

The Neural Mechanisms of Speech Comprehension: fMRI studies of Semantic Ambiguity

Jennifer M. Rodd¹, Matthew H. Davis² and Ingrid S. Johnsrude^{2,3}

¹Department of Psychology, University College London, UK,
²MRC Cognition and Brain Sciences Unit, Cambridge, UK and
³Department of Psychology, Queen's University, Kingston, Canada

A number of regions of the temporal and frontal lobes are known to be important for spoken language comprehension, yet we do not have a clear understanding of their functional role(s). In particular, there is considerable disagreement about which brain regions are involved in the semantic aspects of comprehension. Two functional magnetic resonance studies use the phenomenon of semantic ambiguity to identify regions within the fronto-temporal language network that subserve the semantic aspects of spoken language comprehension. Volunteers heard sentences containing ambiguous words (e.g. 'the shell was fired towards the tank') and well-matched low-ambiguity sentences (e.g. 'her secrets were written in her diary'). Although these sentences have similar acoustic, phonological, syntactic and prosodic properties (and were rated as being equally natural), the high-ambiguity sentences require additional processing by those brain regions involved in activating and selecting contextually appropriate word meanings. The ambiguity in these sentences goes largely unnoticed, and yet high-ambiguity sentences produced increased signal in left posterior inferior temporal cortex and inferior frontal gyri bilaterally. Given the ubiquity of semantic ambiguity, we conclude that these brain regions form an important part of the network that is involved in computing the meaning of spoken sentences.

Keywords: frontal cortex, semantics, sentences, spoken language, temporal lobe

Introduction

Understanding natural speech is ordinarily so effortless that we often overlook the complex computations that are necessary to make sense of what someone is saying. Not only must we identify all the individual words on the basis of the acoustic input, but we must also retrieve the meanings of these words and appropriately combine them to construct a representation of the whole sentence's meaning. When words have more than one meaning, contextual information must be used to identify the appropriate meaning. For example, for the sentence 'The boy was frightened by the loud bark', the listener must work out that the ambiguous word 'bark' refers to the sound made by a dog and not the outer covering of a tree. This process of selecting appropriate word meanings is important because most words are ambiguous. At least 80% of the common words in a typical English dictionary have more than one dictionary definition (Parks *et al.*, 1998; Rodd *et al.*, 2002), and some words have a very large number of different meanings — there are 44 different definitions listed for the word 'run' in the WordSmyth Dictionary (Parks *et al.*, 1998). Therefore, selecting appropriate word meanings is likely to place a substantial load on the neural systems involved in computing sentence meanings. The functional magnetic resonance imaging (fMRI) studies reported

here use this phenomenon of semantic ambiguity to identify the brain regions that are involved in the semantic aspects of speech comprehension, in particular in the processes of activating, selecting and integrating contextually appropriate word meanings.

The modality-specific cortical substrates of the early stages of speech processing are relatively well studied. Evidence from anatomical and neurophysiological studies in non-human primates, and anatomical, electrophysiological and neuroimaging studies in humans supports the idea of hierarchically organized processing streams extending outwards from primary auditory regions with more complex processes localized further away from primary regions. In neuroimaging studies of speech perception, this is manifest as non-specific activation on Heschl's gyrus for all sounds, with increasing specificity for speech and speech-like sounds in surrounding regions. Activation specific to speech is typically observed in the superior temporal sulcus (STS): a region that anatomical evidence suggests is several stages of processing removed from primary auditory cortex (Kaas and Hackett, 2000). The areas that surround Heschl's gyrus, closer to primary auditory cortex than the STS, probably subserve a series of early stages in speech comprehension such as auditory analysis and mapping sounds onto speech units (i.e. phonemes, syllables, etc.; Binder *et al.*, 2000; Scott *et al.*, 2000; Davis and Johnsrude, 2003; Narain *et al.*, 2003; see Scott and Johnsrude, 2003, for a review).

In contrast to the anatomical organization of early processing stages, the brain bases of higher-level, meaning-based processes in language comprehension are more controversial. Although a century or more of neuropsychological studies indicate that the left inferior frontal gyrus and the left temporoparietal junction (often referred to as Broca's area and Wernicke's area) play important roles in speech comprehension, there is still considerable debate about the functional roles of these brain regions, and about the importance of other regions. In particular, there is disagreement about which brain regions are involved in processing the semantic aspects of speech. For instance, lesion analyses of patients with progressive deficits in processing the meaning of words (e.g. Hodges *et al.*, 1992), emphasize the role of the anterior temporal lobes in semantic processing (Mummery *et al.*, 2000). In contrast, analyses of lesions in stroke patients with comprehension deficits suggest that posterior and inferior regions of the temporal cortex (Bates *et al.*, 2003; Dronkers *et al.*, 2004) are critical for spoken language understanding. Similarly, neuroimaging studies of healthy adults have shown activation specific to meaningful speech in both anterior and posterior regions of the left superior and middle temporal gyri (e.g. Scott *et al.*, 2000; Binder *et al.*, 2000; Davis and Johnsrude, 2003). Although it

is likely that these posterior and anterior regions may participate in separate functionally specialized processing streams, there is little consensus about what these functional specializations might be. For example, Scott and Johnsrude (2003) suggest that an anterior system might be important for mapping acoustic-phonetic cues onto lexical representations, whereas a posterior-dorsal system might process articulatory-gestural representations of speech acts. Hickok and Poeppel (2000, 2004) propose a similar role for a posterior-dorsal system, but they emphasize the role of posterior inferotemporal regions in accessing meaning-based representations.

The role of the left inferior frontal gyrus (LIFG) in speech comprehension is also controversial. Several studies have shown elevated activation in the LIFG when an explicit semantic judgement task, requiring semantic information about single words to be explicitly retrieved or selected, is used (Thompson-Schill *et al.*, 1997; Wagner *et al.*, 2001; Thompson-Schill, 2003). At present, however, it remains unclear whether LIFG regions are also important for processing semantic information in the context of natural speech comprehension. Some authors have proposed that frontal regions in general, and the LIFG in particular, is not ordinarily recruited for the comprehension of natural speech, at least not for sentences that are easily and automatically understood (Crinion *et al.*, 2003). Other authors have suggested that within the language domain, frontal brain regions play an important role in processing syntax and morphology (and possibly phonology), but that processing word meanings relies on temporal lobe regions (Ullman, 2001).

