Listening to Narrative Speech after Aphasic Stroke: the Role of the Left Anterior Temporal Lobe

The dorsal bank of the primate superior temporal sulcus (STS) is a polysensory area with rich connections to unimodal sensory association cortices. These include auditory projections that process complex acoustic information, including conspecific vocalizations. We investigated whether an extensive left posterior temporal (Wernicke’s area) lesion, which included destruction of early auditory cortex, may contribute to impaired spoken narrative comprehension as a consequence of reduced function in the anterior STS, a region not included within the boundary of infarction. Listening to narratives in normal subjects activated the posterior--anterior extent of the left STS, as far forward as the temporal pole. The presence of a Wernicke’s area lesion was associated with both impaired sentence comprehension and a reduced physiological response to heard narratives in the intact anterior left STS when compared to aphasics with persisting temporal lobe damage and normal controls. Thus, in addition to the loss of language function in left posterior temporal cortex as the direct result of infarction, posterior ablation that includes primary and early association auditory cortex impairs language function in the intact anterior left temporal lobe. The implication is that clinical studies of language on stroke patients have underestimated the role of left anterior temporal cortex in comprehension of narrative speech.

Keywords: fusiform gyrus, narrative speech comprehension, superior temporal sulcus

Introduction

Infarction of Wernicke’s area, the posterior left superior temporal cortex, impairs speech comprehension. If the lesion is large, the prognosis for recovery is poor (Selnes et al., 1983; Naeser et al., 1987; Hart and Gordon, 1990). The inference is that the unimodal processing of speech proceeds in a ventral and posterior direction prior to processing for meaning (Hickok and Poeppel, 2000, 2004). A similar conclusion comes from studies of transcortical sensory aphasia (TCSA) after stroke. Although these patients can still process the speech signal to allow accurate repetition, the route to word comprehension is impaired. Typically, but not always, the condition is associated with lesions to, or disconnection of, inferolateral temporal and inferior parietal cortex (Berthier, 1999). Recovery of speech comprehension has been equated with restitution of function within or around the lesion following Wernicke’s area reversible ischemia (Hillis et al., 2001) or infarction (Heiss et al., 1999).

However, the literature on lesion location and aphasia is selective. Infarction confined to anterior temporal cortex is rare, as the artery supplying this region often arises proximally, protecting it from emboli lodging at the more distal trifurcation of the middle cerebral artery. Anterior left temporal lobectomy to treat temporal lobe epilepsy is usually reported to impair language comprehension only mildly, but the effects of recurrent left temporal lobe seizures on normal language organization limits the inferences that can be made (Janzsky et al., 2003). In contrast, patients with the temporal variant of frontotemporal dementia commonly known as semantic dementia (Snowden et al., 1989, 1992; Hodges et al., 1992; Edwards-Lee et al., 1997; Gorno-Tempini et al., 2004), have an anterior-posterior gradient of temporal lobe atrophy. These patients gradually lose single word comprehension although syntactic comprehension, within the limits imposed by the loss of knowledge about single word meaning, is usually well preserved (see e.g. Gorno-Tempini et al., 2004). Their atrophy is most marked at the temporal pole and within both inferotemporal cortex and medial temporal lobe (limbic) structures (Chan et al., 2001; Galton et al., 2001).

In the non-human primate, it has been proposed that anterior and posterior projections from early auditory association cortex have different functions (Rauschecker and Tian, 2000; Kaas and Hackett, 2000) — an hypothesis supported by evidence that posterior and anterior auditory processing ‘streams’ project to different prefrontal targets (Romanski et al., 1999). Further, neurons tuned to conspecific vocalizations are encountered more often in primate anterior auditory association cortex (Tian et al., 2001), sounds that also result in a lateralized difference in the rate of glucose metabolism in temporal polar cortex (Poremba et al., 2004). In neuroimaging studies of human auditory perception, acoustic features that contribute to discriminating speech sounds activate auditory association cortex in both anterior and posterior directions (Scott and Johnsrude, 2003). Sounds made by the human vocal tract, whether intelligible or unintelligible, increase activity as far as the dorsal bank of both left and right superior temporal sulci (STS) (Belin et al., 2000). Listening to intelligible sentences after subtraction of activity attributable to acoustic complexity and paralinguistic features has been specifically related to activity in the left STS, including its anterior extent (Scott et al., 2000; Narain et al., 2003).

