Lesion localization of speech comprehension deficits in chronic aphasia

Sara B. Pillay, PhD
Jeffrey R. Binder, MD
Colin Humphries, PhD
William L. Gross, MD, PhD
Diane S. Book, MD

Correspondence to
Dr. Pillay:
sapillay@mcw.edu

ABSTRACT

Objective: Voxel-based lesion-symptom mapping (VLSM) was used to localize impairments specific to multiword (phrase and sentence) spoken language comprehension.

Methods: Participants were 51 right-handed patients with chronic left hemisphere stroke. They performed an auditory description naming (ADN) task requiring comprehension of a verbal description, an auditory sentence comprehension (ASC) task, and a picture naming (PN) task. Lesions were mapped using high-resolution MRI. VLSM analyses identified the lesion correlates of ADN and ASC impairment, first with no control measures, then adding PN impairment as a covariate to control for cognitive and language processes not specific to spoken language.

Results: ADN and ASC deficits were associated with lesions in a distributed frontal-temporal parietal language network. When PN impairment was included as a covariate, both ADN and ASC deficits were specifically correlated with damage localized to the mid-to-posterior portion of the middle temporal gyrus (MTG).

Conclusions: Damage to the mid-to-posterior MTG is associated with an inability to integrate multiword utterances during comprehension of spoken language. Impairment of this integration process likely underlies the speech comprehension deficits characteristic of Wernicke aphasia.

Neurology® 2017;88:970–975

GLOSSARY

ADN = auditory description naming; **ASC** = auditory sentence comprehension; **MTG** = middle temporal gyrus; **PN** = picture naming; **pSTG** = posterior left superior temporal gyrus; **SMG** = supramarginal gyrus; **STG** = superior temporal gyrus; **VLSM** = voxel-based lesion-symptom mapping.

The notion that speech comprehension depends on the posterior left superior temporal gyrus (pSTG)[1] is a persistent one, despite 3 major problems with this model. First, evidence suggests that the principal role of the pSTG is in processing phonologic information for short-term memory and speech production tasks.[2–4] Second, language comprehension depends on a broadly distributed network spanning frontal, temporal, and parietal lobes.[5,6] Thus, any model that proposes a strict localization of language comprehension to the pSTG can no longer be seen as tenable. Third, speech comprehension is not a unitary process, but rather encompasses speech sound recognition, retrieval of individual word meanings, and various processes supporting integration of meaning across multiple words. All of these components also depend on the ability to maintain information in short-term memory and on general executive control processes. Given this complexity, which process or processes, if any, are likely to be localized to the superior temporal region?

Multiword comprehension requires the rapid combination of individual word meanings, guided by syntactic information.[7] Both functional imaging and lesion data suggest extensive cortical involvement in sentence comprehension[8–11]; however, these studies have not attempted to separate the combinatorial processes unique to multiword comprehension from executive control processes. On the basis of several prior functional imaging studies,[12–16] we hypothesized that temporal lobe regions inferior to the classic Wernicke area may contain a critical site for combinatorial processing of phrase-level and sentence-level language. We tested this hypothesis

Supplemental data
at Neurology.org

© 2017 American Academy of Neurology

using lesion-deficit correlation in chronic stroke while incorporating controls for general executive and other deficits not specific to comprehension tasks.

METHODS **Participants.** Participants were 51 patients (25 women) with focal encephalomalacia from chronic left hemisphere stroke, included regardless of language ability, but without comorbid neurologic or major psychiatric disease (e.g., bipolar disorder, dementia, epilepsy, major depression, schizophrenia). Participant summary statistics are listed in table 1 and individual demographic information is provided in table e-1 at Neurology.org. All participants were at least 180 days poststroke, native English speakers, and premorbidly right-handed according to the Edinburgh Handedness Inventory.[17] Estimated premorbid general intelligence was average according to the Barona intelligence quotient.[18] Lesion descriptions are provided in table e-1. Patients were enrolled prospectively.

Standard protocol approvals, registrations, and patient consents. The Medical College of Wisconsin Institutional Review Board approved the use of human subjects for this study. Written informed consent was obtained from all participants.

Behavioral tasks. As described previously,[3] testing was conducted using a computer connected to a touch-sensitive LCD monitor. Manual and vocal responses were digitally recorded for offline scoring.

