Surgery in the language-dominant hemisphere may be associated with injury to areas involved with language function.\textsuperscript{19,57} Cortical stimulation mapping is typically used to identify and protect language regions; consequently, long-term language outcomes after resective surgery in the dominant hemisphere are excellent overall in experienced hands.\textsuperscript{24,28,34–36,43,46,49,50,57,58,63,69} The only negative impact that has been consistently documented is a modest decline in naming after anterior temporal lobectomy in patients whose symptoms began later in life.\textsuperscript{23,44,47,65,67} However, while long-term language outcomes are positive, many patients present with transient aphasias in the first few days after resective surgery in the left hemisphere.\textsuperscript{10,27,29,35,41,49,50,57,58,61,63} The reported incidence of aphasia in the immediate postoperative period has ranged from 17\% to 100\% in the cited studies. This variability probably...
reflects differences in presurgical morbidity, differences in the sensitivity of the various measures for testing language, as well as differences in resection sites in the various case series. While acute deficits can occur in any or all language domains, including naming, comprehension, repetition, reading, and writing, no study has characterized transient postsurgical aphasias using a comprehensive aphasia battery.

The majority of transient aphasias appear to resolve within a month, except for modest naming deficits, which often persist as mentioned above. The transient nature of postsurgical language deficits raises the question of what neural mechanism(s) are responsible for the deficits. Proposed theories have included the generalized effects of surgery (“neuroparalytic edema”), diaschisis (dysfunction of adjacent or functionally connected regions given a lack of normal inputs from the resected region), transient edema in regions adjacent to the resection, and resection of functional tissue with subsequent recovery due to the reorganization of functional networks.

Our study had 2 primary aims. The first was to determine the incidence and nature of transient aphasias after left (dominant)-hemisphere resections in a large series of patients with a wide variety of lesion locations by using a comprehensive and validated aphasia battery 2–3 days postsurgery. Our second aim was to determine, using voxel-based lesion-symptom mapping (VLSM), whether the specific pattern of transient postoperative language deficits is related to the location of the surgical site. If so, this neurosurgical patient cohort would have the potential to provide a unique perspective on the neural organization of language, since resections are discrete and their locations are not limited by vascular distribution or patterns of neurodegeneration as in stroke and primary progressive aphasia, respectively.

Methods

Patient Population

One hundred ten patients undergoing left-hemisphere resective surgery were included in this study. Inclusion criteria were 1) a left hemisphere resection in perisylvian language regions, including the anterior temporal lobe; 2) left hemisphere dominance for language as confirmed with a Wada test, presurgical language deficits, or magnetoencephalography lateralization; 3) fluency in English; 4) availability of postsurgical FLAIR and diffusion-weighted imaging (DWI) suitable for delineation of resections and any associated infarction; and 5) Western Aphasia Battery (WAB) administered 2–3 days postsurgery.

One hundred fifty-six consecutive patients met the first 3 criteria and were considered for study inclusion between September 2010 and October 2013. While most patients underwent follow-up MRI at 2–3 days postsurgery, 33 patients did not and thus were excluded. Of the remaining 123 patients, there were 13 to whom the WAB was not administered for various situational reasons, leaving 110 patients in the study.

Demographic information and the distribution of etiologies are shown in Table 1. In most patients, the goal of surgery was to resect low- or high-grade gliomas, epileptogenic foci, or vascular malformations. Four patients whose etiology was classified as “other” underwent surgery for metastases (2 patients), meningioma (1 patient), and hemorrhage (1 patient).

The study was approved by the UCSF Human Research Protection Program, and all participants gave written informed consent. Analysis of deidentified data took place at the University of Arizona.

Surgical Procedures

All patients underwent craniotomy using monitored anesthesia care without intubation and general anesthesia. Generous local anesthetic infiltration was applied to create a scalp block. Surgical exposure was tailored for each case, depending on the target lesion and/or seizure focus. Patients were sedated with either propofol or dexmedetomidine at the start of the procedure. Intraoperative language mapping with electrical stimulation was performed in the majority of patients. The patients were fully awake for the mapping, and intraoperative electrocorticography was used to monitor for stimulation-induced afterdischarges. The intraoperative language tasks included counting, confrontation naming, and occasionally reading. After mapping, patients were resedated with either propofol or dexmedetomidine for the remainder of the procedure. Essential language sites were identified with stimulation mapping and defined as those resulting in a loss of function in at least 2 of 3 stimulations. The majority of sites were at least 1 cm from the resection margin. The resection was usually performed with an ultrasonic aspirator guided by intraoperative neuronavigation. Subpial resection was used where possible.

