Differentiating Hemispheric Contributions to Syntax and Semantics in Patients with Left-Hemisphere Lesions

Paul Wright,1 Emmanuel A. Stamatakis,2 and Lorraine K. Tyler1

1Centre for Speech, Language and the Brain, Department of Experimental Psychology, University of Cambridge, Cambridge, CB2 3EB, United Kingdom, and 2Division of Anaesthesia, School of Clinical Medicine, University of Cambridge, Cambridge, CB2 2QQ, United Kingdom

Understanding the relationship between brain and cognition critically depends on data from brain-damaged patients since these provide major constraints on identifying the essential components of brain–behavior systems. Here we relate structural and functional fMRI data with behavioral data in 21 human patients with chronic left hemisphere (LH) lesions and a range of language impairments to investigate the controversial issue of the role of the hemispheres in different language functions. We address this issue within a dual neurocognitive model of spoken language comprehension in which core linguistic functions, e.g., syntax, depend critically upon an intact left frontotemporal system, whereas more general communicative abilities, e.g., semantics, are supported by a bilateral frontotemporal system and may recover from LH damage through normal or enhanced activity in the intact right hemisphere. The fMRI study used a word-monitoring task that differentiated syntactic and semantic aspects of sentence comprehension. We distinguished overlapping interactions between structure, neural activity, and performance using joint independent components analysis, identifying two structural–functional networks, each with a distinct relationship with performance. Syntactic performance correlated with tissue integrity and activity in a left frontotemporal network. Semantic performance correlated with activity in right superior/middle temporal gyri regardless of tissue integrity. Right temporal activity did not differ between patients and controls, suggesting that the semantic network is degenerately organized, with regions in both hemispheres able to perform similar computations. Our findings support the dual neurocognitive model of spoken language comprehension and emphasize the importance of linguistic specificity in investigations of language recovery in patients.

Introduction

Although some cognitive functions are impaired following brain damage, others are preserved. Understanding the nature of these preservations and declines is critical for elucidating the relationship between cognition and brain and for developing effective interventions. Conservation of cognitive functions can be achieved by means of functional reorganization, in which regions not typically involved in a specific function are recruited, or by residual functionality of the damaged normal system (Rosen et al., 2000; Zahn et al., 2004; Crinion and Price, 2005; Saur et al., 2006). For language function, it remains unclear which linguistic components can be preserved following left hemisphere (LH) language system damage and which neural mechanisms underlie their preservation. Here we investigate the differential contributions of left-lateralized and bilateral language systems to syntactic and semantic aspects of spoken language comprehension in patients with LH damage using a multivariate method combining structure, function, and performance (Calhoun et al., 2006).

We frame our research within a neurocognitive model of spoken language comprehension (Bozic et al., 2010) that differentiates between core linguistic functions, such as syntax, which depend upon the integrity of a LH frontotemporal (FT) network (Tyler et al., 2011), and broader communicative semantic/pragmatic interpretation supported by a bilateral FT system. Semantic and pragmatic inference may be preserved in patients with LH damage and syntactic deficits (Ostrin and Tyler, 1995; Hagoort, 1997; Tyler and Marslen-Wilson, 2008), raising the possibility that the right hemisphere (RH) residue of the bilateral system may suffice to preserve performance due to degenerate organization (Friston and Price, 2003). This mechanism relies on normally bilateral activity, as distinct from adaptive recruitment of novel RH regions (Saur et al., 2006).

To test this hypothesis, we collected fMRI data while patients listened to spoken sentences in which we manipulated syntactic and semantic information, and we measured patients’ comprehension abilities using various tasks. Syntactic information activates left inferior frontal gyrus (IFG) and posterior middle temporal gyrus (MTG) (Caplan et al., 1996; Tyler et al., 2011). Damage to either region, or the connectivity between them, disrupts syntactic performance (Papoutsi et al., 2011) leaving semantic processing, associated with bilateral MTG, relatively unimpaired. Given the interdependencies between syntax and semantics (Marslen-Wilson and Tyler, 1980) and the two pro-
cesses’ instantiation in partially overlapping neural networks involving left MTG, univariate analyses may not distinguish them, nor can they describe the three-way relationship between performance, tissue integrity, and neural activity that may reflect functional reorganization.

We therefore used joint independent component analysis (jICA) (Calhoun et al., 2006) to combine structural and functional MR images and to differentiate spatially distinct structural–functional networks, each with a variable level of expression across patients. We identified networks relevant to syntax after accounting for regions involved in semantics, and vice versa, by correlating components with performance in two tasks. We predicted that preserved syntax in the patients would be associated with activity and tissue integrity in the LH frontotemporal network (Tyler et al., 2010b, 2011), whereas semantics—normally supported bilaterally—may correlate with normal or enhanced activity in right MTG.

