Spatiotemporal imaging of cortical activation during verb generation and picture naming

Erik Edwards[a,b,*], **Srikantan S. Nagarajan**[c], **Sarang S. Dalal**[c], **Ryan T. Canolty**[a], **Heidi E. Kirsch**[c,d], **Nicholas M. Barbaro**[e], and **Robert T. Knight**[a]

[a] Department of Psychology and Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA

[b] Computer Science & Engineering, University of Washington, Seattle, WA, USA

[c] Department of Radiology, University of California, San Francisco, CA, USA

[d] Department of Neurology, University of California, San Francisco, CA, USA

[e] Department of Neurological Surgery, University of California, San Francisco, CA, USA

Abstract

One hundred and fifty years of neurolinguistic research has identified the key structures in the human brain that support language. However, neither the classic neuropsychological approaches introduced by Broca (1861) and Wernicke (1874), nor modern neuroimaging employing PET and fMRI has been able to delineate the temporal flow of language processing in the human brain. We recorded the electrocorticogram (ECoG) from indwelling electrodes over left hemisphere language cortices during two common language tasks, verb generation and picture naming. We observed that the very high frequencies of the ECoG (high-gamma, 70–160 Hz) track language processing with spatial and temporal precision. Serial progression of activations is seen at a larger timescale, showing distinct stages of perception, semantic association/selection, and speech production. Within the areas supporting each of these larger processing stages, parallel (or "incremental") processing is observed. In addition to the traditional posterior vs. anterior localization for speech perception vs. production, we provide novel evidence for the role of premotor cortex in speech perception and of Wernicke's and surrounding cortex in speech production. The data are discussed with regards to current leading models of speech perception and production, and a "dual ventral stream" hybrid of leading speech perception models is given.

Keywords

Introduction

Human language can be studied only indirectly in animal models, and therefore linguistic neuroscience depends critically on methods of human neuroimaging. Human intracranial studies, using indwelling electrodes in neurosurgical patients, provide a rare opportunity to

*Corresponding author. University of Washington, 566 Allen Center, Box 352350, Seattle, WA 98195-2350, USA. Fax: +1 206 685 1732. ee4@u.washington.edu (E. Edwards).

achieve both high spatial and temporal resolution. Recent ECoG studies in awake patients have shown that the high-gamma band (γ_{high}, ~60–300 Hz, typically studied from ~70–160 Hz) provides a powerful means of cortical mapping and detection of task-specific activations (Crone et al., 1998, 2001; Edwards et al., 2005; Canolty et al., 2007; Towle et al., 2008; Edwards et al., 2009). Furthermore, γ_{high} has emerged as the strongest electrophysiological correlate of cortical blood-flow (Logothetis et al., 2001; Brovelli et al., 2005; Mukamel et al., 2005; Niessing et al., 2005; Lachaux et al., 2007), often showing even higher correlations with blood-flow measures than multi-unit spiking activity. The γ_{high} band also exhibits excellent signal-to-noise ratio (SNR), with event-related increases clearly seen at the single-trial level or after averaging only a few trials (see Figs. 2d–5d below). The present study uses these advantages of the γ_{high} band to study the topography and temporal sequence of cortical activations during two common language tasks, verb generation and picture naming.

Verb generation (Petersen et al., 1988, 1989) is a semantic association task where a noun is presented and the subject responds with an associated verb. Recent neurosurgical studies emphasize the verb generation task for use in preoperative and intraoperative language mapping (Herholz et al., 1997; Thiel et al., 1998; Ojemann et al., 2002). Picture naming is a common language task employed in neurosurgical language mapping (Penfield and Roberts, 1959; Ojemann et al., 1989; Sinai et al., 2005) and aphasia assessment (Goodglass et al., 2000). This task has similar speech production requirements as verb generation (speaking a single word), but may not activate semantic areas as strongly (Herholz et al., 1997). For picture naming, there is usually only one correct answer, and the task of associating a concretely presented object with a noun is one of the earliest linguistic skills mastered developmentally. In verb generation, the patient must creatively link a noun to a verb with freedom to choose amongst competing response alternatives, requiring more executive-level semantic selection (Thompson-Schill et al., 1998). The available evidence consistently implicates inferior and middle frontal gyri (IFG and MFG), in addition to other widely distributed cortical areas, in the semantic association/selection aspects of the verb generation task (Petersen et al., 1988; Herholz et al., 1997; Thiel et al., 1998; Thompson-Schill et al., 1998; Ojemann et al., 2002). The ECoG method does not allow complete or arbitrary spatial coverage, since the electrode placements are determined for clinical purposes, so we will not be able to study the full network involved in "semantics", and will limit our scope to the areas that do find exposure in our cohort.