Language Assessments

A speech-language pathologist, neuropsychologist, or trained research assistant administered the WAB, administered 2–3 days postsurgery.

| TABLE 1. Demographic and etiological information in 110 patients undergoing left hemisphere resective surgery |
| Parameter | No. |
| Mean age in yrs (range) | 44.2 ± 15.2 (19–81) |
| Sex (M/F) | 55/55 |
| Handedness | |
| Right-handed | 101 |
| Left-handed | 8 |
| Ambidextrous | 1 |
| Mean yrs of education (range) | 15.2 ± 2.7 (11–20) |
| Native speaker of English | 102 native, 8 nonnative but fluent |
| Mean time since 1st symptoms (range) | 1479 ± 3255 days (4 days–40 yrs) |
| Etiology | |
| Low-grade glioma | 39 |
| High-grade glioma | 49 |
| Epileptogenic focus | 13 |
| Vascular malformation | 5 |
| Other | 4 |
days postsurgery. The WAB is a comprehensive and validated aphasia battery that yields an Aphasia Quotient (AQ) quantifying overall aphasia severity, as well as subscores in 5 language domains: fluency, information content, comprehension, repetition, and naming. Fluency, a subjective measure rated by the clinician, is based on the patient’s spontaneous speech. The information content measure assesses the functional, communicative value of the patient’s speech and reflects the patient’s ability to convey correct answers to 6 basic questions. Comprehension includes yes/no questions, auditory word recognition, and sequential commands, capturing lexical and syntactic aspects of language comprehension. The repetition measure includes the repetition of words, phrases, and sentences. The naming subscore includes object naming, semantic fluency, sentence completion, and responsive speech tasks. In our study, we used the AQ as well as the fluency, information content, comprehension, and repetition measures from the WAB. However, we did not use the naming subscore of the WAB because we considered the object-naming task too easy and therefore potentially insensitive (that is, subject to ceiling effects), whereas the other tasks that contribute to the naming subscore make substantial demands on cognitive and executive processes and language processes other than naming.

In place of the WAB naming subscore, we administered a 15-item version of the Boston Naming Test (BNT) to 106 of the 110 patients. The BNT contains difficult, low-frequency items, such as “tripod,” “sphinx,” and “palette”; therefore, it is less subject to ceiling effects and so quantifies naming more effectively than the naming component of the WAB.

The WAB and BNT were also administered to most patients a few days prior to surgery. Of the 110 patients, 84 completed the WAB and 85 completed the BNT prior to surgery. The limiting factors in acquiring presurgical data were situational rather than systematically related to language function, so the patients for whom data were acquired represented an unbiased sample of the 110 included patients.

Follow-up testing was attempted 1 month postsurgery, except in patients whose AQ was at least 93.8 and whose BNT score was at least 12 when tested 2–3 days postsurgery. Twenty-nine patients met those criteria. Of the remaining 81 patients, 36 were administered the WAB and 44 the BNT at 1 month postsurgery; 2 were administered the WAB only, and 6 were administered the BNT only. The main limiting factors in acquiring data at this time point were geographical distance and health concerns other than aphasia, so the patients who were studied 1 month postsurgery again represented an unbiased sample of the 110 included patients.

Presurgical and/or 1-month postsurgical time points. Any obtained subscores were included in analyses of language data, but patients with these scores were not included in the lesion analysis since they lacked an AQ (which was to be used as a covariate). Paired t-tests were used to compare scores on each language measure between presurgical and 2- to 3-day postsurgical assessments, between 2- to 3-day and 1-month postsurgical assessments, and between presurgical and 1-month postsurgical assessments. Patients were classified as aphasic according to the WAB when their AQ was less than 93.8, which is the standard cutoff based on the mean of a patient group with brain damage but no clinical aphasia. Aphasia by this definition can range from mild to severe, and in this study we did not distinguish “aphasia” from “dysphasia” (that is, a milder language disturbance). We considered WAB subscores and/or BNT scores as abnormal when they were 2 or more SDs below the means for non–brain-damaged control participants. Because of the way that AQ is calculated, it is possible for a patient to be impaired on one or more language domains without being classified as aphasic, so long as their AQ is greater than or equal to 93.8.

Neuroimaging and Lesion Mapping

Using images acquired 2–3 days postsurgery, a trained research assistant with input from a neurosurgeon and neuroradiologist delineated surgical site locations for each patient (Fig. 1A). A mask of the surgical site was manually created by delineating the extent of resection and any adjacent infarct on each axial slice using MRICron (http://www.mccauslandcenter.sc.edu/mricro/mricron/). Adjacent infarcts were defined as regions showing abnormal FLAIR signal as well as a > 10% reduction in the apparent diffusion coefficient (ADC) relative to the contralateral region. This mask was used for lesion-symptom mapping at the 2- to 3-day and 1-month postsurgery time points.

Two additional masks were also created. The first was a mask of the resection only; any adjacent infarction was not included. This mask was used for lesion-symptom mapping at the presurgical time point (to capture relationships between presurgical language measures and the intended surgical site). The second was a mask of the surgical site as well as any associated edema, defined as tissue that was abnormal on FLAIR but did not show restricted diffusion. This mask was used only for cost function masking in the registration process since much of this edema was typically present preoperatively and its functional status was not known.