Materials and Methods

Participants

We recruited 21 chronic brain-damaged patients with LH damage [mean (SD) age, 60.0 (11.5) years; controls vs patients: abilities varied (see Results, Behavioral results, below). and to understand task instructions but their language comprehension to produce single words. All patients were able to give informed consent to produce language: some patients were fluent while others were only able (RWO) consisted of strings of words with no grammatical or sentential

rect, with their grammatical structure matched to the NP sentences, but

respective contributions to sentence processing (Table 1). Normal Prose (NP) sentences were both grammatical and semantically meaningful. Anomalous Prose (AP) sentences were syntactically correct, with their grammatical structure matched to the NP sentences, but lacked coherent sentential/pragmatic meaning. Random Word Order (RWO) consisted of strings of words with no grammatical or sentential

meanings. Half the RWO strings were derived from NP sentences and half from AP sentences.

We measured the processing of sentential information by manipulating the position of the target word in the sentence. In NP and AP, target words occurring later in the sentence elicit faster response times (RTs) than earlier targets, but in RWO, target word position does not affect RT (Marslen-Wilson and Tyler, 1980; Tyler et al., 2010a). Faster monitoring latencies to later compared with earlier words [the word position effect (WPE)] in NP reflects the on-line construction of a syntactically structured, meaningful representation spanning the sentence, whereas the WPE in AP is underpinned primarily by syntactic processing, without the contribution of sentential or pragmatic meaning (Marslen-Wilson and Tyler, 1975, 1980). The absence of a WPE in RWO confirms this interpretation and rules out effects based solely on anticipation. We calculated the WPE for each condition using the formula  

\[
\text{WPE} = \frac{\text{early RT}}{\text{late RT}}/\text{(mean RT, Single-trial RTs were inverse transformed before averaging to reduce the influence of outliers (Ratcliff, 1993).}
\]

The task was presented in the scanner using a blocked design because patients may have difficulty rapidly switching between tasks. Presentation was divided into two scans, with a short break in between, with equal numbers of each trial type in each scan. Each scan comprised 15 trials of NP (165 s), followed by 12 trials of silent baseline (132 s), 15 trials of RWO (165 s), 12 trials of acoustic baseline (132 s, not used in current analysis), and 15 trials of AP (165 s). Stimulus presentation was controlled by a PC running in-house software, via pneumatic insert earphones (Etymotic Research) and an LCD projector viewed with a mirror inside the MRI head coil. Participants responded using an MRI-compatible button box.

Sentence–picture matching task (out of scanner). We obtained additional measures of language impairment using a sentence–picture matching task, which is frequently used to assess syntactic and semantic abilities in impaired populations (Caramazza and Zurif, 1976; Ostrin and Tyler, 1995; Tyler et al., 2002, 2009, 2011; Wilson et al., 2010). Patients heard a sentence and were asked to match it to one of three line drawings,
Each sentence described an activity involving an agent and a patient (e.g., “The boy chases the horse”) and was semantically reversible, in that either entity could perform the action (e.g., “The horse chases the boy”). Of the three line-drawings presented to the patient, one correctly depicted the scene described in the sentence (e.g., a boy chasing a horse), a second depicted a similar scene but with the agent and patient reversed (a reverse role distractor, e.g., a horse chasing a boy) and the third involved a change in the action/verb (a lexical distractor, e.g., a boy riding a horse). Patients with both syntactic impairments make reverse role errors with few lexical distractor errors (Black et al., 1991), whereas patients with semantic impairment show the opposite pattern.

**Image acquisition**

MR images were acquired using a 3T Tim Trio MRI scanner (Siemens Medical Solutions) at the MRC Cognition and Brain Sciences Unit, Cambridge, UK, as reported previously (Tyler et al., 2010a,b). Functional scanning used sparse sampling with 2 s scans interspersed with 9 s of silence to prevent scanner noise interfering with spoken stimuli (Hall et al., 1999). Target cues were presented 1 s after scan onset, and auditory stimuli were presented 1.1 s later, 100 ms after the onset of silence. The variable duration between the end of the spoken stimulus and the following scan ensured sampling of a range of time points around the time of the likely peak hemodynamic response. Functional imaging parameters were as follows: 32 oblique axial slices angled away from the eyes, each 3 mm thick with interslice gap of 0.75 mm; in-plane resolution of 3 mm; FOV of 192 × 192 mm; total TR = 11 s (2 s acquisition + 9 s silence); TE = 30 ms; and flip angle = 78°. Structural imaging parameters were as follows: T1-weighted MPGR image sequence, 1 mm isotropic resolution in the sagittal plane, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, and flip angle = 9°.

**Imaging processing**

Functional and structural images were preprocessed using SPM5 (Wellcome Department of Imaging Neuroscience, University College London, London, UK). Functional images were realigned to correct for head movement and coregistered with the structural images. Structural images were normalized to Montreal Neurological Institute space using unified segmentation-normalization with high warping regularization to prevent distortion of lesions (Crinion et al., 2007). Unified normalization failed in one patient, so cost function masking was used for normalization (Brett et al., 2001). Images from controls were also normalised using unified normalization, but default warping regularization was used. Normalization parameters from structural images were applied to functional images, which were then smoothed using a Gaussian kernel of 8 mm FWHM.

