

Areas of left perisylvian cortex mediate auditory–verbal short-term memory

Michael Koenigs^{a,*}, Daniel J. Acheson^b, Aron K. Barbey^c, Jeffrey Solomon^d, Bradley R. Postle^{a,b}, Jordan Grafman^{e,*}

^a Department of Psychiatry, University of Wisconsin–Madison, 6001 Research Park Blvd., Madison, WI 53719, USA

^b Department of Psychology, University of Wisconsin–Madison, 1202 West Johnson St., Madison, WI 53706, USA

^c Cognitive Neuroscience Section, National Institutes of Neurological Disorders and Stroke, National Institutes of Health, 10 Center Dr., Bethesda, MD 20892, USA

^d Medical Numerics, Inc., 20410 Observation Dr., Suite 210, Germantown, MD 20876, USA

^e Traumatic Brain Injury Research Laboratory, Kessler Foundation Research Center, 1199 Pleasant Valley Way, West Orange, NJ 07052, USA

ARTICLE INFO

Article history:

Received 30 June 2011

Received in revised form 8 September 2011

Accepted 11 September 2011

Available online 16 September 2011

Keywords:

Short-term memory

Language

Lesion

Neuropsychology

ABSTRACT

A contentious issue in memory research is whether verbal short-term memory (STM) depends on a neural system specifically dedicated to the temporary maintenance of information, or instead relies on the same brain areas subserving the comprehension and production of language. In this study, we examined a large sample of adults with acquired brain lesions to identify the critical neural substrates underlying verbal STM and the relationship between verbal STM and language processing abilities. We found that patients with damage to selective regions of left perisylvian cortex – specifically the inferior frontal and posterior temporal sectors – were impaired on auditory–verbal STM performance (digit span), as well as on tests requiring the production and/or comprehension of language. These results support the conclusion that verbal STM and language processing are mediated by the same areas of left perisylvian cortex.

© 2011 Elsevier Ltd. All rights reserved.

Verbal short-term memory – the ability to actively maintain verbal information “in mind” for brief periods – is a cornerstone of human cognition. The neural substrates of verbal short-term memory (STM) have been a topic of empirical study and theoretical debate for decades; yet to date there is no clear consensus regarding the brain areas or cognitive subprocesses that underlie verbal STM. A central theme of the debate has been whether verbal STM depends on a neural system dedicated specifically to short-term maintenance of information (e.g. the “phonological store”) (Baddeley, 1986, 2007; Baddeley & Hitch, 1974; Shallice & Warrington, 1970) or is instead mediated predominantly by the brain’s language comprehension and production architecture (Acheson & MacDonald, 2009; Allport, 1984; Buchsbaum & D’Esposito, 2008; Martin & Saffran, 1997).

Initial support for the dedicated system view of verbal STM was provided by patients who exhibit marked deficits in verbal STM (e.g. impaired digit span), while exhibiting relatively intact language production and comprehension (Shallice & Butterworth, 1977; Shallice & Warrington, 1970; Vallar & Baddeley, 1984; Warrington, Logue, & Pratt, 1971; Warrington & Shallice, 1969). Early neuroanatomical studies associated such STM deficits with brain damage involving the inferior parietal lobe (Warrington

et al., 1971; Warrington & Shallice, 1969), possibly extending into superior temporal regions (Shallice & Vallar, 1990). Many early neuroimaging studies corroborated this anatomical locus for STM storage by demonstrating STM load sensitivity in parietal areas superior to the Sylvian fissure (Petersen, Fox, Posner, Mintun, & Raichle, 1988; Smith & Jonides, 1997). Importantly, the parietal regions identified in these studies do not fit within the more classically defined “language” regions (Geschwind, 1965; Wernicke, 1874), thus supporting the idea of independent neural substrates for language and verbal STM.

Despite the apparent dissociation between verbal STM and language processing suggested by these studies, substantial evidence exists in favor of a common neural locus to both. For instance, patients with relatively isolated STM deficits are quite rare, and typically exhibit some form of concomitant language deficit when tested thoroughly (Allport, 1984; Caplan & Waters, 1990; Papagno, Cecchetto, Reati, & Bello, 2007). And in addition to STM patients exhibiting language deficits, the converse also holds, in that patients with language deficits often exhibit deficits in STM task performance (Martin & Saffran, 1997). Furthermore, many neuroimaging studies have failed to support the hypothesis that regions of the inferior parietal lobe serve as the locus of verbal STM maintenance (reviewed in Buchsbaum & D’Esposito, 2008), and instead implicate regions of the posterior superior temporal cortex (Buchsbaum & D’Esposito, 2008) and inferior frontal cortex (Chein & Fiez, 2001). Moreover, strong support for a common neural locus for STM and language processing is provided by two recent studies,

* Corresponding authors.

E-mail addresses: mrkoenigs@wisc.edu (M. Koenigs), jgrafman@kesslerfoundation.org (J. Grafman).

one involving stroke patients (Leff et al., 2009) and the other utilizing repetitive transcranial magnetic stimulation (Acheson, Hamidi, Binder, & Postle, 2011). Both studies demonstrate that the proper function of posterior superior temporal gyrus is critical for both STM and language processes. Results such as these challenge the dedicated “phonological store” model, and suggest that the same brain regions (in particular, left inferior frontal and left posterior temporal cortices) may mediate both language processing and verbal STM.

In the present study we address the neural substrates underlying verbal STM by analyzing neuropsychological and neuroimaging data from a unique sample of almost 200 individuals with acquired brain lesions. We perform two complementary sets of analyses: one in which we classify patients on the basis of verbal STM ability and look for systematic differences in lesion location (with no a priori anatomical hypothesis), and another in which we classify patients on the basis of lesion location and look for systematic differences in verbal STM ability (in order to test the specific hypothesis that left inferior frontal cortex and left posterior temporal cortex are critical for both verbal STM and language abilities).

1. Materials and methods

1.1. Participants

Participants were drawn from the Phase 3 Vietnam Head Injury Study (VHIS) registry, which includes American veterans who suffered brain damage from penetrating head injuries in the Vietnam war ($n = 199$), as well as neurologically healthy Vietnam veterans ($n = 55$).