All three masks were smoothed with a 4-mm full width at half maximum (FWHM) Gaussian kernel to account for boundary uncertainties due to anatomical variability and imperfect registration. FLAIR images were warped to Montreal Neurological Institute (MNI) space using the unified segmentation procedure in SPM5 and cost function masking using the surgical site plus edema image dilated by 4 mm. The surgical site and resection-only masks were then warped to MNI space by applying the derived transformation and were resampled with 2 × 2 × 2-mm voxels.

The mean resection volume (including any adjacent in-
fraction) was 33.6 ± 24.4 cm³ (range 2.7–133.1 cm³). Overlays of all surgical site masks (Fig. 1B) showed that the anterior temporal lobe was the most commonly resected region in this patient cohort, followed by the posterior inferior frontal gyrus.

Voxel-based lesion-symptom mapping (VLSM) was performed to identify any systematic relationships between locations of resection and/or damage and language measures by using VLSM 2.55 (http://neuroling.arizona.edu/resources.html). We created t-maps for each of the 5 language variables (fluency, information content, comprehension, repetition, and naming) at each of the 3 time points (presurgery, 2–3 days postsurgery, and 1 month postsurgery).

A general linear model was fit at each voxel with the lesion status of the voxel (0 = intact, 1 = lesioned, with intermediate values near lesion borders due to smoothing) as the independent variable and the language measure in question as the dependent variable. Total resection volume was included as a covariate to avoid showing spurious effects in voxels that were more likely to be involved in large lesions. Aphasia quotient was also included as a covariate to identify voxels that were differentially important for each specific language measure rather than associated with global language deficits. Voxels were only included when the sum of the lesion overlay was at least 5 to avoid spurious effects driven by small numbers of patients in regions that are not frequently resected. Statistical power in VLSM analyses is not stationary but depends on the ratio of lesioned/intact patients in each voxel; in this cohort, power was greatest in the anterior temporal lobe, followed by the posterior inferior frontal gyrus. No conclusions can be drawn regarding regions where the sum of the overlay was less than 5, since these regions were not included in the analyses.

To determine the statistical significance of the t-maps resulting from each analysis, a permutation procedure was used. The behavioral scores were randomly reassigned to the patients 1000 times, then each resulting t-map was thresholded at voxelwise p < 0.001, and the cluster size of
the largest cluster (if any) was recorded. After thresholding the real data similarly, the p value of any observed clusters was determined with reference to the null distribution of maximum cluster sizes from the permuted data sets. Statistical significance was defined as $p < 0.05$ based on this permutation analysis. The covariate of resection volume was not permuted, whereas the AQ covariate was permuted along with each language measure of interest. We also report correlations between resection volume and each language measure of interest as well as the overall AQ.

To ensure that VLSM results did not reflect a confound of etiology, we also repeated all VLSM analyses, restricting the data set to the subset of 88 patients with gliomas, excluding those with epilepsy or other etiologies.

Results

Before surgery, most patients’ language was normal or near normal (Fig. 2 and Table 2). Only 15 (18%) of 84 patients were classified as aphasic based on the WAB criterion. Percentages of patients showing impairments in the 5 language domains—fluency, information content, comprehension, repetition, and naming—ranged from 8% on comprehension to 31% on fluency. Of the 15 patients classified as aphasic, 8 were subclassified as anomic and 7 did not meet criteria for any subtype (see Table 2 footnote for details on unclassifiable patients).

At 2–3 days postsurgery, many patients showed sharp decreases in 1 or more language domains and in overall language function as assessed by the AQ (Fig. 2 and Table 2). At the group level, all language domains and overall language function were significantly impaired relative to presurgical levels (all $p < 0.001$). Seventy-eight (71%) of 110 patients were now classified as aphasic based on the WAB, and each of the 5 language domains was impaired in a majority of patients, ranging from 59% on naming to 86% on fluency. In all, 97 (88%) of 110 patients were classified as either aphasic or impaired on at least 1 language domain. Of the 78 patients classified as aphasic, the aphasia type was anomic in 34, conduction in 11, global in 10, Broca’s in 9, Wernicke’s in 7, and transcortical motor in 1; aphasia in 6 patients did not meet criteria for any subtype.

At 1 month postsurgery, most patients showed substantial improvements in language domains that had been impaired 2–3 days postsurgery and in overall language function as assessed by the AQ (Fig. 2 and Table 2). At the group level, all language domains and overall language function (AQ) were significantly improved relative to levels at 2–3 days postsurgery (all $p < 0.001$). While significant improvements were observed, 13 (25%) of 51 patients still met the cutoff for aphasia based on the WAB criterion. Percentages of patients showing impairments on the 5 language domains ranged from 11% on information content to 52% on fluency, although these impairments were dramatically reduced in severity. Of the 13 patients classified as aphasic, the aphasia type was anomic in 8 and transcortical sensory in 1; 4 of the patients did not meet the criteria for any subtype. Compared with presurgical evaluations, there was a significant decrease in confrontation naming as assessed by the BNT ($t(46) = 2.66, p = 0.011$), marginal decreases in AQ ($t(43) = 1.70, p = 0.096$) and comprehension ($t(43) = 1.91, p = 0.063$), but no significant changes in fluency ($t(50) = 0.82, p = 0.42$), information content ($t(50) = 0.70, p = 0.49$), or repetition ($t(48) = 0.62, p = 0.54$) scores.