Functional images for each participant were entered into the analysis as covariates, and auditory stimuli were presented 1 s after scan onset, and auditory stimuli were presented 1.1 s later, 100 ms after the onset of silence. The variable duration between the end of the spoken stimulus and the following scan ensured sampling of a range of time points around the time of the likely peak hemodynamic response. Functional imaging parameters were as follows: 32 oblique axial slices angled away from the eyes, each 3 mm thick with interslice gap of 0.75 mm; in-plane resolution of 3 mm; FOV of 192 × 192 mm; total TR = 11 s (2 s acquisition + 9 s silence); TE = 30 ms; and flip angle = 78°. Structural imaging parameters were as follows: T1-weighted MPGR image sequence, 1 mm isotropic resolution in the sagittal plane, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, and flip angle = 9°.

Functional images for each participant were entered into the analysis as covariates, and auditory stimuli were presented 1 s after scan onset, and auditory stimuli were presented 1.1 s later, 100 ms after the onset of silence. The variable duration between the end of the spoken stimulus and the following scan ensured sampling of a range of time points around the time of the likely peak hemodynamic response. Functional imaging parameters were as follows: 32 oblique axial slices angled away from the eyes, each 3 mm thick with interslice gap of 0.75 mm; in-plane resolution of 3 mm; FOV of 192 × 192 mm; total TR = 11 s (2 s acquisition + 9 s silence); TE = 30 ms; and flip angle = 78°. Structural imaging parameters were as follows: T1-weighted MPGR image sequence, 1 mm isotropic resolution in the sagittal plane, TR = 2250 ms, TI = 900 ms, TE = 2.99 ms, and flip angle = 9°.

Trading values between corresponding components (r) was interpretable in the respective maps have high loading values and vice versa. Patients with high tissue integrity and/or activity values in the areas described in the respective maps have high loading values and vice versa. The loading values were interpreted post hoc by testing their correlations with performance. Our post hoc analyses test the relative variability in the loading values within each component since the interpretation of the absolute magnitude of the loading values is not straightforward.

We performed jICA analysis using the Fusion ICA Toolbox (FIT v2.0b, The MIND Research Network, http://icatb.sourceforge.net) with the following parameters. We entered three features into the analysis: tissue integrity, activity for NP-silence (semantics + syntax), and activity for AP-silence (syntax only). Since our focus was on sentential semantics and syntax, we did not include the RWO condition, which comprised word strings with neither grammar nor meaning. Tissue integrity was represented by smoothed, desekled structural images (see Imaging processing, above) and activity represented by the contrast images (weighted betas) generated by the GLM.

Univariate analyses with structural images typically include global signal as a confound (Tyler et al., 2005a), but this was not possible within the jICA framework. We therefore tested the effect of including global signal by running a univariate, voxelwise correlation between structural images and WPE AP both with and without global signal as a covariate of no interest in the model. We used WPE AP because this score correlated with tissue integrity in the jICA results. Both analyses produced a cluster centered on the left IFG (LIFG), with the cluster without global signal correction contained entirely within the cluster with global signal correction. We therefore concluded that omitting global signal correction gave similar but more conservative results for our set of structural images, making it unlikely that the jICA results would be driven by intersubject variability in global signal. We also tested the effect of image inhomogeneity by running a separate jICA using bias-corrected images produced by the SPM5 segmentation algorithm (Ashburner and Friston, 2005). The results of the two analyses were highly similar, with virtually indistinguishable patterns in the feature images and highly correlated loading values between corresponding components (r > 0.98). We report results using non-bias-corrected images.

We set the number of jICA components to seven, which was one-third of the sample size (n = 21). This heuristic was previously used to set the number of components when the minimum description length algorithm failed to converge (Specht et al., 2009), as was the case for our data. ICA was performed using the Infomax algorithm. The output feature images were scaled to z scores. Images were thresholded at z = 3.09 (p < 0.001) and a cluster size of 47 voxels (1.27 cm³). Cluster size threshold was set to the average needed to achieve p < 0.05, corrected for multiple comparisons, in the standard random effects analyses for NP-silence and AP-silence in the 21 patients (see Additional imaging analyses, below).