Participants completed an auditory description naming (ADN) task to measure comprehension of brief verbal descriptions and a picture naming (PN) task to control for diverse language and domain-general processes not specific to spoken language comprehension. Each naming task contained equal numbers of living things (e.g., cow) and nonliving objects (e.g., scissors), and the tasks were matched on number of letters, frequency, and imageability of the target names (table e-2). In the 60-item ADN task, participants named a verbally described object (e.g., "crown" in response to "what a king wears on his head") drawn from previously normed items.[19] In the 80-item PN task, participants named line drawings.[20] No time constraints were imposed; however, trials were halted and scored 0 when participants took longer than 20 seconds to respond. The first complete response was scored for accuracy, with errors receiving a 0 even if subsequently self-corrected. Valid alternative names given by healthy controls were accepted.[19]

Our rationale for pairing the ADN and PN tasks is that these tasks are relatively matched on word retrieval, speech production, and general executive demands, whereas the ADN task alone requires comprehension of spoken language. To validate this model, participants also completed a 56-trial auditory sentence comprehension (ASC) task, in which they decided whether a spoken sentence (5–12 words) accurately described a concurrently presented video animation.[21] This task, while emphasizing the same auditory comprehension component engaged during ADN, does not require name retrieval or speech production. Data for this task were missing in 7 participants.

MRI. High-resolution (1 × 1 × 1 mm³) T1-weighted MRI were obtained in the chronic stage in all patients. MRI was performed at 3T in 48 patients and 1.5T in 3 patients. Details regarding lesion tracing and nonlinear image registration methods are described elsewhere.[3,22] Normalized total lesion size (in template voxels) was obtained in each patient from the template-registered lesion map.

Voxel-based lesion-symptom mapping. Voxel-based lesion-symptom mapping (VLSM) uses lesion status at each voxel as a grouping variable, and then compares the lesioned and nonlesioned groups on any given dependent measure, to produce an effect size statistic for each voxel.[23] VLSM was implemented as an analysis of covariance, using a custom MATLAB script to account for within-group variance of no interest. Only voxels lesioned in at least 3 patients were included. Although this cutoff is more inclusive than in some VLSM studies, our goal was to maximize coverage in the anterior and inferior temporal lobe (see figure e-1E for details on how changing this cutoff affects the results). Five analyses were performed, as outlined below.

General comprehension network. The first 2 analyses identified lesion correlates of ADN and ASC accuracy, with no covariates included in the model (see Ref. 9 for a similar analysis). The aim was to identify all components of the comprehension network that are critical for normal performance of the tasks, including components that are common to both tasks.

ADN-specific processes. The third analysis examined lesion correlates of ADN accuracy independent of PN by including PN score as a covariate. The aim was to remove components of

Table 1	Demographic, clinical, and performance data						
			Correlation with ADN accuracy		**Correlation with ASC accuracy**		
Variable	**Mean ± SD**	**Range**	*r*	*p* **Value**	*r*	*p* **Value**	
Age	59.2 ± 12.6	28-84	−0.17	0.234	−0.12	0.432	
Education	14.3 ± 3.0	8-20	0.27	0.053	0.23	0.127	
Days post onset	948 ± 1,212	180-6732	0.13	0.354	0.01	0.970	
Lesion size, mL	58.6 ± 52.5	0.8-2986.7	−0.58	<0.001	−0.61	<0.001	
HQ	87.8 ± 21.9	4-100	−0.13	0.355	0.23	0.141	
Barona IQ	109.7 ± 9.0	85-132	0.28	0.047	0.24	0.118	
PN accuracy	71.7 ± 31.5	0-100	0.89	<0.001	0.62	<0.001	
ASC accuracy	75.7 ± 16.3	33-96	0.72	<0.001			
ADN accuracy	62.7 ± 35.9	0-100					

Abbreviations: ADN = auditory description naming; ASC = auditory sentence comprehension; HQ = Edinburgh Handedness Inventory[17] handedness quotient; PN = picture naming.

the general language network, such as executive (search and retrieval), attention, lexical semantic, phonologic encoding, and articulatory processes, that are not specific to phrase and sentence comprehension, thereby highlighting the multiword comprehension component of the ADN task. Lesion volume was included as an additional covariate in this analysis to more specifically isolate the regions critical for ADN (table 1).

Auditory sentence comprehension. We predicted that the ADN analysis incorporating PN as a covariate to remove processes common to both tasks would identify regions that are specifically necessary for auditory comprehension. We tested this hypothesis further using the ASC task as a dependent variable in a final VLSM. This task does not require name retrieval or speech production, therefore any overlap with the ADN result is likely to reflect language comprehension processes common to both tasks. For comparability with the ADN vs PN analysis, PN and lesion volume were used as covariates to remove variance due to general executive, attention, lexical semantic, and phonologic processes.

