior and posterior superior temporal cortex from those in the central STS by manipulating the demands on phonological and semantic processing independently.

### Semantic Processing

The effect of semantic input was not as robust as the effect of phonological input. This may be because the syllable condition elicited a certain degree of semantic processing despite our efforts to curtail it (see Introduction and Methods). Alternatively, it could be because semantic processing is not strongly activated when subjects name sounds, repeat words, or say “okay” to either. Nevertheless, increased activation in the left dorsal posterior frontal cortex and two regions of the left inferior temporal cortex was observed for words and familiar sounds relative to syllables when the statistical threshold was lowered to  $p < .001$  uncorrected. Explanations for the differential frontal activation have been discussed in the previous section. Here, we focus on the two inferior temporal regions, both of which have been associated with semantic processing in previous studies. For instance, our posterior inferior temporal area (coordinates:  $-68, -42, -14$ ), was lateral to an area reported by Binder et al. (2000) for listening to words relative to pseudowords (coordinates:  $-46, -40, -17$  at  $p < .05$  uncorrected) and by Démonet et al. (1992, 1994a, 1994b) for semantic relative to phonological decisions (coordinates:  $-40, -38, -16$ ). The more anterior inferior temporal area (coordinates:  $-54, -8, -16$ ) is less likely to be detected in fMRI studies because it is subject to macroscopic magnetic susceptibility artifacts (Lipschutz, Ashburner, Friston, & Price, in press; Devlin et al., 2000; Veltman, Friston, Sanders, & Price, 2000; Ojemann et al., 1997). In positron emission tomography (PET) studies, however, this region has been associated with semantics in a wide range of tasks (Gorno-Tempini et al., 1998; Mummery et al., 1998; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996; see Table 4); most notably, it has been associated with the intelligibility of speech even when acoustic processes are controlled (Scott, Blank, Rosen, & Wise, 2000). In

short, our semantic conditions did indeed enhance activity in areas previously associated with semantic processing even though activation did not reach conservative levels of significance.

### Word Specificity

The only areas found to be more active for words than all other conditions were located in bilateral posterior temporo-parietal junctions. These areas have been associated with modality independent semantic processing that is not specific to phonological input or output (Gorno-Tempini et al., 1998; Mummery et al., 1998; see Table 4). Therefore, it appears that semantic activation was greatest during the word condition. Possibly, word repetition and sound naming trigger divergent semantic associations but further experiments that tap semantic processing more explicitly are required to specify what the corresponding neural specialization might be. It is not simply the case that these regions are specific for lexical processing because a study by Engelein et al. (1995) reported posterior inferior parietal activation in response to semantic decisions on environmental sounds.

In terms of cognitive models, our results suggest that word specificity arose from coactivation of areas specialized for phonological input and those specialized for semantic processing. Damage to the connections between these regions may result in word-specific deficits, for instance, impaired auditory word processing with intact repetition of nonwords and semantics (Valdois et al., 1995). In other words, our imaging data are consistent with the function of the auditory input lexicon arising from the integration of phonological and semantic processing. Nevertheless, many more functional imaging studies are required to validate our findings and we cannot exclude the possibility that there are word-specific areas that our study has failed to detect. For instance, there may be clusters of word-specific neurons located in the same vicinity as those that respond to other auditory stimuli and the sensitivity and spatial resolution of our technique (8 mm) may have been insufficient to tease such activations apart. However, if this were the case, it would be extremely unlikely that naturally occurring lesions, which tend to be much larger than our spatial resolution, would selectively remove word-specific neuronal populations. It is far more likely that naturally occurring lesions would remove word-specific connections between two distant regions (i.e., those associated with phonological input and semantic processing).

### CONCLUSION

We draw two distinct conclusions. At an anatomical level, we demonstrate for the first time that central regions in bilateral STS, inferior to the primary auditory

cortices, respond to familiar environmental sounds as well as speech. These areas are therefore not specific to the acoustic features in speech (Binder et al., 2000) or vocal sounds (Belin et al., 2000). The areas that were speech-specific in our study were located in the anterior and posterior parts of the left superior temporal gyri. With respect to cognitive models of speech processing, we conclude that there is no current evidence for a segregated cortical system that is dedicated to lexical processing of auditory words. Our findings suggest that word specificity arises from the pattern of distributed activity in phonological and semantic systems. Further studies are required to investigate whether effects that are specific to words, sounds, and syllables emerge from changes in the effective connectivity between shared processing regions. In other words, the emphasis should move towards functional integration rather than functional segregation.