We used VLSM to identify relationships between surgical sites and language variables. Although most patients’ language was normal or near normal before surgery, we first investigated relationships between presurgical language measures and the location of resections to be subsequently performed (Fig. 3 and Table 3). Naming and comprehension deficits were associated with forthcoming surgery to the anterior temporal lobe (naming $p < 0.001$, comprehension $p = 0.003$); there were no significant relationships involving the other language measures. Thus, patients in whom anterior temporal resections were already planned showed poorer naming and comprehension than the patients in whom other brain regions were to be resected. The volume of the resection to be performed was not correlated with AQ or any of the language measures (all $p \geq 0.18$).

To identify the neural correlates of transient postsurgical aphasias, we next used VLSM to identify relationships between language variables at 2–3 days postsurgery and the locations of surgical sites (Fig. 4 and Table 3). Reduced fluency was associated with resection of the precentral gyrus and adjacent inferior frontal cortex ($p = 0.019$). Reduced information content was associated with resection of the ventral precentral gyrus and the pars opercularis of the inferior frontal gyrus ($p = 0.035$). Repetition deficits were associated with resection of the posterior superior temporal gyrus ($p = 0.024$). Naming deficits were associated with resection of the middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, and parahippocampal gyrus, with the strongest relationship observed in the fusiform gyrus approximately 6 cm posterior to the temporal pole ($p < 0.001$). There was no significant relationship between comprehension scores and resection location, but the most significant voxel was located in the white matter underlying the temporal lobe (MNI coordinates $x = -46, y = -18, z = -12$; uncorrected $p = 0.0012$). The systematic relationships between damage to specific parts of the language network and deficits in particular language domains show that transient postsurgical aphasias reflect localized effects of resection rather than a global effect on brain function. However, resection volume was modestly negatively correlated with AQ ($r = -0.25, p = 0.0087$) and every language measure: fluency ($r = -0.18, p = 0.058$), information content ($r = -0.24, p = 0.010$), repetition ($r = -0.21, p = 0.031$), naming ($r = -0.27, p = 0.0060$), and comprehension ($r = -0.27, p = 0.0045$).

The cluster associated with naming deficits at 2–3 days postsurgery extended along the axis of the temporal lobe all the way from the temporal pole almost to the occipital lobe, so we divided it into anterior temporal, midtemporal, and posterior temporal regions of interest (ROIs) by coronal planes at MNI $y = 0$ and $y = -36$. The midtemporal ROI was the most strongly associated with naming deficits, as stated above. To determine the relative importance of the anterior temporal and posterior temporal ROIs, we fit 2 general linear models with the naming score (BNT) as the dependent variable. The first model included dam-

Unauthenticated | Downloaded 09/17/23 03:19 AM UTC
s. m. wilson et al.

Age to the midtemporal ROI and damage to the anterior temporal ROI as independent variables, with lesion size and AQ as covariates. The midtemporal ROI predicted naming scores ($t = 3.83, p < 0.001$), but the anterior temporal ROI did not ($t = -0.20, p = 0.84$). The second model included damage to the midtemporal ROI and damage to the posterior temporal ROI as independent variables, with the same covariates. Again, the midtemporal ROI predicted naming scores ($t = 5.39, p < 0.001$), but in this analysis, the posterior temporal ROI did so as well ($t = 3.79, p = 0.005$). These analyses suggest that midtemporal and posterior temporal regions make independent contributions to naming, whereas anterior temporal damage does not impact naming after midtemporal damage is taken into account.

At the relatively early time point of 1 month postsurgery (Fig. 5 and Table 3), an association between inferior frontal resections and fluency deficits persisted ($p = 0.025$), although most of these fluency deficits were now mild. An association between ventral temporal resections and naming deficits persisted as well ($p < 0.001$). Associations between information content and resection location and between repetition and resection location were no