In the following experiments, we use fMRI to identify the brain regions that are recruited when listeners need to disambiguate an ambiguous word on the basis of its sentence context (e.g. the *bark of the dog/tree*). Despite an extensive psycholinguistic literature that investigates the cognitive mechanisms that underlie semantic disambiguation (e.g. Swinney, 1979; Rayner and Duffy, 1986; Gernsbacher and Faust, 1991), little is known about how the brain is organized to support this ability. These results will have important consequences for two current debates about the neuroanatomy of speech comprehension. First, it will allow us to identify regions of the temporal lobe that are involved in processing the semantic aspects of speech. Second, it will allow us to determine whether the LIFG, which is known to be important for selecting/retrieving semantic information in the context of explicit semantic judgement tasks, is also involved in selecting/retrieving the appropriate meaning of semantically ambiguous words.

We therefore compare the brain responses to two types of spoken sentences that maximally contrast the demands on the brain regions involved in resolving semantic ambiguity. A set of high-ambiguity sentences was constructed to contain two or more words that have more than one meaning (e.g. 'The *shell* was *fired* towards the *tank*'). These sentences are compared with a matched set of low-ambiguity sentences (e.g. 'Her secrets were written in her diary') that contain minimal semantic ambiguity. Importantly, the two types of sentences were matched for syntactic structure and rated naturalness. In addition, modulated noise with the same spectral profile and amplitude envelope as speech was used as a low-level baseline condition. We conducted two fMRI experiments using these stimuli. In the first study, in order to ensure that participants attended to the sentences, they were required to judge whether a word, presented visually after the sentence, was semantically related to the sentence's meaning. In the second study, this

secondary task was removed and participants were instructed to listen attentively without making any behavioural response.

Experiment 1

Materials and Method

Materials

There were two experimental conditions (high-ambiguity sentences and low-ambiguity sentences) and a low-level noise baseline condition. There were 59 items in each of these three conditions (see Appendix for example sentences). The high-ambiguity sentences all contained at least two ambiguous words (e.g. 'There were *dates* and *pears* in the fruit bowl'). The ambiguous words were either homonyms¹ (which have two meanings that have the same spelling and pronunciation, e.g. 'bark'), or homophones (which have two meanings that have the same pronunciation but have different spelling, e.g. 'knight'/'night'). Each high-ambiguity sentence was matched to a low-ambiguity sentence that had the same number of words and the same syntactic structure but contained words with minimal ambiguity (e.g. 'There was beer and cider on the kitchen shelf'). The duration of the individual sentences ranged from 1.2 to 4.3 s. The two sets of sentences were matched for the number of syllables, physical duration, rated naturalness, rated imageability and the log-transformed mean frequency of the content words in the CELEX database (Baayen *et al.*, 1995) (see Table 1).² The imageability and naturalness scores came from pretests in which two groups of 18 participants who did not take part in the fMRI studies listened to the sentences and rated how imageable or natural they were on a 9-point Likert scale (where 9 is highly imageable or natural).

A set of 59 sentences that were not used in the experiment and were matched for number of syllables, number of words and physical duration to the experimental sentences, were converted to signal-correlated noise (scn; Schroeder, 1968) using Praat software (<http://www.praat.org>). These stimuli have the same spectral profile and amplitude envelope as the original speech, but because all spectral detail is replaced with noise, they are entirely unintelligible. Although the amplitude envelope of speech (which is preserved in signal correlated noise) can, in theory, provide cues to some forms of prosodic and phonological information (Rosen, 1992), such cues are insufficient for the listener to recognize lexical items and therefore to extract any information about the sentence's meaning (Davis and Johnsrude, 2003; see Supplementary Material).

For each real sentence, a probe word was generated that was either semantically related (50%) or unrelated (50%) to the sentence's meaning. These words were matched across condition for length, frequency and their semantic relatedness to the sentence (rated on a 7-point scale) (see Table 2 for descriptive statistics). These probe words were used in the relatedness judgement task in which participants were required to decide whether the word was related to the meaning of the sentences (see Procedure). The aim of this task was to ensure attention to the sentences, and it was intended to be an easy task, of equal difficulty in the two conditions. The probes were selected to be clearly related or unrelated to the meaning of the sentences, and were never related to the unintended meanings of the ambiguous words.

Participants

Fifteen right-handed volunteers (10 females, aged 18–40) were scanned. All participants were native speakers of English, without any history of neurological illness, head injury or hearing impairment. The study was approved by the Addenbrooke's Hospital's Local Research Ethics Committee and written informed consent was obtained from all participants.

Table 1
Descriptive statistics for sentences

	N	Syllables	Words	Length (s)	Naturalness	Imageability	Word frequency
High-ambiguity	59	11.6	9.3	2.2	6.3	5.2	4.5
Low-ambiguity	59	11.8	9.3	2.2	6.4	5.3	4.7
Noise	59	11.8	9.3	2.2			

Table 2

Descriptive statistics for relatedness probes

Ambiguity	Relatedness	Example	Relatedness	Length	Frequency
High	Related	the shell was fired towards the tank. . . battle	6.0	5.7	4.3
Low	Related	her secrets were written in her diary. . . hidden	5.9	5.7	4.1
High	Unrelated	the waist of the jeans was very narrow. . . amuse	1.3	5.7	4.1
Low	Unrelated	the pattern on the rug was quite complex. . . mistake	1.3	5.7	4.0

Procedure

Volunteers were instructed to listen carefully to the sentences. At the end of each sentence a probe word was presented visually. Volunteers were instructed to make a button press to indicate whether this word was related to the meaning of the sentence or not (right index finger or thumb for related, left index finger or thumb for unrelated). For the baseline noise condition, the word 'right' or 'left' appeared on the screen and volunteers simply pressed the appropriate button.

We used a sparse imaging technique (Hall *et al.*, 1999), to minimize interference from scanner noise. Volunteers heard a single sentence (or noise equivalent) in the 8 s silent period before a single 3 s scan (Fig. 1). The timing of stimulus onset and offset was jittered relative to scan onset by temporally aligning the midpoint of the stimulus item (0.6–2.2 s after sentence onset) with a point 5 s before the midpoint of the subsequent scan (i.e. 4.5 s into the 8 s silent period). Based on previous estimates of haemodynamic responses to auditory stimuli (Hall *et al.*, 2000), it is likely that this timing ensures that scans are maximally sensitive to the initial online processing of the sentences. The visual relatedness probes appeared at the start of the scans, ensuring that very little of the haemodynamic response to the probe word could be observed in the scan.

