BRAIN
A JOURNAL OF NEUROLOGY

The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke

Alexander P. Leff,[1] Thomas M. Schofield,[1] Jennifer T. Crinion,[2] Mohamed L. Seghier,[1] Alice Grogan,[1,3] David W. Green[4] and Cathy J. Price[1]

1 Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London WC1N 3BG, UK
2 Institute of Cognitive Neuroscience, UCL, London, WC1N 3AR, UK
3 Department of Human Communication Sciences, University of Sheffield, S10 2TA, UK
4 Research Department of Cognitive, Perceptual and Brain Sciences, UCL, London WC1E 6BT, UK

Correspondence to: Alexander P. Leff,
Wellcome Trust Centre for Neuroimaging,
Institute of Neurology,
UCL, 12 Queen Square,
London WC1N 3BG,
UK
E-mail: address: a.leff@fil.ion.ucl.ac.uk

Competing theories of short-term memory function make specific predictions about the functional anatomy of auditory short-term memory and its role in language comprehension. We analysed high-resolution structural magnetic resonance images from 210 stroke patients and employed a novel voxel based analysis to test the relationship between auditory short-term memory and speech comprehension. Using digit span as an index of auditory short-term memory capacity we found that the structural integrity of a posterior region of the superior temporal gyrus and sulcus predicted auditory short-term memory capacity, even when performance on a range of other measures was factored out. We show that the integrity of this region also predicts the ability to comprehend spoken sentences. Our results therefore support cognitive models that posit a shared substrate between auditory short-term memory capacity and speech comprehension ability. The method applied here will be particularly useful for modelling structure–function relationships within other complex cognitive domains.

Introduction

Normal speech consists of sequences of spoken representations of words, conveyed from the speaker to the listener. Successfully comprehending the meaning of discourse depends upon the ability of the listener to maintain these representations for sufficient time to allow for their combination into meaningful sentences.

Spoken discourse therefore may depend on auditory short-term memory. Most models of auditory short-term memory make a distinction between the processes involved in representing the to-be-remembered items and the processes involved in maintaining these representations. Baddeley's influential model of short-term memory dedicates discrete components to each of these processes; the 'phonological short-term store' acting as a

time-limited representational storage buffer, and the 'articulatory loop' acting to refresh or 'revivify' the representations in the store via a process of active rehearsal (Baddeley, 2003). According to the Baddeley model, there is no necessary relationship between auditory short-term memory and language comprehension, the emphasis being instead upon the necessity of these systems for language acquisition (Baddeley, 2003). A competing model, first proposed by Cowan (1999), does not specify a dedicated storage buffer, but portrays short-term memory as an attentionally modulated activation of existing long-term representations of to-be-remembered items (Chen and Cowan, 2009). This model predicts a relationship between auditory short-term memory capacity and performance on spoken comprehension tasks that require multiple lexical representations to be active. Despite the fundamental differences between these (and other) models of short-term memory, all models accept that there are capacity limits to auditory short-term memory; only a limited number of words or other items can be actively represented in the brain at any one time. The aim of this study was to identify the anatomical structures that support auditory short-term memory capacity and to determine if these regions also support speech comprehension ability.

Auditory short-term memory capacity can be measured by digit span. In the digit-span task, subjects are presented with a list of spoken digits and immediately after the end of presentation are required to repeat back this list in the same order it was presented. The maximum number of items correctly recalled by the subject is taken as representative of the capacity of that subject's short-term auditory memory store. The usual score for neurologically normal subjects is in the range of 5–9 items (Miller, 1956).

The capacity of the short-term store is thus easily measured in normal subjects, but it is problematic to look for a relationship between this measure and speech comprehension abilities. This is primarily due to ceiling effects, i.e. although there is some variability within the normal population in terms of capacity, there is little variability within this population in terms of speech comprehension ability. Normal subjects, by definition, are able to understand the reasonably long and complex sentences that are a feature of normal spoken discourse.

Within a population of brain-damaged individuals there is typically greater behavioural variability than in healthy subjects in both digit-span score and the ability to comprehend spoken sentences. The problem with using this population is that we can no longer rely upon digit-span scores alone as a reliable index of store capacity. This is because patients may perform below normal levels on digit-span tests for reasons unrelated to memory capacity. Some subjects may have a normal memory capacity but have other cognitive impairments that lead to difficulties perceiving the stimuli or forming and producing the correct spoken response. To avoid this problem we factored out scores on a range of tests of other components required for performing digit span but not related to the capacity of auditory short-term memory, e.g. speech production, speech perception and executive control processing. Our aim was to isolate the capacity of auditory short-term memory as much as possible.

Previous patient studies have found dissociations between working memory and speech comprehension ability after brain damage (Warrington and Shallice, 1969; Shallice and Butterworth, 1977). These landmark studies have been used to argue for the independence of speech comprehension from auditory memory capacity, on the basis that if patients exist who have 'pure' auditory short-term memory difficulties in the presence of normal speech comprehension, then normal speech comprehension does not rely upon intact auditory memory. However, although these patients perform within the normal range for comprehension of single words and simple sentences, they do have problems comprehending longer, more complex sentences. To our knowledge there is only one example in the literature of a patient with impaired digit span who retained the ability to comprehend long, complex sentences within the normal range (Butterworth *et al.*, 1986). Importantly, this patient's memory impairment was not particularly severe (a span of four items) and was of developmental origin rather than acquired in later life as a result of brain injury. While such cases are fascinating they also appear to be exceptional. The majority of patients with acquired short-term memory impairment are reported to have accompanying speech comprehension deficits (Shallice and Vallar, 1990; Burgio and Basso, 1997; Dronkers *et al.*, 2004; Martin and Ayala, 2004). This is probably because speech comprehension and the capacity of auditory short-term memory are either structurally or functionally related. In the former case, neural structures supporting speech comprehension are distinct from those supporting auditory short-term memory but are situated close enough to be damaged by neurological disorders such as stroke. The alternative explanation is that the regions supporting auditory short-term memory capacity are a subset of those involved in conversational speech comprehension.