**FIG. 2.** Language measures before surgery, 2–3 days postsurgery, and 1 month postsurgery. A: Fluency. B: Information content. C: Comprehension. D: Repetition. E: Naming. F: AQ. All measures were significantly affected at 2–3 days postsurgery, and all measures recovered between 2–3 days and 1 month, so that the only score that differed from presurgical scores at 1 month postsurgery was naming. Boxes represent the interquartile range; whiskers, the range not including outliers; pluses, the outliers; thick horizontal lines, the medians; asterisks (within the body of each graph), the means. ***$p < 0.001$; *$p < 0.05$. Light gray lines indicate the trajectories of each patient. Red lines indicate the control mean; shaded red areas, within two standard deviations of the control mean. Where multiple patient lines coincide, the thickness of the lines is proportional to the square root of the number of patients. Distributions at 1 month postsurgery include imputed scores, but lines show only actually obtained data.
The temporal isthmus, which had been nonsignificantly associated with comprehension deficits in the immediate postsurgical period, was now significantly associated with comprehension deficits (p = 0.023). Resection volume was significantly negatively correlated with information content 1 month postsurgery (r = -0.30, p = 0.023), and there were trends toward negative correlations with AQ (r = -0.27, p = 0.057) and comprehension (r = -0.27, p = 0.055).

The VLSM analyses restricted to the 88 patients with gliomas yielded very similar results to the analyses of the whole cohort of patients. In particular, at 2–3 days postsurgery, the same regions shown in the main analyses were significantly associated with impaired fluency (p = 0.029), information content (p = 0.025), repetition (p = 0.028), and naming (p = 0.001), and no region was significantly associated with comprehension deficits.

Discussion

We found that transient aphasias were very common after resective surgery to the language-dominant left hemisphere, that all language domains could be affected, and that all major aphasia types occurred. There were systematic relationships between the location of surgical sites and deficits in specific language domains. Recovery was rapid, and by 1 month postsurgery, almost all language deficits had resolved.

Incidence and Nature of Transient Postoperative Aphasias

Seventy-one percent of patients were classified as aphasic based on the WAB when tested 2–3 days postsurgery, and for each language domain investigated, a majority of patients were impaired, with 88% of patients classified as aphasic and/or impaired on at least 1 language domain.

The incidence of language deficits in the immediate postoperative period has varied widely in previous studies, probably because of the differences in the sensitivity of the various means of testing aphasia, and the different distributions of resection sites. The largest series was that of Roberts, who reported transient aphasias after 144 (58%) of 246 left hemisphere resections in a variety of different brain regions. In another large series, Falconer and Serafetinides reported postoperative aphasias in 29 (52%) of 56 patients after surgery to the left temporal lobe. The considerably higher incidence of

**Table 2. Language measures before and after surgery**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Presurgery</th>
<th>2–3 Days Postsurgery</th>
<th>1 Mo Postsurgery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Abnormal</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Fluency</td>
<td>9.5 ± 1.2</td>
<td>27/88 (31%)</td>
<td>7.0 ± 3.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.3 ± 0.9</td>
</tr>
<tr>
<td>Information</td>
<td>9.9 ± 0.3</td>
<td>8/88 (9%)</td>
<td>7.6 ± 3.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9.8 ± 0.6</td>
</tr>
<tr>
<td>Comprehension</td>
<td>197.2 ± 5.1</td>
<td>7/86 (8%)</td>
<td>160.2 ± 54.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>194 ± 12.1</td>
</tr>
<tr>
<td>Repetition</td>
<td>94.9 ± 8.0</td>
<td>19/89 (21%)</td>
<td>69.2 ± 34.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>93.2 ± 8.3</td>
</tr>
<tr>
<td>Naming (BNT)</td>
<td>13.1 ± 2.4</td>
<td>25/85 (29%)</td>
<td>9.2 ± 5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.5 ± 3.1</td>
</tr>
<tr>
<td>Aphasia quotient</td>
<td>96.6 ± 5.2</td>
<td>15/84 (18%)</td>
<td>72.5 ± 30.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>94.9 ± 6.4</td>
</tr>
</tbody>
</table>

* All instances of unclassifiable patients were due to WAB naming subscores being too high for the variant with which the patient would otherwise have been classified. If WAB naming subscores were ignored presurgery, there would be an additional 4 cases classified as anomic, 1 Broca’s, 1 transcortical motor, and 1 conduction. At 2–3 days postsurgery, there would be an additional 3 anomic, 2 Broca’s, and 1 conduction. At 1 month postsurgery, there would be an additional 3 anomic and 1 conduction.

**Fig. 3.** Neural correlates of language deficits presurgery. Subsequent resection of cyan or warm-colored (yellow/orange/red) voxels was significantly correlated with presurgical comprehension or naming measures, respectively.
language deficits found in our study is probably a consequence of using a comprehensive aphasia battery, which is more sensitive in detecting milder language deficits.

We found that anomic aphasias were most common, but that Broca’s, Wernicke’s, conduction, and global aphasias all occurred quite frequently too. Previous studies have documented deficits in a range of language domains, including fluency, comprehension, repetition, and naming, but none have used aphasia batteries, so the precise nature of transient postsurgical aphasias was not clear from prior research.

**Neuroanatomical Correlates of Specific Transient Postoperative Language Deficits**

Voxel-based lesion-symptom mapping revealed systematic relationships between location of surgical sites and transient language deficits in specific domains. To our knowledge, our study is the first to show that the specific profiles of transient postoperative aphasias are related to resection site.

