The peri-Sylvian cortical networks underlying single word repetition revealed by electrocortical stimulation and direct neural recordings

Matthew K. Leonard[*,1,2], **Ruofan Cai**[*,1,2], **Miranda C. Babiak**[*,1,2], **Angela Ren**[*,1,2], and **Edward F. Chang**[*,1,2,3]

[1]Department of Neurological Surgery, University of California, San Francisco

[2]Center for Integrative Neuroscience, University of California, San Francisco

[3]Department of Physiology, University of California, San Francisco

Abstract

Verbal repetition requires the coordination of auditory, memory, linguistic, and motor systems. To date, the basic dynamics of neural information processing in this deceptively simple behavior are largely unknown. Here, we examined the neural processes underlying verbal repetition using focal interruption (electrocortical stimulation) in 59 patients undergoing awake craniotomies, and neurophysiological recordings (electrocorticography) in 8 patients while they performed a single word repetition task. Electrocortical stimulation revealed that sub-components of the left peri-Sylvian network involved in single word repetition could be differentially interrupted, producing transient perceptual deficits, paraphasic errors, or speech arrest. Electrocorticography revealed the detailed spatio-temporal dynamics of cortical activation, involving a highly-ordered, but overlapping temporal progression of cortical high gamma (75–150Hz) activity throughout the peri-Sylvian cortex. We observed functionally distinct serial and parallel cortical processing corresponding to successive stages of general auditory processing (posterior superior temporal gyrus), speech-specific auditory processing (middle and posterior superior temporal gyrus), working memory (inferior frontal cortex), and motor articulation (sensorimotor cortex). Together, these methods reveal the dynamics of coordinated activity across peri-Sylvian cortex during verbal repetition.

Keywords

electrocorticography; electrocortical stimulation; verbal repetition; peri-Sylvian cortex; neurosurgery; speech perception; speech production

Corresponding author: Edward F. Chang, 675 Nelson Rising Lane, Room 511, San Francisco, CA 94158, Phone: (415) 502-7346, ChangEd@neurosurg.ucsf.edu.
[*]All authors contributed equally.

1. Introduction

Hearing a word and repeating it is among the simplest verbal behaviors, yet it requires the coordination of multiple neural systems thought to subserve auditory, memory, linguistic, and motor functions. Recent evidence from functional imaging and detailed lesion behavior studies has implicated multiple peri-Sylvian regions in verbal repetition, and has provided a basis for unifying the functions of a variety of functionally distinct cortical regions in linguistic behaviors (Anderson et al., 1999; Baldo, Klostermann, & Dronkers, 2008; Herman, Houde, Vinogradov, & Nagarajan, 2013; Hope et al., 2014; Majerus, 2013; Moritz-Gasser & Duffau, 2013; Parker Jones et al., 2014; Pei et al., 2011; Price et al., 1996; Vallarb, Di Bettac, & Silveri, 1997; Warburton et al., 1996).

Activity in these regions must be coordinated in a precise temporal fashion for verbal repetition, as illustrated by the disorder known as conduction aphasia (Goodglass & Kohn, 1992). Patients with conduction aphasias have largely fluent natural speech production and perception capabilities, but with pervasive phonological paraphasias and a strong impairment in verbatim repetition (Buchsbaum et al., 2011). This pattern of deficits suggests that perceptual, linguistic, and motor representations are in some way disconnected from one another. The classical Wernicke/Geschwind model of language posits that damage to left hemisphere white matter tracts is responsible for conduction aphasia, best described as a disconnection syndrome. In particular, the arcuate fasciculus, which connects inferior frontal cortex with posterior superior temporal cortex, is hypothesized to be the primary pathway for integrating the various sub-processes involved in tasks like verbal repetition (Geschwind, 1974; Kümmerer et al., 2013; Parker Jones et al., 2014). However, there are several interesting case studies indicating that damage to the arcuate fasciculus is not a necessary condition for conduction aphasia (Quigg & Fountain, 1999; Quigg, Geldmacher, & Elias, 2006), and that it may not play a specific role in sensorimotor translation (Dick & Tremblay, 2012). There is also recent evidence that the sensorimotor transformations required for verbal repetition are bilateral (Cogan et al., 2014). Furthermore, recent quantitative morphometric neuroimaging analyses have suggested that damage to cortical gray matter in posterior temporo-parietal regions is more closely associated with the repetition deficits observed in a variety of aphasia disorders (Rogalsky et al., 2015).

These findings suggest that despite its apparent behavioral simplicity, verbal repetition abilities arise from the coordination of a complex set of cognitive computations supported by multiple brain regions. The sub-processes involved must transform acoustic input into motor output. Recently, we have gained a much deeper understanding of many of these sub-processes. A schematic hypothesis for repetition behavior posits the following flow of information through cortical circuits:

1. Auditory processing in posterior superior temporal gyrus extracts spectrotemporal features of spoken input (Mesgarani, Cheung, Johnson, & Chang, 2014), which are abstracted to phonemic representations at the population level (Chang et al., 2010).

2. Information about the phonological sequence is held in working memory, possibly in inferior frontal and posterior temporo-parietal cortex (Buchsbaum, Olsen, Koch, & Berman, 2005; Paulesu, Frith, & Frackowiak, 1993).

3. Articulatory processes in prefrontal and ventral sensorimotor cortex coordinate the movement of the lips, tongue, jaw, and larynx (Bouchard, Mesgarani, Johnson, & Chang, 2013).

4. Finally, auditory feedback is a crucial component of speech production, allowing fine-scale modulations of speech output in real time (Chang, Niziolek, Knight, Nagarajan, & Houde, 2013; Hickok, Houde, & Rong, 2011; Houde & Jordan, 1998).

The specific neural computations and representations at each successive stage are also known to be influenced by non-sensorimotor processes. For example, repetition performance is modulated by lexical factors including word or sentence length, word frequency, and phonological neighborhood density (Buchsbaum, Hickok, & Humphries, 2001; Graves, Grabowski, Mehta, & Gupta, 2008). Since both the speech signal and these lexical factors unfold in real-time, there must be parallel and overlapping interactions among the various stages of the sensorimotor transformation and other learned representations of linguistic knowledge (McClelland & Elman, 1986).

To date, the real-time coordination and organization of these sub-systems is largely unknown. Recent work has suggested that posterior temporal and inferior parietal cortex may be directly involved in the phonological working memory aspects of sensorimotor transformations (Buchsbaum et al., 2011; Cogan et al., 2014), consistent with the hypothesis that these regions link input and output functions in speech (Hickok & Poeppel, 2007). However, it is likely that many of the underlying representations involved in the various stages of verbal repetition involve distributed neural populations with precisely coordinated timing, which has not been described. To understand how auditory, phonological working memory, motor, and other linguistic sub-processes are coordinated to allow rapid verbal repetition, we utilized two complementary methods that allow us to identify regions critical for specific functions and provide high-resolution spatiotemporal pictures of neural activity while these regions process information. To establish causal roles for specific local brain regions in each sub-process in the behavior, we used electrocortical stimulation (ECS) to transiently interrupt local processing during verbal repetition (Boatman, Lesser, & Gordon, 1995; G. A. Ojemann, 1978; Penfield & Roberts, 1959). To examine the fine-scale spatiotemporal dynamics of neural activity during verbal repetition, we recorded electrocorticographic (ECoG) signals directly from the cortical surface Together, direct cortical stimulation and neurophysiological recordings provide unparalleled precision in mapping the peri-Sylvian regions involved in repetition, both to the extent that there is discrete functional anatomy, and also overlapping processes within cortical areas.

2. Materials and Methods

2.1 Patients: ECS and ECoG and cohorts

Fifty-nine patients undergoing awake craniotomy at the University of California San Francisco Medical Center volunteered as participants in this study (33 male, 26 female; age range 19–71; Table 1). Of the 50 patients with dominant hemisphere craniotomy, 42 were part of the ECS-only cohort, 3 were part of the ECoG-only cohort, and 5 patients had both ECS and ECoG. ECS and ECoG were not routinely completed together because of time limitations during surgery. Awake craniotomy with language mapping was performed on all patients due to the potential risk of post-operative speech and language difficulties and was used to identify and preserve eloquent cortex relating to motor control, naming, and repetition. All but three patients were right-handed (Table 1a). A pre-operative speech and language battery was routinely performed. With the exception of three patients who were slightly below normal range and exhibited mild anomia, we did not observe pre-operative aphasia symptoms according to the Western Aphasia Battery (Kertesz, 1982). Patients provided written informed consent, and the study was approved by the UCSF Institutional Review Board.

2.2 Electrocortical stimulation (ECS)

Patients underwent ECS mapping to determine essential sensory, motor, counting, naming, and repetition sites located in the exposed left hemisphere lateral cortex (Table 1a; 9 patients were mapped on the non-dominant hemisphere, Table 1b). ECS was carried out using the Ojemann Cortical Stimulator (Integra LifeSciences, Plainsboro NJ) with typical settings (60Hz, bipolar, biphasic). Stimulation threshold was determined on an individual basis, typically between 2–5mA, and set at a level that elicited speech arrest without causing after-discharges as determined by intraoperative electrocorticography. Detailed descriptions of intraoperative counting and picture naming tasks using this procedure have been described previously (Corina et al., 2010; Haglund, Berger, Shamseldin, Lettich, & Ojemann, 1994; G. A. Ojemann, 1978; G. Ojemann, Ojemann, Lettich, & Berger, 1989). Speech arrest was defined as the involuntary cessation of speech output during counting without any observed movement of the vocal tract articulators. Patients typically reported a feeling of trying to speak but "cannot get any words out".