We identified functionally relevant components by testing the correlation of their loading values with semantic and syntactic performance scores. We tested correlations with WPE for NP and AP, and with both reverse role and lexical errors from sentence picture matching. We did not include WPE for RWO for the same reasons we did not include activity for RWO. The significance of each Pearson’s correlation was
calculated as one-tailed, since our a priori prediction was that improved performance would be associated with greater tissue integrity/activity. Because our data comprised patients with heterogeneous lesions, jICA was liable to produce components dominated by one or two patients whose lesion and/or activity varied a great deal from the group. The loading values from these components are unsuitable for correlations because their distributions are highly skewed. Moreover, a significant correlation would not be interpretable, as it would be driven by one or two patients, rather than a consistent trend across the group. We therefore removed outlying loading values before testing correlations with performance. Outliers were defined as data points that were outside the interquartile range by >1.5 interquartile ranges (Tukey, 1977). Previous work using jICA selected only components with effects predominantly in the brain rather than in ventricles or around the edge of the brain. The latter were assumed to describe movement or pulsation artifacts (Specht et al., 2009). In our analysis, we found no components with effects predominantly outside the brain, so none were excluded under these criteria.

Additional imaging analyses

Independent to the main analysis, we defined damaged voxels in each patient by using a lesion-detection algorithm as previously published (Stamatakis and Tyler, 2005). Briefly, T1-weighted structural images from each LH-damaged participant were entered into individual two-sample t tests with images from a group of age-matched controls. Voxels with a significant drop in T1-weighted image intensity were defined as damaged. The threshold for significance was manually adjusted, with the statistical image overlaid on the patient’s structural image, to minimize classification of healthy tissue or CSF in enlarged ventricles as lesion. The resulting lesion images were used to mask contrast images (as described in Imaging processing, above) and were combined to give a lesion frequency map for the group (Fig. 1) as previously described (Tyler et al., 2010b, 2011). The tissue integrity data used in the jICA was independent of automatic lesion identification.

We compared functional activation in controls and patients using group-level (random effects) one- and two-sample t tests in SPM. One-sample t tests were used to describe activity significantly greater than silent baseline in each group by comparing contrast images for each group with zero. Two-sample t tests were used to identify regions with significantly greater activity in controls than patients, and vice versa. In the two-sample t tests, the two sets of contrast images were defined as independent groups with unequal variance; global mean scaling was not used.

Results

Behavioral results

Patients showed a wide range of syntactic abilities, with some very impaired and others performing similarly to controls. On sentence–picture matching, patients’ reverse role error rate ranged from 0 to 47% (mean, 17%; SD, 16%), extending well outside the range for controls (0–9%; mean, 2.9%; SD, 2.6%). In contrast, they made few lexical errors (0–6%; mean, 1.9%; SD, 2.2%) and controls made none. On the word-monitoring task performed in the scanner, patients’ overall ability to perform the task was as good as controls. Patients generated similar numbers of missing or inappropriate (e.g., anticipatory) responses as controls (percent missing/inappropriate responses for patients: mean, 5%; SD, 6%; controls: mean, 3%; SD 3%; t test, patients vs controls: p > 0.05). Patients and controls also showed the typical reaction time advantage to NP over AP, indicating their overall ability to comprehend normal speech (RT difference AP – NP for patients: mean, 91 ms; SD, 54 ms; controls: mean, 70 ms; SD, 35 ms; t test, patients vs controls: p > 0.05). Patients’ range of performance on syntax (WPE AP) was numerically, but not significantly, lower than controls: patients, −0.17–0.32 (mean, 0.09; SD, 0.13); controls, −0.09–0.38 (mean, 0.14; SD, 0.12). Both groups showed a similar range of performance on semantics (WPE NP): patients, 0.08–0.58 (mean, 0.33; SD, 0.13) and controls, −0.05–0.62 (mean, 0.29; SD, 0.17). (Note: sentence–picture matching scores were not available for the scanned controls, so scores are taken from a separate group, also age-matched to the patients. These scores are used in this section only to illustrate the normal range, and do not appear in any further analyses.)

Selection of components relevant to performance

We used jICA to combine structural and functional images from 21 patients and then decompose these data into seven independent components. From these, we selected functionally relevant components on the basis of their correlation with performance, using the word-position effect for normal prose (WPE NP) as a measure of successful processing of sentential semantics, WPE for anomalous prose (WPE AP) as a measure of successful syntactic processing, and reverse-role errors in the sentence–picture matching task as a measure of syntactic comprehension impairment. Correlations with performance were significant in two of the seven components (we report one-tailed significance, since we predict that larger component loading should only correlate with better performance). Component 1 had loading values that correlated positively with WPE NP (r = 0.412, p < 0.05) but not with either the WPE AP (r = −0.329, p > 0.05) or the reverse-role errors in sentence–picture matching (r = −0.320, p > 0.05). Because this component only correlated with performance for the NP sentences that carried both semantic and syntactic information, and not with AP sentences that contained only syntactic information, it was designated the semantics component. Component 2 had loading values that correlated positively with WPE AP (r = 0.580, p < 0.01) and negatively with reverse role errors in sentence–picture matching (r = −0.774, p < 0.001) and was designated the syntax component. The syntax component did not correlate with WPE NP (r = 0.048, p > 0.05). Figure 2 illustrates the dissociation of the two components by plotting the relationship between each component’s loading values and both corresponding and noncorresponding performance measures. Each score correlated only with loading values for its corresponding ICA component. None of the reverse correlations were significant (|r| < 0.39, p > 0.05, one-tailed).