There were 59 trials of each stimulus type and an additional 21 silent trials for the purpose of monitoring data quality. The experiment was divided into three sessions of 66 scans/trials. Stimulus items were pseudo-randomized to ensure that the three experimental conditions and rest scans were evenly distributed among the three sessions, and that each condition occurred equally often after each of the other conditions. Session order was counterbalanced across participants. Stimuli were presented to both ears using a high-fidelity auditory stimulus-delivery system incorporating flat-response electrostatic headphones inserted into sound-attenuating ear defenders (Palmer *et al.*, 1998). To further attenuate scanner noise, participants wore insert earplugs. DMDX software running on a Windows 98 PC (Forster and Forster, 2003) was used to present the stimulus items and record button-press responses. Volunteers were given a short period of practice in the scanner with a different set of sentences.

The imaging data was acquired using a Bruker Medspec (Ettlingen, Germany) 3 T MR system with a head gradient set. Echo-planar image volumes (198 for each volunteer) were acquired over three 12 min sessions. Each volume consisted of 21 × 4 mm thick slices with an interslice gap of 1 mm; FOV: 25 × 25 cm; matrix size, 128 × 128, $T_E = 27$ ms; acquisition time 3.02 s; actual $T_R = 11$ s. Acquisition was transverse-oblique, angled away from the eyes, and covered all of the brain except in a few cases the very top of the superior parietal lobule, the anterior inferior temporal cortex, and the inferior aspect of the cerebellum.

Results

Behavioural Results

There were no significant differences between participants' button-press responses for the high- and low-ambiguity sentences in either the mean response times (813 and 816 ms) or the error rates (2.3% and 3.1%) (both $P > 0.2$). Therefore, any differences in the imaging data are unlikely to relate to differences in task difficulty or response generation in the two conditions. Participants' responses were significantly faster (577

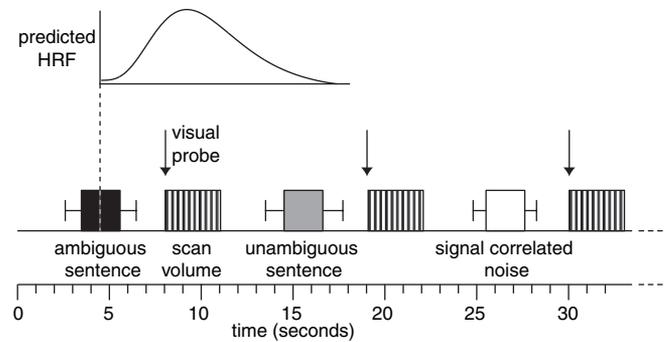


Figure 1. Details of the sparse imaging procedure (see Materials and Methods) in which a single stimulus item was presented in the silent periods between scans. The mid-point of the sentence was timed such that the predicted BOLD response to each sentence (based on the canonical haemodynamic response function in the SPM software) would be maximal at the time of the scan. Error bars show the range of sentence durations used. Visual probes (in experiment 1) occurred at the onset of the scan, minimizing sensitivity to the BOLD effect of these events.

ms) and more accurate (0.7%) in the baseline noise condition compared with both of the speech conditions (all $P < 0.001$).

fMRI Results

The fMRI data were preprocessed and analysed using Statistical Parametric Mapping software (SPM99, Wellcome Department of Cognitive Neurology, London, UK). Pre-processing steps included within-subject realignment, spatial normalization of the functional images to a standard EPI template, masking regions of susceptibility artefact to reduce tissue distortion (Brett *et al.*, 2001), and spatial smoothing using a Gaussian kernel of 12 mm, suitable for random-effects analysis (Xiong *et al.*, 2000). Analysis was conducted using a single general linear model for each volunteer in which each scan within each session (after excluding two initial dummy volumes) was coded for whether it followed the presentation of signal correlated noise, a low-ambiguity or a high-ambiguity sentence. Each of the three scanning runs was modelled separately within the design matrix. Additional columns encoded subject movement (as calculated from the realignment stage of preprocessing) as well as a constant term for each of the three scanning runs. Images containing the contrast of parameter estimates for each single-subject analysis were entered into second-level group analyses in which intersubject variation was treated as a random effect (Friston *et al.*, 1999).

The contrast between the low-ambiguity speech and the baseline noise condition showed a large area of activation in left and right superior and middle temporal gyri, extending in the left hemisphere into posterior inferior temporal cortex and the left fusiform gyrus. Activation was also observed bilaterally in lingual gyrus. We also observed activation in left frontal cortex centred on the dorsal part of the inferior frontal gyrus (pars triangularis) (see Table 3).

For the high-ambiguity sentences compared to low-ambiguity sentences there was increased signal in three brain regions (significant at $P < 0.05$ corrected for multiple comparisons); these were the left and right inferior frontal gyrus (pars triangularis), and a region of the left posterior inferior temporal cortex (see Table 4).

The results of this experiment provide clear answers to the theoretical questions raised in the introduction. Within the temporal lobe, activation was confined to inferior, posterior

Table 3

Low-ambiguity versus noise: all significant activation peaks >8 mm apart ($P < 0.05$ corrected for multiple comparisons). Additional peaks within a single activation cluster are indented, following the most significant peak from each cluster.