1.2. Lesion analysis

CT data were acquired during the Phase 3 testing period. Axial CT scans without contrast were acquired at Bethesda Naval Hospital on a GE Medical Systems Light Speed Plus CT scanner in helical mode. Images were reconstructed with an in-plane voxel size of $0.4 \text{ mm} \times 0.4 \text{ mm}$, overlapping slice thickness of 2.5 mm and a 1 mm slice interval. Lesion location and volume were determined from CT images using the Analysis of Brain Lesion (ABLE) software (Makale et al., 2002; Solomon, Raymont, Braun, Butman, & Grafman, 2007) contained in MEDx v3.44 (Medical Numerics, Germantown, MD) with enhancements to support the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). Lesion volume was calculated by manual tracing of the lesion in all relevant slices of the CT image in native space, then summing the traced areas and multiplying by slice thickness. The CT image of each subject's brain was then spatially normalized to a CT template brain image in MNI space. A trained neuropsychiatrist performed the manual tracing, which was then reviewed by J.G., who was blind to the results of the neuropsychological testing.

1.3. Neuropsychological tests

1.3.1. Short-term memory

Digit Span is the quintessential neuropsychological test of verbal STM. In Digit Span (Wechsler, 1997), the subject hears a sequence of digits and attempts to repeat the sequence in order. Sequence length is increased until the subject can no longer correctly repeat the sequence. The maximum sequence length for each subject is interpreted as the subject's Digit Span. In Spatial Span (Wechsler, 1997), the subject watches the examiner tap a sequence of locations on a board and attempts to repeat the tapping sequence in order. Sequence length is increased until the subject can no longer correctly repeat the sequence. The maximum sequence length for each subject is interpreted as the subject's Spatial Span.

1.3.2. Verbal processing

In the Boston Naming Test (Goodglass & Kaplan, 2000), the subject views a series of 60 ink drawings of objects and attempts to name aloud each item. In the Token Test (Spellacy & Spreen, 1969), the subject attempts to follow a series of 16 oral commands (e.g. “Put the green square between the red square and the blue circle”). Each command is valued 6–8 points, corresponding to the complexity of the command, for a total possible score of 100. In the Vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1997), the subject is asked to orally define a series of 33 words (e.g. “What does [blank] mean?”). The subject receives a score 0–2 for each item depending on the accuracy and precision of the response.

1.3.3. Non-verbal processing

In Matrix Reasoning (Wechsler, 1997), the subject completes a series of visual pattern completion problems. In the Tower Test (Delis, Kaplan, & Kramer, 2001), subjects attempt to rearrange discs on a set of pegs to match a model.

2. Results

To determine which areas of the brain are critical for auditory–verbal STM capacity, we performed a series of complementary analyses.

2.1. Lesions associated with Digit Span impairment

First we generated a descriptive (i.e. non-statistical) representation of lesion–deficit relationships by selecting all patients with significant impairments in Digit Span – scores of 4 or less ($n = 12$; Fig. 1C). This threshold identifies individuals with Digit Spans that are at least two standard deviations below the mean Digit Span of the neurologically healthy comparison group. (It is important to note that all VHIS participants had a Digit Span of at least 3, indicating that Digit Span impairment in these individuals is not simply due to a gross inability to comprehend or produce speech.) As can be seen in Fig. 1C, the most severe impairments in auditory–verbal STM capacity were associated with relatively large lesions involving left perisylvian cortex.

2.2. Voxel-based lesion symptom mapping (VLSM)

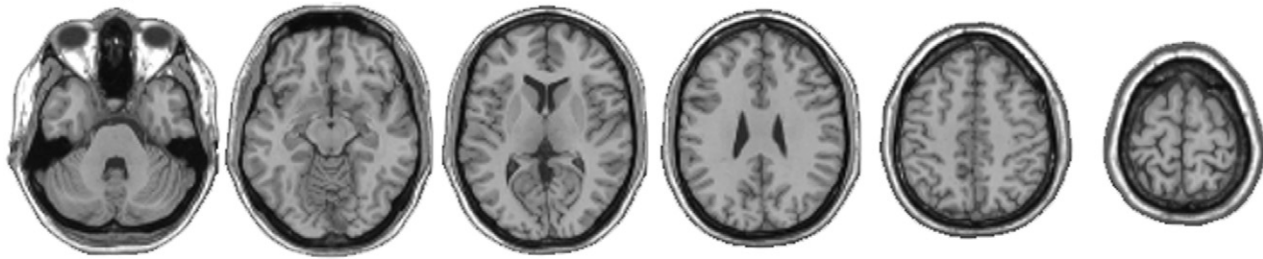
To test for statistically significant relationships between lesion location and Digit Span performance, we performed a follow-up analysis employing voxel-based lesion symptom mapping (VLSM) (Bates et al., 2003). For this analysis we computed a map of z values, where the value of each voxel represents the z score based on a Mann–Whitey U -test that compares the Digit Spans among individuals with damage to that voxel to the Digit Spans among individuals without damage to that voxel. To determine the critical z value for statistical significance, we used a false discovery rate (FDR) correction for multiple comparisons with $q = 0.03$ (Genovese, Lazar, & Nichols, 2002), which resulted in a critical z value of 3.5. This analysis revealed statistically significant voxels only in left perisylvian cortex (Fig. 1D). The largest cluster of significant voxels was located in superior and middle temporal gyri in the mid- to posterior region of left temporal lobe, while smaller clusters were located in left inferior frontal gyrus and left inferior parietal lobule.

These first two analyses, neither of which required any a priori anatomical hypothesis, both identified areas of left perisylvian cortex as critical for Digit Span performance. The extent of left perisylvian damage identified by these analyses raises the question of whether subregions in this area may play independently critical roles in verbal STM.