2.3 Single word repetition task for ECS

All patients were familiarized with a 50-item word list[1] prior to surgery, which was used to verify patients' baseline ability to repeat single words. During the intraoperative ECS repetition task, in addition to the 50-item word list (List 1), we also used a four-syllable word list (List 2). List 2 contained 184 low-frequency, four-syllable words (log-transformed HAL frequencies were approximately ~6 for all items; note that not all words were presented for every patient in the intraoperative setting). These four-syllable words also provided the opportunity to examine the effect of ECS on syllable position, which allowed

---

[1]The 50-item word list consisted of two and four syllable words, which varied in length and articulatory difficulty. Difficult words contained consonant clusters and large articulatory movements from front to back; e.g., "statistical". This list also contained pseudowords that were derived from the real words by scrambling the segments, e.g. "delight" → "ledite", "ceremony" → "moanessery".

us to evaluate specific hypotheses regarding memory subprocesses in verbal repetition. Both lists were highly effective for ECS induced errors at sites that showed any disruptions (List 1: 58.39%, List 2: 77.96%), and we therefore included data from both lists in the analyses. Intraoperative video recording of patients' single word repetition responses were scored as correct or incorrect to determine positive error sites. An intraoperative photograph of the brain with markers on the stimulation sites was taken after mapping was completed. The stereotactic coordinates of the site locations were also acquired using BrainLab neuronavigation software.

Patients were instructed to repeat individual words, each time following the short prompting phrase, "say…" (e.g., "say 'questionnaire'"). Bipolar ECS was manually applied to the cortex at the end of the prompting phrase; each stimulation lasted approximately 1–3 seconds. For the purposes of clinical language mapping, the majority of the exposed posterior peri-Sylvian cortical surface was sampled during the repetition task. If stimulation at a given cortical site caused an error in repetition, the site was tested at least two more times, some up to 20 times depending on individual patients. If greater than 50% of the stimulations caused errors, that site was demarcated.

2.4 ECS data analysis

Repetition errors were coded as one of six categories: perceptual, phonological, no response, neologism, perseveration, motor speech, and offset (described in Table 2). All of the errors were coded by two experts and were analyzed based on three variables: cortical site, word type, and error type (Figure 1a). During the repetition task, we found two anterior STG sites in one patient that reliably caused disruptions, but we excluded these sites because this patient had a lesion in the posterior superior temporal gyrus, which anatomically displaced the local anatomy. To examine how word length affected repetition behavior, we further analyzed phonological errors that occurred on four-syllable words, where we hypothesized that different types of errors would occur for different syllable positions.

Stimulation sites were marked on individual MRI brain reconstructions based on intraoperative photographic markers (Figure 1a–b). To combine data across subjects, anatomical landmarks were used to map the data onto a common brain reconstruction. For each subject, the central sulcus and Sylvian fissure were traced on the individual anatomy. These traces were then matched as closely as possible to the same traces on the common space. Visual inspection of individual subjects showed that this procedure maintained the positions of stimulation sites relative to gyral anatomy, and also relative to other subjects. For descriptive purposes, we divided the STG into posterior, middle, and anterior regions by referencing the boundary of the pre-central and post-central gyri.

2.5 Electrocorticography (ECoG)

Eight patients underwent temporary, acute intraoperative placement of a high-density subdural electrode array (256 channels in 7 patients, 128 channels in 1 patient; 4 mm center-to-center spacing; Figure 1c) over the exposed peri-Sylvian cortex. The ECoG signal was recorded with a multichannel amplifier optically connected to a digital signal processor (TuckerDavis Technologies, Alachua USA) at a sampling rate of 3015 Hz. The array

placement was registered using frameless stereotactic neuronavigation software (BrainLab, Munich Germany) and verified manually using intra-operative photographs.

2.6 Single word repetition task for ECoG

A subset of the 50-item list (List 1) containing 28 real words was presented via free-field loudspeakers at approximately 80 dB SPL. At the onset of a trial, an auditory stimulus cue (slide projector sound) was delivered to prepare the patient for the trial, followed by the acoustic presentation of the word, a two second delay period, and another auditory cue (short beep sound) to signal the patient to repeat the target word (Figure 4a). We used a short delay to be able to disambiguate perception and production processes, and also to examine hypothesized working memory processes during the delay period.

2.7 ECoG data analysis

ECoG data were downsampled to 400 Hz, and each channel time series was visually and quantitatively inspected for artifacts or excessive noise. Our analysis focused on the high-gamma (HG) frequency component of the local field potential (70–150 Hz), which has been found to correlate well with extracellular multi-unit action potentials (Ray & Maunsell, 2010; Steinschneider, Fishman, & Arezzo, 2008). For each electrode, the time-varying high-gamma amplitude was extracted using the Hilbert transform, and was downsampled to 100 Hz for plotting purposes. The common average reference was removed from the signal on 16 channel blocks, corresponding to inputs to the amplifier. Event-related averages of HG analytic amplitude were taken relative to stimulus and response onsets, and were normalized relative to baseline statistics.

Two events were used to time-align the average HG event related potential (ERP) for each electrode: word presentation onset (audio speakers) and word production onset (recorded by microphone). Time epochs around each event (defined as 3 seconds before and after word presentation onset, which includes neural responses to slide sound and word presentation), and 3 seconds before response onset to 3 seconds after response onset (which includes preparatory activity and articulation of the word) were concatenated for each trial on each electrode.

To examine the response similarity across electrodes without assuming functional organization based solely on anatomy, we used two unsupervised clustering techniques. The first is convex non-negative matrix factorization (NMF), which decomposes the set of neural timecourses into basis functions that describe canonical response types (Ding, Li, & Jordan, 2010). A major advantage of this method is that it is a 'soft' clustering technique, which does not force a given data point into a single cluster. Instead, the neural data weighted by the NMF bases can be interpreted as probabilities of belonging to a given cluster. This is critical because we hypothesized that activity within circumscribed cortical areas is heterogeneous. NMF clustering allows us to examine the extent to which the activity at a single electrode is similar to a variety of response types. The basic form of the matrix factorization is:

$$X \approx FG^T$$

where X is the normalized matrix of neural data timeseries, F is the weight matrix, and G are the factor loadings, which can be used as spatial indicators to plot weights on the cortical surface.

The second clustering technique we used is k-means clustering. Unlike NMF, this approach forces each data point into a single cluster membership. However, it can be shown that k-means and NMF have the same form and interpretation (where G is orthogonal in k-means). Therefore, we used k-means to validate the results from NMF, and to provide a 'hard' clustering of the same data. The clustering criterion was the sample correlation between the concatenated timecourses of the perception- and production- locked neural data on each electrode. The *k-means* algorithm was initialized five times with random cluster centroid starting points, and the set of clusters that minimized the within-cluster sums of point-to-centroid distances was used. Therefore, electrodes in distributed brain regions that show similar response profiles are clustered, producing functionally relevant groups of electrodes, which are not necessarily spatially adjacent.

While there is no universally accepted method for choosing the number of clusters, we evaluated the explanatory power of different numbers of clusters. For NMF, since the data can be approximately reconstructed using the weights and factor loadings (see Equation 1), it is possible to measure the similarity of the reconstructed timecourses and the original timecourses on each electrode. This metric is derived from the ratio of the residual sum of squares and the total sum of squares of the difference between the original and reconstructed timecourses. Subtracting this ratio from 1 gives the coefficient of determination, or R^2 value.

We calculated the NMF solution for a range of clusters from 2 to 15, and examined the percent explained variance of each model. We found that 5 clusters represented an 'elbow' in the curve, suggesting that additional clusters did not contribute significant variance (Figure S1). Furthermore, 5 clusters explained ~90% of the variance in the data. We also examined the similarity of additional clusters to the original 5 for both NMF and k-means, and found that these additional clusters were either largely redundant with the first 5, or represented noise-like responses that did not diverge significantly from the baseline. Therefore, we examined the spatiotemporal patterns in the neural data described by 5 functionally-defined clusters using both NMF and k-means.

Once each timecourse was multiplied by the NMF weights (or assigned to a cluster in k-means), we calculated the mean response for each cluster. This meant that in some cases, anatomically distant regions could contribute to the same mean cluster timecourse, and likewise anatomically proximal regions showed significantly different response profiles and contributed to different clusters. To assess if each timepoint of the averaged ERPs was significantly greater than baseline, a one-way t-test was performed and False Discovery Rate (FDR) corrected ($p<0.01$).

All electrode positions were mapped on the surface of a common brain by measuring the distance along the Sylvian fissure as well as the dorsal-ventral distance from the Sylvian fissure along the central sulcus from images of each subject's brain reconstruction. We then

remapped them using the measured distances in reference to the Sylvian fissure of the common brain (Figure 1d).

3. Results

3.1 Intraoperative electrocortical stimulation

To establish which cortical sites are essential to the specific functions associated with verbal repetition, we applied electrocortical stimulation (ECS) to a wide range of temporal, frontal and parietal lobe sites while participants heard and repeated words aloud. While counting and picture naming are commonly used for clinical language and motor mapping, verbal repetition is not commonly used as a task in the ECS setting. Since repetition involves a transformation from auditory to motor output, stimulation during this task provides a unique opportunity to understand the essential or causal components of the full speech-motor cortical circuit.