## METHODS

### Design

The experiment was designed around six types of auditory stimuli: words, meaningless syllables, environmental sounds, and 3 white noise conditions which controlled for duration and amplitude of each word, syllable or sound. A  $2 \times 2$  factorial design emerged because the words and syllables (but not environmental sounds or noises) involved phonological input and the words and environmental sounds (but not syllables or noises) had semantic associations. The third (between subjects) factor was task. Half the subjects said “okay” in response to each stimulus irrespective of condition. The other half said “okay” to the stimuli presented in the white noise conditions but repeated the words and syllables and named the source of the environmental sounds. The order of the 6 conditions was counterbalanced both within and between subjects. There were two replications of each condition (see stimuli below).

All subjects, in all conditions, were instructed to articulate the response without generating any sound. Subjects’ responses were monitored using a video camera and lip reading. The silent response ensures that auditory processing from the sound of the subjects’ own response did not contribute to the activation patterns observed.

### Stimuli

Twelve different sets of digitized auditory stimuli were generated for the 12 different scans. This included 4 distinct steps. In the first step, two sets of environmental sounds were selected including either animate (animal sounds) or inanimate (telephone ringing, drill, car, etc.) stimuli (see Appendix). In the second step, the source

names of the selected sounds were recorded in a male voice (e.g., the word “dog” is the source name of the sound of a dog barking). The words were selected like this to match, as far as possible, the word condition to the sound condition with respect to the phonological output and semantic associations (although the semantic associations for words and sounds will not be identical). Differences in the duration of words and sounds were controlled by including stimulus-specific baselines (see Step 4, below) and irrespective of condition, the rate of presentation was held constant across all conditions (1/4 sec).

In the third step, two sets of syllable stimuli were generated that matched the words in terms of number of syllables. However, in order to make the syllabic stimuli as unwordlike as possible, no attempt was made to match the syllables to the words for phonemic content or complexity; even the bi- and trisyllabic combinations were composed of repetitions of the same syllable/phoneme (e.g., “ta,” “tata,” “tatata,” see Appendix). To further minimize lexical processing in response to the syllables, (1) syllabic combinations with familiar associations (e.g., “dada”) were excluded; and (2) blocked presentation of all conditions ensured that during the syllable condition, subjects were not expecting to hear words—they simply attended to the simplest phonological stimuli. The male voice used to record the syllables was the same as that used to record the words. These constraints aimed to curtail lexical processing and make the stimuli as easy to repeat as possible, thereby maximizing the possibility of detecting word-specific effects. However, any word-specific effect observed may also relate to phonemic complexity or word-specific semantic associations. These can be distinguished in reference to previous studies.

Finally, in the fourth step, we attempted to partially control for the inevitable acoustic differences between speech and environmental sounds, by generating a burst of white noise for every word, syllable, and sound stimulus. Noise bursts were created by applying the low-pass temporal envelope of each experimental stimulus (cut-off 4 Hz) to a broad-band white noise (bandwidth 0–20000 Hz). The envelope was derived using a low-pass filter (cut-off 4 Hz). The critical feature was that the duration, temporal envelope, and average amplitude of each noise matched that of the corresponding experimental stimulus. This resulted in condition-specific baseline conditions corresponding to each of the 6 experimental conditions described above (2 sounds, 2 words, 2 syllables) with a total of 12 sets of stimuli (6 experimental conditions and 6 baseline conditions). Within each condition, stimuli were presented at a constant rate of 0.25 sec and the order of conditions was counterbalanced within and between subjects. The words, syllables, and environmental sounds correspond to the experimental conditions and the noises correspond to the baseline conditions.

## Subjects

There were 12 participants. Those who said “okay” to each stimulus had a mean age of 36.3 years; those who repeated the speech stimuli and named the sounds had a mean age of 36.6. There were 4 men and 2 women in each group. All subjects were neurologically normal, right-handed, used English as their first language, and gave written informed consent after the rationale of the studies had been explained.

## Data Acquisition

PET was used to generate 12 measurements of regional cerebral blood flow (rCBF) assessed with intravenous injection of water labelled with  $^{15}\text{O}$ . The dose received was 9 mCi per measurement. Images were acquired using a Siemens CTI III camera. The study was approved by the local hospital ethics committee and the UK Administration of Radioactive Substances Advisory Committee (ARSAC).