	<i>P</i> (corrected)	Z-score	co-ordinates		
			<i>x</i>	<i>y</i>	<i>z</i>
<i>Experiment 1</i>					
L anterior MTG	0.001	6.0	-62	-10	-8
L MTG	0.001	6.0	-60	-30	-6
L posterior MTG	0.001	6.2	-56	-38	0
L anterior MTG	0.001	5.7	-58	-12	-20
L anterior MTG	0.023	5.0	-54	6	-28
L anterior STG	0.035	4.8	-54	10	-16
L fusiform gyrus	0.002	5.6	-30	-8	-30
L fusiform gyrus	0.000	5.4	-34	-18	-22
L fusiform gyrus	0.018	5.0	-22	-32	-20
L parahippocampal gyrus	0.022	5.0	-30	-24	-16
R parahippocampal gyrus	0.004	5.4	26	-16	-20
L Hippocampus	0.006	5.3	-14	-28	-8
R anterior MTG	0.007	5.3	56	2	-24
R anterior MTG	0.015	5.1	60	-2	-18
R anterior STS	0.015	5.1	62	-8	-12
R MTG	0.021	5.0	56	-28	-6
R anterior STG	0.023	5.0	50	12	-20
R STG	0.029	4.9	64	-14	-6
L fusiform gyrus	0.013	5.1	-36	-34	-26
L superior lateral cerebellar hemisphere	0.018	5.0	-30	-40	-24
L lingual gyrus	0.016	5.0	-16	-44	0
L IFG (pars triangularis)	0.034	4.9	-54	26	12
R lingual gyrus	0.035	4.8	10	-32	-2
<i>Experiment 2</i>					
L MTG	0.005	5.3	-58	-30	-0
L MTG	0.008	5.2	-64	-22	-2
L MTG	0.014	5.0	-62	-12	-6
R STG	0.007	5.2	66	-8	-8
<i>Experiments 1 and 2 combined</i>					
L MTG	0.001	7.8	-62	-10	-8
L MTG	0.001	7.6	-58	-34	0
L anterior MTG	0.001	7.0	-58	-10	-20
L anterior MTG	0.001	6.3	-54	6	-24
L posterior MTG	0.001	5.3	-48	-62	18
L posterior MTG	0.019	5.3	-58	-58	14
L fusiform gyrus	0.001	6.1	-32	-40	-22
L lingual gyrus	0.001	6.0	-12	-30	-6
L lingual gyrus	0.001	6.0	-14	-46	-2
L collateral sulcus	0.001	6.0	-26	-34	-14
R STS/STG	0.001	7.1	62	-8	-10
R anterior MTG	0.001	6.7	56	2	-24
R anterior STS	0.001	6.3	52	10	-24
L fusiform gyrus	0.001	6.4	-32	-10	-28
R parahippocampal gyrus	0.001	5.9	26	-16	-22
R superior lateral cerebellar hemisphere	0.009	4.9	26	-34	-22
L IFG (pars triangularis)	0.003	5.2	-54	26	10
L IFG (pars triangularis)	0.003	5.2	-42	20	20
R lingual gyrus	0.004	5.1	16	-48	0
R superior colliculus	0.006	5.0	6	-30	-4
L superior temporal pole	0.045	4.5	-44	24	-18
R superior lateral cerebellar hemisphere	0.046	4.5	36	-62	-24

STG, superior temporal gyrus; MTG, middle temporal gyrus; ITG, inferior temporal gyrus; STS, superior temporal sulcus; IFG, inferior frontal gyrus; IFS, inferior frontal sulcus.

regions, confirming the importance of this region in the semantic aspects of speech comprehension. Additionally, the LIFG activation confirms that this region is recruited to select/retrieve relevant semantic information in the context of speech comprehension.

In this experiment, events were timed so as to ensure that scans were maximally sensitive to initial processing of the sentences and insensitive to the relatedness judgements. Together with the absence of performance differences on the task, this makes it unlikely that the observed haemodynamic differences between high- and low-ambiguity sentences are a direct consequence of this task. However, given the role of the LIFG in explicit semantic decision tasks (e.g. Thompson-Schill *et al.*,

Table 4

High-ambiguity versus low-ambiguity sentences: all significant peak activations >8 mm apart ($P < 0.05$ corrected for multiple comparisons). Additional peaks within a single activation cluster are indented, following the most significant peak from each cluster.

	<i>P</i> (corrected)	Z-score	co-ordinates		
			<i>x</i>	<i>y</i>	<i>z</i>
<i>Experiment 1</i>					
1. L posterior ITG	0.016	5.0	-50	-48	-10
2. L IFG (pars triangularis)	0.046	4.7	-52	28	20
3. R IFG (pars triangularis)	0.006	5.2	52	22	12
R IFG (pars triangularis)	0.033	4.8	52	34	14
<i>Experiment 2 (ROI analysis)</i>					
1. L posterior MTG	0.013	4.0	-56	-56	0
L posterior MTG	0.029	3.9	-64	-48	-2
L fusiform	0.035	3.7	-38	-40	-14
2. L posterior IFS	0.031	3.8	-40	12	32
3. R anterior insula	0.025	3.8	32	26	2
<i>Experiments 1 and 2 combined</i>					
L posterior ITG	0.001	5.9	-52	-50	-10
L IFS	0.001	5.5	-42	14	32
L IFG (pars triangularis)	0.001	5.3	-50	30	20
L IFG (pars triangularis)	0.010	4.8	-56	16	22
R opercular IFG	0.006	5.0	36	26	4
R IFG (pars triangularis)	0.016	4.7	50	36	16
L STG	0.048	4.4	-58	-8	-6

1997; Wagner *et al.*, 2001), it is important to confirm that the frontal activations seen in experiment 1 reflect sentence comprehension processes that are recruited even in the absence of an explicit task. We therefore conducted a second experiment using the same sentences, but without the task: participants were told simply to listen carefully to the sentences.

Experiment 2

Materials and Method

Fifteen right-handed volunteers, from the same population as experiment 1 but who had not participated in experiment 1, took part in this experiment. The sentences, scanning procedure and analysis method were identical to that in experiment 1 except that the relatedness probes were removed, and the silent period between scan offset and stimulus onset was shortened by 2 s. This reduced the duration of the experiment while preserving the temporal relationship between each sentence and the subsequent scan. Participants were instructed to lie still and listen carefully to all the sentences. At the end of the experiment, ~15–45 min after initial presentation of the sentences, participants were given a surprise recognition memory test in which they were presented with a printed list of 60 sentences from the experiment (30 ambiguous and 30 unambiguous) along with an equal number of filler sentences that had not been presented in the experiment. Their task was to indicate for each sentence whether they had heard it in the experiment or not. Due to technical difficulties, 2 of the 15 participants only completed two of the three scanning sessions.

Results

Behavioural Results

The recognition-memory test showed above-chance performance in recognizing sentences from the experiment and rejecting novel foil sentences for both the ambiguous [mean $d' = 1.2$, $t(14) = 5.1$, $P < 0.001$] and the unambiguous [mean $d' = 1.1$, $t(14) = 4.8$, $P < .001$] conditions, indicating that participants listened attentively to the sentences. There was no difference between the recognition-memory performance for the two types of sentences [$t(14) = 0.8$, $P > 0.4$]. In subsequent debriefing, only 2 out of 15 participants spontaneously reported noticing any ambiguity in the sentences. Even when the

ambiguity manipulation was explained to them, participants consistently underestimated the proportion of high-ambiguity sentences. This suggests that the presence of ambiguous words within the sentences goes largely undetected. Any additional brain activation in perceiving ambiguous sentences is therefore unlikely to reflect overt awareness of ambiguity.