At 2–3 days postsurgery, fluency deficits were associated with resection of the precentral gyrus and inferior frontal junction, where the inferior frontal sulcus meets the ventral precentral sulcus. Previous studies have reported associations between reduced fluency and left inferior frontal damage in poststroke aphasia and primary progressive aphasia. Similarly, left inferior frontal regions have been implicated in functional imaging studies of speech production. The specific regions reported in these neuropsychological and neuroimaging studies have varied and have included the inferior frontal gyrus and sulcus, middle frontal gyrus, precentral gyrus, rolandic operculum, anterior insula, and underlying white matter (anterior part of the arcuate fasciculus). The relatively circumscribed resections in the patients in our study and the presumed lack of any functional reorganization given the short time frame may be taken to support our specific finding that a region spanning the precentral gyrus and inferior frontal junction, posterior and dorsal to Broca’s area, is the region most strongly associated with decreased fluency.

Reduced information content was associated with resection of the ventral precentral gyrus and posterior inferior frontal gyrus (pars opercularis). This region includes part of Broca’s area, along with caudally adjacent tissue, and is ventral to the region associated with fluency deficits. The information content measure assesses the functional, communicative value of a patient’s speech and reflects the patient’s ability to convey correct answers to 6 basic questions. Information content is a rather nonspecific measure in that deficits in comprehension, word finding, and/or speech production could all result in a failure to answer basic questions. However, it is noteworthy that damage to Broca’s area, long considered one of the most important brain regions for language, has the greatest impact on the functional, communicative value of patients’ speech.

Repetition deficits were associated with resection of the posterior superior temporal gyrus. This finding is consistent with studies of repetition deficits in poststroke aphasia and primary progressive aphasia. Functional imaging studies of healthy controls and cortical stimulation studies. Repetition deficits have also been associated with damage to or disruption of the adjacent inferior parietal regions and the arcuate component of the superior longitudinal fasciculus, which connects posterior temporal and inferior parietal regions to the frontal lobe.

Naming deficits were associated with resection of the middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, and parahippocampal gyrus. The strongest relationship was observed in the fusiform gyrus approximately 6 cm posterior to the temporal pole, in a region referred to as the basal temporal language area (BTLA). The BTLA

---

**TABLE 3. Brain regions showing significant correlations between resection location and language variables**

<table>
<thead>
<tr>
<th>Language Measure</th>
<th>Brain Region(s)</th>
<th>MNI Coordinates</th>
<th>Max t</th>
<th>Extent (mm³)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presurgery (resection only)</td>
<td>Anterior inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus</td>
<td>−32 −5 −34</td>
<td>5.57</td>
<td>11,328</td>
<td>0.003</td>
</tr>
<tr>
<td>Naming</td>
<td>Anterior middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus</td>
<td>−38 −5 −34</td>
<td>5.89</td>
<td>25,080</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

2–3 days postsurgery (resection plus infarction)

<table>
<thead>
<tr>
<th>Language Measure</th>
<th>Brain Region(s)</th>
<th>MNI Coordinates</th>
<th>Max t</th>
<th>Extent (mm³)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>Precentral gyrus, inferior frontal junction</td>
<td>−36 −4 33</td>
<td>4.08</td>
<td>3192</td>
<td>0.019</td>
</tr>
<tr>
<td>Information content</td>
<td>Precentral gyrus, posterior inferior frontal gyrus (pars opercularis)</td>
<td>−51 4 8</td>
<td>4.11</td>
<td>2576</td>
<td>0.035</td>
</tr>
<tr>
<td>Repetition</td>
<td>Posterior superior temporal gyrus</td>
<td>−52 −41 12</td>
<td>5.17</td>
<td>4560</td>
<td>0.024</td>
</tr>
<tr>
<td>Naming</td>
<td>Middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus</td>
<td>−43 −12 −30</td>
<td>6.09</td>
<td>44,456</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

1 mo postsurgery (resection plus infarction)

<table>
<thead>
<tr>
<th>Language Measure</th>
<th>Brain Region(s)</th>
<th>MNI Coordinates</th>
<th>Max t</th>
<th>Extent (mm³)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluency</td>
<td>Precentral gyrus</td>
<td>−52 6 16</td>
<td>4.89</td>
<td>3048</td>
<td>0.025</td>
</tr>
<tr>
<td>Comprehension</td>
<td>Temporal isthmus</td>
<td>−40 −12 −8</td>
<td>6.65</td>
<td>2648</td>
<td>0.023</td>
</tr>
<tr>
<td>Naming</td>
<td>Middle temporal gyrus, inferior temporal gyrus, fusiform gyrus, parahippocampal gyrus</td>
<td>−40 −4 −32</td>
<td>7.93</td>
<td>34,088</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
Transient aphasias after left hemisphere resective surgery was identified as most predictive of naming deficits, even though the VLSM analysis had more power in the temporal pole since more patients had resections there. Although the BLTA was the most highly predictive region of naming deficits, the cluster associated with naming deficits extended from the temporal pole all the way along the axis of the temporal lobe, almost reaching the occipital lobe. A follow-up ROI analysis showed that midtemporal and posterior temporal regions made independent contributions to naming deficits, whereas the anterior temporal region did not. Naming involves a series of processing stages including conceptual processing, word selection, phonological retrieval, phonological encoding, and articulation; hence, it relies on numerous brain regions and is vulnerable in all types of aphasia. The midtemporal and posterior temporal regions we identified may be relatively more involved in word selection and phonological retrieval, respectively.