To investigate these alternative accounts, we used high-resolution lesion mapping techniques to identify where the structural integrity of cortex determines auditory short-term memory capacity in a large group ($n = 210$) of stroke patients. We did not classify patients as impaired or unimpaired on the basis of any behavioural measures because, in our opinion, it is possible to draw more robust conclusions about normal cognitive function by making full use of the behavioural variability present within a large, heterogeneous group. Our only selection criterion was that patients had a digit span of at least two. This criterion was used to exclude patients with severe speech output difficulties. Cortical areas associated with digit span were identified by using a multiple-regression analysis that allowed us to focus on auditory short-term memory capacity while controlling for co-occurring deficits in single word speech perception, comprehension and production. Spatial resolution was increased by (i) including patients with a wide range of lesion sites, (ii) acquiring MRI T$_1$ weighted images with small, isotropic voxels (1 mm^3) and (iii) using advanced spatial normalization routines.

Multi-patient lesion-mapping analyses have been used previously to investigate relationships between brain structure and other cognitive functions (Bates *et al.*, 2003; Dronkers *et al.*, 2004; Glascher *et al.*, 2009), but our method differs from these approaches in three critical ways: First, we include a series of behavioural scores from different language tasks in order to control for key 'general' cognitive components that are involved in performing digit span, such as speech perception and production,

as well as central executive processes. Second, rather than classify each voxel in a binary manner as either a lesioned voxel or a normal one, we preserve the variability inherent in structural data, treating grey matter density as a continuous variable. Third, for a qualitative display of lesion overlap maps and estimation of lesion volume, we identified each patient's lesion(s) using an automated process (Seghier et al., 2008) rather than attempting to identify the lesion boundary 'by eye' as in previous studies, a method that is both laborious and operator-dependent (Filippi et al., 1995; Ashton et al., 2003). This automated process relies on an outlier detection algorithm (i.e. after brain segmentation, each grey matter voxel is compared with the mean and standard error of a group of 64 healthy controls in standard space).

Having identified brain regions where damage covaries with auditory short-term memory capacity, we then investigated whether the structural integrity of these regions also determines spoken discourse comprehension performance. Our hypothesis was that grey matter voxel values in regions supporting auditory short-term memory capacity would correlate with performance on tests of spoken sentence comprehension. The alternative hypothesis was that there would be no significant correlation between voxel values in these regions and tests of spoken sentence comprehension.

Methods

Subjects

Two hundred and ten patients were identified from a large, ongoing research database compiled for studying the structure–function relationships between language and brain integrity. All patients on the database selected for this study had: (i) a history of stroke (single or multiple); (ii) a high-resolution T_1 MRI brain scan; (iii) behavioural measures pertinent to the main study question. They had no other neurological or major psychiatric disorder. One hundred and twenty-eight (61%) were male and 186 (89%) were right-handed, with a mean age of 59 years (SD = 14). They were not selected on the basis of their lesion site, therefore the sample included patients with a variety of left and right hemisphere strokes, although left hemisphere stroke predominated (Fig. 1). There was considerable variability in lesion size between subjects, with a median lesion volume of 25 cm^3 (interquartile range = 2–98 cm^3). English was either their first (84%) or dominant (16%) language; that is, they were all premorbidly fluent in English. The median time after stroke upon testing was 35 months (interquartile range = 12–74 months), with a skew towards relatively recent stroke (26% ⩽ 12 months) The minimum time since stroke was >1 month to avoid testing patients in the acute and subacute phases. All participants gave written consent and the study was approved by the National Hospital for Neurology and Neurosurgery and Institute of Neurology joint research ethics committee.

Behavioural assessments

All 210 patients' language abilities were assessed using the comprehensive aphasia test (Swinburn et al., 2004).

Digit span

Subjects were presented with a list of spoken digits (at a rate of ~1 digit per second) and immediately after presentation had to repeat back the list in the same order. The score for each subject was the maximum number of digits correctly repeated. The maximum score for this test is seven. All 210 patients included in the study scored two or more. Patients with a capacity of zero or one were not selected for the study as their digit span could have been caused by other reasons (e.g. severe speech output disorder or failure to comprehend the test). Digit span engages speech perception and production; in terms of Baddeley's model it involves the phonological short-term store, but probably not articulatory rehearsal processes as there is no evidence that rehearsal occurs at this presentation rate in the absence of a retention interval (Tan and Ward, 2008). In Cowan's model, digit span involves the attentionally modulated activation of multiple long-term representations of numbers. According to both models the item capacity of auditory short-term memory is examined by this test.
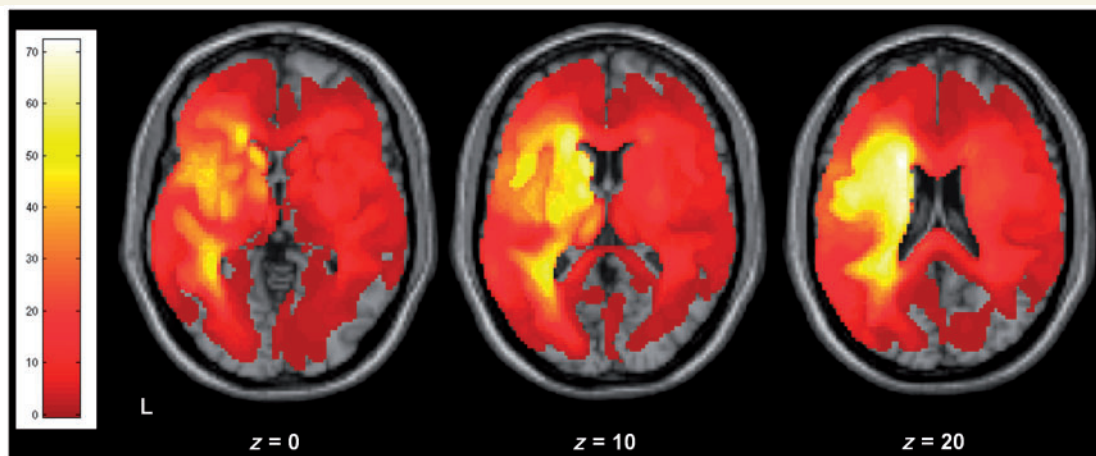
Figure 1 Lesion overlap map for the 210 patients shown for three different axial slices in standard MNI space. The colour bar on the left shows the number of patients who have a lesion including that particular voxel (range 1–73).