The syntax component

Preserved syntax was associated with intact structure and function in a left frontotemporal network, consistent with previous studies (Caplan et al., 1996; Tyler et al., 2010b, 2011). Significant tissue integrity effects formed a single, long cluster extending from the gray matter of LIFG, through the white matter of the superior longitudinal and arcuate fasciculi, to the white matter underlying the planum temporale (Table 2; Fig. 3, green). Consistent with the structural data, the functional images in the syntax component described activity in LIFG and in left posterior superior and middle temporal gyri (LpSTG/MTG; Table 2; Fig. 3, blue and red). Effects in LIFG were seen only for AP, the condition that loaded most strongly on syntactic processing. Effects in LpSTG/MTG were seen for both NP and AP. Taken as a whole, these findings are consistent with studies showing that syntactic comprehension depends critically upon the integrity of a left frontotemporal network, particularly the LIFG and left posterior temporal lobe (Dronkers et al., 2004; Caplan et al., 2008; Tyler and Marslen-Wilson, 2008; Tyler et al., 2010a; Papoutsi et al., 2011).

The semantics component

Performance on sentential semantics (i.e., WPE NP but not WPE AP) correlated with activity in right pSTG/MTG (RpSTG/MTG)
for NP and, to a lesser extent, AP (Table 3, Fig. 4). Activity for NP included a small cluster in right temporal pole. The location of activity in RpSTG/MTG was consistent across both prose types (Fig. 4); however, the significance and extent of activity was greater for NP than AP (NP: peak $z = 10.18$, size $= 16.1 \text{ cm}^3$; AP: peak $z = 5.04$, size $= 3.4 \text{ cm}^3$). Unlike the syntax component, the semantics component showed no significant relationship with tissue integrity, indicating that support of semantics by activity in RpSTG/MTG was not related to damage to a particular region in LH. Although jICA combines tissue integrity and activity such that they show the same pattern of expression over subjects in a given component, their patterns of expression over the brain remain independent. This allows jICA to explain variance that is

Table 2. Brain regions involved in syntax component

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Cluster size ($\text{cm}^3$)</th>
<th>$x$</th>
<th>$y$</th>
<th>$z$</th>
<th>$z$ score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal prose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LpSTG/MTG</td>
<td>21,22</td>
<td>8.2</td>
<td>−60</td>
<td>−24</td>
<td>9</td>
<td>8.65</td>
</tr>
<tr>
<td>Left medial SFG</td>
<td>10</td>
<td>1.7</td>
<td>−3</td>
<td>66</td>
<td>18</td>
<td>5.56</td>
</tr>
<tr>
<td>Medial precentral</td>
<td>7, 19</td>
<td>1.6</td>
<td>3</td>
<td>−78</td>
<td>39</td>
<td>5.22</td>
</tr>
<tr>
<td>Anomalous prose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIFG</td>
<td>47, 45</td>
<td>1.7</td>
<td>−51</td>
<td>18</td>
<td>−3</td>
<td>4.87</td>
</tr>
<tr>
<td>LpSTG/MTG</td>
<td>21, 22</td>
<td>1.4</td>
<td>−60</td>
<td>−27</td>
<td>12</td>
<td>4.19</td>
</tr>
<tr>
<td>Tissue integrity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIFG, insula, SLF, AF</td>
<td>44, 45, 47</td>
<td>53.2</td>
<td>−26</td>
<td>10</td>
<td>18</td>
<td>4.97</td>
</tr>
</tbody>
</table>

Peak voxel coordinates ($x$, $y$, $z$) in Montreal Neurological Institute space (in mm), BA, Brodmann area; SFG, superior frontal gyrus; SLF, superior longitudinal fasciculus; AF, arcuate fasciculus.

Figure 2. The syntax and semantics components show dissociated relationships with performance. A, WPE NP correlated positively with loading on the semantics component (blue) but showed no relationship with the syntax component (red). B, WPE AP correlated positively with loading on the syntax component only (red) and showed a reversed but nonsignificant relationship with the semantics component (blue). C, Reverse-role error rate from the sentence–picture matching task correlated negatively with loading on the syntax component only (red, larger loading value = fewer errors); the relationship with the semantics component was in the same direction but was not significant. $^*$p < 0.05; $^{**}$p < 0.01; $^{***}$p < 0.001; n.s., not significant; a.u., arbitrary units.