Across the 47 patients who completed the ECS verbal repetition task on the dominant hemisphere, we applied ECS on 2225 trials[2]. In total, we found 163 sites where stimulation caused reproducible behavioral errors (649 total stimulations at these sites). Of all stimulations at these 163 sites, 477 trials disrupted speech, for an overall error rate of 73.5%. Overall, combining the two word lists, the majority of repetition error sites were located on the posterior and middle superior temporal gyrus (STG; n=104), and spanned a large portion of the gyrus from posterior to anterior (Figure 1b). Additionally, there were 9 out of 48 (19%) sites on the anterior STG, 5 out of 133 (4%) sites on the middle temporal gyrus (MTG), 25 out of 239 (10%) sites on the supramarginal gyrus (SMG), 3 out of 17 (18%) sites on the angular gyrus (AG), 10 out of 32 (31%) sites on the pre-central gyrus, and 7 out of 28 (25%) sites on the post-central gyrus. For cases with frontal lobe exposure, we stimulated 31 sites total and found 6 positive sites on the pars opercularis of the inferior frontal gyrus (IFG). For each individual error site, we calculated the number of errors according to specific behavioral subtypes (Table 2): perceptual, phonological, no response, neologism, perseveration, motor speech, and offset. The stimulations per patient that resulted in each error type were pooled and plotted on normalized anatomy (Figure 2).

This analysis revealed that across patients, specific anatomical sites were associated with different error subtypes. We observed both overlapping and distinct clusters of anatomical areas associated with particular functions. Perceptual errors (where the patient "did not hear" the target word) were most associated with sites along the entire span of STG that we stimulated, particularly in posterior and anterior regions. We also observed a large number of perceptual errors evoked by stimulation to the pars opercularis of the inferior frontal gyrus (Figure 2a).

Phonological errors (where the patient produced the target word with changes in phonemic content) were evoked by stimulation throughout the STG and part of SMG (Figure 2b). We examined whether there was a relationship between error rate and posterior-anterior location

---

[2]The numbers of trials and sites are estimates due to the limitations of video recording angles. In addition, subjects 80 and 101 do not have any brain video recordings, so they are excluded from this count.

along the STG, and found none (r=−0.32, p>0.37), indicating that phonological errors were evoked along the entire extent of the STG.

We also observed three other types of errors that were evoked primarily by stimulation to superior temporal cortex. No response errors, in which the patient did not produce any verbal response, were evoked by stimulation to posterior and middle STG, with a small number of additional errors in anterior STG and SMG (Figure 2c). Neologism errors, in which the patient produced a non-English word-like form, were evoked by stimulation to anterior STG, with individual errors in the pre-central and inferior frontal gyri (Figure 2d). Finally, perseveration errors, in which the patient simply repeated the word from the previous trial, were evoked by stimulation to posterior and middle STG, and SMG (Figure 2e).

These errors evoked by stimulation to the superior temporal cortex contrast with the types of errors associated with supra-Sylvian stimulation. Motor speech errors, where the patient produced a phonologically correct word with slurred or distorted speech, were most associated with two specific supra-Sylvian regions on middle pre-central gyrus and ventral post-central gyrus, and were less often evoked by stimulation to middle STG and IFG. In contrast, offset errors, where the patient does not produce a response until after stimulation, were caused by ECS in the pre-central and angular gyri, with some additional errors in the ventral IFG and along the STG.

This analysis revealed that phonological errors were the most common error type during single word repetition (Table 2). This is of particular interest given recent findings of phonological representations in STG (Mesgarani et al., 2014), and we therefore asked whether specific types of phonological errors were associated with specific stimulus features. We coded errors for each syllable on the four-syllable words according to the following categories: correct, syllable deletion, vowel deletion, consonant deletion, syllable addition, vowel addition, consonant addition, syllable substitution, vowel substitution, and consonant substitution. This analysis is meant to examine whether phonological errors evoked by stimulation to specific sites is a generalized effect for the whole word, or whether there is an interaction between syllable position and error type.

An omnibus repeated measures analysis of variance (ANOVA) with syllable number and error type as within-subject factors revealed a significant overall effect [$F(35,1155) = 24.66$, $p < 0.001$]. This motivated further analyses to examine how error type and syllable number explain the phonological errors we observed. There was a significant decrease in the number of correct responses over syllables 1–4 [$F(3,132) = 4.4862$, $p = 0.004$] (Figure 3). The tendency to produce errors on later syllables was driven by two specific types of phonological errors: syllable deletion [$F(3,132) = 7.66$, $p < 0.0001$], and syllable addition [$F(3,132) = 4.9652$, $p = 0.0027$]. This indicates that subjects were more likely to both delete and add syllables later in the word. We examined whether each phonological error type at each syllable position was associated with specific anatomical sites. We observed that the increase in syllable deletions over syllable positions 1–4 was associated with stimulation to sites across the extent of the STG, while syllable additions and the non-significant increase

in syllable substitutions were mostly associated with middle and anterior STG sites (Figure S2).

We also found that syllables were more likely to be disrupted compared to individual phonemes. We performed one-way ANOVAs on data averaged over syllable number to compare error rates for deletions, additions, and substitutions. In all cases, we observed main effects of syllable/phoneme: (deletions: $F(2,99) = 32.92$, $p < 10^{-10}$; additions: $F(2,99) = 4.65$, $p = 0.012$; substitutions: $F(2,99) = 3.13$, $p = 0.048$). For deletions and substitutions, this was due to higher error rates for syllables compared to both vowels and consonants. For additions, syllables and consonants had higher error rates compared to vowels. See Table S1 for post-hoc t-test results for each comparison.

We also tested 9 patients who had craniotomies exposing the non-dominant hemisphere (Table 1b). Across all patients and stimulation trials, we observed only a single site on middle STG where there were two instances of motor-speech errors.

3.2 Electrocorticography of single word repetition

The results from the stimulation experiments indicate that a network of spatially overlapping peri-Sylvian brain regions is involved in the various sub-processes associated with verbal repetition. ECS has allowed us to identify regions that are essential for hearing and producing the correct sequence of phonemes and syllables, however it does not provide an account of how these regions interact with one another during the complicated series of cognitive and sensorimotor processes required to repeat words aloud. To characterize the network dynamics in detail, we recorded neural activity from high-density ECoG arrays implanted directly on the lateral surface of the left hemisphere in 8 patients undergoing awake craniotomy. Each array consisted of a grid of 256 electrodes (128 in one patient) spaced 4mm apart, providing high spatial resolution for examining activity both within and across cortical regions (Figure 1c, d). This method also allows us to examine the fine-scale temporal dynamics of neural activity during each sub-process of verbal repetition.

A standard approach for determining which brain regions are involved in a task or behavior is to define regions of interest (ROIs) and average the activity of all measurements within each brain region. This approach assumes homogeneous activity within spatially adjacent brain regions, usually encompassing several centimeters of cortex. The results from the ECS experiments suggest that greater resolution within brain regions is necessary to understand the neural basis of verbal repetition. Furthermore, cortical evoked response patterns tend to be relatively heterogeneous across individuals (Cogan et al., 2014; Flinker, Chang, Barbaro, Berger, & Knight, 2011; Leuthardt et al., 2011), making spatial averaging difficult, even once individual subjects' anatomy is transformed into a standard space. Therefore, in addition to the fine-scale resolution of the ECoG recording method, we employed an analysis approach that does not make assumptions about the spatial homogeneity of neural responses. Activity at each ECoG electrode was clustered using both 'soft' (non-negative matrix factorization, [NMF](Ding et al., 2010)) and 'hard' (*k*-means) clustering techniques that find similar patterns of activity. This approach allows us to examine which local cortical regions contribute to the various stages of verbal repetition.

The verbal repetition task had two major behavioral stages: Participants heard a cue (slide projector sound) followed by a word (Figure 4a). Then, after a two-second delay, they were given another cue to repeat the word they had heard previously (Figure 4b). Neural activity on each electrode was segmented relative to the onset of the word during the listening phase, and during the acoustic onset of the participants' speech output. These two timecourses were concatenated and NMF was used to find basis functions that describe the 5 canonical response types (Figure S1). The neural timecourses for each electrode were then projected onto each basis and averaged to generate mean and standard error timecourse plots. Finally, the factor loadings from NMF (see Equation 1) were used to plot each electrode's probability of cluster membership on the cortical surface.

Cluster 1 showed a timecourse that indicated non-specific auditory processing. Two peaks in the HG signal were apparent, one after the listen cue, and another after the presentation of the spoken word (Figure 4c). Both peaks occurred early relative to the acoustic onset (170ms and 210ms, respectively). One-way t-tests compared to a baseline response of zero (since the data were z-scored relative to a rest block) revealed that for the response to the cue, activity was significantly greater than baseline from 20ms before cue onset to 410ms after cue onset (FDR corrected $P < 0.05$). The fact that activity was significant slightly before the cue onset likely reflects small deviations in the baseline activity from the assumed value of zero. The response to the word was significant from 10ms after word onset to 570ms after word onset (FDR corrected $P < 0.05$). Cluster 1 also showed a small response to the onset of the speaking cue, and from ~1600–2900ms after the onset of the participants' own speech (FDR corrected $P < 0.05$; Figure 4d). There was a small peak after the onset of the participants' own speech, however it was not significantly different from baseline. Electrodes with weights greater than 20% contributing to cluster 1 were localized primarily to posterior and middle STG, with a small number in MTG and SMG, and a single electrode was localized to dorsal vSMC (Figure 4e). In summary, cluster 1 represents short-latency responses to acoustic input, primarily generated in high-order auditory cortex.