In the version of the verb generation task used here, the stimulus is an auditorily presented noun, allowing the study of single-word speech processing. In the Supplementary Information (SI), we also present results for other auditory and speech processing tasks performed in the same patients. Receptive speech has traditionally been associated with Wernicke's (1874) area of the posterior superior temporal gyrus (STG) and surrounding areas. Current models of speech perception include greater Wernicke's area, but also implicate a "dorsal" stream that includes regions also involved in speech motor processing (Hickok and Poeppel, 2007). The involvement of motor areas in speech perception was foreshadowed in a long tradition of motor theories of speech perception (Liberman et al., 1967), but neuroimaging evidence was only recently obtained (Wilson et al., 2004; Meister et al., 2007).

Verb generation and picture naming involve similar speech production requirements, allowing within-subject confirmation of single-word speech production activations with two distinct "lead-in" tasks. One of the outstanding questions for speech production is the possible role of posterior (Wernicke's and surrounding) areas traditionally associated with speech perception. There is now strong evidence for a motor speech area of the posterior superior temporal plane (STP) near to or overlapping with the temporal–parietal junction (TPJ), a finding foreshadowed in earlier studies of conduction aphasia (Wise et al., 2001; Bates et al., 2003; Buchsbaum et al., 2005; Hickok and Poeppel, 2007). Another outstanding question in speech production is

which cortical regions are involved in preparing the phonological representation used to guide articulation. A leading model (Levelt, 1999; Indefrey and Levelt, 2004) implicates the posterior STG (pSTG), but we present here the first evidence combining high spatial and temporal resolution and arrive at a different conclusion.

Methods

Patients

A total of 4 patients (3 ♀, 1 ♂) with intractable epilepsy were studied several days after surgery for implantation of subdural electrodes. All patients were right-handed, native English speakers, and all were left-hemisphere dominant as assessed by a pre-operative intracarotid amobarbital test using picture naming (full details in Loring et al., 1997). The age range was 15–45 y.o. (Patient 1: 39 y.o. ♀; Patient 2: 37 y.o. ♀; Patient 3: 15 y.o. ♂; Patient 4: 45 y.o. ♀). The seizures had begun 5–31 years prior to surgery (seizure onsets at age 13, 32, 8, and 12 y.o., respectively). The purpose of the electrode implant was to monitor/localize seizure activity, so all patients had been tapered off of antiseizure medication in order to record ictal events (>24 h before the reported tasks). The electrodes were also used for electrical stimulation mapping (ESM) of language and motor sites. We present no formal comparison of our ECoG results to the clinical ESM results (e.g., Sinai et al., 2005, 2009; Towle et al., 2008), but we note that the ESM-identified mouth motor sites (ventral peri-Rolandic cortices) were also identified with excellent correspondence by ECoG activations during speech production. Patients consented to participate in the present study, which posed no additional risk, and were cooperative in performing the verb generation task and a variety of other tasks; task order was random across patients. All testing was at bedside in the post-operative monitoring rooms.

The focus of the main text is on two subjects (Patients 1 and 2) of normal or better intelligence and language ability. Both patients underwent comprehensive neuropsychological evaluation (including WAIS-III, WRAT-3, D-KEFS, WCST, etc.) by the clinical team, and performed average or above-average on all measures. In the research study, both subjects performed well on a variety of auditory, motor and language tasks (e.g., phoneme identification, single-word processing, mouth-motor mapping, picture naming, and verb generation). In both patients the seizure focus was in the medial temporal lobe and only restricted portions of the temporal lobe exhibited interictal epileptiform activity. In Patient 1, a group of 14 anterior temporal electrodes showed interictal epileptiform activity and was excluded from the present analyses. In Patient 2, a group of 4 inferior temporal electrodes showed interictal epileptiform activity and was excluded (gray–white circles in Fig. 1). It should be emphasized that these two patients are ideal patients for language studies amongst the larger cohort of epilepsy patients that generally have more widespread epilepsy and/or compromised cognitive abilities. They are also above average amongst those epilepsy patients considered as candidates for language mapping and resection, who generally have more focal epilepsy and preserved language function than the total epilepsy population.

In the SI, two additional patients (Patients 3 and 4) are shown to reinforce the main findings from the speech perception and production intervals only. Patient 3 was of approximately normal cognitive ability, and Patient 4 was below normal although still more than able to understand and speak single words. Both performed well for picture naming and adequately for verb generation. Both exhibited more widespread epileptic activity across the grid as compared to the restricted foci of interictal activity exhibited by Patients 1 and 2. It is for these reasons (and to save space in the main text) that these results are relegated to the SI, but we believe the results to be fully valid with respect to the two main findings: location of speech perception activations on the STG, and the strictly post-response activation from the STG during speech production.