All *t* statistic maps were thresholded at voxel-wise $p < 0.005$ and cluster-corrected at a family-wise α of $p < 0.05$ using a minimum cluster size criterion of 5,500 μL, as determined by randomization testing with 10,000 permutations. A potential concern in VLSM analyses is that areas with higher lesion overlap, which occur in regions where vascular occlusions are most likely, can show factitious localization due to greater statistical power.[24] We addressed this issue quantitatively by correlating lesion overlap and *t* value in the ADN-specific analysis at the voxel level.

RESULTS Behavioral data. Table 1 summarizes descriptive data for demographic variables and behavioral measures. Performances on the 3 tasks were strongly correlated (table 1), suggesting the presence of nonspecific factors (e.g., executive control deficits) affecting performance on all tests. After residualizing the ADN and PN scores against each other to isolate task-specific variance on these tests, a clear dissociation emerged in which ASC correlated with ADN ($\beta = 0.306$, $r_{semi-part} = 0.24$, $p = 0.002$), but not with PN ($\beta = 0.019$, $r_{semi-part} = 0.01$, $p = 0.871$), consistent with our assumption that the ADN and ASC tasks, though different in many ways, share a dependence on processes specific to spoken language comprehension.

Lesion volume was correlated with all of the behavioral measures, with larger lesions associated with poorer performance (table 1). Education and premorbid IQ estimates were correlated with ADN accuracy in the expected direction. There was no relationship between education or premorbid IQ estimates and ASC or PN accuracy. Age, days post onset of stroke, and handedness quotient were not correlated with any task (table 1). Individual participant performance data are provided in table e-1.

VLSM results. Figure 1A shows the lesion overlap map thresholded to show only voxels damaged in at least 3 patients. When no covariates were included in the analysis, ADN impairment was associated with cortical and subcortical damage in the left inferior frontal cortex, superior temporal gyrus (STG), middle temporal gyrus (MTG), supramarginal gyrus (SMG), and angular gyrus (figure e-1A; see table e-3 for peak coordinates). When no covariates were included in the analysis, ASC impairment was associated with damage in similar regions (figure e-1B; see table e-3 for peak coordinates), replicating previous results using similar methods.[9] Notably, PN impairment was also associated with damage in similar frontal, inferior parietal, and posterior STG regions (figure e-1C), consistent with our hypothesis that all 3 tasks require nonspecific executive control and phonologic processes.

The main analyses included performance on the PN task and lesion volume as covariates to identify voxels that were specifically associated with ADN and ASC impairments independent of nonspecific factors. As shown in figure 1B, ADN impairment was correlated with damage primarily in the middle and posterior MTG, extending into the superior temporal sulcus. White matter beneath the MTG and STG was also part of the main cluster. Notably, there were no voxel clusters in the frontal lobe, SMG, or angular gyrus that were significantly associated with ADN independent of PN impairment. Including education or Barona IQ as third covariates did not appreciably change the results (figure e-1, C and D).

The final analysis included accuracy on the PN task and lesion volume as covariates to identify voxels that were independently associated with ASC impairment. Uncorrected, a small cluster survived the $p < 0.005$ voxel-level significance threshold in the posterior superior temporal sulcus; however, this cluster did not survive the cluster threshold correction. When only PN accuracy was included as a covariate, voxels correlated with ASC impairment were identified in the posterior MTG and adjacent lateral occipital lobe. As shown in figure 1E, the voxel cluster associated with ASC impairment included most of the voxels associated with ADN impairment, indicating a shared neural substrate critical for both tasks.

The region of highest lesion overlap in our sample was in the perisylvian cortex and insula rather than the MTG (figure 1A). Across all voxels, lesion overlap values and *t* values in the ADN-specific analysis (figure 1B) showed a small correlation ($r = -0.026$, $p = 0.007$); however, the sign was negative, indicating that as lesion overlap increases, the corresponding *t* value actually tends to be slightly smaller. There was, therefore, no indication that the MTG cluster in figure 1B was an artifact of high lesion overlap.