In non-human primates, auditory projections from the subcortical medial geniculate complex are distributed widely throughout the ipsilateral superior temporal gyrus. There are major parallel projections to primary auditory cortex and adjacent periauditory cortex, known as belt and parabelt cortex (Kaas and Hackett, 2000). Onward projections pass in series from primary auditory cortex to belt and from belt to parabelt. Parallel projections from belt and parabelt pass to other regions in unimodal auditory association cortex, to polysensory (heteromodal) cortex and to prefrontal cortex. This hierarchical cortical organization applies to other sensory
The manner whereby encoded sensory information passes progressively through a number of synaptic levels in unimodal association cortex to polysensory cortex, to paralimbic and limbic structures, and to prefrontal cortex has been proposed as an anatomical account of how language, spoken or written, is processed (Mesulam, 1998).

We reasoned that infarction of primary and adjacent associative auditory cortex as part of Wernicke’s area infarction, by disrupting many of the subcortical-cortical auditory projections, impairs the outward polysynaptic transmission of encoded speech information to structurally intact regions in auditory association and polysensory temporal lobe neocortex. Further, if the infarct destroys the posterior part of auditory unimodal association and polysensory cortex, the additional loss of posterior–anterior connections along the lateral temporal neocortex, the additional loss of posterior–anterior connections along the lateral temporal neocortex as, for example, have been demonstrated in the STS of the non-human primate (see e.g. Seltzer and Pandya, 1989; Padberg et al., 2003) may contribute to the behavioral consequences of a Wernicke’s area infarct. This would accord with the many functional imaging studies on normal subjects that have shown anterior temporal responses to spoken or written single words, sentences and narratives (e.g. Mazoyer et al., 1993; Bavelier et al., 1998; Perani et al., 1998; Schlosser et al., 1998; Stowe et al., 1998; Mummery et al., 1999; Binder et al., 2000; Scott et al., 2000; Ferstl and von Cramon, 2001; Wise et al., 2001; Vandenbergh et al., 2002; Davis and Johnsrude, 2003; Friederici et al., 2003; Narain et al., 2003; Scott and Wise, 2004).

Particularly compelling is a recent magnetoencephalographic (MEG) study on normal subjects, which combined excellent temporal resolution with good spatial resolution (Marinkovic et al., 2003). An explicit task based on knowledge about object size in response to hearing or reading object nouns resulted in activity spreading forwards from early auditory cortex (to heard words) and early visual cortex (for seen words). A strong response to both modalities of word perception was observed several hundreds of milliseconds later in the anterior STS and temporal pole, left greater than right. Thus, the ability in this study to follow cortical responses over time gave clear evidence of a posterior–anterior response along the temporal lobe during the processing of words for meaning. There was additional activity in the left inferior frontal gyrus and medial prefrontal cortex, which we have previously related to an explicit semantic task demand (Scott et al., 2003; Sharp et al., 2004). In response to ‘passive’ (i.e. automatic) processing of speech, after ‘subtraction’ of the signal related to early prelexical acoustic processing by using a baseline of spectrally rotated or reversed speech, we have also observed activity at anterior extent of the left temporal neocortex, centered on the STS (Scott et al., 2000; Crinion et al., 2003).

To further confirm that one route for the processing of heard speech involves projections to polysensory cortex in the anterior STS, we recruited aphasic stroke patients and grouped them according to specific sites of pathological ablation. We used positron emission tomography (PET) to investigate whether Wernicke’s area infarction reduced normal anterior left temporal activation, even though the lesion had not directly involved this area of cortex. We further investigated whether relative activity within this area correlated with a measure of the patients’ residual spoken language comprehension, using the data both from the patients with Wernicke’s area infarction and from patients with aphasic stroke that spared the left temporal lobe.