Cluster 2 showed a timecourse that was similar to cluster 1, except that responses were selective to hearing speech, and did not respond to either non-speech cue sound. Relative to the onset of the spoken stimulus, the neural response peaked at 380ms, and was significantly greater than baseline from 90–910ms (FDR corrected $P < 0.05$; Figure 4f). There was also a significant response to hearing the participants' own speech (auditory feedback), peaking at 280ms, and significantly greater than baseline from 80–890ms after acoustic onset (FDR corrected $P < 0.05$; Figure 4g). There were no significant responses to either the listening or speaking cues. Electrodes with weights greater than 20% contributing to cluster 2 were localized primarily to the entire extent of the STG, extending into posterior MTG and SMG (Figure 4h). There were also electrodes over the pars triangularis of the IFG and the dorsal vSMC. In summary, cluster 2 represents relatively short-latency responses to hearing speech, both external and self-produced.

Verbal repetition involves holding the item to be repeated in working memory during the delay period. Cluster 3 exhibited neural activity that is consistent with ongoing neuronal firing during the delay period. Relative to stimulus onset, there was a small peak at 570ms, which was significantly greater than baseline from 20–1930ms (FDR corrected $P < 0.05$;

Figure 4i). During the delay period and continuing through speech output, cluster 3 was significantly active (−3000ms to 1130ms relative to speaking onset; FDR corrected $P < 0.05$; Figure 4j). This significant activity continued intermittently until ~2600ms after speaking onset. Electrodes with weights greater than 20% contributing to cluster 3 were localized in a diffuse network including the entire extent of the STG, posterior MTG, inferior posterior parietal cortex, pars orbitalis and pars opercularis of the IFG, and both ventral and dorsal vSMC (Figure 4k). In summary, cluster 3 represents ongoing activity during the delay period in regions involved in phonological working memory.

Cluster 4 closely resembled the timecourse observed in cluster 2. There were prominent peaks in response to speech input (both listening and self-produced speech), but not to the non-speech cue sounds. However, both peaks occurred later than the corresponding peaks in cluster 2 (listening: 500ms, Figure 4l; speaking: 540ms, Figure 4m). Furthermore, activity was significantly greater than baseline for longer periods during both the listening and speaking phases. In particular, relative to speaking onset, activity was significant from −450ms to 2790ms (FDR corrected $P < 0.05$). There was also intermittent activity during the initial portion of the delay period (Figure 4l). The speech-evoked response was also smaller in amplitude compared to the analogous peak in cluster 2. Electrodes with weights greater than 20% contributing to cluster 4 were localized primarily to the entire extent of the STG, posterior MTG, pars orbitalis and pars triangularis of the IFG, and both ventral and dorsal vSMC (Figure 4n). In summary, cluster 4 represents activity evoked by hearing speech, and also some contribution of working memory processes during the delay period.

Finally, cluster 5 displayed activity that was consistent with speech production processes. There were no visible peaks during the listening phase, and only limited intermittent significant activity during the delay period (FDR corrected $P < 0.05$; Figure 4o). During the speaking phase, there was a prominent peak 190ms after acoustic speech onset, which was significant from −1110ms to 1290ms (FDR corrected $P < 0.05$; Figure 4p). Electrodes with weights greater than 20% contributing to cluster 5 were localized primarily to ventral, middle, and dorsal vSMC, with a small number of electrodes on posterior STG, MTG, and SMG (Figure 4q). In summary, cluster 5 represents motor planning and speech output during the speaking phase.

The NMF clustering method does not force each electrode to belong to a single cluster, but instead assigns probabilities of cluster membership. To provide a description of these clusters as discrete – although not necessarily spatially contiguous – functional groups, we replicated the clustering analysis using k-means clustering. We found that the average cluster timecourses and electrode groupings were highly similar between the two methods (Figure S3, Figure S4). This suggests that verbal repetition sub-processes are highly distributed across cortical regions, but may have important hubs in local neural populations.

Finally, we examined how these clusters process higher-order linguistic information such as lexical frequency. To examine the influence of higher-order linguistic information on the cortical network involved in verbal repetition, we performed a regression analysis that explained neural activity on each electrode at each timepoint during the listening phase as a linear combination of each stimulus' lexical frequency and length. Across all 438 electrodes,

the regression R^2 ranged from ~0.1 to ~0.75. The majority of significant effects occurred between ~500–1000ms after word onset (Figure 5a). During that time window, we evaluated whether any of the clusters derived from *k*-means showed stronger effects of lexical frequency and word length. There was considerable variability in R^2 values within each cluster, however the mean R^2 was not significantly different across clusters ($P > 0.05$; Figure 5b). Across the cortex, there were significant R^2 values at electrodes belonging to every cluster, with no clear dominance of any cluster (Figure 5c).

In summary, these results reveal a distributed set of cortical regions that process specific types of information during verbal repetition. However, there is large overlap within cortical regions encompassing multiple functions. Furthermore, while each functional group plays a distinct role in verbal repetition, all are equally (and strongly) influenced by higher-order linguistic knowledge.

4. Discussion

Repetition is a fundamental language ability. It is an important basis for language acquisition, where learning new words requires mapping acoustic inputs to motor outputs, and serves as a foundation for sensorimotor integration in speech behavior. There are multiple processes and representations involved in verbal repetition, including auditory, phonological, working memory, linguistic, and motor processes, and it is largely unknown how these functions are coordinated to produce a fluid and deceptively simple behavior.

The present study aimed to determine the cortical circuitry underlying single word repetition by using the complementary approaches of electrocortical stimulation (ECS) mapping and electrocorticography (ECoG) of left hemisphere peri-Sylvian cortex. ECS demonstrated that a range of overlapping functions are attributable to individual cortical regions in the superior temporal, inferior frontal, and inferior parietal lobes, while ECoG revealed a dynamic visualization of the real-time flow of cortical activation throughout the peri-Sylvian cortex.

It is well known that there is great variability in cortical functional anatomy between individuals (Cogan et al., 2014; Flinker et al., 2011; Leuthardt et al., 2011). Our goal in the present study of examining the fine-scale dynamics of verbal repetition required this variability to be accounted for, rather than averaged out, as is typically the case in ROI analyses of functional brain data. Therefore, we used an unsupervised machine learning approach to cluster ECoG data according to functional properties, rather than anatomy. This allowed us to more clearly delineate the time course of cortical activity during specific phases of the verbal repetition task. Combined with the results from the ECS portion of the study, our findings provide new evidence for a unified and highly dynamic circuit involved in both the input and output of speech in peri-Sylvian cortex (Flinker et al., 2011).

During the listening phase, cortical activity was primarily localized to the STG, with some activity in the motor and frontal cortex (Edwards et al., 2010; Wilson, Saygin, Sereno, & Iacoboni, 2004). However, this anatomical organization reflects the selective responses of specific STG neural populations. While some STG neural populations respond at an early latency to any sound input (cluster 1), other populations respond selectively to speech at a

similar latency (cluster 2; only two non-speech sounds were tested, therefore it is possible that this cluster would respond to other non-speech stimuli). These functions in pSTG correspond with the results from ECS, in which stimulation at some sites caused perceptual (general auditory) disruptions, while stimulation at other sites caused phonological (linguistic) disruptions. While previous work has shown stimulus-selective responses in pSTG (Chang et al., 2010; DeWitt & Rauschecker, 2012; Steinschneider et al., 2011), our findings demonstrate that this region is organized according to a complex network-level functional anatomy (ECoG), which involve population-specific essential processing hubs (ECS).

While STG neural populations respond to any auditory input, we found evidence that they are selective to the source and type of sound. Cluster 1 showed robust ECoG responses to the listening cue, the auditory word, and the production cue. It did not respond to the participant's production of the target word. Furthermore, cluster 2 showed responses only to speech auditory stimuli, and not to either cue sound. However, the response to auditory input from self-produced speech was smaller compared to hearing external speech. These results are consistent with the literature on self-monitoring during vocal production in both monkeys (Eliades & Wang, 2008) and humans (Greenlee et al., 2011), and likely reflect feedback control processes (Hickok et al., 2011; Houde & Jordan, 1998).

We also observed longer latency perceptual and auditory-evoked responses to words in the anterior aspects of the superior temporal lobe and in frontal regions including the ventral sensorimotor cortex (e.g., clusters 2 and 4). There is accumulating evidence that later activity in more anterior STG neural populations reflects the higher-order hierarchical processing of speech as part of the ventral stream of word recognition (Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003; DeWitt & Rauschecker, 2012; Hickok & Poeppel, 2007; Leonard & Chang, 2014; Rauschecker & Scott, 2009). In contrast, auditory-evoked activity in frontal motor areas has generated a lively debate regarding the role of motor regions in speech perception (Hickok et al., 2011; Saur et al., 2008). However, there were some inferior frontal sites where ECS disrupted perceptual and phonological processes, suggesting that whatever role these frontal regions play in speech perception, it is critical for understanding auditory input (Cheung, Hamiton, Johnson, & Chang, 2016).

We used a verbal repetition paradigm to examine the multiple parallel and overlapping parts of the speech circuit, including perception and production. The time between these two end stages (the delay period) provided an opportunity to examine the cortical processes involved in phonological, working memory, and sensorimotor representations. We found that there is a set of frontal and parietal regions encompassing the ventral sensorimotor cortex and inferior frontal gyrus, which show sustained activity during the delay period. While it is possible that this activity reflects a passive rehearsal process, the short time window (2 seconds) may not necessitate such a conscious process, and instead may reflect phonological buffering in verbal working memory (Buchsbaum et al., 2011) and the subsequent sensorimotor transformation required by the task (Cogan et al., 2014). There has been a strong interest in understanding the relationships between acoustic and motor representations in speech, and there are a variety of proposals regarding where the sensorimotor transformation takes place (Cogan et al., 2014; Hickok, 2012; Rauschecker &

Scott, 2009). Our results suggest that multiple regions are responsible for this process, and that they are both spatially and temporally distributed.