## Data Analysis

Standardized procedures were used for data acquisition and data analysis. Realignment, normalization, and statistics were performed with SPM99 ([www.fil.ion.ucl.ac.uk/spm](http://www.fil.ion.ucl.ac.uk/spm); Friston et al., 1995, 1996). Images were spatially smoothed with a 16-mm Gaussian filter. In the statistical analysis, data were collapsed over semantic category (animate or inanimate) because an initial preview of the results indicated that this did not alter the results. All subjects were analyzed in a single design matrix and six conditions were specified per subject: words, word control, sounds, sound control, syllables, syllable control.

The contrasts used to isolate different processes were as follows:

1. Areas associated with “speech and sound perception” were those engaged by all experimental conditions (words, syllables, sounds) irrespective of response. They were identified with a conjunction analysis of (i) words relative to word-specific noise (the effect of words); (ii) syllables relative to syllable-specific noise (the effect of syllables); and (iii) environmental sounds relative to sound-specific noise (the effect of sounds), for each group of subjects.

2. Areas associated with “phonological input” required that two criteria were made. First, there must be a main effect of words and syllables relative to noise and sounds. Second, there must be greater activation for (i) words relative sounds and (ii) syllables relative to sounds; (iii) words relative to word noise; and (iv) syllables relative to syllable noise. This was achieved using conjunctions and the inclusive masking option in SPM. See below for thresholds used.

3. Areas associated with “semantic input” were identified as for phonological input. First, there must be an effect of words and sounds relative to noise and syllables. Second, there must be greater activation for (i) words relative syllables; (ii) sounds relative to syllables; (iii) words relative to word noise; (iv) sounds relative to sound noise.

4. Areas associated with “word-specific activations” were those where there were effects of words relative to (i) sounds, (ii) syllables, and (iii) noise.

5. Areas associated with “sound-specific activations” were those where there were effects of sound relative to (i) words, (ii) syllables, and (iii) noise.

6. Areas associated with “syllable-specific activations” were those where there were effects of syllables relative to (i) words, (ii) sounds, and (iii) noise.

The six different effects described above were identified across subjects (i.e., irrespective of task). In addition, we also looked at the effects for each subject group independently.

We do not report differences between subject group because this would require a random effect analysis based on between-subject variance and we had insufficient subjects in each group to make this viable.

## Setting the Threshold for Significance

We used two different statistical thresholds. First, we set the threshold for the SPMs to  $p < .05$  corrected for multiple comparisons and the masks were thresholded at  $p < .01$  uncorrected. Then, we lowered the threshold of the SPMs to  $p < .001$  uncorrected and the threshold for the masks to  $p < .05$  uncorrected. The corrected level of significance will increase false-negatives. The uncorrected level of significance will increase false-positives.

## Appendix: List of the Stimuli

### *Words and Sounds*

elephant	telephone
cockerel	motorbike
baby	siren
donkey	doorbell
cow	car
dog	bell
sheep	train
duck	hammer
pig	water
horse	whistle
owl	airplane
bird	engine

### Syllables (with phonetic symbols for vowels)

va	æ	terter	ə	va	æ	terter	ə
to	a	deedee	i	to	a	deedee	i
dee	i	veivei	e	dee	i	veivei	e
tu	ʌ	terterter	ə	tu	ʌ	terterter	ə
bi	ɪ	toto	o	bi	ɪ	toto	o
va	æ	to	a	va	æ	to	a
toto	o	teitei	e	toto	o	teitei	e
ter	ə	tatata	æ	ter	ə	tatata	æ
to	a	bi	ɪ	to	a	bi	ɪ
tata	æ	vava	æ	tata	æ	vava	æ
tetete	ɛ	tei	e	tetete	ɛ	tei	e
ter	ə	tata	æ	ter	ə	tata	æ

### Acknowledgment

This work was supported by the Wellcome Trust.

Reprint requests should be sent to Dr. Cathy Price, Institute of Neurology, Wellcome Department of Cognitive Neurology, 12, Queen Square, WC1N 3BG London, UK, or via e-mail: c.price@fil.ion.ucl.ac.uk.