The motivation was to reconcile those anatomical models of language processing, based predominantly on lesion-deficit analyses of aphasic stroke patients, that depict speech mapping onto to meaning via projections to the left posterior temporal and inferior parietal cortex and those, based on neuropsychological and anatomical imaging studies of patients with semantic dementia, that emphasize the role of anterior temporal cortex.

Materials and Methods

Subjects

Eleven normal subjects (nine males, aged 37–76 years) and 24 patients (18 males, aged 32–85 years) were studied. All were right-handed and had English as their first language. Each gave informed consent to participate in the study. The local ethics committee approved the project. Permission to administer radioisotopes was given by the Department of Health, UK. The results from the normal subjects have been previously reported (Crinion et al., 2003). The patients became aphasic following a single left hemisphere stroke. The time from the onset of stroke was very variable, between 2 and 204 months (mean 32 months). Each had a volumetric MRI to co-register with the PET scan and to identify the site of their lesion. A neurologist (RJSW) classified each patient according to whether the infarct involved the left temporal lobe and, if it did, the posterior–anterior distribution of the temporal lobe damage. At the time of classification, he was not aware of the behavioral and PET data that corresponded to the individual MR images. There were six patients with a left hemisphere lesion sparing the left temporal lobe (T- group), as judged from the T1-weighted volumetric images. The remaining 18 patients had temporal lobe damage (T+ group). Of these, the anterior lateral temporal cortex was preserved in nine patients but the volume of their infarcts was judged to include part or all of primary and adjacent association auditory cortex in the supratemporal plane (Wernicke’s group), based on the probabilistic atlas of Rademacher et al. (2001) as reference. In this atlas, the anterior–posterior extent of primary auditory cortex in the supratemporal plane is from \( y = -4 \) mm at its lateral border to \( y = -35 \) mm at its medial border, using the stereotactic coordinate system of Talairach and Tournoux (1998). All nine patients had the left STS intact anterior to \( y \) coordinate \(-20 \) mm; the STS extends from \( y = -60 \) mm at its posterior end to \( y = 15 \) mm, approximately. The distribution of infarction in the sub-groups of patients is shown in Figure 1.

Behavioral Assessment

Twenty-two aphasic patients underwent neuropsychological testing using a range of tests. For the analysis of the PET imaging results, the results from the Comprehensive Aphasia Test (CAT) were used (Swinburn et al., 2004). This includes a range of language comprehension and production tests. One patient in the T+ group died and another in the T- group had a second left hemisphere stroke before behavioral testing was completed.

Scanning Stimuli

All patients had at least some preservation of spoken single word comprehension, and most had at least some preservation of spoken sentence comprehension. To maximize comprehension by the patients, we chose simple children’s stories designed for a normal reading age of 4–6 years, obtained from the Oxford Reading Tree Scheme (www.oup.co.uk/oxed/primary/or/): nouns of high frequency and imageability are incorporated within simple sentence structures. Using SoundEdit16\textsuperscript{TM} computer software the waveform for each story was also reversed (RSt), which destroys intelligibility while retaining the overall spectrotemporal complexity of the original speech signal. The reversed versions of the narratives were expected to control for much of the prelexical acoustic processing of the speech signal in both left and right superior temporal cortex. Each version of St and RSt were used only once during each subject’s study, across 12–16 scans. The order of presentation was randomized, both within and between subjects. The subjects were asked to simply listen and try to understand the St.
They were aware that the RSt were unintelligible but were asked to pay attention to the sound.

**PET Scanning**

All normal subjects and 16 aphasic patients were studied at the MRC Clinical Sciences Centre, Hammersmith Hospital using a CTI-Siemens ECAT EXACT HR++/966 PET scanner operated in high-sensitivity 3D mode. The remaining eight aphasic patients were studied at the Wolfson Brain Imaging Centre, Addenbrooke’s Hospital using the General Electric Medical Systems (Milwaukee, WI) Advance PET Scanner in 3D mode. $^{15}$O, administered i.v., was used to estimate regional cerebral blood flow (rCBF) while the subjects either heard St or RSt. The stories, and thus the reversed stories, were delivered at a normal rate of speaking (~120–150 words/min), with 5–16 words per sentence. The most sensitive period of each scan is during the rapid uptake of tracer into cerebral tissue, which begins ~10–15 s after the head counts begin to rise and continues for about 20 s. The stimuli were begun ~10 s before the head counts began to rise and continued throughout the period of data acquisition during each scan. Measured attenuation correction was applied to all emission scans.