DISCUSSION Spoken language comprehension impairment, as measured by deficits on the ADN and ASC tasks, was associated with lesions in a widely distributed frontal, temporal, and parietal language network. When PN impairment was included as a covariate to remove components of the network that

Figure 1   Voxel-based lesion-symptom mapping (VLSM) results
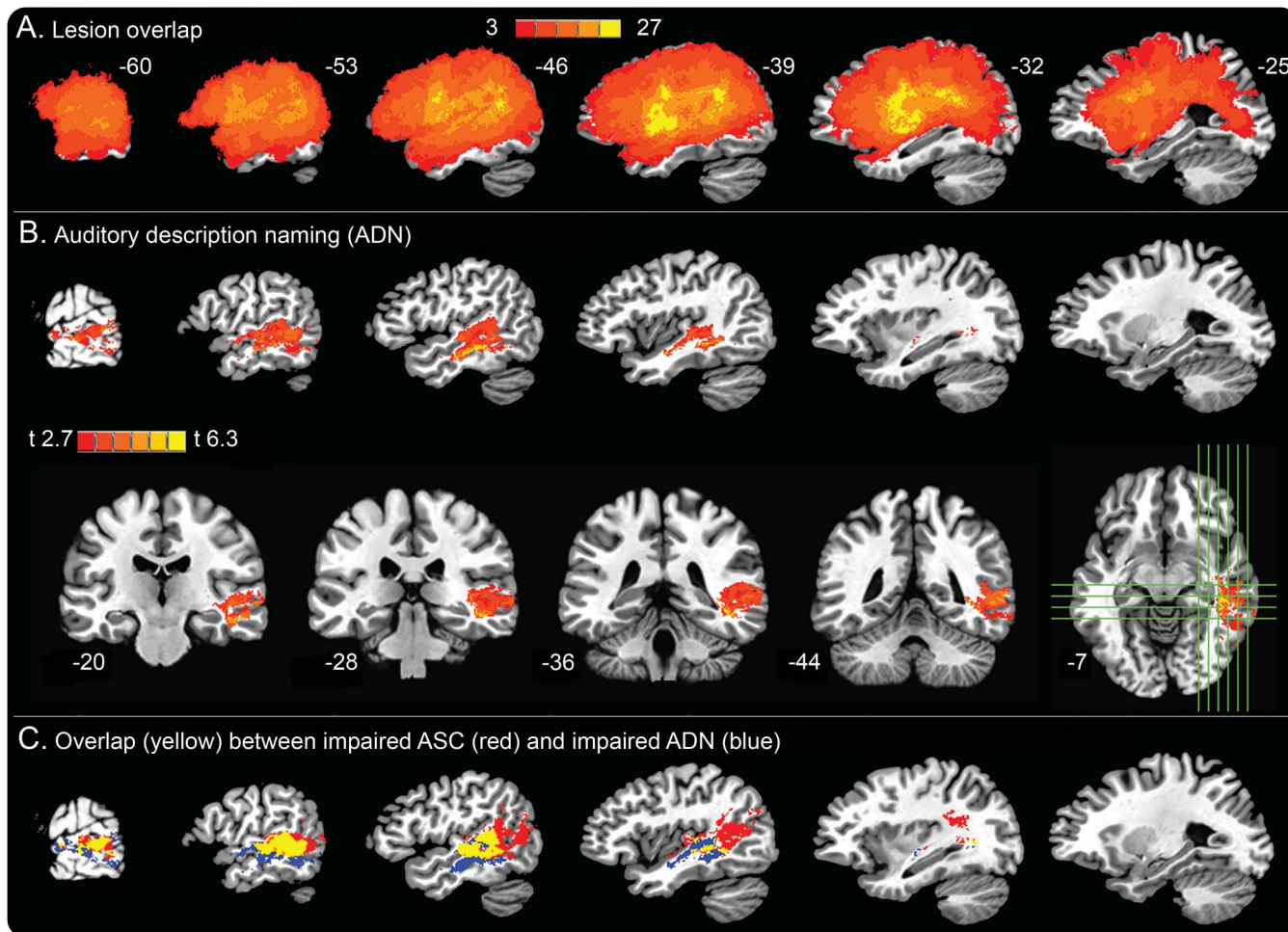
(A) Lesion overlap across all 51 patients, thresholded to include only voxels that were lesioned in at least 3 patients. Colors indicate the degree of overlap. Numbers beside each image indicate the x-axis location of the slice in standard space. The same locations are used for the other sagittal series. (B) VLSM analysis of auditory description naming (ADN) accuracy independent of picture naming accuracy and total lesion volume, shown in both sagittal and coronal sections. Subscripts below the coronal images indicate the y-axis location of the slices. Green lines on the axial image indicate the locations of orthogonal slices. (C) Overlap (yellow) between regions associated with impaired auditory sentence comprehension (ASC: red) and impaired ADN (blue).

are not specific to spoken language comprehension, such as single-word semantic, phonologic, executive (search and selection), and articulatory processing, both ADN and ASC deficits were uniquely correlated with overlapping lesion areas localized to the mid-to-posterior portion of the MTG. This region appears to be critically necessary for integration of multiword combinations during spoken language comprehension.