### REFERENCES

- Allport, D. A., & Funnell, E. (1981). Components of the mental lexicon. *Philosophical Transactions of the Royal Society of London, Series B*, *295*, 397–410.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice selective areas in human auditory cortex. *Nature*, *403*, 309–312.
- Bench, C. V., Frith, C. D., Grasby, P. M., Paulesu, E., Frackowiak, R. S. J., & Dolan, R. J. (1993). Investigations of the functional anatomy of attention using the Stroop test. *Neuropsychologia*, *31*, 907–922.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*, 512–528.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, *17*, 353–362.
- Bischoff-Grethe, A., Proper, S. M., Mao, H., Daniels, K. A., & Berns, G. S. (2000). Conscious and unconscious processing of nonverbal predictability in Wernicke's area. *Journal of Neuroscience*, *20*, 1975–1981.
- Blumstein, S. E. (1995). The neurobiology of the sound structure of language. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 915–929). Cambridge: MIT Press.
- Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé suivie d'une observation d'aphémie. *Bulletin de la Société Anatomique de Paris*, *6*, 330–357.
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–222.
- Caplan, D. (1992). *Language. Structure, processings and disorders*. Cambridge: MIT Press.
- Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L., & Chollet, F. (1999). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage*, *9*, 135–144.
- Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-D., Wise, R., Rascol, A., & Frackowiak, R. S. J. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753–1768.
- Démonet, J. F., Price, C. J., Wise, R., & Frackowiak, R. S. J. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: A positron emission tomography study. *Neuroscience Letters*, *182*, 25–28.
- Démonet, J. F., Price, C. J., Wise, R., & Frackowiak, R. S. J. (1994). A PET study of cognitive strategies in normal subjects during language tasks: Influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain*, *117*, 671–682.
- Devlin, J. T., Russell, R. P., Davis, M. H., Price, C. J., Wilson, J., Mathews, P. M., Moss, H. E., & Tyler, L. K. (2000). Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. *Neuroimage*, *11*, 589–600.
- Dronkers, N. F. (1996). A new brain region for co-ordinating speech articulation. *Nature*, *384*, 159–161.
- Engelien, A., Silbersweig, D., Stern, E., Huber, W., Doring, W., Frith, C., & Frackowiak, R. S. J. (1995). The functional anatomy of recovery from auditory agnosia. A PET study of sound categorisation in a neurological patient and normal controls. *Brain*, *118*, 1395–1409.
- Friston, K. J. (1995). Functional and effective connectivity in neuro-imaging: A synthesis. *Human Brain Mapping*, *2*, 56–78.
- Friston, K. J., Ashburner, J., Poline, J.-B., Frith, C. D., Heather, J. D., & Frackowiak, R. S. J. (1996). Spatial realignment and normalization of images. *Human Brain Mapping*, *2*, 165–189.
- Friston, K. J., Holmes, A., Worsley, K. J., Poline, J.-B., Frith, C. D., & Frackowiak, R. S. J. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*, 189–210.
- Gaskell, G. M., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613–656.
- George, M. S., Ketter, T. A., Parekh, P. I., Rosinsky, N., Ring, H., Casey, B. J., Trimble, M. R., Horwitz, B., Herscovitch, P., & Post, R. M. (1994). Regional brain activity when selecting response despite interference: An H2150 PET study of the Stroop task and an emotional Stroop task. *Human Brain Mapping*, *1*, 194–209.
- Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R. S. J., & Kleinschmidt, A. (2000). Temporal envelope representation in the human auditory cortex. Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*, *84*, 1588–1598.
- Gorno-Tempini, M. L., Price, C. J., Vandenberghe, R., Josephs, O., Cappa, S., Kapur, N., & Frackowiak, R. S. J. (1998). The neural systems sustaining face and proper name processing. *Brain*, *121*, 2103–2118.
- Howard, D., Patterson, K., Wise, R. J. S., Brown, W. D., Friston, K., Weiller, C., & Frackowiak, R. S. J. (1992). The cortical localization of the lexicons: Positron emission tomography evidence. *Brain*, *115*, 1769–1782.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Lipschutz, B., Ashburner, J., Friston, K. J., & Price, C. J. (2001). Assessing study-specific regional variations in fMRI signal. *Neuroimage*, *13*, 392–398.