Control variables

We wanted to control for the various cognitive components of digit span that are not specific to auditory short-term memory capacity such as: auditory perception of speech, central linguistic processing, speech output and executive control processing. To this end we included data for the following measures from the comprehensive aphasia test: auditory repetition (word and non-word), verbal fluency and picture naming. Below we discuss the cognitive processes likely to be associated with each task, and the testing procedures. The range of possible scores is indicated in brackets.

Auditory word repetition (0–16)

Auditory repetition of 16 heard whole words (syllable range 1–3). Verbal, phonemic, neologistic and dyspraxic errors were marked as incorrect. The lowest score across the group was one. This test is probably most closely matched to digit span; it corresponds to the repetition of a single digit. Auditory input and speech production processes are similarly engaged across the two tasks; the main difference is that for single word repetition, unlike digit span, the capacity of auditory short-term memory is not stretched.

Auditory non-word repetition (0–5)

Auditory repetition of five heard non-words (syllable range 1–2). Scoring was as for word repetition. Unlike word repetition, repetition of non-words cannot be facilitated by word recognition or semantic processing; it is entirely reliant on phonological processing. The memory load may therefore be higher than that required for auditory word repetition but it is lower than that required for digit span.

Verbal fluency (0–49)

This score is a sum of two component tests: category fluency, ''Name as many animals as you can'' and phonological fluency, ''Name words beginning with the letter 's''. The subject is allowed 60 s for each test. Subjects are allowed to make articulatory errors but repeated items (perseverations) are not counted. The lowest score across the group was zero. There is no auditory perceptual component to this task (other than self-monitoring). It is designed primarily to test word retrieval and is commonly used as a test of central executive processing (Baddeley, 1996). We included this measure to control for the executive demands and sequencing necessary to perform digit span.

Picture naming (0–48)

Twenty-four black-and-white line drawings were presented for the subjects to name. Correct items, promptly named, were given a score of two. If the subject correctly named the item but only after a delay (5 s) or after self-correcting a previous response, they scored one. Incorrect responses were scored as zero. The lowest score across the group was zero. This test controls for anomia, elements of word retrieval and phonological access/sequencing (i.e. 'linguistic' processes) that may be present in performing digit span.

Comprehension variables

Written sentence comprehension (0–16)

A written sentence was presented and subjects had to select the matching line drawing from amongst three distractor drawings. Sixteen sentences were presented. The sentences were of increasing complexity, the number of words in each sentence ranged from 3 to 7. Performance was scored in the same way as for spoken sentences. This test controls for the syntactic processing demands present in spoken sentence comprehension; however as the cue sentence is available for inspection throughout each trial, any memory demands are lower than for spoken sentence comprehension.

Spoken word comprehension (0–15)

Subjects were presented with four black-and-white line drawings and a spoken word was presented. Subjects had to point to the corresponding target drawing. Alongside the target drawing there were three distractors. One was phonologically related to the target, one was semantically related and one was unrelated. Subjects could request that the word was repeated without penalty. Subjects scored one point if they pointed to the correct target. There were 15 presentations in total. This test controls for word-level lexico-semantic factors that would be present in spoken sentence comprehension, as well as response-related task demands (such as selecting a target from amongst three competing alternatives). Unlike spoken sentence comprehension, any auditory short-term memory load is low (one item) and only the lexical (noun) level of processing is assessed.

Spoken sentence comprehension (0–16)

A spoken sentence was presented and subjects had to select the matching line drawing from amongst three distractor items. Sixteen sentences were presented. The sentences were of increasing syntactic complexity, the number of words in each sentence ranged from 3 to 7. Subjects scored one point for each correct target selected. Subjects could request that the sentence was repeated without penalty.

MRI data acquisition

All subjects were scanned either on the same day or within one week of their behavioural testing. Focal grey matter density was estimated on the basis of T_1-weighted anatomical whole brain images acquired with a Siemens 1.5 T Sonata MRI scanner (Siemens Medical Systems, Erlangen, Germany). A T_1-weighted 3D modified driven equilibrium Fourier transform sequence was used to acquire 176 sagittal slices with an image matrix of 256×224 yielding a final resolution of 1 mm^3: repetition time/echo time/inversion time, 12.24/3.56/530 ms (Deichmann *et al*., 2004). The same scanner and scanning parameters were used for the acquisition of all structural images.

MRI data analysis

Structural images were pre-processed with Statistical Parametric Mapping software (SPM5: Wellcome Trust Centre for Neuroimaging: http://www.fil.ion.ucl.ac.uk/spm) running under MATLAB 7.0.4 (MathWorks, Natick, MA, USA). The images were spatially normalized into standard Montreal Neurological Institute (MNI) space using a unified segmentation algorithm optimized for use in patients with focal brain lesions. The unified segmentation algorithm is a generative model that combines tissue segmentation, bias correction and spatial normalization in the inversion of a single unified model (Ashburner and Friston, 2005). This algorithm was developed to deal with normal subjects' brains but with patients it outperforms the previous 'gold-standard' of cost-function masking (Crinion *et al*., 2007). More recently, a modified version of the tissue segmentation component has been developed to further improve identification and spatial normalization of 'brain' as opposed to 'non-brain' components by adding in an extra tissue class, 'lesion', into which outlier voxels can be classified (Seghier *et al*., 2008). These images were then smoothed with an isotropic kernel of 8 mm at full-width half maximum to increase the chance that regional effects are expressed at a spatial scale in which homologies in structural anatomy are shared over subjects. After smoothing, the value in each voxel represents the probability that

the tissue belongs to the grey matter class and not one of the three others (white matter, non-brain or lesion). Higher scores indicate higher (normal) grey matter density.