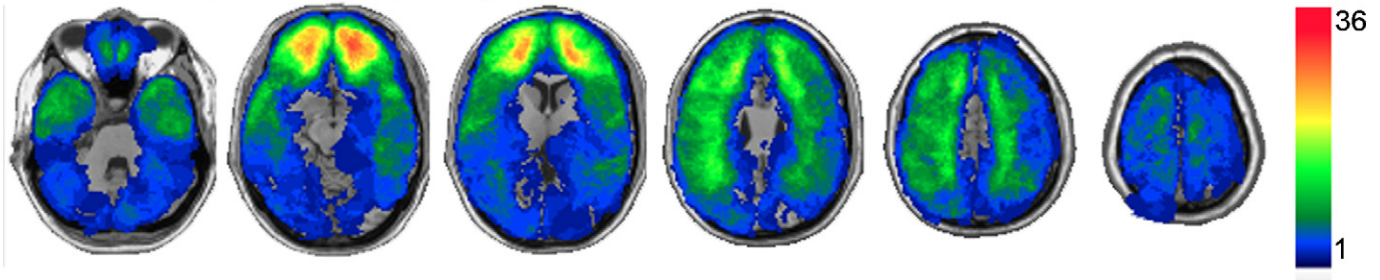
2.3. ROI analyses: Left inferior frontal and posterior temporal cortices

To assess the relative importance of discrete perisylvian subregions for auditory–verbal STM capacity, we performed a hypothesis-driven ROI analysis in which we grouped patients on the basis of lesion location and compared Digit Span performance between groups. We first selected patients with lesions primarily involving left inferior frontal cortex (L InfFront lesion group; $n = 10$; Fig. 2) or left posterior temporal cortex (L PostTemp lesion group; $n = 12$; Fig. 3). Patients with damage involving both areas (L InfFront and L PostTemp; $n = 3$) were excluded from this analysis. To determine whether any observed impairments in Digit Span performance in the L PostTemp and L InfFront groups are specific to lesions in those areas, rather than damage to adjacent areas, we then selected several additional comparison groups. As can be observed in Fig. 2, individuals in the L InfFront lesion group had damage that frequently involved adjacent areas of prefrontal cortex (anterior, medial, and/or dorsal to IFG). Therefore, for a comparison group, we selected individuals with left prefrontal lesions adjacent

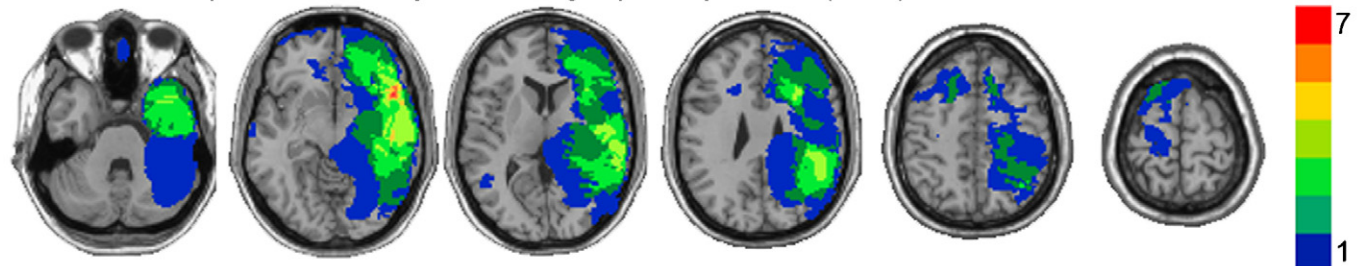
A. Healthy adult brain



B. Lesion overlap: All subjects (n=196)



C. Lesion overlap: Subset of subjects with digit span impairment (n=12)



D. VLSM analysis: Areas where damage is associated with significantly lower digit span

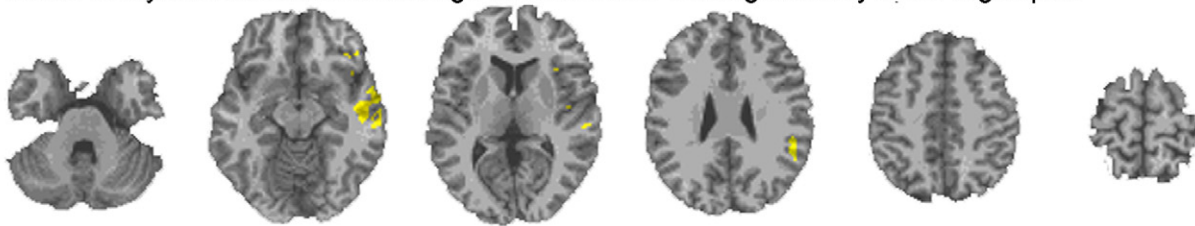


Fig. 1. (A) Transverse slices of a healthy adult brain, for reference. Corresponding slices are shown in (B–C). In all slices the left hemisphere is on the reader's right. (B) Lesion overlap of all subjects ($n = 196$) who completed the Digit Span test and CT imaging. The color bar indicates the number of overlapping lesions at each voxel. Maximal overlap occurs in ventral prefrontal cortex bilaterally. (C) Lesion overlap of brain-injured subjects with the most severe impairments in auditory–verbal working memory capacity ($n = 12$ with Digit Span ≤ 4). The color bar indicates the number of overlapping lesions at each voxel. Maximal overlap occurs in left perisylvian areas. (D) Mann–Whitney statistical map (VLSM analysis). Color indicates voxels where damage is associated with significantly lower Digit Span (corrected for multiple comparisons). Significant voxels are found only in left perisylvian cortex. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to, but not including, IFG (L nonIFG PFC lesion group; $n = 11$; Fig. 2). As can be observed in Fig. 3, individuals in the L PostTemp lesion group had damage that frequently involved anterior temporal or parietal cortex. Therefore, for additional comparison groups, we selected individuals with lesions that included left anterior temporal lobe (L AntTemp lesion group; $n = 8$; Fig. 3) or left parietal cortex (L Par lesion group; $n = 5$; Fig. 3), but spared left posterior temporal cortex.

We compared the Digit Span scores of each left hemisphere lesion group with a demographically matched group of neurologically healthy subjects ($n = 53$), as well as with groups of patients with lesions involving the respective homotopic areas of right

hemisphere (R InfFront [$n = 8$], R PostTemp [$n = 10$], R nonIFG PFC [$n = 14$], R AntTemp [$n = 8$], and R Par [$n = 8$]) (Fig. 4). Demographic and background data for each group are presented in Table 1. Because of the relatively small sample sizes in the lesion groups, planned comparisons were conducted with the Mann–Whitney U -test. The L InfFront group exhibited significantly lower Digit Span scores than the neurologically healthy group ($z = -3.1$; $p = 0.002$), the R InfFront group ($z = -2.4$; $p = 0.02$), and the L nonIFG PFC group ($z = -2.6$; $p = 0.01$). Similarly, the L PostTemp group exhibited significantly lower Digit Span scores than the neurologically healthy group ($z = -3.7$; $p < 0.001$), the R PostTemp group ($z = -2.7$; $p = 0.007$), the L AntTemp group ($z = -2.5$; $p = 0.01$), and the L Par

Table 1

Demographic and background data for subject groups. Unless otherwise noted, group means are presented with standard deviations in parentheses. Age, age at the time of the Phase 3 evaluation; Yrs Edu, years of education; Pre-combat AFQT, percentile scores on the Armed Forces Qualification Test, a battery of tests measuring basic cognitive function at the time of enlistment (pre-injury). For each pair of lesion groups (right and left hemisphere lesions), there were no significant differences on any of the measures (all p values > 0.15).