Imaging Results

As in experiment 1, the contrast between the low-ambiguity speech and noise showed a large area of activation in the left and right temporal cortex, with significant peaks in the superior temporal sulcus bilaterally (see Table 3). The reduced extent of this activation (compared with experiment 1) most likely reflects additional variability in the data resulting from variations in participants' attentional state, since they did not have a task to keep them alert and focused on the sentences. A region of interest (ROI) analysis was then conducted for this contrast, where the ROI is based on the activation for this contrast in experiment 1 (thresholded at $P < 0.001$ uncorrected). This analysis revealed additional peaks in left anterior STS (-62 -18 -4), right STS (50 -30 -2), right anterior STS (56 6 -22), and the left fusiform gyrus (-34 -14 -28).

For the contrast between high-ambiguity sentences and low-ambiguity sentences, no voxels reached significance at the conservative threshold of $P < 0.05$ corrected for multiple comparisons. However, when the power of this contrast was increased by restricting the analysis to a region of interest defined as all the voxels that showed a significant ambiguity effect in experiment 1 (thresholded at $P < 0.001$, uncorrected), the comparison showed significantly increased signal for the high-ambiguity sentences at voxels within all three of the clusters of activation from experiment 1 ($P < 0.05$ corrected for multiple comparisons using small volume correction: Worsley *et al.*, 1996; see Table 4). The Euclidean distances between corresponding peaks in the two experiments (listed in Table 4) range between 14.1 and 24.6 mm. The variation in the precise location of the peak voxels in these contrasts (compared with experiment 1) is likely to be a consequence of the relatively high level of variability in the data in experiment 2. Importantly, the three peak voxels that showed an ambiguity effect in experiment 1 (listed at the top of Table 4) all showed a significant effect of ambiguity in experiment 2 ($Z = 3.1, 2.0, 3.4, P < .05$ in each voxel, see plots in Fig. 2). Furthermore, the magnitude of these differences between the high- and low-ambiguity sentences was similar for the two experiments in these three voxels at least (see Fig. 2).

Experiments 1 and 2: Combined Analysis

The results of the two experiments were compared (modelling subjects as a random effect) for the two critical contrasts (low-ambiguity versus noise and high-ambiguity versus low-ambiguity). These contrasts showed no significant differences between the two experiments in any cortical regions (all $P > 0.1$). This suggests that, as intended, the results of experiment 1 primarily reflect participants' initial processing of the spoken sentences and not their performance of the visual probe task. The absence of differences between the results from the two experiments allows us to combine the data in an analysis that includes all 30 participants (with Experiment as a between-groups factor), in order to increase power and thus sensitivity to subtle condition-

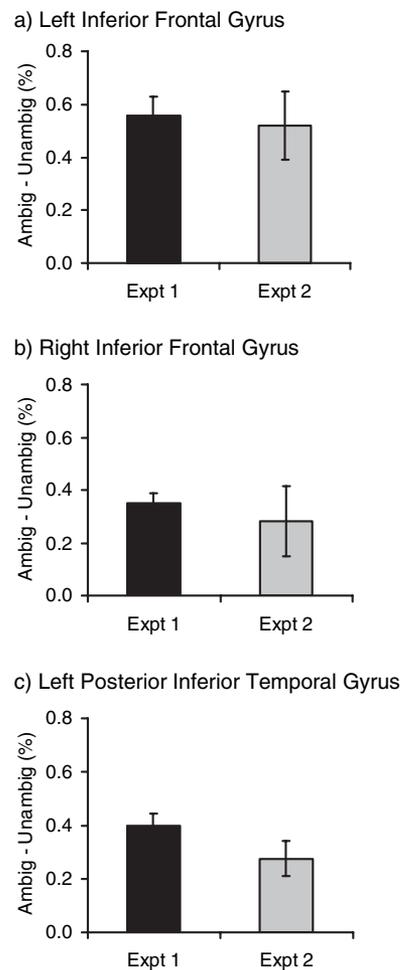


Figure 2. Percent signal change for high-ambiguity versus low-ambiguity sentences in experiments 1 and 2 for peak voxels from experiment 1.

related changes in activity level. It also allows us to identify the locations of activation peaks with greater reliability.³

In this combined analysis, the contrast between the low-ambiguity speech and the baseline noise condition revealed a large area of activation in left and right superior and middle temporal gyri, extending in the left hemisphere into posterior inferior temporal cortex and the fusiform gyrus. Activation was also observed bilaterally in lingual gyrus. We also observed activation in left frontal cortex centred on the dorsal part of the inferior frontal gyrus, extending into the inferior frontal sulcus (shown in blue and yellow in Fig. 3; see Table 3). High-ambiguity sentences produced greater activity than low-ambiguity sentences in the lateral frontal cortex, centred on the inferior frontal sulcus bilaterally but greater in the left than in the right. Greater activity was also observed in the left temporal region including posteriorly in inferior temporal cortex, middle temporal gyrus and superior temporal sulcus (shown in yellow and red in Fig. 3; see Table 4). A single voxel in the mid superior temporal gyrus was also significantly activated.