**Analyses of PET data**

The Statistical Parametric Mapping (SPM99) software package (Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK; www.filion.ucl.ac.uk/spm/) was used for the analysis of the PET data in this study. The SPM approach is voxel based: images are realigned, spatially normalized into a standard Montreal Neurological Institute (MNI) stereotactic space, and smoothed using an isotropic 10 mm full-width, half-maximum Gaussian kernel to account for variation in gyral anatomy and to improve the signal-to-noise ratio.

With stroke patients, distortions of normal anatomy ipsilateral to the volume loss accompanying mature infarcts may reduce the certainty that identical brain regions are being sampled across all subjects, despite image smoothing. To further minimize the error the spatial normalization of the patients’ scans used the cost function masking method, in which a hand-drawn mask of the infarct (derived from the MRI scan) was used, preventing the normalization algorithm from defining the edge of the infarct as part of the brain surface (Brett et al., 2001).

Within SPM, parametric statistical models are assumed at each voxel, using the General Linear Model GLM to describe the data in terms of experimental and confounding effects, and residual variability. Classical
statistical inference is used to test hypotheses that are expressed in terms of GLM parameters. This uses an image whose voxel values are statistics, a Statistic Image, or Statistical Parametric Map (SPM(Z), SPM(F)). For such classical inferences, the multiple comparisons problem is addressed using continuous random field theory RFT, assuming the statistic image to be a good lattice representation of an underlying continuous stationary random field. This results in inference based on corrected P-values.

In all SPM analyses, unless otherwise stated, a blocked AnCova with global counts as confound was used to remove the effect of global changes in perfusion across scans in each subject. The type of scanner on which the subject was studied was also entered as confound. The statistical threshold for the imaging analyses was set at $P < 0.05$, corrected for whole brain volume, using the correction for familywise error rate (FWE), which, although conservative, is reliable when both detecting an experimentally induced signal and locating this effect in the brain (Peterson et al., 1999).

**Normal Subjects**
The normal subjects were entered into a single-group, multi-subject design matrix to investigate the contrast of St with RSt, using a fixed effects statistical model to initially investigate the results in this group alone.

**Comparing Patients and Normal Subjects**
Random-effect analyses in SPM99 which emulate full mixed-effects analyses were used to investigate any differences between the normal and patient groups, T-- and T+ groups. This involved taking the contrasts of parameters estimated from the first-level (fixed-effect) analysis ST--RSt and entering them into a second-level (random-effect) analysis. This ensured that there is only one observation (i.e. contrast) per subject in the second-level analysis and that the error variance was computed using the subject to subject variability of estimates from the first level. The second-level design matrix simply tests the null hypothesis that the contrasts are zero (implementing a single-sample t-test). Random-effects analyses are usually more conservative but allow the inference to be generalized to the T-- and T+ populations from which the patients were selected.

**ROI Analysis**
As we had an a priori hypothesis about the role of the left anterolateral temporal lobe in speech comprehension, a region of interest (ROI) analysis, using the MARSBAR toolbox within SPM99, was performed on all subjects who did not have infarction there, i.e. the normal, T-- and Wernicke’s groups. The ROI was defined using the binary activation image obtained from the contrast of St with RSt in the normal group alone. To ensure that that any changes observed in this region were not just part of a general reorganization of the language system, two additional ROI were defined, using the same technique, for the left ventral and right anterolateral temporal regions. The mean voxel values for individual patient’s ROIs, from the contrast of St with RSt were extracted. First, these ROI values from the T-- and Wernicke’s groups were correlated with the sentence comprehension data from the CAT using a Pearson correlation. Second, the ROI values from the T--, Wernicke’s and normal control groups were contrasted using independent samples t-test, equal variances not assumed.