Group	<i>n</i>	Age	Yrs Edu	% right-handed	Pre-combat AFQT	Lesion size (cm ³)
Inffront						
L	9	56.9 (1.5)	13.9 (1.9)	78	54.7 (27.8)	71.0 (51.9)
R	8	58.6 (2.6)	13.9 (2.1)	88	60.0 (21.6)	51.6 (39.5)
nonIFG PFC						
L	11	57.8 (1.7)	15.0 (1.9)	91	63.8 (25.5)	22.0 (11.0)
R	14	57.1 (2.9)	14.4 (2.1)	100	58.6 (22.2)	28.4 (23.1)
PostTemp						
L	12	58.5 (2.0)	14.6 (2.6)	83	62.3 (23.8)	49.8 (56.2)
R	10	58.1 (2.5)	15.4 (2.3)	70	66.6 (22.4)	43.5 (32.4)
AntTemp						
L	8	59.4 (2.2)	15.9 (2.6)	88	70.1 (27.0)	22.0 (18.0)
R	8	58.9 (3.1)	14.5 (4.5)	88	58.3 (32.4)	23.4 (12.8)
Par						
L	5	57.4 (1.1)	14.0 (2.3)	60	76.2 (10.3)	59.2 (37.3)
R	8	57.3 (1.8)	13.9 (2.1)	88	56.5 (30.4)	37.7 (24.8)
No lesion	53	59.0 (3.4)	15.2 (2.5)	77	65.4 (22.9)	n/a

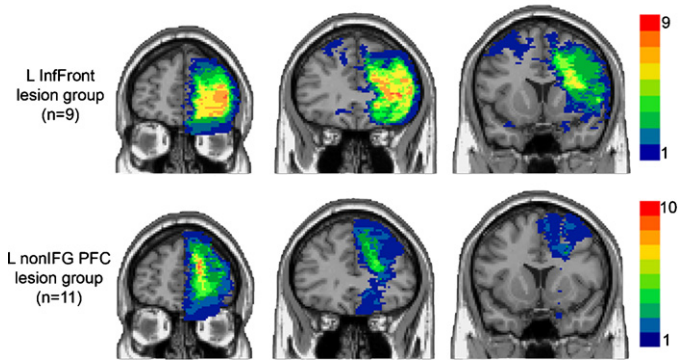


Fig. 2. Lesion overlaps for the L Inffront group (top row) and L nonIFG PFC group (bottom row). Coronal slices, from left to right: $y = 54, y = 34, y = 14$. Color indicates the number of overlapping lesions at each voxel. The key difference between groups is the involvement of left IFG, which is damaged in all L Inffront subjects but spared in all L nonIFG PFC subjects. The left hemisphere is on the reader's right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

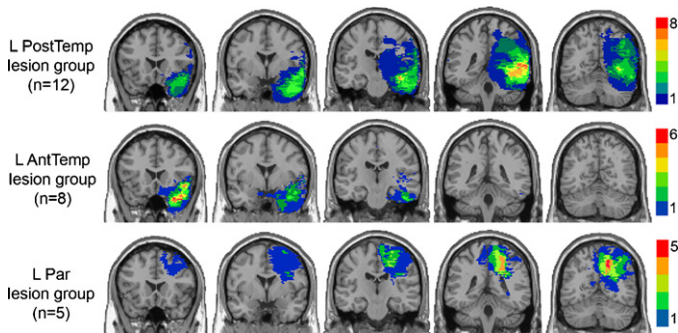


Fig. 3. Lesion overlaps for the L PostTemp group (top row), L AntTemp group (middle row), and L Par group (bottom row). Coronal slices, from left to right: $y = 22, y = 2, y = -18, y = -38, y = -58$. Color indicates the number of overlapping lesions at each voxel. The key difference between groups is the involvement of left posterior temporal cortex, which is damaged in all L PostTemp subjects but spared in all L AntTemp and L Par subjects. The left hemisphere is on the reader's right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

group ($z = -2.3; p = 0.02$). Both the L PostTemp and L Inffront groups included individuals with significant Digit Span impairments (Digit Span ≤ 4 , see Fig. 1C): $n = 4$ in the L PostTemp group and $n = 2$ in the L Inffront group. Digit Span scores for the L nonIFG PFC group did not significantly differ from either the neurologically healthy group ($z = -0.3; p = 0.80$) or the R nonIFG PFC group ($z = 1.8; p = 0.08$). Digit Span scores for the L AntTemp group did not significantly differ from either the neurologically healthy group ($z = -0.2; p = 0.85$) or the R AntTemp group ($z = 1.2; p = 0.22$). Digit Span scores for the L Par group did not significantly differ from either the neurologically healthy group ($z = -0.4; p = 0.70$) or the R Par group ($z = 0.8; p = 0.44$). These results thus demonstrate that damage specifically involving left inferior frontal or left posterior temporal cortex is associated with reduced Digit Span performance.

2.4. Association between Digit Span and language abilities

Given our hypothesis that the same brain regions which support language production and comprehension may also support verbal STM, we next sought to determine whether the brain areas involved in Digit Span performance are also involved in the production and/or comprehension of language. We addressed this issue in three ways.