A study of naming deficits in semantic dementia and another study of naming deficits in 3 variants of primary progressive aphasia both observed the strongest correlations between atrophy and naming deficits in the anterior fusiform gyrus, probably corresponding to the BLTA. The only language measure for which we found no association with specific surgical sites at 2–3 days postsurgery was comprehension. The most significant voxel in the VLSM t-map for comprehension was located in the white matter of the temporal isthmus. Tracts in this region include the inferior longitudinal fasciculus and the extreme capsule fiber system, both of which have been associated with auditory comprehension. A previous VLSM study indicated that comprehension deficits in chronic stroke patients were most associated with damage to the posterior middle temporal gyrus. Our study had little to no power in this region (Fig. 1) because resections to the posterior temporal cortex were relatively rare in this study cohort; this is probably why we did not find any significant lesion correlates of comprehension deficits.

**FIG. 4.** Neural correlates of language deficits 2–3 days postsurgery. A lateral projection (upper) and 2 slices (lower) are shown. Resection or infarction of colored voxels was significantly correlated with language measures 2–3 days postsurgery: blue = fluency; magenta = information content; green = repetition; yellow/orange/red = naming.
to surgery, forthcoming surgery to anterior ventral temporal regions was correlated with naming deficits. Therefore, it is certain that those patients who still showed deficits at 1 month postsurgery would subsequently show further improvements, but our study was not designed to examine long-term outcomes.

Rapid Recovery by 1 Month Postsurgery

At 1 month postsurgery, almost all language deficits that had been apparent 2–3 days postsurgery were now resolved in terms of severity and frequency, and no language measure except for naming differed significantly from its presurgical level. We must emphasize that 1 month postsurgery does not represent a stable end point, especially compared with the 3-month or 1-year postsurgery time points examined in previous outcome series. It is certainly likely that those patients who still showed deficits at 1 month postsurgery would subsequently show further improvements, but our study was not designed to examine long-term outcomes.

There was a slight but significant decrease in naming scores relative to presurgical levels, and resection of ventral temporal regions continued to be associated with naming deficits. This result is consistent with a number of studies documenting persistent naming deficits after anterior temporal lobectomy, in particular in patients with a later onset of symptoms.25,44,47,63 One study showed that naming deficits 6–12 months after anterior temporal lobectomy were worse when the basal temporal language area was resected than when it was not, even after accounting for resection volume.44 It should be noted that even prior to surgery, forthcoming surgery to anterior ventral temporal regions was correlated with naming deficits. Therefore, the association between ventral temporal damage and persistent naming deficits probably reflects a combination of damage due to the tumor (or other abnormality requiring surgery) and damage due to the resection and any associated infarction.

The few deficits that persisted at 1 month postsurgery should not obscure the striking recovery that took place in every language domain. The relatively rapid recovery from these transient postsurgical aphasias that we have documented raises the question of the mechanism(s) of the transient aphasia and its recovery. As mentioned above, there are several possibilities, which are not mutually exclusive.

Penfield and Roberts speculated that transient aphasias may be caused by nonspecific mechanisms after exposing the cortex to air and ultraviolet rays, and/or the electrical stimulation performed for language mapping. They referred to this phenomenon as “neuroparalytic edema” (p 141). The specific lesion-deficit correlations that we observed argue against a mechanism such as this, since a global mechanism would presumably affect different domains of language without regard to the specific location of the resection.

A more likely potential mechanism is edema in brain regions adjacent to the resection,29,49 with attendant cerebral hypoperfusion and neuronal dysfunction. If brain regions surrounding the resection are dysfunctional in the immediate postsurgical period, then the aphasia in this period would essentially reflect the effect of a larger lesion. The specific relationships we observed between lesion locations and aphasic symptoms would still be compatible with an explanation in terms of edema, since edema is generally localized to regions in the vicinity of resection. Edema would resolve in these surrounding regions and normal function would return, with aphasia resolving accordingly. Note that edema may or may not be visible on MRI. Moreover, many patients already showed substantial edema (probably of a different nature) in the vicinity of tumors prior to surgery. For these reasons, we were unable to quantify the extent of postsurgical edema. However, one observation suggests that edema of surrounding regions must be an important mechanism underlying transient aphasias in at least some patients: We observed that many patients’ language was actually less impaired immediately after surgery than when they were formally assessed after 2–3 days. This time course approximately matches the time course of edema following traumatic brain injury6 and clearly rules out a direct effect of the resection itself.