Spoken language comprehension is a complex and multifaceted task requiring phoneme perception, semantic, syntactic, working memory, and attention processes. In this study, we focused on processes that are relatively specific to spoken language comprehension, as opposed to those that are common to nearly all language tasks, such as working memory, attention, and response selection. We also argue for an important distinction between processing the meaning of a single word—sufficient for lexical tasks such as picture naming and word–picture matching—and processing the

meaning of multiword combinations. In the latter case, not only must the meaning of individual words be understood, but they must also be combined using syntactic (i.e., grammatical role), relational semantic, and pragmatic information. Consider, for example, the difference between a single-word comprehension task that requires pointing to a mirror on hearing the word "mirror" (word–picture matching) and a multiword comprehension task that requires pointing to a mirror on hearing the phrase "a reflective glass for seeing yourself." The former task requires only that a single word be mapped to its meaning. The latter task requires identifying the grammatical role of each word (e.g., identifying "reflective" as a modifier, "glass" as a subject noun, "seeing" as a verb), retrieving and simultaneously maintaining several word concepts, combining these meanings to form more specific concepts (e.g., "seeing yourself" is a more specific concept than "seeing"), and integrating the whole into

a conceptual representation of a set of circumstances (i. e., seeing oneself in a reflective glass). Such combinatorial processes are ubiquitous in everyday spoken language comprehension, which consists almost entirely of multiword phrases and sentences.

Despite the critical importance of multiword integration processes in everyday language, the brain networks that support these processes remain unclear. One longstanding view is that they depend mainly on frontal lobe systems specialized for processing syntactic or semantic information.[8,25–27] Our data strongly support a role for the inferior frontal lobe in spoken language comprehension (figure e-1, A and B); however, we were unable to demonstrate that damage in this region impairs spoken language comprehension more than picture naming (figure e-1C), a task that depends little, if at all, on multiword integration processes. Damage to the inferior frontal lobe, at least at the spatial scale of typical vascular lesions, more likely affects language processing in a general way, by impairing working memory, selection, attention, and phonologic processes that support a wide range of language tasks.

Relatively few functional imaging or lesion correlation studies have attempted to separate the multiword integration processes specific to sentence comprehension from more general executive processes. The most relevant evidence comes from a small number of functional imaging studies in which participants simply listened to or read sentences without the need to make a complex decision, thereby minimizing attention and working memory demands.[12–16] Frontal lobe activation is minimal in such cases, whereas extensive activation typically occurs in the lateral temporal lobe. These data and the current results converge on the conclusion that the lateral temporal cortex plays a central role in multiword integration independent of frontal or parietal executive control systems. Given the temporal lobe location of this system and its activation even under passive listening conditions, these combinatorial processes may be relatively automatic and outside the control of frontal networks.

Anatomic considerations also suggest a central role for the MTG in spoken language comprehension. The MTG is bordered superiorly by speech perception and phonologic systems in the STG, and anteriorly and posteriorly by anterior temporal and inferior parietal regions supporting semantic memory.[5,28] Interactions with all of these systems might be necessary for the efficient combinatorial integration on which multiword speech comprehension depends. Structural and functional connectivity analyses have highlighted the broad connectivity of the MTG with not only these regions but also with inferior frontal and dorsomedial frontal areas implicated in language and semantic control processes.[29]

The present VLSM results suggest that damage to the mid-to-posterior MTG is specifically associated with spoken language comprehension impairment. This region appears to be critically necessary for the ability to integrate multiword utterances during comprehension. The methods used in the present study do not allow conclusions to be drawn concerning regions outside the left middle cerebral artery territory that are not well-represented in this patient sample. This limits the ability to test the contribution of structures such as the medial frontal and parietal lobe and the ventral and medial temporal lobe. Future studies on this topic should include a wider variety of neurologic patients with more varied lesion locations to allow more complete assessment of these areas.