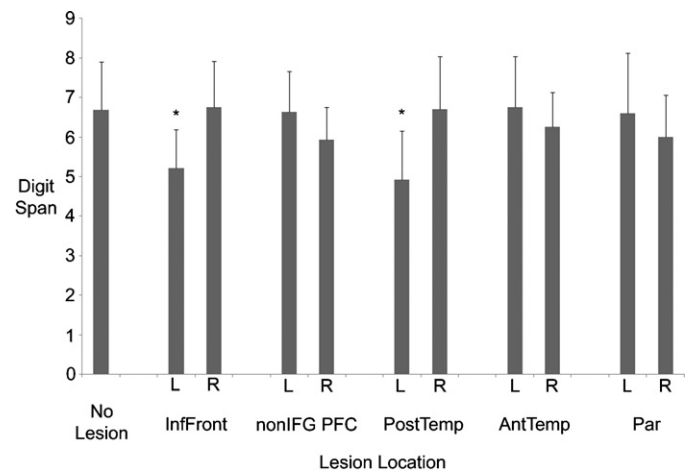


Fig. 4. Mean Digit Span for each subject group. Error bars indicate standard deviations.

Table 2
Neuropsychological test data for subject groups. Group means are presented with standard deviations in parentheses. Scores significantly lower than the neurologically healthy group ($p < 0.05$) are italicized.

Group	Verbal tests			Non-verbal tests		
	Boston Naming Test	Token Test	WAIS Vocabulary	Spatial Span	Matrix Reasoning	Tower Test
Inffront						
L	48.2(8.1)	94.7(4.4)	32.6(19.1)	6.9(2.3)	13.3(5.4)	14.9(6.4)
R	55.1(3.0)	97.6(3.9)	36.0(12.1)	7.7(1.8)	10.9(6.8)	17.0(3.9)
nonIFG PFC						
L	55.7(3.4)	99.6(0.7)	46.6(12.1)	8.1(1.7)	16.5(5.3)	17.9(3.3)
R	55.3(2.4)	98.9(1.0)	40.4(14.1)	8.1(2.0)	13.8(4.3)	16.4(3.7)
PostTemp						
L	46.6(11.8)	94.9(9.1)	36.7(16.1)	7.3(1.4)	15.4(5.6)	16.5(4.5)
R	56.2(3.3)	98.7(2.1)	49.2(10.5)	8.9(1.7)	14.6(7.5)	16.9(4.8)
AntTemp						
L	53.5(8.0)	98.3(3.5)	45.4(16.4)	7.3(2.0)	16.3(5.5)	18.1(4.5)
R	54.6(4.5)	99.7(0.5)	44.0(12.2)	8.6(1.8)	13.3(6.4)	18.0(6.2)
Par						
L	56.8(1.5)	97.8(3.8)	47.0(5.6)	7.5(1.8)	17.0(1.6)	15.8(4.0)
R	54.4(3.5)	97.8(3.1)	48.3(10.1)	7.5(2.8)	11.4(5.5)	17.3(5.0)
No Lesion	55.4(4.7)	98.8(1.6)	47.8(9.7)	8.0(1.5)	16.6(5.3)	17.7(4.0)

2.5. Association between Digit Span and language abilities: ROI analyses

First, we examined performance by each lesion group (as well as comparison groups) on three canonical tests of language processing that vary with respect to the involvement of speech production and verbal comprehension (Table 2). The Boston Naming Test (BNT) requires speech production, but places minimal (if any) demand on auditory–verbal comprehension and short-term memory. By contrast, the Token Test requires the subject to comprehend a simple oral command and produce the appropriate non-verbal motor response; this task involves auditory–verbal comprehension and short-term retention, but no speech production. The WAIS Vocabulary subtest requires both speech production and comprehension in that the subjects are asked to produce the definitions of auditorily presented words. The only lesion groups with significantly lower scores on the BNT relative to the neurologically healthy group were the L Inffront group ($z = -2.8$; $p = 0.004$) and the L PostTemp group ($z = -2.4$; $p = 0.02$) – the same two groups with Digit Span performance deficits (all other p values > 0.90). Similarly, the only two lesion groups with significantly lower Token Test scores relative to the neurologically healthy group were again the L Inffront group ($z = -3.1$; $p = 0.002$) and the L PostTemp group ($z = -2.5$; $p = 0.01$; all other p values > 0.45). The L Inffront and L PostTemp groups also exhibited significantly lower-than-normal scores on the WAIS Vocabulary subtest ($z = -2.4$; $p = 0.02$ and $z = -2.2$; $p = 0.03$, respectively), as did the R Inffront and R nonIFG PFC groups ($z = -2.5$; $p = 0.01$ and $z = -2.1$; $p = 0.04$, respectively; all other p values > 0.40).

It is important to point out that the Digit Span and language processing deficits observed in the L Inffront and L PostTemp lesion groups are not simply manifestations of a more general cognitive impairment. Neither the L Inffront group nor the L PostTemp exhibited significantly lower scores than the neurologically healthy group on Spatial Span, Matrix Reasoning, or the Tower Test (all p values ≥ 0.10 ; see Table 2). In addition, the Digit Span and language processing deficits observed in the L Inffront and L PostTemp lesion groups cannot be simply attributed to lesion size; lesion sizes in these two groups were not significantly different than lesion sizes in their respective right hemisphere comparison groups (Table 1). Furthermore, across patients in this analysis, overall lesion size was not strongly correlated with Digit Span ($r = -0.17$; $p = 0.10$). Thus, lesion location (rather than overall lesion size) was the critical determinant of Digit Span performance in this analysis.

2.6. Association between Digit Span and language abilities: VLSM analyses

Next, we used conjunction VLSM analyses to determine if deficits in Digit Span are associated with damage to the same areas of the brain as are deficits in each of the three canonical language tests (BNT, Token Test, and WAIS Vocabulary; Figs. 5–7, respectively). These analyses demonstrate similar and extensively overlapping substrates for Digit Span, BNT, Token Test, and WAIS Vocabulary performance. In particular, the critical areas for Digit Span are almost entirely contained within the critical areas for BNT performance.

2.7. Association between Digit Span and language abilities: Regression analysis

Finally, as a more general test of the language–verbal STM relationship, we conducted a regression analysis using data from both the brain-injured ($n = 93$) and neurologically healthy subjects ($n = 53$) from our previous analyses (see Table 1). The regression included Digit Span as the criterion variable and three measures of language processing (Token Test, BNT, and WAIS Vocabulary Test) as predictor variables. In addition to these language measures, we also included two variables related to general cognitive function that might predict Digit Span performance: number of years of education and scores on the Armed Forces Qualification Test, a battery of tests measuring basic cognitive function at the time of enlistment (i.e. pre-injury). Subjects were included only if they had data for all six measures, which resulted in a sample of 112 individuals. Results of the regression are presented in Table 3. Overall, the model predicted 32% of the variance in Digit Span performance, with two of the variables as significant predictors: the Boston Naming Test and the WAIS Vocabulary Test. What is noteworthy about this regression is not simply that language processing measures predict Digit Span performance, but rather, that the measures that more heavily involve language production processes are significant predictors, whereas the more pure measure of language comprehension (i.e. the Token Test) is not. The implications of this finding in regard to the overall pattern of results are discussed below.