Thus, the ROI was an irregularly shaped region extracted from the functional data on normal subjects and not based on anatomical landmarks. It transpired that its anatomical distribution was centered on the most anterior extent of the STS (the peak voxel lay at Talarach and Tournoux stereotactic coordinates --52, 8, -20 mm), and encompassed a volume of -12 cc within its boundary.

**Results**

**Behavioral Data**
Overall, spoken and written single word-picture matching were better than spoken and written sentence–picture matching in the patients (Fig. 2). Although heard single word–picture matching was below normal in many patients, their scores all remained well above chance; by contrast, heard sentence–picture matching was more variable, being at chance in one patient (7/30) and close to chance in a further three (10--12/30). As can be inferred from Figure 2, the T-- group was significantly better than the T+ group on the CAT tests of both auditory and written sentence comprehension (unpaired Student’s t-test, $P < 0.01$). Across all patients, heard single word comprehension correlated with heard sentence comprehension. Otherwise, measures of picture naming, single word repetition and semantic fluency (free retrieval of animal names in one minute) was very variable, both within groups and across all patients (Fig. 2).

**PET Data – Normal Subjects**
In the normal group, St contrasted with RSt (Fig. 3) demonstrated activation of the left temporal cortex, with three voxel clusters: in the anterior STS, extending into the temporal pole; in the posterior STS as far as the temporo-parietal junction; and in the anterior left fusiform gyrus. In the right temporal cortex
PET Data – Comparison of Patients and Normal Subjects

There were no significant differences between the patients, as one group, and the normal subjects within the voxels that were preserved in all patients. Thus, activations common to the two groups were observed in the anterior left fusiform gyrus and the anterior right STS. When the patients were analysed separately as T– and T+ groups, again there were no significant differences between either of them and the normal group. The pattern of activation in the T– group was identical to that in the normal group, and there were no differences between the T+ and normal groups within the voxels that were preserved in all patients (Fig. 3).

PET Data – ROI Analysis

The brain tissue within the anterior left temporal ROI, encompassing the anterior third of the STS and extending into the lateral temporal pole, was intact in the T– and Wernicke’s groups. Within this ROI, activity was identified as the percentage increase in activity in response to hearing speech relative to the unintelligible baseline auditory condition. Behavioral data was available for only 5/6 patients in the T– group. In one of these five patients whose performance on tests of single word and sentence comprehension were within the normal range, activity within the ROI was identified as an outlier within SPSS (> 3 SD above the mean for the group). His data were therefore excluded from the analysis. Across the 13 patients included in the analysis, ROI activity correlated significantly with the scores on heard sentence comprehension ($r = 0.57$, $P < 0.05$; Pearson’s correlation). Including the outlying data point from the fourteenth subject improved the correlation coefficient a little ($r = 0.62$), and so excluding the data from this one subject in the final analysis did not bias the result. Importantly, there was no correlation between activity with the ROI and time since the onset of stroke ($r = 0.06$, $P > 0.8$), and so the correlation cannot simply be attributed to recovery over time after acute focal brain injury. Inspection of Figure 4 demonstrates that the four patients from the T– group had both high ROI activity and good sentence comprehension relative to the nine patients from the Wernicke’s group. This was confirmed with independent $t$-tests (mean ROI activity ± SEM: 6.7 ± 1.1 and 2.7 ± 0.8, $P < 0.05$; mean auditory sentence comprehension ± SEM: 25.5 ± 2.2 and 14.9 ± 1.8, $P < 0.01$), although with data from only four patients in one group these results can only be considered reliable for this study group. ROI activity within the normal group was 6.3 ± 1.4, significantly greater than the activity in the Wernicke’s group ($P < 0.05$). By contrast, ROI activity did not correlate significantly with the scores on heard single word comprehension.
The study demonstrated that activity in the intact anterior left STS in a contrast of listening to narratives with listening to reversed narratives was reduced in patients with Wernicke’s area infarction. Expanding the group of aphasics to include a number without left temporal lobe infarction demonstrated that left anterior STS activity correlated with a measure of sentence comprehension.