Lesion overlap map

The lesion of each patient was automatically identified using a modified unified segmentation and an outlier detection algorithm using default parameters (see procedure in Seghier et al., 2008). An outlier image was thus generated that coded the degree of abnormality of each voxel (i.e. how far the value at a given voxel is from the normal range of our 64 healthy controls). Each outlier image was thresholded into a binary image and then all the binary images were overlapped (i.e. summed across subjects) to generate the lesion overlap map. Therefore, the lesion overlap map indicates the number of patients who have a lesion at any given voxel. Note that this procedure (overlapping binary outlier images) is used here only for illustrating the spatial distribution of lesions over our 210 stroke patients, as all statistical analyses (see below) were carried out on the grey matter images using a continuous range of grey matter values.

Statistical analysis

All statistical analyses used voxel-based morphometry; that is a whole-brain, unbiased, semi-automated technique for characterizing regional differences in structural magnetic resonance images (Ashburner and Friston, 2000). Statistical analyses were performed on the smoothed grey matter images using the general linear model as implemented in SPM5. The grey matter images from the 210 patients were entered into a multiple regression model. In all analyses the linear and non-linear effects of age were excluded by including them as covariates. Time since stroke (in months) was also included as a covariate of no interest. In three of our analyses, we also entered the estimated volume of tissue loss (stroke volume) as a regressor. This was measured using the automated lesion-identification algorithm cited earlier (Seghier et al., 2008). These variables are likely to impact on cognitive performance and we wished to minimize their associated variance so that we could focus our analyses on the relationship between digit-span score and grey matter density. Tests of regression coefficients in multiple regression models are equivalent to testing corresponding partial correlations.

Figure 2 illustrates the design matrices of three of the four statistical analyses conducted here. Analysis 1 was designed to show where cortical damage leads to reduced digit span. We included digit span as the sole behavioural measure. Analysis 2 was designed to identify neural structures that determine auditory short-term memory capacity. We included five extra regressors: four scores from the comprehensive aphasia test (auditory word repetition, auditory non-word repetition, verbal fluency and picture naming) and one structural parameter (stroke volume). Analyses 3 and 4 were designed to test the structural relationship between auditory short-term memory and language comprehension. Analysis 3 included the same regressors as Analysis 2, with the addition of two comprehension scores; spoken single word comprehension and written sentence comprehension. Finally Analysis 4 (not shown) included the same regressors as Analysis 3, with the addition of spoken sentence comprehension.

We only report and discuss regions that showed significant positive effects at $P < 0.05$ after correction for multiple comparisons across the whole brain at the height (voxel) level as well as an extent threshold for each cluster of 25 voxels (i.e. a minimum cortical volume of 0.2 cm^3).

Results

Lesion overlap map

The overlap of the lesion volumes (as defined by the automated lesion detection algorithm) from the 210 patients are shown in Fig. 1, in standard MNI space.

Analysis 1: digit span as the sole behavioural variable

The variability in the structural data that are modelled by digit span, with the non-linear effects of age, and time since stroke removed, is extensive within the left hemisphere (Fig. 3 upper panel). Perisylvian regions typically associated with speech perception and production are identified, as are subcortical regions (particularly the caudate) associated with executive control processes (Table 1).

## Analysis 2: auditory short-term memory capacity

This analysis utilized the same scans and regressors as the first analysis but with five additional regressors: auditory word repetition, auditory non-word repetition, verbal fluency, picture naming and stroke volume (Fig. 2). We included these to control for processes involved in digit span but unrelated to auditory short-term memory capacity. The grey matter values within one cluster covary significantly ($P < 0.05$ after family error-wise correction for multiple comparisons across the whole brain) with auditory short-term memory capacity. This region forms part of the left posterior superior temporal gyrus and sulcus (Fig. 3 lower panel and Table 1). To illustrate the relationship between grey matter values in this region and the patient's performance on the digit-span task the principal eigenvariate of the voxels within an 8 mm radius of the peak voxel of this cluster was extracted and plotted against digit-span scores using the Statistical Package for the Social Sciences software (Fig. 4 left panel).

Analysis 3: auditory short-term memory capacity after controlling for spoken word and written sentence comprehension

This analysis utilized the same scans and regressors as Analysis 2 but with two additional ones: spoken word comprehension and written sentence comprehension (Supplementary Figure). The cluster identified by Analysis 2 remained significant (Table 1).

Analysis 4: the relationship between auditory short-term memory and spoken discourse comprehension

In order to investigate this relationship we added in spoken sentence comprehension as an additional regressor; this resulted in the loss of a significant effect of digit span.
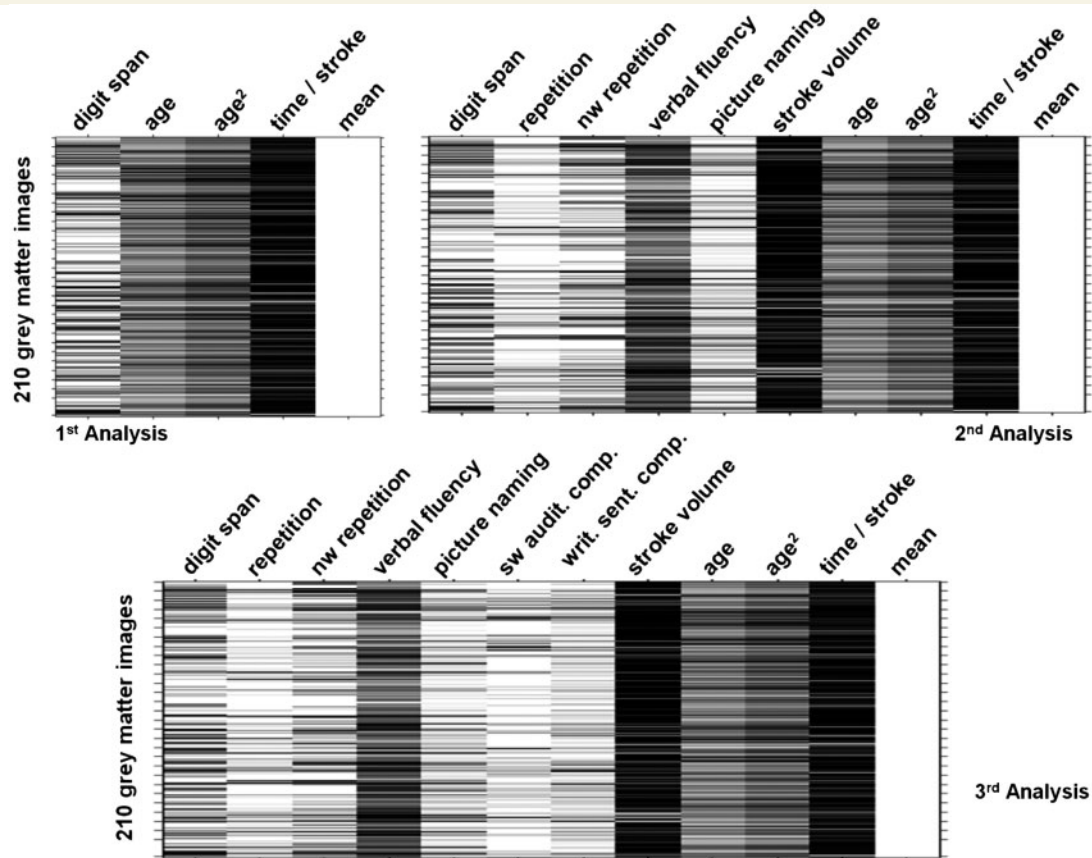
Figure 2 The three design matrices used in the regression analyses. The 210 grey matter volumes were entered as the dependent variable or data = y. The subject-specific independent variables were entered in columns (one value for each subject). For the first analysis the only behavioural score entered was digit span (first column), along with key demographic variables (age, age^2 and time since stroke); each variable was mean centred. The last column (mean) modelled the constant term of the multiregression analyses. In the second analysis four additional variables were added (repetition, verbal fluency, picture naming and stroke volume). In the third analysis two additional variables were added (spoken single word comprehension, written sentence comprehension). The order the variables were entered into the design matrix had no effect on how the residual variance was modelled.