3. Discussion

In this study, we examined a large sample of adults with acquired brain lesions to identify the critical neural substrates

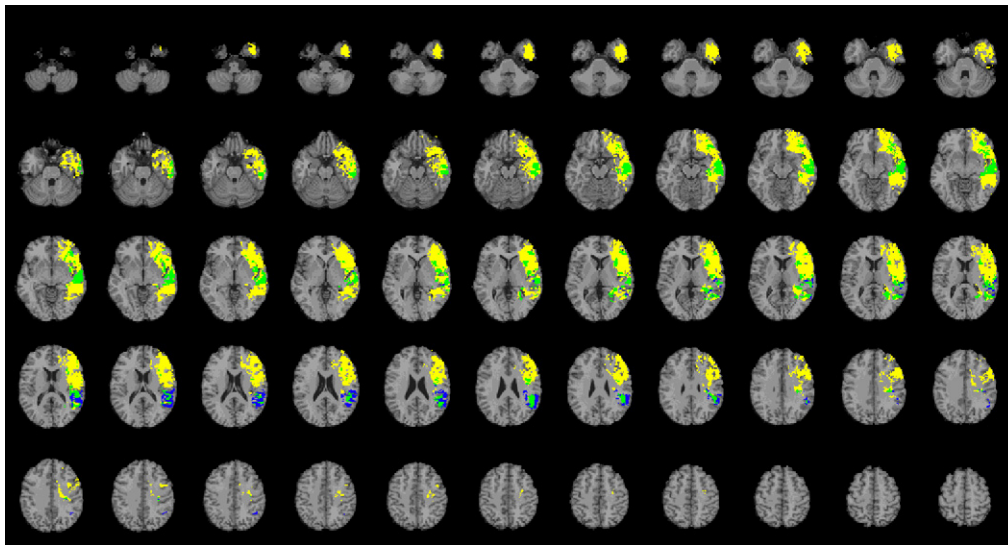


Fig. 5. Conjunction VLSM analysis for Digit Span and BNT. Blue indicates areas where damage was exclusively associated with significantly lower Digit Span (Mann–Whitney *U*-test, $q = 0.05$). Yellow indicates areas where damage was exclusively associated with significantly lower BNT score (Mann–Whitney *U*-test, $q = 0.05$). Green indicates areas of overlap, where damage was associated with significantly lower Digit Span and BNT scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

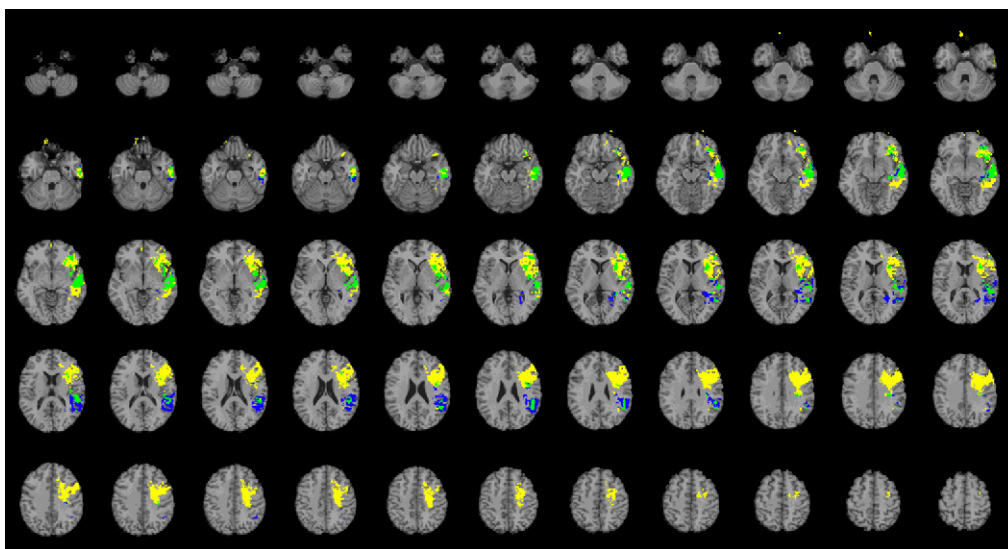


Fig. 6. Conjunction VLSM analysis for Digit Span and Token Test. Blue indicates areas where damage was exclusively associated with significantly lower Digit Span (Mann–Whitney *U*-test, $q = 0.05$). Yellow indicates areas where damage was exclusively associated with significantly lower Token Test score (Mann–Whitney *U*-test, $q = 0.05$). Green indicates areas of overlap, where damage was associated with significantly lower Digit Span and Token Test scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

underlying verbal STM as well as the relationship between verbal STM and language processing abilities. Our initial analyses associate Digit Span impairment with lesions involving areas of left perisylvian cortex (Fig. 1). More focused anatomical analyses demonstrate that, within left perisylvian cortex, the inferior frontal and posterior temporal regions were critical for Digit Span

performance. Damage involving either area was associated with significantly shorter Digit Span, whereas damage to adjacent regions in frontal, parietal, and temporal cortex was associated with normal Digit Span (Figs. 2–4). Moreover, Digit Span performance was linked to language processing abilities; patients with damage involving the left inferior frontal or left posterior

Table 3

Multiple regression of language and pre-injury measures on Digit Span performance across the entire sample. Statistically significant predictors are italicized.