Of the 1877 voxels that were significantly more active for high-ambiguity compared with low-ambiguity sentences, 1810 (96%) are also significantly more active for the high-ambiguity sentences compared with the noise baseline (whole-brain analysis, $P < 0.05$ uncorrected). Of the 67 voxels that did not

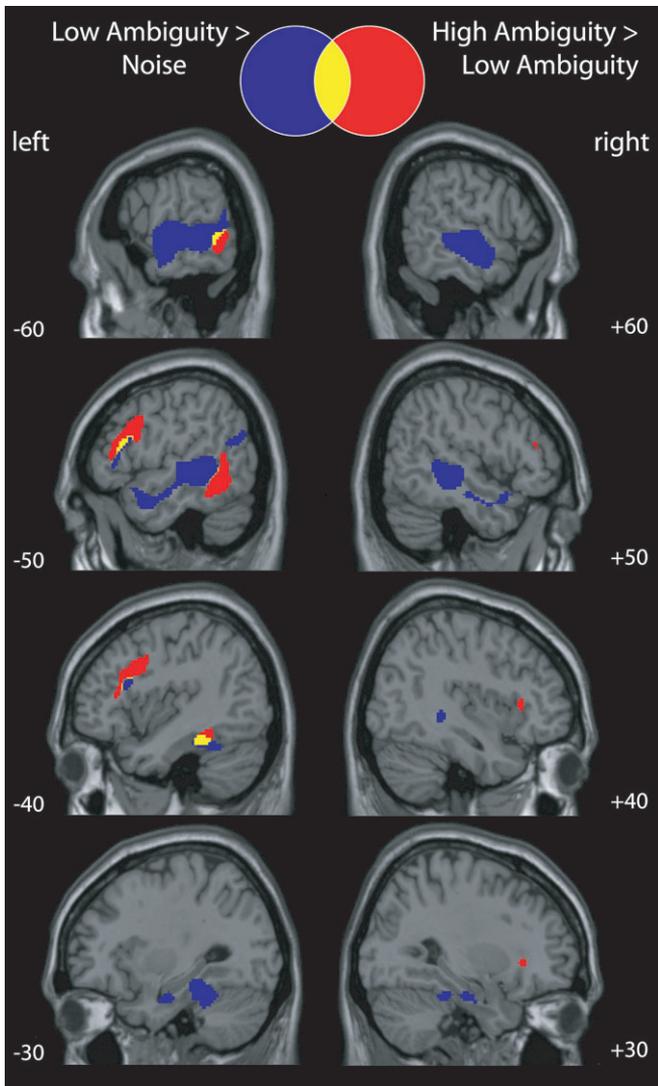


Figure 3. Combined analysis of experiments 1 and 2. Activations shown at $P < 0.05$ corrected for multiple comparisons.

show this effect, all but one are in right frontal cortex. Therefore, while the left hemisphere regions that showed a significant ambiguity effect also showed increased activation for high-ambiguity sentences compared with the noise baseline (at uncorrected levels), only 51% of the right hemisphere voxels that showed an ambiguity effect (66/130 voxels) were more active for high-ambiguity sentences than for the noise baseline.

Discussion

The results of two fMRI experiments show that when volunteers listen to sentences that contain semantically ambiguous words, activity increases in both temporal and frontal brain regions. This confirms the involvement of these regions in the semantic aspects of sentence comprehension (i.e. activating, selecting or integrating word meanings). The most striking aspect of the temporal-lobe ambiguity effect is that it is largely restricted to the inferior and posterior regions that are associated with comprehension deficits in stroke patients (Bates *et al.*, 2003) and not the more anterior temporal regions that have been activated in other imaging studies of sentence

comprehension (Scott *et al.*, 2000; Davis and Johnsruide, 2003) or that appear to be associated with semantic dementia (Mummery *et al.*, 2000). Therefore, our results are consistent with models in which *posterior* inferior temporal regions play an important role in the semantic aspects of speech comprehension (Hickok and Poeppel, 2000, 2004).

Although the inferior temporal gyrus (ITG) is often associated with visual processing, activations in left posterior ITG have been reported in previous studies of speech perception. For example, Binder *et al.* (2000) found greater activation for spoken words compared with pseudowords. In addition, a meta-analysis of studies involving both linguistic and non-linguistic stimuli presented in visual, auditory and tactile (Braille) modalities (Price, 2000) identified a region of posterior inferior temporal cortex (together with posterior inferior frontal cortex) that was significantly activated in all experiments involving word retrieval, independent of modality. Increasing evidence therefore indicates that this posterior inferior region (extending into the posterior middle temporal gyrus) is involved in semantic processing across a range of input modalities (including speech). Our activation cluster encompasses a large region that is, as yet, poorly anatomically defined, and is likely to incorporate several anatomically differentiable areas subserving different functions. Nevertheless our data are consistent with the view that structures within this region play a role in establishing the meaning of spoken utterances and in other, related, semantic processes. The region of temporal cortex that shows an effect of semantic ambiguity (shown in red/yellow in Fig. 3) lies posterior and inferior to the area of STS/STG where activity was seen for low-ambiguity speech, compared with unintelligible noise (shown in blue/yellow in Fig. 3). This pattern of results is consistent with the view that processing pathways radiate out from auditory cortex, and that successively more distant areas of the temporal lobe may be recruited for higher levels of processing (Binder *et al.*, 2000; Davis and Johnsruide, 2003; Scott and Johnsruide, 2003), and that the pathway that extends posteriorly and inferiorly plays a critical role in meaning-based comprehension processes (Hickok and Poeppel, 2000, 2004).

Increased activation for high-ambiguity sentences was also seen in an extensive area of left inferior frontal cortex (Fig. 3). Within this cluster, there are significant peaks within the LIFG pars triangularis (BA 45) (Amunts *et al.*, 1999) and more posterior portions of the left inferior frontal sulcus (LIFS), extending onto the surface of the middle frontal gyrus (BA 9/46, 8) (Petrides and Pandya, 1999). This cluster (shown in red/yellow in Fig. 3) lies dorsal to, but overlaps with, the region that was significantly more active for the low-ambiguity speech compared with unintelligible noise (shown in blue/yellow in Fig. 3). The presence of this frontal ambiguity effect in the absence of an explicit task (experiment 2, see Fig. 2) confirms the involvement of frontal brain regions in speech comprehension, and is problematic for the view that frontal brain regions are not ordinarily recruited for the comprehension of sentences that are easily and automatically understood (Crinion *et al.*, 2003). The high- and low-ambiguity sentences in this study were rated as being equally natural and the ambiguity went largely unnoticed.