Figure 3. Syntax component. Syntactic performance correlated with tissue integrity and activity in a left frontotemporal network. A, Surface view showing tissue integrity (TI) in LIFG (green) activity to AP in LIFG and LSTG (red) and activity to NP in LpSTG/MTG (blue). Central legend shows colors of the feature images and their overlap in A and B. B, Sagittal sections (at the indicated MNI $x$-coordinate) showing tissue integrity (green) in white matter in the superior longitudinal and arcuate fasciculi, extending from LIFG to the planum tempora. All effects shown at voxel-level $p < 0.001$ ($z > 3.09$) and cluster size $> 1.27 \text{ cm}^3$ (approximately $p < 0.05$ corrected). C, Higher loading values from the syntax component correlated with larger WPE AP (red) but not for WPE NP (blue). D, Higher loading values correlated with fewer reverse role errors on sentence-picture matching (RR, red) but not with lexical errors (Lex, blue). $^*$p < 0.05; $^{**}$p < 0.01; $^{***}$p < 0.001; n.s., not significant; a.u., arbitrary units.
correlated across subjects but spatially independent between image types, which in this case included a pattern of subject-wise variance in activity in RpSTG/MTG that was not expressed in tissue integrity in any region.

Investigating the mechanism underlying right hemisphere support of performance

Semantic performance depended upon activity in RpSTG/MTG in the context of LH damage. Since semantics/pragmatics typically involves bilateral pSTG/MTG regions (Tyler et al., 2010a), the most straightforward explanation is that function was preserved because of degeneracy in the normal system, i.e., residual function in the intact part of the normally bilateral system was sufficient to support semantic processing. Alternatively, activity within the bilateral network may have shifted to become more right-lateralized as a response to LH damage, on the hypothesis that right hemisphere (RH) regions may be recruited to perform either new computations or play an enhanced role in existing computations. At first glance, the jICA results do not support the recruitment hypothesis, since this would predict a negative correlation between activity in RH and tissue integrity in LH, which we did not observe.

However, since these effects may have been too small to reach statistical threshold, and since recruitment of contralesional regions is a critical issue in poststroke recovery, we ran follow-up analyses looking for any increase in RH activity in patients with LH damage, over normal levels of activity in controls. If activity did not differ between groups, then we could rule out the adaptive recruitment hypothesis in favor of the degeneracy hypothesis. We contrasted whole-brain activity during NP in our 21 LH-damaged patients with a group of 21 age-matched controls. In one-sample t tests, both groups activated bilateral pSTG/MTG during NP, but in the patients, activity on the left appeared to be weaker than controls (Table 4, Fig. 5A, B). Confirming this, two-sample t tests contrasting the two groups showed significantly lower activity for patients than controls in LSTG at the voxel-level threshold of p < 0.001, extending to MTG at p < 0.01 (Table 4, Fig. 5C), but no significant differences in RpSTG/MTG, even at a liberal voxel-level threshold of p < 0.01 without correction for cluster size. The differences in LSTG reflected reduction of activity by underlying damage in LpSTG/MTG. The absence of differences in RpSTG/MTG, even at a low threshold, argues against adaptive recruitment of RH, and suggests that RpSTG/MTG can support semantics in patients with left-hemisphere damage because the normally bilateral semantic network is a degenerate system, such that RH is sufficient for function after damage to LH (Friston and Price, 2003).

Discussion

We used multivariate analysis to identify distinct structural–functional networks supporting syntactic and semantic aspects of speech comprehension in patients with left hemisphere damage. First, we found that syntactic performance depends on tissue integrity in a network comprising LIFG and LpMTG. Second, we identified a region in RpSTG/MTG supporting semantic processing, which emerged when using jICA to account for variance in recruitment, which emerged when using jICA to account for variance in structure and function related to syntax and semantics separately. Post hoc analyses showed that activity in RpSTG/MTG was not enhanced in patients as a consequence of their LH damage, but was at similar levels to controls, and thus reflected remaining function in the intact residue of a normally bilateral network (Bozic et al., 2010). These results shed light on the distinct neural mechanisms by which syntax and semantics may be preserved following brain injury. These findings, moreover, support a dual neurobiological model for language, with a left-lateralized system supporting core linguistic functions such as syntax (Capan et al., 1996; Friederici et al., 2003; Dronkers et al., 2004; Snijders et al., 2009; Pallier et al., 2011), and a bilateral system supporting the semantic and pragmatic interpretation of spoken language (Dobel et al., 2001; Bookheimer, 2002; Friederici et al., 2003; Indefrey and Cutler, 2004; Jung-Beeman, 2005; Humphries et al., 2006; Bozic et al., 2010; Rodd et al., 2010).

Performance on two separate tests of syntactic comprehension correlated with tissue integrity in left IFG, superior longitudinal and arcuate fasciculi and pSTG/MTG, as well as activity...
Table 4. Brain regions identified by univariate fMRI analyses in patients and controls