To confirm the relationship between the structural integrity of this region of cortex and language comprehension, we used the values from the principal eigenvariate extracted from the significant cluster identified in Analysis 2 (see above) and performed partial regression against each of the measures of language comprehension (controlled variables were as Analysis 2). The right panel of Fig. 4 illustrates the significant correlation between the structural integrity of this region and spoken sentence comprehension score, $r(199) = 0.22$, $P = 0.002$ (right panel Fig. 4), but not single word comprehension, $r(199) = -0.03$, $P = 0.66$ or written sentences, $r(199) = 0.05$, $P = 0.45$. Thus grey matter density in a region important for auditory short-term memory capacity correlated more strongly with comprehension of spoken sentences than with comprehension of either single words or written sentences.

Discussion

Using the behavioural variability present within a large, heterogeneous group of stroke patients and a sensitive lesion mapping

technique that preserves variability present in structural brain images, we identified a region of the left posterior superior temporal gyrus that covaries with auditory short-term memory capacity. We found that the structural integrity of this region affects a subject's ability to comprehend spoken sentences. From this we conclude that the capacity of auditory short-term memory is an important factor in determining sentence-level speech comprehension ability.

As with every neuropsychological test, digit span calls upon several cognitive components any one of which, if impaired, could lead to a low or abnormal score. Some components are generic, e.g. speech production, speech perception and executive 'controlled' processing; that is, they are required for the task but are not related to the capacity of auditory short-term memory. Our aim was to isolate this latter factor as much as possible. The first analysis identified many regions in the left hemisphere that correlated with impaired digit span (Table 1 and top panel of Fig. 3), but these could be supporting the generic functions just alluded to, rather than auditory short-term memory function *per se*. The cortical regions identified as supporting digit span are confluent and extensive, including most of the key Perisylvian areas
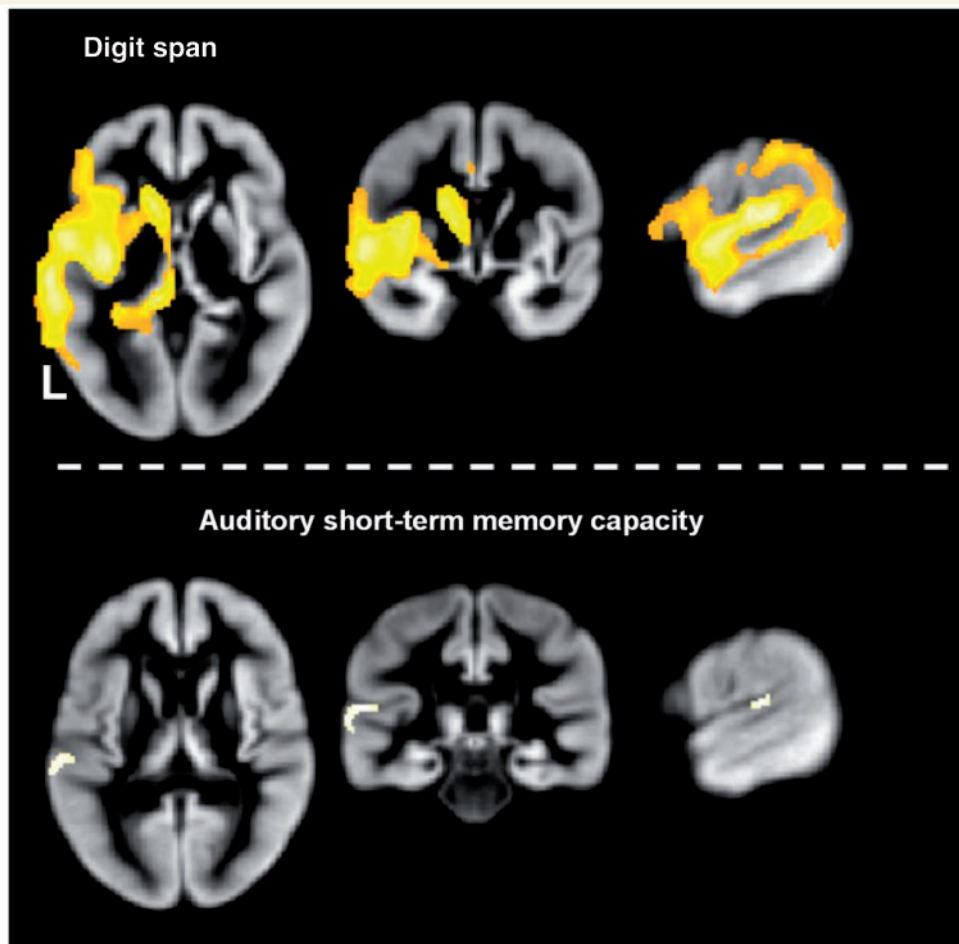
Figure 3 Results from the first (top) and second analyses (bottom) depicting where variance in the structural data significantly correlated with digit-span score. All images thresholded at a peak of $P < 0.05$ (family-wise error corrected) and extent of 25 contiguous voxels.