It is noteworthy that we observed only two instances of disrupted repetition behavior when ECS was applied to the non-dominant hemisphere in nine patients. Recent reports (Cogan et al., 2014) have suggested that the kinds of sensorimotor transformations required by a verbal repetition task are mediated by processes in bilateral inferior frontal, inferior parietal, superior temporal, and sensorimotor cortical areas. However, the causal role of these right hemisphere areas is less clear (Tate, Herbet, Moritz-Gasser, Tate, & Duffau, 2014). Our results suggest that even for explicit sensorimotor transformation tasks like verbal repetition, these areas are not critical, or at least do not perform the same functions as their left hemisphere homologues.

It is also interesting to note that we observed a unique role for syllables in the ECS experiment. It was much more common for ECS to cause subjects to delete or add syllables compared to phonemes such as vowels and consonants (and syllable substitutions were more likely than vowel substitutions), suggesting their possible role as a fundamental unit of speech production. The ECS results further support this notion because they show that specific cortical sites (including ventral parietal and superior temporal cortex) are critical for particular phonological transformations. For example, whereas stimulation at certain sites causes vowel substitutions, stimulation at other sites causes syllable deletions. This suggests that these distributed areas represent not only different types of phonological information (e.g., vowels vs. consonants), but also different hierarchical levels of speech (e.g., phonemes vs. syllables).

We also examined the neural processes associated with producing speech during the verbal repetition task. As has been demonstrated previously, a set of regions encompassing the ventral sensorimotor cortex showed responses that begin before and peak just after the acoustic onset of speech (clusters 4–5; (Bouchard et al., 2013)). Intermingled among these areas, we found electrodes that responded to both perception and production (clusters 2–4), which indicates responses to auditory feedback of self-produced speech. Interestingly, there were some supra-Sylvian electrodes that responded to both perception and production (cluster 4), further suggesting a high-level role for frontal motor regions in sensorimotor transformations.

Finally, recent work has also demonstrated that distributed cortical networks encode abstract statistical information about lexical and other linguistic features, including lexical frequency (Cibelli, Leonard, Johnson, & Chang, 2015; Leonard, Bouchard, Tang, & Chang, 2015). These studies challenge the notion of a 'mental lexicon' that is simply a dictionary of all known words (Gow Jr, 2012; Ullman et al., 1997), and provide evidence that many aspects of speech processing including acoustic-phonetic, phonemic, and lexical processing are deeply influenced by language statistics. The present work supports these conclusions in demonstrating that lexical frequency has generalized effects across all functional clusters involved in verbal repetition.

It is important to consider the fact that the present data were obtained in a patient population undergoing invasive surgery for tumor or epileptogenic focus resection. While previous research in these clinical circumstances has yielded findings that are consistent and replicable in normal subjects using non-invasive methods (Leonard & Chang, 2015), the conclusions of the present study must be interpreted in the context of the limitations of potentially abnormal brains. We did not observe any large-scale cortical reorganization that differs from what is normally seen in healthy subjects, however there may be more subtle differences in some aspects of the fine-scale organization of functions involved in verbal repetition. At a more basic level, there may be anatomical displacement of local anatomy due to lesions in peri-Sylvian cortex. We attempted to minimize these effects by doing initial analyses on individual subject anatomy, taking such displacement into account, however this may make the projection onto a common anatomical space less accurate for the group-level analyses. Ultimately, these data confirm and extend previous work in healthy subjects and patient populations that do not share these potential issues.

4.1 Conclusions

Our results contribute to the characterization of the complete speech circuit. The Geschwind model (Geschwind, 1974) is still commonly referenced in literature reviews and textbooks as the framework for the neural basis of repetition. This model posits that the arcuate fasciculus acts as a relay between a posterior perceptual site (Wernicke's) and an anterior production site (Broca's), and that a "disconnection" between these two areas results in a characteristic disruption of repetition. However, as recent studies have shown (Anderson et al., 1999; Boatman et al., 1995; Buchsbaum et al., 2011; Dronkers, 2000; Quigg & Fountain, 1999) disruption to the arcuate fasciculus alone does not accurately describe the detailed neural processes that are involved in the seemingly simple task of single word repetition. While some researchers highlight pSTG's role in accessing phonology, others argue that pSTG and the temporal planum (area Spt) are responsible for the sensory-motor integration from an auditory input code to an articulatory output code (Buchsbaum et al., 2011; Hickok & Poeppel, 2007). Our data support both theories. On one hand, our ECoG data heavily implicate cortical areas involved in sensorimotor processing throughout all phases of perception and production. On the other hand, ECS in posterior temporal regions demonstrates complete and partial disruption (perceptual and phonological errors) to online acoustic and linguistic information in the absence of disruption to motor planning and production. Thus, even though ECoG data suggest that subjects are simultaneously accessing sensorimotor codes during perceptual phases of repetition, ECS disruption during early perceptual phases does not result in difficulties of motor production of target words. Taken together, our ECoG and ECS data demonstrate that verbal repetition is not simply a multi-node, serial processing event. Rather it involves distributed parallel processing with coordination between many peri-Sylvian and sensorimotor regions that differentially respond to stimulus presentation and word production during the single word repetition task.

From a clinical perspective, patients undergoing neurosurgical procedures in the language-dominant hemisphere are at high risk for acquiring post-operative speech and language deficits (Haglund et al., 1994; Penfield & Roberts, 1959). ECS during awake craniotomy surgery is currently the gold standard of clinical brain mapping within this population

(Berger, 1994; Berger & Ojemann, 1992; Duffau et al., 2003; G. Ojemann et al., 1989). Established intraoperative tasks during ECS mapping include counting, picture naming, and reading (Berger, 1994; Corina et al., 2010; Penfield & Roberts, 1959). Despite its routine use in neuropsychological and speech-language evaluations (Goodglass & Kohn, 1992), however, repetition has rarely been reported as a clinical tool during ECS mapping (Boatman et al., 2000, 1995; Duffau et al., 2003). The present results suggest that verbal repetition is a highly informative behavioral task for mapping multiple sensory, cognitive, linguistic, and motor functions using both stimulation and neurophysiological recording techniques.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Anderson J, Gilmore R, Roper S, Crosson B, Bauer R, Nadeau S, … Kortencamp S. Conduction aphasia and the arcuate fasciculus: a reexamination of the Wernicke Geschwind model. Brain and Language. 1999; 70(1):1–12. [PubMed: 10534369]

Baldo JV, Klostermann EC, Dronkers NF. It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace. Brain and Language. 2008; 105(2):134–140. [PubMed: 18243294]

Berger M. Lesions in functional (" eloquent") cortex and subcortical white matter. Clinical Neurosurgery. 1994; 41:444. [PubMed: 7842620]

Berger M, Ojemann G. Intraoperative brain mapping techniques in neuro-oncology. Stereotactic and Functional Neurosurgery. 1992; 58(1–4):153–161. [PubMed: 1439333]

Boatman D, Gordon B, Hart J, Selnes O, Miglioretti D, Lenz F. Transcortical sensory aphasia: revisited and revised. Brain. 2000; 123(8):1634–1642. [PubMed: 10908193]

Boatman D, Lesser RP, Gordon B. Auditory speech processing in the left temporal lobe: an electrical interference study. Brain and Language. 1995; 51(2):269–290. [PubMed: 8564472]

Bouchard KE, Mesgarani N, Johnson K, Chang EF. Functional organization of human sensorimotor cortex for speech articulation. Nature. 2013; 495(7441):327–332. [PubMed: 23426266]

Buchsbaum BR, Baldo J, Okada K, Berman KF, Dronkers N, D'Esposito M, Hickok G. Conduction aphasia, sensory-motor integration, and phonological short-term memory an aggregate analysis of lesion and fMRI data. Brain and Language. 2011; 119(3):119–128. [PubMed: 21256582]

Buchsbaum BR, Hickok G, Humphries C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. Cognitive Science. 2001; 25(5):663–678.

Buchsbaum BR, Olsen RK, Koch P, Berman KF. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron. 2005; 48(4):687–697. [PubMed: 16301183]

Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. Human cortical sensorimotor network underlying feedback control of vocal pitch. Proceedings of the National Academy of Sciences. 2013; 110(7):2653–2658.

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. Nature Neuroscience. 2010; 13:1428–1432. [PubMed: 20890293]

Cheung C, Hamiton LS, Johnson K, Chang EF. The auditory representation of speech sounds in human motor cortex. eLife. 2016; 5:e12577. http://doi.org/10.7554/eLife.12577. [PubMed: 26943778]

Cibelli ES, Leonard MK, Johnson K, Chang EF. The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening. Brain and Language. 2015; 147:66–75. [PubMed: 26072003]

Cogan GB, Thesen T, Carlson C, Doyle W, Devinsky O, Pesaran B. Sensory-motor transformations for speech occur bilaterally. Nature. 2014; 507(7490):94–98. [PubMed: 24429520]

Corina DP, Loudermilk BC, Detwiler L, Martin RF, Brinkley JF, Ojemann G. Analysis of naming errors during cortical stimulation mapping: implications for models of language representation. Brain and Language. 2010; 115(2):101–112. [PubMed: 20452661]

Crinion JT, Lambon-Ralph MA, Warburton EA, Howard D, Wise RJ. Temporal lobe regions engaged during normal speech comprehension. Brain. 2003; 126(5):1193–1201. [PubMed: 12690058]

DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream. Proceedings of the National Academy of Sciences. 2012; 109:E505, E514.