The activity we observed was dependent on a simple ‘subtractive’ contrast, using auditory stimuli that were processed automatically by the subjects. The contrast ‘subtracted’ early processing of the speech signal, evidenced by the absence of signal in primary auditory cortex and the supratemporal plane; these regions are also activated by sounds that are acoustically less complex than speech (Scott and Johnsrude, 2003). The observed signal reliably reflected ‘bottom-up’ implicit processing of intelligible sentences, connected by a narrative plot. In reality, the observed signal will have reflected less than this. One pervasive problem in functional imaging is that there is no universal baseline state of cognitive ‘neutrality’ with which to compare the activation conditions under investigation. This has been highlighted by studies that have shown the regional activity associated with the ‘rest’ or ‘passive’ state (Binder et al., 1999; Gusnard and Raichle, 2001), a period when the subject in the scanner is likely to be engaged in stimulus independent thoughts. Such thoughts may draw on declarative memories, episodic and semantic, and, possibly, imagery. Therefore, stimulus independent thoughts during ‘passive’ listening to reversed narratives will have reduced or abolished signal in systems also engaged during language comprehension.

PET and functional magnetic resonance imaging (fMRI) studies in normal subjects have demonstrated activity in anterior temporal cortex in response to single words, sentences and narratives (e.g. Mazoyer et al., 1993; Bavelier et al., 1998; Perani et al., 1998; Schlosser et al., 1998; Stowe et al., 1998; Mummery et al., 1999; Binder et al., 2000; Scott et al., 2000; Ferstl and von Cramon, 2001; Wise et al., 2001; Vandenbarghe et al., 2002; Davis and Johnsrude, 2003; Friederici et al., 2003; Narain et al., 2003; Scott and Wise, 2004). A few study designs demonstrated that these anterior regions are activated more strongly when processing language at the sentence rather than the single content word level (e.g. Mazoyer et al., 1993; Stowe et al., 1998; Vandenbarghe et al., 2002). Vandenbarghe et al. (2002) specifically propose that left anterior temporal cortex is involved in binding together the components of a sentence into one message, and differentiate this from the controlled performance of specific syntactic operations. On this account, activity will be greater for sentences and narratives than word lists. This interpretation appears to be equivalent to that taken by electrophysiologists studying the N400, an event-related potential sensitive to content word meaning within the context of a sentence. This view envisages that meaning builds gradually and continuously throughout the processing of sentences (for review, see Kutas and Federmeier, 2000).

An alternative view considers that the rule-based grammar system and the acquired lexical system are anatomically as well as functionally distinct. In the functional neuroimaging studies that have attempted to distinguish these processes (e.g. Kuperberg et al., 2000; Ni et al., 2000; Friederici et al., 2003), there is little consistency between them in terms of results. For example, Friederici et al. (2003) demonstrate both common and distinct activations in the supratemporal planes of both hemispheres for syntactic and semantic processes. By contrast, Kuperberg et al. (2000) could not distinguish between semantic and syntactic processes, other than activations specific to semantics in the right temporal lobe alone, but they did relate a common activation in left inferotemporal cortex to pragmatic, syntactic and semantic processes. Both clinical neuropsychological and some fMRI studies provide evidence that activity in the superior temporal gyrus is associated with syntactic operations during speech processing, including the supratemporal plane (Humphries et al., 2001; Friederici and Kotz, 2003; Dronkers et al., 2004). However, in the present study on normal subjects there was no detectable difference in activity (P > 0.1, uncorrected) between listening to stories and reversed stories within either supratemporal plane. This lack of consistency raises a number of issues. Localizing function in clinical neuropsychological studies using the ‘lesion overlap’ technique, in which a common behavioral deficit is related to a common lesioned site across patients with very differently distributed...
lesions, has its limitations (Hillis et al., 2004). Further, the deficit may be attributable as much to white matter damage disconnecting intact regions of cortex as to the local cortical damage. In functional imaging studies, if neural networks for the acoustic, lexical and syntactic processing of connected speech are intermingled in the supratemporal plane, it is entirely possible that the signal from acoustic processing of both speech and speech-like stimuli may predominate and mask additional co-localized signal from the automatic linguistic processing of sentences. Nevertheless, the use of explicit tasks to isolate syntactic operations must engage, as well as language systems, prefrontal executive processes not used in normal language comprehension. Altered fronto-temporal functional interactions and differences in attentional processes may modulate the signal in language-related temporal cortex in an unpredictable manner. Any of these factors may confound a simple ‘subtraction’ experimental design (Friston et al., 1996) and potentially alter the relationship between an observable hemodynamic response and the localization of a specific psychological operation. This encapsulates some of the arguments contained in a review by Kaan and Swaab (2002). They conclude that syntactic processing is distributed between the posterior-anterior extent of superior and lateral temporal neocortex and the left inferior frontal gyrus (Broca’s area), the latter region having been observed in a number of other studies of controlled processing of syntax (reviewed in Kaan and Swaab, 2002). They argue that variations in processing load required of the subjects may alter the balance between observed frontal and temporal lobe activations, with or without the appearance of additional activations associated with the recruitment of parallel processes for the execution of a study-specific task. They take the view that process-specific language networks may be very broadly distributed, and it may be the dynamic interactions between language-related brain regions, rather than loci of activations, that are specific to the processes of extracting lexical and syntactic information from the speech stream; to the extent that right hemisphere regions may be recruited in the comprehension of complex sentences (e.g. Caplan et al., 1996; Just et al., 1996).