known to be involved in speech perception and production. Lateral frontal and parietal regions of this cortical network were identified by a study using similar techniques applied to patients with a variety of degenerative diseases (Amici *et al.*, 2007a, b). They found a correlation between grey matter density and performance on a 'digits backwards' task, which probes executive function as well as auditory short-term memory. We also identified a subcortical component involved in mediating executive control processes such as task-switching: the caudate nucleus (Cummings, 1993; Crinion *et al.*, 2006; Abutalebi *et al.*, 2008).

In the second analysis we attempted to partial out the variance in the structural data accounted for by generic cognitive factors by including five further regression variables. This allowed us to identify grey matter regions that covary with the key psychological component of digit span: the capacity of auditory short-term memory. Of the behavioural control variables added, auditory word and non-word repetition are probably the closest to digit span, in that they rely on auditory perception and recognition of speech, auditory verbal memory (albeit for a single word) and speech production. The psychological factor that digit span

almost certainly engages more than single word or non-word repetition is the capacity of auditory short-term memory. Verbal fluency requires that subjects be able to retrieve lexico-semantic representations of items and produce and sequence them. This variable was included to control for the executive control processes necessary for performing digit span, including sequential speech production and keeping track of recently spoken items. Picture naming requires no auditory input but does engage 'central' language processes including the retrieval of lexical information (i.e. retrieving the abstract linguistic representation of the object to be named), and phonological output. This task was included to increase control for anomia and any linguistic component of digit span. The final variable included in the second analysis was stroke volume. This is a structural parameter but it has behavioural ramifications. Cognitive impairments are clearly related to the site of a stroke, with relatively small strokes able to cause disproportionate impairment if they occur in strategically important brain regions; there is, however, something to be said for the old adage that ''size matters'' (Knopman *et al.*, 1984; Hillis *et al.*, 2000).

The second analysis identified a cortical region where grey matter density covaried with digit span despite the inclusion of these five extra control regressors, suggesting that the structural integrity of this region determines auditory short-term memory

Table 1 Brain regions significantly correlated with digit span (all in the left hemisphere)

Analysis 1	Co-ordinates (peak voxels)			Z-score
Digit span	x	y	z	
Posterior				
Superior temporal gyrus (posterior)	–66	–32	6	Inf
Heschl's gyrus (A1)	–54	–20	14	7.84
Superior temporal gyrus (anterior)	–58	0	–2	7.67
Superior temporal sulcus (posterior)	–60	–44	6	7.19
Middle temporal gyrus	–64	–46	–8	6.97
Mid				
Posterior insula (long insula gyrus)	–38	–10	4	7.83
Anterior insula (short insula gyrus)	–36	16	–8	6.56
Anterior				
Inferior frontal gyrus (ventral pre-motor)	–62	6	18	6.36
Inferior frontal gyrus (pars orbitalis)	–50	42	–10	6.01
Inferior frontal gyrus (pars triangularis)	–48	36	2	5.86
Sub-cortical				
Body of caudate	–20	–28	18	Inf
Head of caudate	–14	16	6	7.46
Analysis 2				
Superior temporal gyrus (posterior)	–66	–32	4	5.26
Analysis 3				
Superior temporal gyrus (posterior)	–56	–26	12	5.15

Coordinates are in MNI space. Z-scores are significant at $P < 0.05$ after correction for multiple comparisons across the whole brain. The location and extent of these activations are illustrated in Fig. 3.

capacity. This region occupies part of the superior temporal gyrus and sulcus lateral and posterior to Heschel's gyrus (Penhune *et al.*, 1996), overlapping with and extending below the most inferior portion of the planum temporale (Westbury *et al.*, 1999). Recent event-related functional MRI investigations of auditory short-term memory showed activation in this region and also in areas of the inferior parietal lobe (Becker *et al.*, 1999; Buchsbaum *et al.*, 2001). Functional imaging studies can be assumed to highlight regions that may be sufficient to perform tasks in the scanner—they measure correlations between activation and assumed task demands—but lesion-deficit studies are able to provide evidence as to which regions may be necessary for the task studied (Price *et al.*, 2006). Our results show that it is the posterior superior temporal gyrus and sulcus rather than the inferior parietal lobe that is necessary for intact auditory short-term memory capacity. Isolating the brain structures that support auditory short-term memory capacity has been relatively straight-forward in our lesion-based analysis. In contrast, identifying anatomical correlates of the 'capacity' of any sensory modality would be much more difficult to do using functional imaging studies where it would be difficult to avoid confounding capacity with task demands.