Dick AS, Tremblay P. Beyond the arcuate fasciculus: Consensus and controversy in the connectional anatomy of language. Brain. 2012:aws222.

Ding C, Li T, Jordan M. Convex and semi-nonnegative matrix factorizations. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2010; 32(1):45–55.

Dronkers NF. The pursuit of brain language relationships. Brain and Language. 2000; 71(1):59–61. [PubMed: 10716807]

Duffau H, Capelle L, Denvil D, Sichez N, Gatignol P, Taillandier L, … Muller J-C. Usefulness of intraoperative electrical subcortical mapping during surgery for low-grade gliomas located within eloquent brain regions: functional results in a consecutive series of 103 patients. Journal of Neurosurgery. 2003; 98(4):764–778. [PubMed: 12691401]

Edwards E, Nagarajan SS, Dalal SS, Canolty RT, Kirsch HE, Barbaro NM, Knight RT. Spatiotemporal imaging of cortical activation during verb generation and picture naming. Neuroimage. 2010; 50(1):291–301. [PubMed: 20026224]

Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature. 2008; 453(7198):1102–1106. [PubMed: 18454135]

Flinker A, Chang E, Barbaro N, Berger M, Knight R. Sub-centimeter language organization in the human temporal lobe. Brain and Language. 2011; 117(3):103–109. [PubMed: 20961611]

Geschwind, N. Disconnexion syndromes in animals and man. Springer; 1974.

Goodglass H, Kohn S. Diagnosis of conduction aphasia. Conduction Aphasia. 1992:39–49.

Gow DW Jr. The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. Brain and Language. 2012; 121:273–288. [PubMed: 22498237]

Graves WW, Grabowski TJ, Mehta S, Gupta P. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. Cognitive Neuroscience, Journal of. 2008; 20(9):1698–1710.

Greenlee JD, Jackson AW, Chen F, Larson CR, Oya H, Kawasaki H, … Howard MA III. Human auditory cortical activation during self-vocalization. PLoS One. 2011; 6(3):e14744. [PubMed: 21390228]

Haglund MM, Berger MS, Shamseldin M, Lettich E, Ojemann GA. Cortical localization of temporal lobe language sites in patients with gliomas. Neurosurgery. 1994; 34(4):567–576. [PubMed: 7516498]

Herman AB, Houde JF, Vinogradov S, Nagarajan SS. Parsing the phonological loop: activation timing in the dorsal speech stream determines accuracy in speech reproduction. The Journal of Neuroscience. 2013; 33(13):5439–5453. [PubMed: 23536060]

Hickok G. Computational neuroanatomy of speech production. Nature Reviews Neuroscience. 2012; 13:135–145. [PubMed: 22218206]

Hickok G, Houde J, Rong F. Sensorimotor integration in speech processing: computational basis and neural organization. Neuron. 2011; 69:407–422. [PubMed: 21315253]

Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience. 2007; 8:393–402. [PubMed: 17431404]

Hope TMH, Prejawa S, Parker Jones 'Oiwi, Oberhuber M, Seghier ML, Green DW, Price CJ. Dissecting the functional anatomy of auditory word repetition. Frontiers in Human Neuroscience. 2014; 8 http://doi.org/10.3389/fnhum.2014.00246.

Houde JF, Jordan MI. Sensorimotor adaptation in speech production. Science. 1998; 279(5354):1213–1216. [PubMed: 9469813]

Kertesz, A. Western aphasia battery test manual. Psychological Corp; 1982.

Kümmerer D, Hartwigsen G, Kellmeyer P, Glauche V, Mader I, Klöppel S, … Saur D. Damage to ventral and dorsal language pathways in acute aphasia. Brain. 2013; 136(2):619–629. [PubMed: 23378217]

Leonard MK, Bouchard KE, Tang C, Chang EF. Dynamic Encoding of Speech Sequence Probability in Human Temporal Cortex. The Journal of Neuroscience. 2015; 35(18):7203–7214. [PubMed: 25948269]

Leonard MK, Chang EF. Dynamic speech representations in the human temporal lobe. Trends in Cognitive Sciences. 2014; 18(9):472–479. [PubMed: 24906217]

Leonard, MK., Chang, EF. Neurobiology of Language. Elsevier; 2015. Neural organization of speech perception: Intracranial Recording.

Leuthardt EC, Gaona C, Sharma M, Szrama N, Roland J, Freudenberg Z, … Schalk G. Using the electrocorticographic speech network to control a brain computer interface in humans. Journal of Neural Engineering. 2011; 8(3):036004. [PubMed: 21471638]

Majerus S. Language repetition and short-term memory: an integrative framework. Frontiers in Human Neuroscience. 2013; 7

McClelland JL, Elman JL. The TRACE model of speech perception. Cognitive Psychology. 1986; 18:1–86. [PubMed: 3753912]

Mesgarani N, Cheung C, Johnson K, Chang EF. Phonetic feature encoding in human superior temporal gyrus. Science. 2014

Moritz-Gasser S, Duffau H. The anatomo-functional connectivity of word repetition: insights provided by awake brain tumor surgery. Frontiers in Human Neuroscience. 2013; 7 http://doi.org/10.3389/fnhum.2013.00405.

Ojemann GA. Organization of short-term verbal memory in language areas of human cortex: evidence from electrical stimulation. Brain and Language. 1978; 5(3):331–340. [PubMed: 656902]

Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere: an electrical stimulation mapping investigation in 117 patients. Journal of Neurosurgery. 1989; 71(3):316–326. [PubMed: 2769383]

Parker Jones, 'Oiwi, Prejawa, S., Hope, T., Oberhuber, M., Seghier, ML., Leff, AP., … Price, CJ. Sensory-to-motor integration during auditory repetition: A combined fMRI and lesion study. Frontiers in Human Neuroscience. 2014; 8 http://doi.org/10.3389/fnhum.2014.00024.

Paulesu E, Frith CD, Frackowiak RS. The neural correlates of the verbal component of working memory. Nature. 1993; 362(6418):342–345. [PubMed: 8455719]

Pei X, Leuthardt EC, Gaona CM, Brunner P, Wolpaw JR, Schalk G. Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. Neuroimage. 2011; 54(4):2960–2972. [PubMed: 21029784]

Penfield, W., Roberts, L. Speech and brain mechanisms. Princeton University Press; 1959.

Price C, Wise R, Warburton E, Moore C, Howard D, Patterson K, … Friston K. Hearing and saying. Brain. 1996; 119(3):919–931. [PubMed: 8673502]

Quigg M, Fountain N. Conduction aphasia elicited by stimulation of the left posterior superior temporal gyrus. Journal of Neurology, Neurosurgery, and Psychiatry. 1999; 66(3):393.

Quigg M, Geldmacher DS, Elias WJ. Conduction aphasia as a function of the dominant posterior perisylvian cortex: Report of two cases. Journal of Neurosurgery. 2006; 104(5):845–848. [PubMed: 16703895]

Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nature Neuroscience. 2009; 12:718–724. [PubMed: 19471271]

Ray S, Maunsell JH. Differences in gamma frequencies across visual cortex restrict their possible use in computation. Neuron. 2010; 67(5):885–896. [PubMed: 20826318]

Rogalsky C, Poppa T, Chen KH, Anderson SW, Damasio H, Love T, Hickok G. Speech repetition as a window on the neurobiology of auditory motor integration for speech: A voxel-based lesion symptom mapping study. Neuropsychologia. 2015; 71:18–27. [PubMed: 25777496]

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry MS, … Abel S. Ventral and dorsal pathways for language. Proceedings of the National Academy of Sciences. 2008; 105:18035–18040.

Steinschneider M, Fishman YI, Arezzo JC. Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. Cerebral Cortex. 2008; 18:610–625. [PubMed: 17586604]

Steinschneider M, Nourski KV, Kawasaki H, Oya H, Brugge JF, Howard MA. Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. Cerebral Cortex. 2011; 21:2332–2347. [PubMed: 21368087]

Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. Brain. 2014; 137(10):2773–2782. [PubMed: 24970097]

Ullman MT, Corkin S, Coppola M, Hickok G, Growdon JH, Koroshetz WJ, Pinker S. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. Journal of Cognitive Neuroscience. 1997; 9(2):266–276. [PubMed: 23962016]

Vallarb G, Di Bettac AM, Silveri MC. The phonological short-term store-rehearsal system: patterns of impairment and neural correlates. Neuropsychologia. 1997; 35(6):795–812. [PubMed: 9204486]

Warburton E, Wise RJ, Price CJ, Weiller C, Hadar U, Ramsay S, Frackowiak RS. Noun and verb retrieval by normal subjects. Brain. 1996; 119(1):159–179. [PubMed: 8624678]

Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. Nature Neuroscience. 2004; 7(7):701–702. [PubMed: 15184903]

Highlights

- Verbal repetition requires the coordination of auditory, memory, linguistic, and motor systems.