	β	SE	<i>t</i>	<i>p</i>	<i>R</i> ²	Model significance
(Constant)	1.95	2.61	0.75	0.46	0.32	<i>F</i> (5,106) = 9.82; $p < 0.001$
Years of education	−0.051	0.048	−1.05	0.30		
Pre-injury AFQT	−0.001	0.006	−0.19	0.85		
Token test	0	0.033	0.012	0.99		
Boston naming test	0.072	0.026	2.79	0.006		
WAIS vocabulary	0.026	0.012	2.17	0.032		

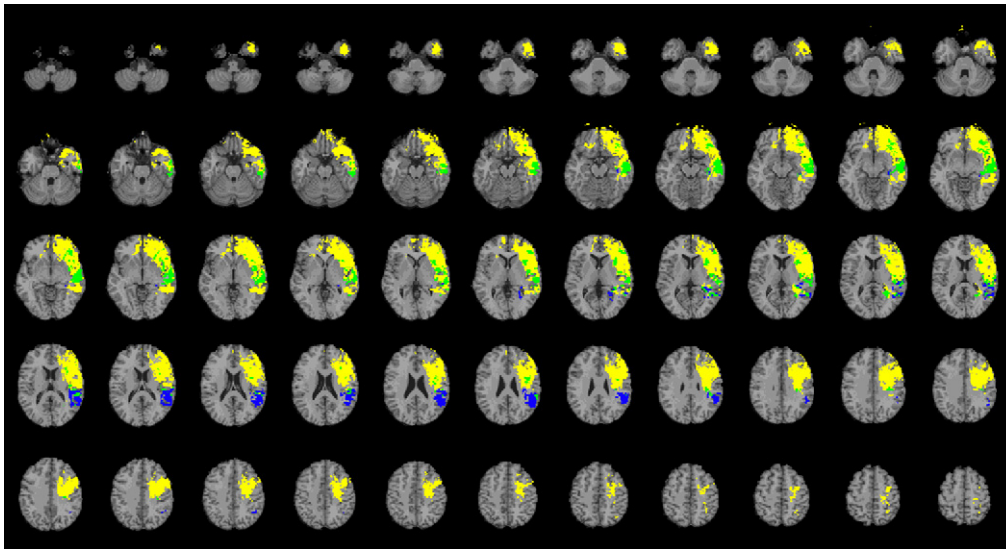


Fig. 7. Conjunction VLSM analysis for Digit Span and WAIS Vocabulary Test. Blue indicates areas where damage was exclusively associated with significantly lower Digit Span (Mann–Whitney U -test, $q=0.05$). Yellow indicates areas where damage was exclusively associated with significantly lower WAIS Vocabulary score (t -test, $q=0.05$). Green indicates areas of overlap, where damage was associated with significantly lower Digit Span and WAIS Vocabulary scores. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

temporal regions also exhibited impairments in tests of language production and comprehension, and a multiple regression analysis revealed that Digit Span performance was significantly predicted by performance on tasks primarily involving language production processes. Furthermore, conjunction VLSM analyses indicate that brain areas critical for Digit Span were largely localized within the areas critical for language production task performance. Taken together, these results suggest that Digit Span performance and language processing may share the same neural substrates – namely the left inferior frontal and left posterior temporal cortices.

These results confirm and elaborate upon those of a recent study of a large sample of stroke patients (Leff et al., 2009). Leff et al. used multiple regression to relate gray matter density to neuropsychological test performance. When Digit Span was entered as the sole neuropsychological variable, the analysis identified an extensive area of left perisylvian cortex, similar to what we report in Fig. 1 of this paper. When Leff et al. included various other auditory and/or verbal neuropsychological test data in their regression, they identified an area of left posterior superior temporal cortex where gray matter density significantly predicted Digit Span as well as speech comprehension performance. In accord with this result, we found that focal lesions involving left posterior temporal cortex significantly impaired Digit Span and language comprehension performance. Our present paper highlights two additional findings. First, the left inferior frontal cortex is also critical for Digit Span performance, and second, measures of language production ability significantly predict Digit Span performance. The difference in findings between the two studies may be due to differences in the methodological approach; while Leff et al. employed a multiple regression strategy to identify key brain areas on the basis of regional variation in gray matter density, we selected patients with circumscribed lesions according to a priori anatomical regions of interest. Furthermore, although Leff et al. included very similar measures of language production processing in their regression analyses, the order in which these measures were entered (i.e. production before comprehension) may have led to an emphasis on the comprehension–STM relationship. Another notable distinction between our study and the Leff et al. (2009) study is the spatial distribution of lesions. While the stroke patients included in the Leff et al. study predominantly feature left hemisphere lesions, our

brain-injured sample exhibits a roughly symmetrical distribution of lesions (Fig. 1B), allowing us to more definitively demonstrate the specific importance of left (vs. right) perisylvian cortices for verbal STM.

In addition, our data demonstrating a link between verbal STM and language production in brain-damaged patients are consistent with other recent studies. One of these studies showed that auditory STM capacity correlates with gray matter density in left posterior superior temporal cortex in normal and dyslexic adults (Richardson et al., *in press*). A second study showed that repetitive transcranial magnetic stimulation applied to a region of left posterior superior temporal gyrus important for the phonological encoding stage of speech production also impairs verbal STM performance (Acheson et al., 2011). This convergence of results, across a variety of disruption models and analytic methods, provides strong evidence for a necessary functional role for the left inferior frontal cortex and left posterior temporal cortex in both language production and verbal STM.

It is noteworthy that the left inferior frontal area described in the present study and the left posterior temporal areas described here, by Leff et al. and by Acheson et al. appear to be directly connected via the arcuate fasciculus (Hong, Kim, Ahn, & Jang, 2009). We see the convergent results as strong support for the general conclusion that verbal STM and language processing share overlapping neural substrates (for similar conclusions see Buchsbaum & D'Esposito, 2008; Martin & Saffran, 1997). We speculate that the underlying role of left perisylvian areas in both verbal STM and language processing may be related to the representation and maintenance of the sequential input and output of speech sounds. In other words, verbal STM may not rely on a dedicated short-term storage buffer. Instead, verbal STM may simply reflect the temporary activation of the same long-term representations of speech sounds that are used to comprehend and generate normal speech (Ruchkin, Grafman, Cameron, & Berndt, 2003). However, we cannot exclude the possibility that Digit Span and language abilities are in fact mediated by distinct neural circuits that lie in extremely close proximity to one another (i.e. both within areas of left perisylvian cortex), such that the observed association between Digit Span and language processing deficits is due to damage commonly encompassing both circuits.