Implications for models of short-term memory

We tested the hypothesis that the capacity of auditory short-term memory correlates with spoken sentence comprehension. By finding a correlation in a single focal brain region (Analysis 2) we confirm this hypothesis. Our finding is inconsistent with models of short-term memory that segregate auditory short-term memory from the on-line comprehension of spoken language (Baddeley, 2003). It is, however, consistent with models that posit a dependency between auditory comprehension and short-term memory capacity (Cowan, 1999). Our finding is also
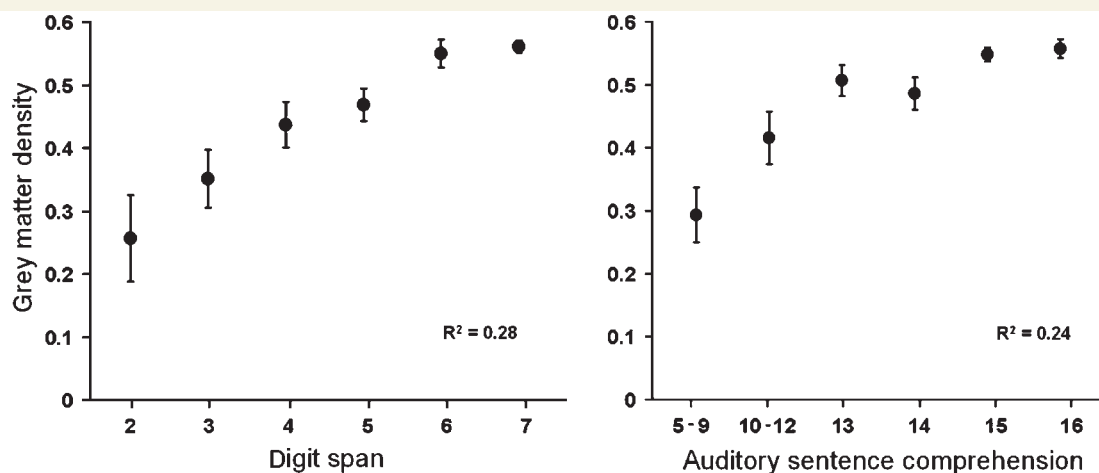
Figure 4 Means (with SEM error bars) of grey matter density plotted against scores of digit span (left panel) and spoken sentence comprehension (right panel) for the 210 patients. Grey matter density values were taken from the principal eigenvariate (i.e. a summary of values) of the cluster identified by the second analysis as supporting auditory short-term memory (MNI co-ordinates: –66 –32 4). Statistical analyses were performed on all 210 data points.

consistent with prior evidence that the posterior portion of the left temporal lobe is an important site for access to, and long-term storage of, phonological and lexical representations (Leff *et al.*, 2008; Schofield *et al.*, 2009).

Our study supports the hypothesis that auditory short-term memory is likely to have a necessary role in normal propositional speech comprehension, insofar as speech comprehension requires multiple auditory lexical representations to be concurrently active. Functional and structural imaging studies show that this region is involved in performing sentence-level spoken comprehension tasks, acting in concert with other important brain regions (Crinion *et al.*, 2003; Capek *et al.*, 2004; Spitsyna *et al.*, 2006; Amici *et al.*, 2007; Friederici *et al.*, 2009). We show that this region is necessary to perform these tasks within the normal range. Damage to this region did not, however, predict subjects' ability to comprehend written sentences of equal syntactic complexity. We therefore also provide evidence for an anatomical dissociation between sentence level auditory and visual processing, driven by the probable anatomically separable storage of the relevant long-term lexical representations.

Given that our patients were in the chronic post-stroke phase (∼36 months on average), the clinical utility of our result is that patients with damage to Perisylvian regions of the left hemisphere are likely to remain impaired on digit span (Analysis 1); although this could be due to a variety of reasons, not least damage to the speech output system. More specifically, patients with damage to the left superior temporal gyrus and sulcus lateral and posterior to primary auditory cortex may well have a deficit due to damaged auditory short-term memory capacity, specifically the ability to maintain multiple auditory lexical representations. This has implications for rehabilitation strategies especially if the patient is also impaired at spoken sentence comprehension, which, as recapitulated in our behavioural data, seems likely.

Methodological points

Our study utilized both continuous behavioural scores and continuous grey matter values as in many previous voxel based morphology studies (Stamatakis *et al.*, 2005; Amici *et al.*, 2007*a, b*; Acres *et al.*, 2009; Butler *et al.*, 2009; Taylor *et al.*, 2009). But in contrast to other approaches where damaged voxels are classified as intact or not in a binary fashion (Bates *et al.*, 2003; Dronkers *et al.*, 2004; Glascher *et al.*, 2009), we chose to employ the voxel-based morphology method for two major reasons. First, identifying damage by hand/eye is subject to human error and dependent upon the subjective opinion of the classifier. Second, it is possible that there will be brain regions that suffer from volume loss but yet may still be functioning, although perhaps not at normal levels (Kanda *et al.*, 2008). This could occur if inputs to and from the region are damaged (Kinoshita *et al.*, 2002), resulting in a loss of dendritic connections and a subsequent reduction in grey matter density (Stadelmann *et al.*, 2008). If this partial reduction in grey matter has an effect on cognitive function, then methods which assign binary values to grey matter voxels will be unable to account for it.

Conclusion

We have identified the left posterior superior temporal gyrus and sulcus as the substrate mediating the capacity of auditory short-term memory. The structural integrity of this region also predicts speech comprehension ability. Our study therefore provides novel data in support of a shared substrate for auditory short-term memory and speech comprehension.

Funding

Supplementary material

Supplementary material is available at *Brain* online.

References

Abutalebi J, Annoni JM, Zimine I, Pegna AJ, Seghier ML, Lee-Jahnke H, et al. Language control and lexical competition in bilinguals: an event-related FMRI study. Cereb Cortex 2008; 18: 1496–505.

Acres K, Taylor KI, Moss HE, Stamatakis EA, Tyler LK. Complementary hemispheric asymmetries in object naming and recognition: a voxel-based correlational study. Neuropsychologia 2009; 47: 1836–43.

Amici S, Brambati SM, Wilkins DP, Ogar J, Dronkers NL, Miller BL, et al. Anatomical correlates of sentence comprehension and verbal working memory in neurodegenerative disease. J Neurosci 2007; 27: 6282–90.

Amici S, Ogar J, Brambati SM, Miller BL, Neuhaus J, Dronkers NL, et al. Performance in specific language tasks correlates with regional volume changes in progressive aphasia. Cogn Behav Neurol 2007; 20: 203–11.

Ashburner J, Friston KJ. Voxel-based morphometry—the methods. Neuroimage. 2000; 11 (Pt 1): 805–21.

Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005; 26: 839–51.

Ashton EA, Takahashi C, Berg MJ, Goodman A, Totterman S, Ekholm S. Accuracy and reproducibility of manual and semiautomated quantification of MS lesions by MRI. J Magn Reson Imaging 2003; 17: 300–8.

Baddeley A. Working memory and language: an overview. J Commun Disord 2003; 36: 189–208.

Baddeley AD. Exploring the central executive. Q J Exp Psychol 1996; 49A: 5–28.

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, et al. Voxel-based lesion-symptom mapping. Nat Neurosci 2003; 6: 448–50.