A third potential mechanism that also relies on the concept of dysfunction of regions other than the resection itself is diaschisis,49 that is, abnormalities in metabolism, neuronal function, or connectivity in intact brain regions as a consequence of the lesion. Diaschisis can be identified with functional imaging methodologies but would not be visible on structural MRI. In this scenario, recovery would involve plasticity in connected regions to adjust for the lack of inputs from the resected region. One argument against this explanation is that it would seem to predict that right-hemisphere resective surgeries would result in transient aphasias via diaschitic effects on homotopic brain regions, since transhemispheric diaschisis is the most well-established form of cortico-cortical diaschisis.18 However, we have not observed transient aphasias after right hemisphere surgery.

A final factor that may contribute to transient aphasias is the loss of functional tissue in the tumor itself (in the case of low-grade gliomas), immediately adjacent to the tumor and within the resection boundaries, or in adjacent areas that are infarcted as a consequence of surgery.27 If this is an important mechanism, then the rapid rate of recovery would be remarkable and would suggest that there is considerable redundancy within language networks, since the time frame of recovery seems too short for explicit relearning.

While the specific lesion-deficit correlations that we observed argue against a global mechanism, our data are
consistent with any of the other explanations we have discussed.

Limitations and Future Directions

Our study has several notable limitations. First, only routine clinical MRI studies were obtained, and they were only obtained 2–3 days postsurgery. Presurgical scans were not used, and follow-up scans 1 month postsurgery were not obtained. It is possible that a more comprehensive battery of scans, including functional MR images in particular, would more effectively identify functional abnormalities beyond the resection in the immediate postoperative period. This could enable us to quantitatively define a larger region of transient dysfunction than the area of the resection.

Second, although the WAB is a more comprehensive aphasia battery than the brief assessments used in previous studies investigating postsurgical aphasias, it still has several weaknesses. In particular, the fluency scale is subjective, maps a multidimensional deficit to an ordinal scale, and does not take into account features such as articulatory agility.33 The comprehension score does not discriminate between lexical and syntactic comprehension, and there is no assessment of nonverbal semantic processing. In future work, we plan to use quantitative analysis of connected speech samples to better quantify different aspects of fluency. Additionally, domain-specific tests could be added to evaluate aspects of language that are not well captured by the WAB.

Third, our cohort harbored diverse pathologies, including high- and low-grade gliomas, epilepsy, vascular malformations, and others. The majority of patients had tumors (88 patients had either low-grade or high-grade glioma), and we showed that the brain-language relationships in the whole group also held in this subset of patients, indicating that our findings are not attributable to any confound of etiology. However, we did not have a sufficient number of patients with any other etiology to determine whether brain-language relationships might differ as a function of etiology.33 There is evidence that language reorganization takes place because of epilepsy, tumors, or other abnormalities in language regions34 and may differ as a function of etiology.33 Any such reorganization of language regions constitutes a source of noise for our study, which would make it more difficult to identify consistent brain-language relationships. The fact that we found robust relationships between resection location and deficits in specific language domains suggests that language reorganization is generally not so dramatic as to obscure the basic functional neuroanatomy of the language network.

Conclusions

Transient aphasias are very common after resective surgery in the language-dominant left hemisphere, and there are systematic relationships between the location of surgical sites and deficits in specific language domains. Most language deficits resolved within 1 month. Our study did not attempt to characterize permanent deficits, as this would require significantly longer follow-up with equally careful testing. This patient cohort provides a unique window into the neural basis of language, since resections are discrete and their locations are not limited by vascular distribution or patterns of neurodegeneration, factors that limit the interpretation of many lesion-symptom mapping studies. Moreover, because we evaluated patient language 2–3 days after surgery, we can observe language function before substantial reorganization has taken place.

Acknowledgments

We thank Jessica DeLeon for her assistance with this project, and we thank all of the patients who participated in the research.

References


J Neurosurg Volume 123 • September 2015 591
Transient aphasias after left hemisphere resective surgery


**Author Contributions**

Conception and design: Wilson, Chang. Acquisition of data: all authors. Analysis and interpretation of data: Wilson, Lam, Babiak, Chang. Drafting the article: Wilson, Chang. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Wilson. Statistical analysis: Wilson. Study supervision: Wilson, Chang.

**Supplemental Information**

**Previous Presentation**

Portions of this work were presented in poster form at the Society for the Neurobiology of Language 2013 Annual Meeting, San Diego, California, November 6–8, 2013.

**Correspondence**

Stephen M. Wilson, Department of Speech, Language, and Hearing Sciences, University of Arizona, P.O. Box 210071, Tucson, AZ 85721. email: smwilson@u.arizona.edu.