- Repetition was examined using electrocortical stimulation and electrocorticography

- Critical repetition areas in left peri-Sylvian cortex were identified with ECS

- ECoG revealed functional clusters of activity associated with specific processes

- Results reveal the dynamics of coordinated activity across peri-Sylvian cortex
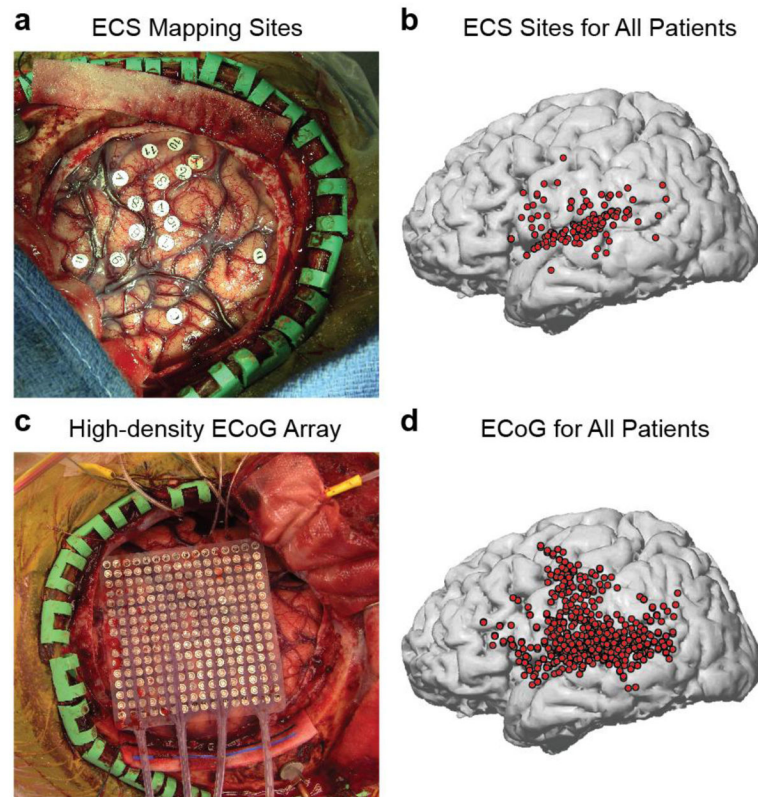
Figure 1. Two complementary methods for examining the cortical networks involved in verbal repetition

(**a**) Example intraoperative photograph showing exposed craniotomy and markers where electrocortical stimulation was performed. (**b**) Reconstructed brain showing all positive stimulation sites across 47 patients, covering the major peri-Sylvian regions hypothesized to be involved in verbal repetition. (**c**) Example intraoperative photograph showing exposed craniotomy and high-density 256-channel ECoG grid covering peri-Sylvian cortex. (**d**) Reconstructed brain showing all electrode locations included in the ECoG analyses for 8 patients.
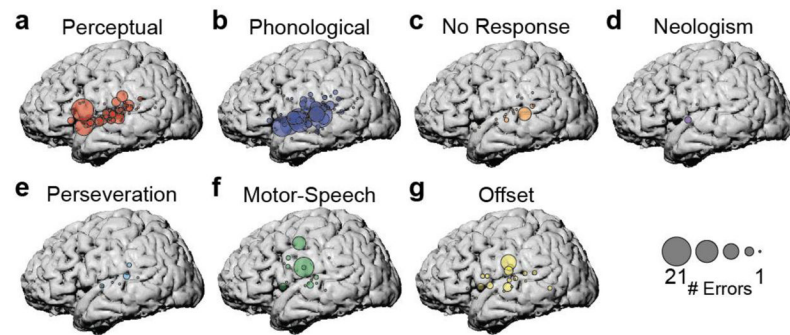
Figure 2. 7 error types elicited by ECS to peri-Sylvian cortex

(**a**) Perceptual errors (patient "did not hear") stimulus. (**b**) Phonological errors (addition, deletion, or change of phonemes/syllables in the target word). (**c**) No response errors. (**d**) Neologism errors (patient says a phonologically plausible pseudoword). (**e**) Perseveration errors (patient repeats previous stimulus). (**f**) Motor speech errors (slurred or distorted speech). (**g**) Offset errors (no response until after ECS).
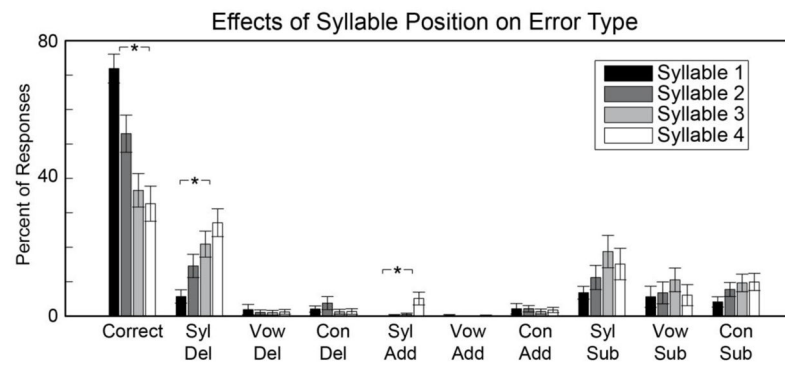
Figure 3. Phonological errors broken down by syllable position

There was a significant decrease in the number of correct responses over the course of each word. This effect was driven by significant increases in syllable deletion and syllable addition across syllables 1–4. There were no significant effects of syllable position on phoneme-related phonological errors ("Vow" = vowels, "Con" = consonants). Asterisks represent statistical significance at $P < 0.05$ (Bonferroni corrected).

Figure 4. Functional clustering of ECoG responses using NMF reveals distributed and overlapping functions underlying verbal repetition

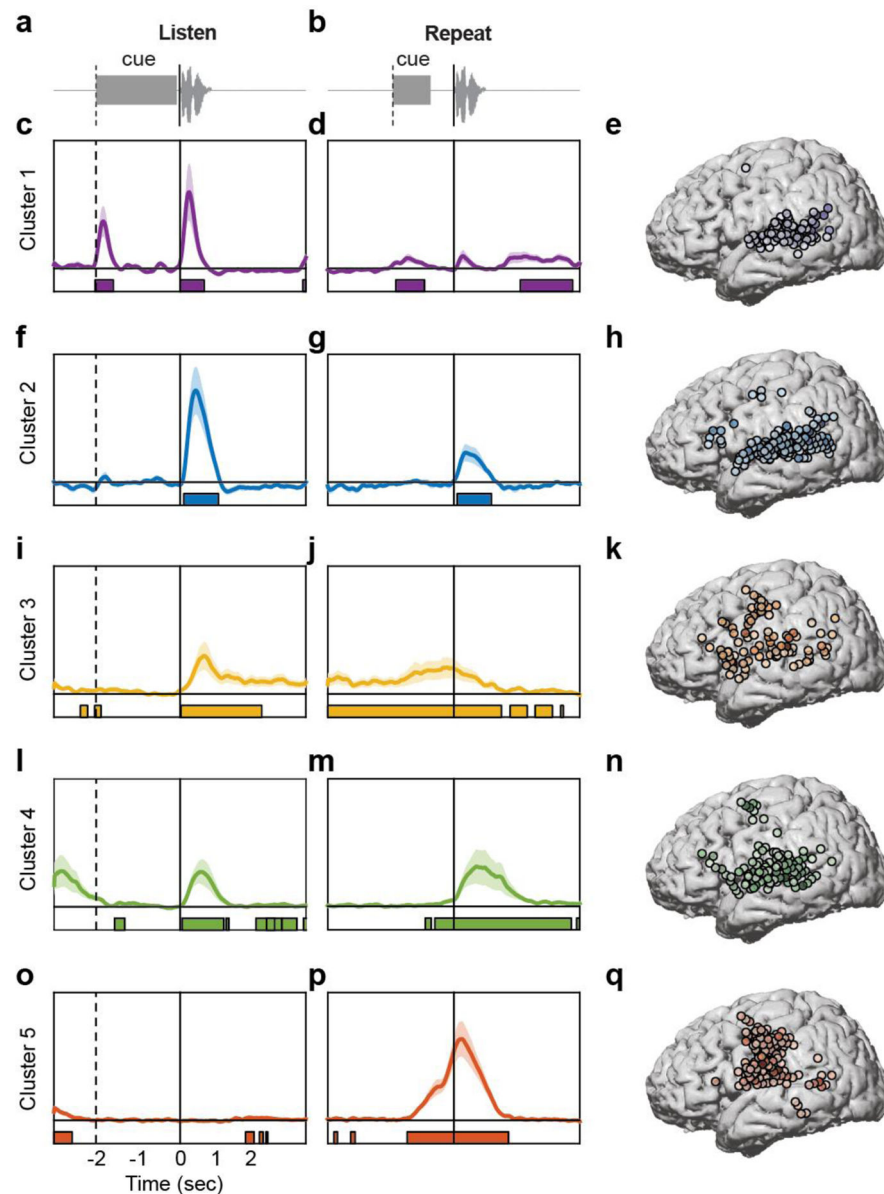(**a**) During the listening phase of the task, participants heard a cue (slide projector sound), followed by a word. (**b**) After a 2 second delay, participants heard a cue (beep) and then repeated the word they had heard in the listening phase. The data from both phases were concatenated and clustered using convex NMF, a 'soft' clustering technique that assigns probabilities for each electrode belonging to one of five clusters. Early auditory responses to all acoustic input during the listening (**c**) and speaking (**d**) phases were localized primarily to posterior STG (**e**). Auditory responses unique to spoken input during listening (**f**) and speaking (**g**) were localized to the entire extent of STG, with additional electrodes in ventral and dorsal frontal areas (**h**). Ongoing activity during the delay period between listening (**i**) and speaking (**j**) was associated with working memory, and was primarily localized to lateral