However, an anatomical separation of lexical semantics and the comprehension of syntactical structure may be evident from clinical studies. A recent example comes from a behavioural and anatomical study of patients with primary progressive aphasia. The anterior temporal lobe atrophy in patients classified as semantic dementia resulted in a ‘striking dissociation’ between their impaired performance on single word comprehension and their intact performance on a test of sentence comprehension comprising different levels of syntactic complexity: the reverse dissociation was observed in patients with what was termed logopenic progressive aphasia, in whom the left posterior temporal and inferior parietal lobes were most atrophied (Gorno-Tempini et al., 2004). Although the combined evidence from the clinical and functional imaging studies suggests that the comprehension of grammatical structure can exist in the presence of anterior temporal cortical atrophy, the correlation of left anterior temporal activity in response to listening to narratives with the patients’ performance on a sentence comprehension task in the present study has to be viewed in the context of infarction elsewhere in the left hemisphere language system. The posterior left STS was activated by the narratives in the normal subjects, but in the patients from the Wernicke’s group this region was either destroyed or disconnected from primary auditory and periauditory cortex. Therefore, the implicit comprehension of sentences is normally dependent on activity along the length of the left STS, but we have only observed a relationship between sentence comprehension and left anterior STS activity in the context of destruction or disconnection of the posterior STS. The complementary study would be to investigate patients with isolated left anterior temporal lobe infarction and correlate activity within the posterior left STS with measures of language recovery, but such lesions are rare for reasons outlined in the introduction.

The normal activation of the anterior left fusiform gyrus of the patients, within the region identified by electrophysiological studies at the time of temporal lobe epilepsy as the basal language area (Burnstine et al., 1990), reflected their relative preservation of single object word comprehension. It has been demonstrated with direct electrical recordings that the anterior fusiform gyrus responds to lists of object words but not function words (Nobre et al., 1994). This region consists of high-order visual association cortex (Insausti et al., 1998). When contrasting the perception of lists of imageable (concrete) nouns with lists of less imageable (abstract nouns), increased activity has been observed in polysensory perirhinal cortex (Wise et al., 2000): part of memory-related temporal lobe lying medial to the fusiform gyrus in the collateral sulcus (Insausti et al., 1998). The equivalent regions in the non-human primate, visual area TE and perirhinal area 36, form a functional unit that operates to learn and remember long-term arbitrary visual–visual associations (Naya et al., 2003). In the human, long-term cross-modal associations between arbitrary sounds (words) and visual objects may be formed in the same way. Once these associations are ‘consolidated’, their automatic recall during implicit narrative comprehension may become less dependent on memory-related (limbic) cortex (for reviews, see Zola-Morgan and Squire, 1993; Murre et al., 2001), which would explain the absence of left perirhinal cortical activation in this study. However, the absence of observable medial temporal lobe activation is not proof of an absence of involvement, especially as the choice of baseline task can be critical when attempting to visualize activity within limbic cortex (Stark and Squire, 2001).