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Cluster size</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>z score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients, NP-silence, $p &lt; 0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R STG, MTG, IFG</td>
<td>22, 21, 44</td>
<td>56.2</td>
<td>51</td>
<td>−12</td>
<td>−3</td>
<td>6.57</td>
</tr>
<tr>
<td>R occipital</td>
<td>18, 19, 37</td>
<td>9.2</td>
<td>36</td>
<td>−87</td>
<td>−6</td>
<td>5.57</td>
</tr>
<tr>
<td>L + R SMA</td>
<td>6, 24, 32</td>
<td>21.5</td>
<td>9</td>
<td>12</td>
<td>39</td>
<td>3.54</td>
</tr>
<tr>
<td>L STG, MTG</td>
<td>22, 21</td>
<td>37.9</td>
<td>−42</td>
<td>−33</td>
<td>12</td>
<td>5.40</td>
</tr>
<tr>
<td>L occipital</td>
<td>19, 18, 37</td>
<td>11.6</td>
<td>−39</td>
<td>−90</td>
<td>0</td>
<td>4.79</td>
</tr>
<tr>
<td>R precentral gyrus, MFG</td>
<td>6, 4, 3</td>
<td>5.9</td>
<td>54</td>
<td>0</td>
<td>45</td>
<td>4.47</td>
</tr>
<tr>
<td>L precentral gyrus</td>
<td>6</td>
<td>1.6</td>
<td>−48</td>
<td>3</td>
<td>39</td>
<td>4.36</td>
</tr>
<tr>
<td>L cerebellum</td>
<td></td>
<td>1.5</td>
<td>−12</td>
<td>−60</td>
<td>−15</td>
<td>4.24</td>
</tr>
<tr>
<td>Controls, NP-silence, $p &lt; 0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R STG, MTG, IFG, precentral</td>
<td>22, 21, 38, 44, 45, 47, 4, 6</td>
<td>79.3</td>
<td>51</td>
<td>−21</td>
<td>−6</td>
<td>&gt; 8.00</td>
</tr>
<tr>
<td>L STG, MTG, IFG, precentral, fusiform</td>
<td>22, 21, 38, 44, 45, 47, 6, 4−1, 18, 19, 37</td>
<td>162.7</td>
<td>−63</td>
<td>−24</td>
<td>9</td>
<td>6.95</td>
</tr>
<tr>
<td>R occipital, fusiform, cerebellum</td>
<td>19, 37, 18</td>
<td>26.7</td>
<td>45</td>
<td>−75</td>
<td>−9</td>
<td>5.58</td>
</tr>
<tr>
<td>Controls &gt; patients, NP-silence, $p &lt; 0.001$</td>
<td>L STG</td>
<td>22</td>
<td>4.2</td>
<td>−51</td>
<td>−18</td>
<td>9</td>
</tr>
<tr>
<td>Controls &gt; patients, NP-silence, $p &lt; 0.01$</td>
<td>L STG, MTG</td>
<td>19.8</td>
<td>−57</td>
<td>−42</td>
<td>6</td>
<td>4.40</td>
</tr>
</tbody>
</table>

Cluster size in cubic centimeters. Peak voxel coordinates ($x$, $y$, $z$) in Montreal Neurological Institute space (in mm). BA, Brodmann area; SMA, supplementary motor area; MTG, middle frontal gyrus. Note: for controls > patients, NP-silence, the cluster at $p < 0.01$ includes a peak voxel that was not present at $p < 0.001$ since at the higher threshold this voxel was part of a separate cluster that failed to meet cluster size threshold.

during syntactic processing (AP) in LIFG and left pSTG/MTG. This combination of structural and functional effects strengthens the claim that syntax depends not only on the LIFG, but also on the pSTG/MTG and white matter pathways connecting the two regions (Papoutsi et al., 2011; Rolheiser et al., 2011).

Syntactic performance correlated with activity in LIFG only during AP and with activity in LpSTG/MTG during both AP and NP. In AP, word-monitoring performance depends on syntax in the absence of sentential semantics, whereas in NP, both syntactic and semantic sentential information are available. Therefore, word monitoring during AP loads upon syntax, whereas during NP semantic sentential information also supports word monitoring and the loading upon syntax is reduced (Tyler et al., 2010b). The observed pattern of activity implicates LIFG in syntactic computations, but suggests that LpSTG/MTG is involved in the normal processes of both syntactic and semantic analysis processes in spoken language. In the case of LSTG/MTG, however, syntactic processing is coupled with activity in LIFG, whereas semantic processing does not require the involvement of the LIFG, except perhaps in cases of high cognitive demands involving processes such as selection (Thompson-Schill et al., 1999; Badre et al., 2005). This relationship between LIFG and LpSTG/MTG complements previous findings that syntactic performance in patients with LH damage depends on effective connectivity between LIFG and LpSTG/MTG, with the LIFG driving performance (Papoutsi et al., 2011). Papoutsi et al. (2011) used a different syntactic manipulation from that used in the current study and did not require listeners to perform an explicit task, thus minimizing cognitive load, but found the same left frontotemporal network as reported here and previously (Snijders et al., 2009).

While the exact nature of the relationship between the LIFG and LpSTG/MTG in syntactic processing is unclear, one possibility is that interactivity between LIFG and MTG may be necessary for the sustained maintenance and updating of syntactic representations (Snijders et al., 2009). Whatever the nature of this functional relationship, it is clear from the current findings that although they may be distinct regions, the interaction between LIFG and LMTG is essential for syntactic processing.