The role of the LIFG in semantic processing is more commonly discussed in the context of explicit semantic judgement tasks using single words (e.g. Thompson-Schill *et al.*, 1997; Wagner *et al.*, 2001). Thompson-Schill *et al.* (1997) argue, on the basis of a number of fMRI studies using

different semantic judgement tasks, that the LIFG is involved in selecting task-relevant semantic information from competing representations. In all three of the experiments that reported by Thompson-Schill *et al.* (1997), the peak LIFG activation is in a voxel that is also significantly activated in this study ($P < 0.05$ corrected). In a related proposal, Wagner *et al.* (2001) suggest that this region guides controlled semantic retrieval by providing top-down signals to guide access to semantic knowledge. Again there is substantial overlap between the activations reported in this study (in particular those peaks described as posterior LIPC) and the activations reported here. This striking degree of overlap across these three very different studies suggests that the frontal activations seen in this study may reflect (in part) the operation of a semantic selection/retrieval mechanism, which operates both in the context of explicit semantic judgement tasks and in natural speech comprehension. Under this view, the neural mechanism that allows particular semantic attributes to be selected/retrieved in order to make a semantic judgment about the meaning of a word may also be involved in selecting the appropriate semantic attributes (i.e. meaning) of an ambiguous word on the basis of the semantic properties of the rest of a sentence. It is also interesting to note that in both these explicit semantic selection/retrieval studies (Thompson-Schill *et al.*, 1997; Wagner *et al.*, 2001) the LIFG activation is accompanied by activation in posterior inferior temporal cortex. These regions of cortex are anatomically distant but probably well connected (e.g. Webster *et al.*, 1994; Miller, 2000), and appear to form part of a semantic processing network.

Increased LIFG activation has also been reported in studies that contrast syntactically complex sentences with syntactically simple sentences (e.g. Just *et al.*, 1996; see Kaan and Swaab, 2002, for a review). The precise location of these syntactic complexity activations is variable (Kaan and Swaab, 2002), but there is considerable overlap between these activations (primarily BA 44/45) and the LIFG activation reported here. These syntactic complexity effects have been used by some authors to support strong claims that the posterior portion of the LIFG (BA 44/45) is specifically involved in computing the syntactic structure of sentences (Friederici, 2002), a view that seems incompatible with our finding that that activation in this region increases as a result of a semantic variable. In contrast, other authors have claimed that the LIFG plays a more general role in sentence processing. For example, Kaan and Swaab (2002) suggest that it has a working memory function, and that it is required to temporarily store information that has not yet been successfully integrated into the ongoing representation of the utterance. Similarly, Muller *et al.* (2003) suggest that the involvement of this region in processing syntactically complex sentences and making explicit semantic judgements suggests that its function is to hold lexical representations in working memory. These views are compatible with our results, although it is as yet unclear whether the function of this region should best be characterized in terms of working memory, or in terms of the selection/retrieval mechanisms discussed earlier.

Significantly elevated activation for ambiguous sentences, although to a much reduced degree, was also observed in right IFG. However this region did not show any significant activation in the contrast between high-ambiguity sentences and the baseline noise condition at a corrected level of significance, and even at uncorrected level only 51% of the voxels were more active for high-ambiguity sentences than for noise. This activity

in response to high-ambiguity sentences may simply reflect interhemispheric, callosal, connections between homologous regions. However, right IFG activation for ambiguous sentences is also consistent with the results of lateral presentation studies which suggest that both hemispheres are important in resolving semantic ambiguities (Burgess and Simpson, 1988; Faust and Chiarello, 1998). Further work is needed to investigate the different role that the two hemispheres might play in ambiguity resolution [and whether the lateralized presentation paradigm does indeed preferentially recruit contralateral processes (Cohen *et al.*, 2002)].

The results of our study demonstrate that a key aspect of spoken language comprehension — the resolution of semantic ambiguity — can be used to identify the brain regions involved in the semantic aspects of speech comprehension (e.g. activating, selecting and integrating word meanings). It is striking that the comparison between these two sets of sentences that were rated as equally natural, and which differ only on a relatively subtle linguistic manipulation should produce such extensive activation differences. Further work is clearly needed to determine the precise function of these regions, but it is clear that they must form an important part of a neural pathway that processes the meaning of spoken language. These results support models of speech comprehension in which posterior inferior temporal regions are involved in semantic processing (Hickok and Poeppel, 2000, 2004), and they demonstrate that the LIFG, which has long been known to be important in syntactic processing of sentences and the semantic properties of single words, also plays an important role in processing the meanings of words in sentences.

Notes

1. In this study no distinction was made between ambiguity among unrelated word meanings and related word senses (Rodd *et al.*, 2002). Ambiguous words of both types were used. In all cases, the ambiguous words had an inappropriate meaning that was in the same syntactic class as the correct interpretation and therefore needed to be disambiguated on the basis of semantic, not syntactic, cues. For example, in the above sentence the word 'shell' is ambiguous between multiple noun meanings. In addition, as with many words in English, some of the ambiguous words (and their unambiguous controls) could also be used in different syntactic classes (e.g. 'shell' can also be used as a verb).
2. To ensure that the sentences could be closely matched, two low-ambiguity sentences were initially constructed for every high-ambiguity sentence. After recording and pretesting, the sentence that most closely matched the high-ambiguity sentence for duration, naturalness and imageability was selected. No manipulation of the sound files was conducted.
3. Although the absence of significant differences between the two experiments can be used to justify a combined analysis, it should be noted that, in isolation, this result does not demonstrate that there are no differences between the neural systems that are recruited for speech comprehension in the presence or absence of an explicit task. It is possible that these null results reflect the limited power of the between-group comparisons, and larger group sizes may be necessary to provide the necessary statistical power to observe subtle effects of the task conditions on speech comprehension. However, this issue does not critically affect the interpretation of our results. The plots shown in Figure 2 provide a clear demonstration that, for all the brain regions for which we make theoretical claims, the effect of ambiguity is present for both experiments.

The authors thank the staff at the Wolfson Brain Imaging Unit and Alexis Hervais-Adelman for their help with the data acquisition and processing. This work was funded by the UK Medical Research Council, and by fellowships to J.M.R. from Peterhouse, Cambridge and the Leverhulme

Trust. Much of this work was carried out when J.M.R. was based at the Centre for Speech and Language, University of Cambridge, which is funded by a UK Medical Council programme grant to Professor Lorraine Tyler. I.S.J. is currently funded by the Canada Research Chairs Program.

Correspondence should be sent to Jennifer Rodd, Department of Psychology, University College London, 26 Bedford Way, London WC1H 0AP, UK. Email: j.rodd@ucl.ac.uk.