REFERENCES
1. Wernicke C. Der aphasische symptomencomplex: eine psychologische studie auf anatomischer basis [The symptom complex of aphasia: a psychological study on an anatomical basis]. In: Eggert GH, ed. Wernicke's Works on Aphasia. The Hague, the Netherlands: Mouton; 1874:91–145.
2. Buchsbaum BR, Baldo J, Okada K, et al. Conduction aphasia, sensory-motor integration, and phonological short-term memory: an aggregate analysis of lesion and fMRI data. Brain Lang 2011;119:119–128.
3. Pillay SB, Stengel BC, Humphries C, Book DS, Binder JR. Cerebral localization of impaired phonological retrieval during rhyme judgment. Ann Neurol 2014;76:738–746.
4. Binder JR. The Wernicke area: modern evidence and a reinterpretation. Neurology 2015;85:2170–2175.
5. Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex 2009; 19:2767–2796.

6. Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann NY Acad Sci 2010;1191: 62–88.

7. Hagoort P, Indefrey P. The neurobiology of language beyond single words. Annu Rev Neurosci 2014;37:347–362.

8. Friederici AD, Rüschemeyer SA, Hahne A, Fiebach CJ. The role of left inferior frontal and superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb Cortex 2003;13:170–177.

9. Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition 2004;92:145–177.

10. Kuperberg GR, Sitnikova T, Lakshmanan BM. Neuroanatomical distinctions within the semantic system during sentence comprehension: evidence from functional magnetic resonance imaging. NeuroImage 2008;40:367–388.

11. Thothathiri M, Kimberg DY, Schwartz MF. The neural basis of reversible sentence comprehension: evidence from voxel-based lesion symptom mapping in aphasia. J Cogn Neurosci 2011;24:212–222.

12. Crinion JT, Lambon-Ralph MA, Warburton EA, Howard D, Wise RJS. Temporal lobe regions engaged during normal speech comprehension. Brain 2003;126:1193–1201.

13. Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS. Converging language streams in the human temporal lobe. J Neurosci 2006;26:7328–7336.

14. Lindenberg R, Scheef L. Supramodal language comprehension: role of the left temporal lobe for listening and reading. Neuropsychologia 2007;45:2407–2415.

15. Okada K, Rong F, Venezia J, et al. Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. Cereb Cortex 2010;20:2486–2495.

16. Pallier C, Devauchelle AD, Dehaene S. Cortical representation of the constituent structure of sentences. Proc Natl Acad Sci USA 2011;108:2522–2527.

17. Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971;9:97–113.

18. Barona A, Reynolds CR, Chastain R. A demographically based index of premorbid intelligence for the WAIS-R. J Consult Clin Psychol 1984;52:885–887.

19. Hammeke TA, Kortenkamp SJ, Binder JR. Normative data on 372 stimuli for descriptive naming. Epilepsy Res 2005;66:45–57.

20. Snodgrass JG, Vanderwart M. A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol Hum Learn 1980;6:174–215.

21. Westbury C. ALFAB: The Alberta Language Function Assessment Battery [online]. 2015. Available at: psych. ualberta.ca/~westburylab/downloads/alfab.download.html.

22. Binder JR, Pillay SB, Humphries CJ, Gross WL, Graves WW, Book DS. Surface errors without semantic impairment in acquired dyslexia: a voxel-based lesion–symptom mapping study. Brain 2016;139:1517–1526.

23. Bates E, Wilson SM, Saygin AP, et al. Voxel-based lesion–symptom mapping. Nat Neurosci 2003;6:448–450.

24. Mah YH, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. Brain 2014;137: 2522–2531.

25. Goodglass H. Agrammatism. In: Whitaker H, Whitaker HA, eds. Studies in Neurolinguistics. New York: Academic Press; 1976.

26. Dapretto M, Bookheimer SY. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 1999;24:427–432.

27. Hagoort P. On Broca, brain, and binding: a new framework. Trends Cogn Sci 2005;9:416–423.

28. Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci 2007;8: 976–987.

29. Turken AU, Dronkers NF. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. Front Syst Neurosci 2011;5:1–20.

Pt Neurology.org Offers Important Information to Patients and Their Families

The *Neurology*® Patient Page provides:

- A critical review of ground-breaking discoveries in neurologic research that are written especially for patients and their families
- Up-to-date patient information about many neurologic diseases
- Links to additional information resources for neurologic patients

All *Neurology* Patient Page articles can be easily downloaded and printed, and may be reproduced to distribute for educational purposes. Click on the 'Patients' link on the home page (Neurology.org) for a complete index of Patient Pages.