Activity in RpSTG/MTG correlated with behavioral performance on stimuli that loaded on sentential semantics (WPE for NP, but not AP). Although normal prose typically elicits bilateral
temporal activity (Fig. 4 B), the jICA analysis separated activity in the two hemispheres into two components with distinct patterns of variability across patients. Activity in LpSTG/MTG covaried with LH damage, whereas activity in RpSTG/MTG was unrelated to damage: there were no significant voxels in the tissue integrity feature in this component and activity in RpSTG/MTG did not differ between patients and controls. This suggests that the latter component therefore reflects the residue of the bilateral semantic network. The ability of a subset of the network to support performance demonstrates that semantic processing is instantiated in a degenerate fashion, with regions in both hemispheres capable of performing similar computations (Friston and Price, 2003; Price and Crinion, 2005). This finding extends previous work relating semantic comprehension in patients with LH stroke to activity in regions that were normally active in controls, without differences in lateralization of activity between patients and controls (Zahn et al., 2004). Our analysis additionally locates activity supporting semantics to a specific node of the normal semantic network and dissociates preserved semantics from impaired syntax in the same group of patients.

The jICA analysis did not reveal any contribution of frontal regions in semantic processing. Activity and tissue integrity in LIFG only correlated with performance on syntax. Although some studies do report LIFG involvement (particularly pars orbitalis/BA 47) (Binder et al., 2009) in semantics, these studies typically involve tasks that involve making explicit decisions about semantic attributes or involve processes of semantic selection/competition or controlled retrieval (Dapretto and Bookheimer, 1999; Kuperberg et al., 2000; Homea et al., 2002; for full list, see Binder et al., 2009). In contrast, the task we used here—the word-monitoring task—on NP sentences involved responding to high-frequency, concrete words without an explicit semantic decision. Moreover, the word-monitoring task reflects the ability of a listener to activate the semantic and syntactic properties of individual words and integrate them into a coherent and structured representation. These processes have been argued to be automatic and obligatory (Marslen-Wilson and Tyler, 1980), and therefore poor candidates for LIFG involvement. When sentences do not carry the normal combination of semantic/pragmatic and syntactic information, as is the case with Anomalous Prose, then performance on the word-monitoring task reflects the listener’s ability to integrate only the syntactic properties of each word into the developing sentential representation, a process that reveals processes of syntactic integration in the frontotemporal network.

Overall, our findings support a dual neurocognitive model for speech comprehension previously described in healthy controls (Bozic et al., 2010) and extend the evidence for this model by demonstrating distinct bilateral and left-lateralized systems in patients with LH damage. In the dual model, bilateral perisylvian regions provide the foundation for interpretation and processing of speech. Within this context, left-lateralized perisylvian regions support core linguistic processing of morpho-syntax and sentential syntax. The dual model is based in part on the observation that agrammatic patients often have preserved semantics (Hagoort et al., 2003; Crinion and Price, 2005; Tyler and Marslen-Wilson, 2008) and predicts RH support of semantics in the context of LH damage that impairs syntax. The current findings confirm this prediction and add to existing neuropsychological findings using a novel analysis that integrates performance, tissue integrity, and neural activity to show evidence of two functionally dissociable neural systems in the same group of patients.

The dual model, supported by the current findings, suggests refinement of a recently proposed hierarchical model of mechanisms for recovery from aphasia in patients with LH damage (Heiss and Thiel, 2006). In this model, the best recovery is afforded by full or partial recuperation of the damaged left hemisphere. With increasing LH damage, RH regions may support function, but with less efficiency than their LH homologues. Our findings show that this model applies differentially to the LH core linguistic network and the bilateral semantic network. Syntactic processing is a strong case of the hierarchical model: performance depends entirely upon an intact LH, with RH recruitment unable to support recovery after damage to LH (Tyler et al., 2010b, 2011). In contrast, semantic processing appears to be supported efficiently by RpSTG/MTG even when the LpSTG/MTG is damaged. This qualification of the hierarchical model underlines the importance of using linguistically specific imaging paradigms when investigating the role of RH in supporting language recovery. A full model for language recovery must account separately for left-lateralized core linguistic processes, such as syntax, and bilaterally supported processes, such as sentential semantics.

In conclusion, using a novel multivariate analysis, we differentiated distinct neural systems supporting syntax and semantics, and each system was associated with a distinct mechanism of recovery from LH damage. Syntax depends critically on spared activity and tissue in LIFG and posterior LpSTG/MTG and the white matter connecting them. In contrast, semantics may be supported by activity in RpSTG/MTG, regardless of LH damage, as a result of bilateral, degenerate representation of semantic processes in the intact brain. These results support a neurocognitive framework with distinct systems supporting core linguistic and general communicative processes, and emphasize the importance of linguistic specificity in imaging investigations of language recovery in patients.

References