Appendix

Example sentences (randomly selected subset)

High-ambiguity	Low-ambiguity
she saw a <i>hare/hair</i> while she was <i>skipping</i> across the field	he met his father while he was walking to the shops
the <i>panel</i> were supposed to ignore the <i>race</i> and sex of the contestants	the woman was hoping to discover the name and address of the culprit
the <i>steak/stake</i> was <i>rare</i> just as the customer had requested	the goat was greedy just as the family had expected
the <i>blind</i> on the window kept out the <i>sun/son</i>	the wife of the priest helped out the elderly
the <i>lock</i> on the <i>chest</i> had been broken with the <i>poker</i>	the picture on the wall had been chosen by his girlfriend
they kept a <i>record</i> of the events in the <i>log</i>	they told the truth about the fight to the teacher
the <i>cymbals/symbols</i> were making a <i>racket/racquet</i>	the king was making many enemies
there was <i>thyme/time</i> and <i>sage</i> in the stuffing	there was milk and sugar in his coffee
the <i>creak/creek</i> came from a <i>beam</i> in the <i>ceiling</i>	the woman laughed at the joke about the dog
his <i>calf</i> was only <i>strained</i> and would heal quickly	their holiday was quite short and would end soon

References

- Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K (1999) Broca's region revisited: cytoarchitecture and intersubject variability. *J Comp Neurol* 412:319-341.
- Baayen RH, Piepenbrock R, Gulikers L (1995) The CELEX Lexical Database (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF (2003) Voxel-based lesion-symptom mapping. *Nat Neurosci* 6:448-450.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET (2000) Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512-528.
- Brett M, Leff AP, Rorden C, Ashburner J (2001) Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14:486-500.
- Burgess C, Simpson GB (1988) Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain Lang* 33:86-103.
- Cohen L, Lehericy S, Chochon F, Lemer C, Rivaud S, Dehaene S (2002) Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain* 125:1054-1069.
- Crinion JT, Lambon Ralph MA, Warburton EA, Howard D, Wise RJS (2003) Temporal lobe regions engaged during normal speech comprehension. *Brain* 126:1193-1201.
- Davis MH, Johnsrude IS (2003) Hierarchical processing in spoken language comprehension. *J Neurosci* 23:3423-3431.
- Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ (2004) Lesion analysis of the brain areas involved in language comprehension. *Cognition* 92:145-177.
- Faust M, Chiarello C (1998) Constraints on sentence priming in the cerebral hemispheres: effects of intervening words in sentences and lists. *Brain Lang* 63:219-236.
- Forster KI, Forster JC (2003) A Windows display program with millisecond accuracy. *Behav Res Methods Instrum Comput* 35:116-124.
- Friederici AD (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78-84.
- Friston KJ, Holmes AP, Worsley KJ (1999) How many subjects constitute a study? *Neuroimage* 10:1-5.
- Gernsbacher MA, Faust ME (1991) The mechanism of suppression: a component of general comprehension skill. *J Exp Psychol Learn Mem Cogn* 17:245-262.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW (1999) 'Sparse' temporal sampling in auditory fMRI. *Hum Brain Mapp* 7:213-223.
- Hall DA, Summerfield AQ, Goncalves MS, Foster JR, Palmer AR, Bowtell RW (2000) Time-course of the auditory BOLD response to scanner noise. *Magn Reson Med* 43:601-606.
- Hickok G, Poeppel D (2000) Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci* 4:131-138.
- Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92:67-99.
- Hodges JR, Patterson K, Oxbury S, Funnell E (1992) Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain* 115:1783.
- Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996) Brain activation modulated by sentence comprehension. *Science* 274:114-116.
- Kaan E, Swaab TY (2002) The brain circuitry of syntactic comprehension. *Trends Cogn Sci* 6:350-356.
- Kaas JH, Hackett TA (2000) Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci USA* 97:11793-11799.
- Miller EK (2000) The prefrontal cortex and cognitive control. *Nat Rev Neurosci* 1:59-65.
- Muller RA, Kleinhans N, Courchesne E (2003) Linguistic theory and neuroimaging evidence: an fMRI study of Broca's area in lexical semantics. *Neuropsychologia* 41:1199-1207.
- Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RS, Hodges JR (2000) A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. *Ann Neurol* 47:36-45.
- Narain C, Scott SK, Wise RJ, Rosen S, Leff A, Iversen SD, Matthews PM (2003) Defining a left-lateralized response specific to intelligible speech using fMRI. *Cereb Cortex* 13:1362-1368.
- Palmer AR, Bullock DC, Chambers JD (1998) A highoutput, high-quality sound system for use in auditory fMRI. *Neuroimage* 7:S359.
- Parks R, Ray J, Bland S (1998) *Wordsmyth English Dictionary-Thesaurus*. University of Chicago, <http://www.wordsmyth.net/>, University of Chicago.
- Petrides M, Pandya DN (1999) Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur J Neurosci* 11:1011-1036.
- Price CJ (2000) The anatomy of language: contributions from functional neuroimaging. *J Anat* 197:335-359.
- Rayner K, Duffy SA (1986) Lexical complexity and fixation times in reading — effects of word-frequency, verb complexity, and lexical ambiguity. *Mem Cognit* 14:191-20.
- Rodd JM, Gaskell MG, Marslen-Wilson WD (2002) Making sense of semantic ambiguity: semantic competition in lexical access. *J Mem Lang* 46:245-266.
- Rosen S (1992) Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci* 336:367-373.
- Schroeder MR (1968) Reference signal for signal quality studies. *J Acoust Soc Am* 44:1735-1736.
- Scott SK, Johnsrude IS (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26:100-107.
- Scott SK, Blank CC, Rosen S, Wise RJ (2000) Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123:2400-2406.
- Swinney DA (1979) Lexical access during sentence comprehension: (re)-consideration of context effects. *J Verb Learn Verb Behav* 18:645-659.
- Thompson-Schill SL (2003) Neuroimaging studies of semantic memory: inferring 'how' from 'where'. *Neuropsychologia* 41:280-292.
- Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ (1997) Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci USA* 94:14792-14797.
- Ullman MT (2001) A neurocognitive perspective on language: the declarative/procedural model. *Nat Rev Neurosci* 2:717-726.

- Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA (2001) Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31:329-38.
- Webster MJ, Bachevalier J, Ungerleider LG (1994) Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb Cortex* 4:470-483.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58-73.
- Xiong J, Rao S, Jerabek P, Zamarripa F, Woldorff M, Lancaster J, Fox PT (2000) Intersubject variability in cortical activations during a complex language task. *Neuroimage* 12:326-339.