- Martin, N., & Saffran, E. M. (1992). A computational account of deep dysphasia: Evidence from a single case study. *Brain and Language*, *43*, 240–274.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*, 467–479.
- McCrorry, E., Frith, U., Brunswick, N., & Price, C. J. (2000). Abnormal functional activation during a simple word repetition task: A PET study of adult dyslexics. *Journal of Cognitive Neuroscience*, *12* (5), 753–762.
- Moore, C. J., & Price, C. J. (1999). Three distinct ventral occipito-temporal regions for reading and object naming. *Neuroimage*, *10*, 181–192.
- Mummery, C. J., Ashburner, J., Scott, S. K., & Wise, R. J. (1999). Functional neuro-imaging of speech perception in six normal and two aphasic subjects. *Journal of the Acoustical Society of America*, *106*, 449–457.
- Mummery, C. J., Patterson, K., Hodges, J. R., Price, C. J. (1998). Functional neuroanatomy of the semantic system: divisible by what? *Journal of Cognitive Neuroscience*, *10*, 766–777.
- Ojemann, J. G., Akbudak, E., Snyder, A. Z., McKinstry, R. C., Raichle, M. E., & Conturo, T. E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo planar fMRI susceptibility artifacts. *Neuroimage*, *6*, 156–167.
- Pardo, J. V., Pardo, P. J., Janer, K. W., & Raichle, M. F. (1990). The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Sciences, U.S.A.*, *87*, 256–259.
- Patterson, K., & Shewell, C. (1987). Speak and spell: Dissociations and word class effects. In M. Coltheart, G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language*. London: Erlbaum.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomography studies of the cortical anatomy of single word processing. *Nature*, *331*, 585–588.
- Price, C. J. (2000). The functional anatomy of language. Contributions from Neuroimaging. *Journal of Anatomy*, *197*, 335–359.
- Price, C. J., Moore, C. J., & Frackowiak, R. S. J. (1996). The effect of varying stimulus rate and duration on brain activity during reading. *Neuroimage*, *3*, 40–52.
- Price, C. J., Moore, C. J., Humphreys, G. W., Frackowiak, R. S. J., & Friston, K. J. (1996). The neural regions sustaining object recognition and naming. *Proceedings of the Royal Society, Series B*, *263*, 1501–1507.
- Price, C. J., Warburton, E. A., Moore, C. J., Frackowiak, R. S. J., & Friston, K. J. (2001). Dynamic diaschisis: Context sensitive human brain lesions. *Journal of Cognitive Neuroscience*, *13*, 419–429.
- Price, C. J., Wise, R. J. S., & Frackowiak, R. S. J. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cerebral Cortex*, *6*, 62–70.
- Price, C. J., Wise, R. J. S., Warburton, E. A., Moore, C. V., Howard, D., Patterson, K., Frackowiak, R. S. J., & Friston, K. J. (1996). Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain*, *119*, 919–931.
- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*, *2*, 1131–1136.
- Scott, S. K., Blank, C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*, *123*, 2400–2406
- Shallice, T. (1988). Language operations: Are input and output processes separate? In *From neuropsychology to mental structure* (pp. 158–183). Cambridge: Cambridge University Press.
- Talairach, J., & Tournoux, P. (1988). *A co-planar stereotactic atlas of the human brain*. Stuttgart: Thieme.
- Taylor, S. F., Kornblum, S., Minoshima, S., Oliver, L. M., & Koeppe, R. A. (1994). Changes in medial cortical blood flow with a stimulus-response compatibility task. *Neuropsychology*, *32*, 249–255.
- Valdois, S., Carbonnel, S., David, D., Rousset, S., & Pellat, J. (1995). Confrontation of PDP models and dual-route models through the analysis of a case of deep dysphasia. *Cognitive Neuropsychology*, *12*, 681–724.
- Vandenberghe, R., Price, C. J., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional anatomy of a common semantic system for words and pictures. *Nature*, *383*, 254–256.
- Veltman, R., Friston, K. J., Sanders, G., & Price, C. J. (2000). Regionally specific sensitivity differences in fMRI and PET: Where do they come from? *Neuroimage*, *11* (6), 575–588.
- Wernicke, C. (1874). *Der Aphasische Symptomenkomplex*. [The aphasias]. Breslau, Poland: Cohen and Weigert.
- Wise, R. J. S., Scott, S. K., Blank, C., Mummery, C. J., Murphy, K., & Warburton, L. A. (2001). Separate neural subsystems within Wernicke's area. *Brain*, *124*, 83–95.
- Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex*, *6*, 21–30.