Becker JT, MacAndrew DK, Fiez JA. A comment on the functional localization of the phonological storage subsystem of working memory. Brain Cogn 1999; 41: 27–38.

Buchsbaum BR, Hickok G, Humphreys G. Role of left superior temporal gyrus in phonological processing for speech perception and production. Cogn Sci 2001; 25: 663–78.

Burgio F, Basso A. Memory and aphasia. Neuropsychologia 1997; 35: 759–66.

Butler CR, Brambati SM, Miller BL, Gorno-Tempini ML. The neural correlates of verbal and nonverbal semantic processing deficits in neurodegenerative disease. Cogn Behav Neurol 2009; 22: 73–80.

Butterworth B, Campbell R, Howard D. The uses of short-term memory: a case study. Q J Exp Psychol A 1986; 38: 705–37.

Capek CM, Bavelier D, Corina D, Newman AJ, Jezzard P, Neville HJ. The cortical organization of audio-visual sentence comprehension: an fMRI study at 4 Tesla. Cogn Brain Res 2004; 20: 111–9.

Chen Z, Cowan N. Core verbal working-memory capacity: the limit in words retained without covert articulation. Q J Exp Psychol (Colchester). 2009; 62: 1420–9.

Cowan N. An embedded-process model of working memory. Models of working memory. Cambridge: Cambride University Press; 1999. p. 62–101.

Crinion J, Ashburner J, Leff A, Brett M, Price C, Friston K. Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. Neuroimage 2007; 37: 866–75.

Crinion J, Turner R, Grogan A, Hanakawa T, Noppeney U, Devlin JT, et al. Language control in the bilingual brain. Science 2006; 312: 1537–40.

Crinion JT, Lambon-Ralph MA, Warburton EA, Howard D, Wise RJ. Temporal lobe regions engaged during normal speech comprehension. Brain 2003; 126 (Pt 5): 1193–201.

Cummings JL. Frontal-subcortical circuits and human behavior. Arch Neurol 1993; 50: 873–80.

Deichmann R, Schwarzbauer C, Turner R. Optimisation of the 3D MDEFT sequence for anatomical brain imaging: technical implications at 1.5 and 3 T. Neuroimage 2004; 21: 757–67.

Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition 2004; 92: 145–77.

Filippi M, Horsfield MA, Bressi S, Martinelli V, Baratti C, Reganati P, et al. Intra- and inter-observer agreement of brain MRI lesion volume measurements in multiple sclerosis. A comparison of techniques. Brain 1995; 118 (Pt 6): 1593–600.

Friederici AD, Makuuchi M, Bahlmann J. The role of the posterior superior temporal cortex in sentence comprehension. Neuroreport 2009; 20: 563–8.

Glascher J, Tranel D, Paul LK, Rudrauf D, Rorden C, Hornaday A, et al. Lesion mapping of cognitive abilities linked to intelligence. Neuron 2009; 61: 681–91.

Hillis AE, Barker PB, Beauchamp NJ, Gordon B, Wityk RJ. MR perfusion imaging reveals regions of hypoperfusion associated with aphasia and neglect. Neurology 2000; 55: 782–8.

Kanda T, Ishii K, Uemura T, Miyamoto N, Yoshikawa T, Kono AK, et al. Comparison of grey matter and metabolic reductions in frontotemporal dementia using FDG-PET and voxel-based morphometric MR studies. Eur J Nucl Med Mol Imaging 2008; 35: 2227–34.

Kinoshita T, Moritani T, Shrier DA, Wang HZ, Hiwatashi A, Numaguchi Y, et al. Secondary degeneration of the substantia nigra and corticospinal tract after hemorrhagic middle cerebral artery infarction: diffusion-weighted MR findings. Magn Reson Med Sci 2002; 1: 175–8.

Knopman DS, Selnes OA, Niccum N, Rubens AB. Recovery of naming in aphasia: relationship to fluency, comprehension and CT findings. Neurology 1984; 34: 1461–70.

Leff AP, Schofield TM, Stephan KE, Crinion JT, Friston KJ, Price CJ. The cortical dynamics of intelligible speech. J Neurosci 2008; 28: 13209–15.

Martin N, Ayala J. Measurements of auditory-verbal STM span in aphasia: effects of item, task, and lexical impairment. Brain Lang 2004; 89: 464–83.

Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. Psychol Rev 1956; 63: 81–97.

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

Price CJ, Crinion J, Friston KJ. Design and analysis of fMRI studies with neurologically impaired patients. J Magn Reson Imaging 2006; 23: 816–26.

Schofield TM, Iverson P, Kiebel SJ, Stephan KE, Kilner JM, Friston KJ, et al. Changing meaning causes coupling changes within higher levels of the cortical hierarchy. Proc Natl Acad Sci USA 2009; 106: 11765–70.

Seghier ML, Ramlackhansingh A, Crinion J, Leff AP, Price CJ. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage 2008; 41: 1253–66.

Shallice T, Butterworth B. Short-term memory impairment and spontaneous speech. Neuropsychologia 1977; 15: 729–35.

Shallice T, Vallar G. The impairment of auditory–verbal short-term storage. Cambridge: Cambridge University Press; 1990.

Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJ. Converging language streams in the human temporal lobe. J Neurosci 2006; 26: 7328–36.

Stadelmann C, Albert M, Wegner C, Bruck W. Cortical pathology in multiple sclerosis. Curr Opin Neurol 2008; 21: 229–34.

Stamatakis EA, Tyler LK. Identifying lesions on structural brain images—validation of the method and application to neuropsychological patients. Brain Lang 2005; 94: 167–77.

Swinburn K, Porter G, Howard D. Comprehensive Aphasia Test. Hove: Psychology Press; 2004.

Tan L, Ward G. Rehearsal in immediate serial recall. Psychon Bull Rev 2008; 15: 535–42.

Taylor KI, Stamatakis EA, Tyler LK. Crossmodal integration of object features: voxel-based correlations in brain-damaged patients. Brain 2009; 132 (Pt 3): 671–83.

Warrington EK, Shallice T. The selective impairment of auditory verbal short-term memory. Brain 1969; 92: 885–96.

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