This study demonstrated an apparent left-lateralization for speech comprehension in inferotemporal cortex, whereas previous electrophysiological and functional imaging studies of automatic single object word processing have shown bilateral involvement (Nobre et al., 1994; Wise et al., 2000). However, the asymmetry we have reported needs to be qualified (Jernigan et al., 2003). We did not investigate contrast × region × hemisphere interactions with multiple ROI analyses, and in fact mean activity in the anterior right fusiform gyrus was greater in the speech than reversed speech condition in the normal subjects, although the Z-score in this region (z = 3.9) was below the threshold we selected for the analyses. Further, automatic speech perception was associated with activity in the anterior right STS, in both the normal subjects and the patients. As reversed speech lacks the normal segmental stress patterns, suprasegmental prosody and voice identity of normal speech, the conservative assumption is that the anterior stream of auditory processing on the right predominantly reflects processing of one or more aspects of non-linguistic speech information (for review, see Wong, 2002). However, activation of the anterior right superior temporal sulcus was observed in, for example, the study by Ferstl and von Cramon (2001) on
written text comprehension, in which paralinguistic speech information is absent. There is also a study on patients with semantic dementia, which provided evidence that both temporal lobes have a role in lexical semantics (Lambon Ralph et al., 2001). Thus, we cannot assume that the right temporal lobe did not contribute to linguistic processing in both the normal subjects and patients. Further, although we did not demonstrate significant differences between the activation patterns in the right hemisphere between normal subjects and patients, reorganization may have occurred in the right temporal lobe of patients without this change necessarily being captured by any difference in hemodynamic response. In other words, lack of evidence of post-stroke neural plasticity in this study is not evidence for its absence.

As a final footnote, this study does not imply that the language functions of the STS are confined to spoken language. Written text comprehension activates both the anterior and posterior left STS (e.g. Ferstl and von Cramon, 2001). This accords with the abundant anatomical evidence from non-human primates that the length of the STS is an important zone of convergence for sensory information from different modalities. Thus, if this study had investigated written rather than spoken narrative comprehension, the result and conclusions might have been very similar.

In summary, the results from this study are best interpreted as evidence that automatic speech comprehension engages the length of the left STS in normal subjects and that language functions in the intact anterior STS are disrupted by a Wernicke’s area lesion. Polysensory cortex in the dorsal bank of the anterior STS in non-human primates, the anterior superior temporal polysensory area, receives reciprocal projections from the highest-order auditory, visual and somatosensory association cortices. Onward reciprocal connections connect the anterior superior temporal polysensory area to limbic, parietal and prefrontal cortex (e.g. Seltzer et al., 1996; Saleem et al., 2000; Padberg et al., 2003). These afferent and efferent connections lend plausibility to the hypothesis that the anterior STS is one of perhaps several 'convergence zones' or 'transmodal areas' in a hierarchy that ultimately maps the encoded speech signal on to long-term memories of attributes, associations and concepts (Damasio, 1989; Mesulam, 1998). Although the primate anterior STS has become equated with visual processing (e.g. Bruce et al., 1981; Oram and Perrett, 1996; Anderson and Siegel, 1999), comparative studies of auditory neurophysiology between primates and humans may indicate a functional homology between the non-human primate and human anterior STS in auditory processing. This may provide insights into the evolution of speech perception and comprehension that parallel inferences being made about mirror neurons in the inferior frontal gyrus and the evolution of speech production (Rizzolatti and Arbib, 1998).

Notes
Address correspondence to Richard Wise, Division of Neurosciences and Mental Health and MRC Clinical Sciences Centre, Imperial, Hammersmith Hospital, London W12 0NN, Uk. Email: richard.wise@csc.mrc.ac.uk.

References


Narrative Speech Comprehension and Left STS