The present results join a large and growing body of work aimed at specifying the neuroanatomical basis of verbal working memory. The lesion data presented here serve to qualify the interpretation of functional imaging results associating verbal STM task performance with activation of particular brain areas. While functional imaging studies have collectively identified an extensive network of cortical and subcortical areas active during verbal STM tasks (reviewed by Wager & Smith, 2003), the lesion results specify a subset of these areas that are actually necessary for basic verbal STM capacities. Along these lines, recent studies demonstrate that posterior parietal cortex is critical for the manipulation or rearrangement of information in working memory, but not the basic short-term retention of information (Champod and Petrides, 2010; Koenigs et al., 2009). Integrating these results with the lesion data presented here and recent related findings (Acheson et al., 2011; Leff et al., 2009; Richardson et al., in press), we propose a neuroanatomical model of verbal STM with two key features: (1) the short-term maintenance of verbal information is mediated by a left hemisphere inferior-frontal/posterior-temporal circuit, and (2) the active manipulation or rearrangement of information within short-term memory additionally recruits posterior parietal cortices. Future research may serve to better characterize how these brain areas interact to facilitate the mental maintenance and manipulation of verbal information.

Funding

This work was supported by funding from the U.S. National Institute of Neurological Disorders and Stroke intramural research program and a project grant from the United States Army Medical Research and Materiel Command administered by the Henry M. Jackson Foundation (Vietnam Head Injury Study Phase III: a 30-year post-injury follow-up study, grant number DAMD17-01-1-0675). B.R.P. received support from MH064498.

Acknowledgements

We thank V. Raymont for assistance with patient testing and lesion tracing, S. Bonifant for data management, and the veterans for their participation.

References

- Acheson, D. J., Hamidi, M., Binder, J. R. & Postle, B. R. (2011). A common neural substrate for language production and verbal working memory. *Journal of Cognitive Neuroscience*, 23(6), 1358–1367.
- Acheson, D. J. & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, 135(1), 50–68.
- Allport, A. D. (1984). *Auditory-verbal short-term memory and conduction aphasia*. Hillsdale: Lawrence Erlbaum Associates.
- Baddeley, A. D. (1986). *Working memory*. New York: Clarendon Press/Oxford University Press.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. New York: Oxford University Press.
- Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory*. Academic Press: New York.
- Bates, E., Wilson, S. M., Saygin, A. P., Dick, F., Sereno, M. I., Knight, R. T., et al. (2003). Voxel-based lesion-symptom mapping. *Nature Neuroscience*, 6(5), 448–450.
- Buchsbaum, B. R. & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, 20(5), 762–778.
- Caplan, D. & Waters, G. S. (1990). Short-term memory and language comprehension: A critical review of the neuropsychological literature. In G. Vallar, & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory*. Cambridge University Press: New York.
- Champod, A. S. & Petrides, M. (2010). Dissociation within the frontoparietal network in verbal working memory: A parametric functional magnetic resonance imaging study. *Journal of Neuroscience*, 30(10), 3849–3856.
- Chein, J. M. & Fiez, J. A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex*, 11(11), 1003–1014.
- Delis, D. C., Kaplan, E. & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS): Examiner's manual*. San Antonio: The Psychological Corporation.
- Genovese, C. R., Lazar, N. A. & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, 15(4), 870–878.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. I. *Brain*, 88(2), 237–294.
- Goodglass, H. & Kaplan, E. (2000). *Boston naming test*. Philadelphia: Lippincott Williams & Wilkins.
- Hong, J. H., Kim, S. H., Ahn, S. H. & Jang, S. H. (2009). The anatomical location of the arcuate fasciculus in the human brain: A diffusion tensor tractography study. *Brain Research Bulletin*, 80(1–2), 52–55.
- Koenigs, M., Barbey, A. K., Postle, B. R. & Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *Journal of Neuroscience*, 29(47), 14980–14986.
- Leff, A. P., Schofield, T. M., Crinion, J. T., Seghier, M. L., Grogan, A., Green, D. W., et al. (2009). The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke. *Brain*, 132(Pt 12), 3401–3410.
- Makale, M., Solomon, J., Patronas, N. J., Danek, A., Butman, J. A. & Grafman, J. (2002). Quantification of brain lesions using interactive automated software. *Behavior Research Methods, Instruments, and Computers*, 34(1), 6–18.
- Martin, N. & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, 14, 641–682.
- Papagno, C., Cecchetto, C., Reati, F. & Bello, L. (2007). Processing of syntactically complex sentences relies on verbal short-term memory: Evidence from a short-term memory patient. *Cognitive Neuropsychology*, 24(3), 292–311.
- Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M. & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331(6157), 585–589.
- Richardson, F. M., Ramsden, S., Ellis, C., Burnett, S., Megnin, O., Catmur, C., et al. Auditory STM capacity correlates with gray matter density in the left posterior STS in cognitively normal and dyslexic adults. *Journal of Cognitive Neuroscience*, in press.
- Ruchkin, D. S., Grafman, J., Cameron, K. & Berndt, R. S. (2003). Working memory retention systems: A state of activated long-term memory. *Behavioral Brain Science*, 26(6), 709–728 (discussion 728–777).
- Shallice, T. & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia*, 15(6), 729–735.
- Shallice, T. & Vallar, G. (1990). The impairment of auditory-verbal short-term storage. In G. Vallar, & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory*. Cambridge University Press: Cambridge.
- Shallice, T. & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *Quarterly Journal of Experimental Psychology*, 22(2), 261–273.
- Smith, E. E. & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive Psychology*, 33(1), 5–42.
- Solomon, J., Raymont, V., Braun, A., Butman, J. A. & Grafman, J. (2007). User-friendly software for the analysis of brain lesions (ABLE). *Computer Methods and Programs in Biomedicine*, 86(3), 245–254.
- Spellacy, F. J. & Spreen, O. (1969). A short form of the token test. *Cortex*, 5(4), 391–397.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1), 273–289.
- Vallar, G. & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behavior*, 23, 151–161.
- Wager, T. D. & Smith, E. E. (2003). Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, and Behavioral Neuroscience*, 3(4), 255–274.
- Warrington, E. K., Logue, V. & Pratt, R. T. (1971). The anatomical localisation of selective impairment of auditory verbal short-term memory. *Neuropsychologia*, 9(4), 377–387.
- Warrington, E. K. & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, 92(4), 885–896.
- Wechsler, D. (1997). *Wechsler adult intelligence scale-III*. San Antonio: The Psychological Corporation.
- Wernicke, C. (1874). *Der Aphasische Symptomenkomplex*. Breslau: Taschen.