frontal cortex, with some electrodes in auditory regions (**k**). A separate cluster of speech-specific auditory responses during listening (**l**) and speaking (**m**) showed similar localization as cluster 2 (**n**), except with longer latency peaks. Finally, a cluster showed no responses during listening (**o**), followed by large production-evoked responses around movement onset (**p**), primarily in vSMC electrodes (**q**).
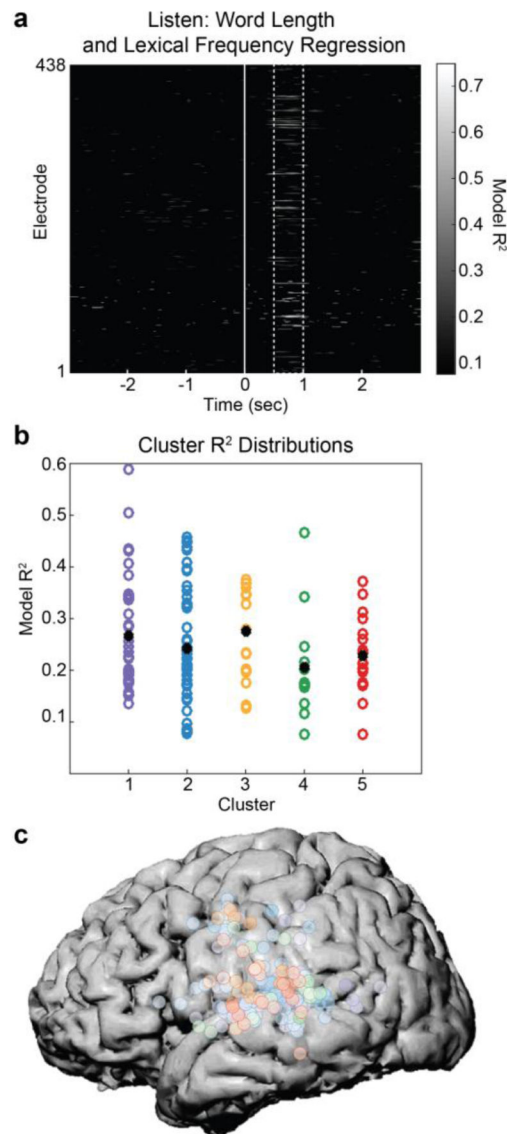
Figure 5. Clusters involved in various stages of verbal repetition are modulated by word length and lexical frequency

(**a**) For each of the 438 electrodes at each time point during the listening phase, high-gamma activity was predicted from a linear combination of word length and lexical frequency using regression. R^2 values for these models ($P < 0.05$) peaked between 500–1000ms after stimulus onset (dashed lines). These two variables accounted for up to ~75% of the variance on some electrodes. (**b**) To determine whether word length and lexical frequency effects differed across the 5 clusters identified in previous analyses, the R^2 distributions for each frequency were plotted. While mean R^2 differed slightly (black dots), there were no significant differences across clusters. (**c**) Electrodes plotted with dominant cluster membership (color) and peak R^2 value (intensity) demonstrate that word length and lexical frequency effects are pervasive throughout peri-Sylvian regions that are involved in verbal repetition.

Table 1a

Patient Demographics

Demographic information for all patients undergoing craniotomy on the dominant hemisphere. White background cells correspond to ECS-only patients. Light gray cells correspond to ECS and ECoG patients. Dark gray cells correspond to ECoG-only patients.

ID	Sex	Age	Handedness	Disease type	Lesion location (left)
16	F	34	R	Epilepsy	anterior temporal lobe, hippocampus, amygdala
17	F	40	R	Epilepsy	posterior temporal cortex
29	M	38	R	Grade I ganglioglioma	anterior fusiform
31	M	41	R	Epilepsy	anterior temporal lobe, hippocampus
35	M	38	R	Grade III anaplastic astrocytoma	posterior superior parietal lobule
38	M	51	L	Grade II oligodendroglioma	posterior superior parietal lobule
39	F	49	R	Grade III anaplastic astrocytoma	anterior temporal lobe, hippocampus
41	F	19	R	Grade III anaplastic astrocytoma	middle superior temporal gyrus
46	M	36	R	Grade II oligoastrocytoma	anterior middle temporal gyrus
48	M	64	R	Gliosis	inferior frontal gyrus
49	M	61	L	Glioblastoma multiforme	anterior middle frontal gyrus
57	M	71	R	Glioblastoma multiforme	middle inferior temporal gyrus
61	F	30	R	Grade III anaplastic astrocytoma	middle inferior temporal gyrus
63	M	25	R	Cavernous malformation	supramarginal gyrus
119	M	30	R	Grade III Anaplastic Astrocytoma	frontotemporal lobe
130	F	30	R	Grade II oligoastrocytoma	frontal
132	F	32	R	Epilepsy	temporal lobe, hippocampus
133	M	28	R	Grade I Low-grade glioneuronal neoplasm	temporal lobe
134	F	40	R	Grade II Oligoastrocytoma	parieto-occipital lobe
136	M	33	R	Grade IV Glioblastoma	temporal lobe
142	M	30	R	Low-grade glioma, favor astrocytoma, Grade II	temporal lobe
147	F	64	R	Epilepsy, Mild hypercellularity	frontal operculum
152	M	37	R	Grade II Oligodendroglioma	superior temporal gyrus
155	F	31	R	Grade III Anaplastic oligoastrocytoma	temporal lobe
158	F	52	R	Epilepsy	temporal lobe, hippocampus

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

ID	Sex	Age	Handedness	Disease type	Lesion location (left)
161	M	61	R	Grade IV glioblastoma	fronto-temporal
163	F	26	R	Arteriovenous vascular malformation	frontal lobe
165	F	22	R	Epilepsy	temporal lobe, hippocampus
167	M	58	R	Grade IV Astrocytoma	temporal lobe
170	F	56	R	Subpial (Chaslin's) gliosis	temporal cortex
172	M	32	R	Grade I Ganglioglioma	temporal lobe
179	F	47	R	Grade IV astrocytoma	temporal lobe
180	M	24	L	Focal cortical dysplasia, ILAE type IIIa, mesial temporal sclerosis	temporal, hippocampus
184	F	62	R	Grade IV astrocytoma	temporal lobe
188	F	33	R	Grade III Anaplastic astrocytoma	temporal, medial temporal
189	M	36	R	Cortical malformation	parietal lobe
208	M	26	R	Grade II oligodendroglioma	frontal lobe
212	M	31	R	Grade IV Glioblastoma	frontal lobe
214	M	58	R	Grade IV Glioblastoma	temporal lobe
217	M	24	R	Grade I Ganglioglioma	temporal lobe
219	M	29	R	Gliosis	parietal lobe
223	M	65	R	Gliosis	amygdala, hippocampus
71	F	39	R	Epilepsy	anterior temporal lobe, hippocampus
76	F	49	R	Grade IV gliosarcoma	middle middle temporal gyrus, hippocampus
97	F	52	R	Cavernous malformation	supplementary motor
112	M	23	R	Cortical dysplasia	anterior temporal lobe, hippocampus
118	M	35	R	Epilepsy	anterior temporal lobe, hippocampus
80	F	29	R	Epilepsy	anterior temporal lobe, hippocampus
101	F	65	R	Cavernous malformation	anterior temporal lobe, hippocampus
103	M	42	R	Grade I ganglioglioma	anterior temporal lobe

Table 1b

Patient Demographics

Demographic information for all patients undergoing craniotomy on the non-dominant hemisphere.

ID	Sex	Age	Handedness	Disease type	Lesion location (left)
ND 1	M	57	R	Grade III anaplastic astrocytoma	right frontoparietal lobe
ND 2	M	44	R	Low grade glial neoplasm	right medial temporal lobe
ND 3	M	30	R	Right temporal pole, hippocampus	right subpial (Chaslin's) and white matter astrogliosis
ND 4	F	47	R	subpial gliosis, grey matter and white matter astrogliosis	right temporal neocortex
ND 5	F	44	R	Subpial and subcortical gliosis	right temporal lobe, hippocampus
ND 6	M	46	R	Epilepsy	right anterior temporal pole, hippocampus
ND 7	M	29	R	Epilepsy	NA
ND 8	F	51	R	Neuronal loss with astrogliosis	Left hippocampus
ND 9	F	20	R	Focal cortical dysplasia, type IIB (ILAE)	right parietal lobe

Table 2

Error codes, criteria, and examples

Seven error types were identified: perceptual, phonological, no response, neologism, perseveration, motor speech, and offset.

Error category	Criteria	Example
Perceptual	Patient appears to have not "heard" the target word	"huh", "what", "I don't know", "can you repeat that?"
Phonological	Addition, deletion, substitution, or transposition of phonemes of the target word	"de" for debate, "opti" for optimism, ""caterpilly" for caterpillar, "fedometer" for speedometer
No response	No verbal response	
Neologism	Patient produces a word that has less than half of the target word's phonemes.	[ɛfəˈbɔɹfn̩] for delicacy
Perseveration	Patient repeats previous stimulus, in whole or in part (could also be a persistent production error)	"superstition" (preceding word) for extinguisher, "commotion" for illustrator while previous word was locomotion, "kæpəˈɹɪʃənɪs;kælvəˈɹɪʃənɪsˈ; ˌæʃəˈɹɪʃənɪs;kæl…ˈɪʃənɪs" for salutation and carryover consecutively
Motor Speech	Phonologically correct, but the phonetic quality is altered (i.e., slurred or distorted)	prosperity [slurred]
Offset	Delay in response until stimulation is terminated	say…delivery [stimulation offset] → "delivery"