Verb generation task

The verb generation task (Petersen et al., 1988, 1989) is a semantic association task where a noun is presented and the subject must respond with an associated verb. In the variation used here, the noun was an auditory stimulus (median duration 470 ms) synthesized using Text Aloud software based on AT&T Natural Voices (http://www.nextup.com/TextAloud/). Although no perceptual studies were done, the female voice was clear and natural sounding (i.e., no choppiness or roughness characteristic of some early speech synthesizers). Stimuli were presented at roughly 60 dB SPL from 2 loudspeakers placed ~1 m in front of the patient on a tray table overlying the patient bed. The patient overtly responded by speaking into a microphone (median response duration, Patient 1: 310 ms; Patient 2: 380 ms; Patient 3: 802 ms; Patient 4: 215 ms). 50 distinct nouns are presented twice each (100 trials total) with 4.1–4.2 s between stimulus onsets. The verbal reaction time (RT) was measured from stimulus onset to response onset. A handful of trials were rejected for a late or missing response (Patient 1: 4 trials; Patient 2: 2 trials).

Picture naming task

The task used here was based on the Boston Naming Test (Kaplan et al., 1983). 60 line drawings (black lines, white background) were presented twice each on an LCD monitor ~1 m in front of the patient. The pictures were displayed for ~4 s, and then were removed briefly (white background for ~60 ms) before the presentation of the next picture (SOA range 3.5–4.5 s, mean 4.0 s). A photodiode on the corner of the monitor was recorded as an analog channel for exact stimulus timing with the ECoG. The patient's task was simply to name the picture, with responses spoken into a microphone. The microphone was also recorded as an analog channel with the ECoG for exact response timing. Most of the pictures were common, easily-recognized objects with essentially one correct answer ('scissors', 'umbrella'), but a few were more difficult or allowed multiple response possibilities ('latch'). Both patients performed this task with no difficulty and neither patient made any errors or late responses (>3.5 s that would have contaminated the baseline of the following trial). Experimental and display (screensaver) errors lost 5 trials in Patient 1 and 4 trials in Patient 2, and the first trial was rejected in all cases. Artifact rejection and other data analyses were identical to those for the verb generation task.

ECoG recordings

Electrodes are embedded in clear silastic in an 8-by-8 array with 1 cm center-to-center spacing. The electrodes are platinum–iridium disks with 2.3 mm contact diameter (impedance ~1–10 kΩ). Additional strips of electrodes (also 1 cm spacing) extended over orbito-frontal, parietal or other cortices. A high-resolution T1 MRI scan was obtained preoperatively for each patient, and a 3D MRI reconstruction was made using BET2 (www.fmrib.ox.ac.uk/analysis/research/bet) and MRIcro (www.sph.sc.edu/comd/rorden/mricro.html). Electrodes were localized on the MRI using intraoperative photographs and post-implant X-rays (Fig. 1; Dalal et al., 2008). Localizations were additionally constrained by the requirement to be on the cortical surface and by the known 1 cm inter-electrode spacing. MNI coordinates were obtained for all electrodes using SPM2. Although there is some error associated with this procedure (roughly ± 0.5 cm), incorrect localization relative to major gyral and sulcal landmarks is unlikely given the clear intraoperative photographs.

ECoG was recorded continuously throughout the task at a sampling rate of 2003 Hz, with an amplifier filter bandpass of 0.01–250 Hz or 0.01–1000 Hz. The original recordings used one electrode from the corner of the grid as the reference, but all data were subsequently re-referenced to the common average (CAR). The CAR consists of the mean across all non-epileptic electrodes, excluding sample points rejected for artifacts. The CAR approximates the activity contributed by the original reference, so subtracting the CAR from each channel largely

removes contributions of the original reference electrode (see Crone et al., 2001 for a discussion of the CAR in ECoG recordings). The original reference was not used and is omitted in the figures.

Data analyses

Time–frequency (TF) analyses used a Gaussian filter bank and the Hilbert transform on each channel separately (Edwards et al., 2009). There is one Gaussian filter (Gaussian-shaped window in the frequency domain) in the filter bank for each of 42 center frequencies (cfs) from 4–250 Hz. For each cf, the result is a time-series of analytic amplitude (AA) of the same sampling rate (2003 Hz), duration (~7 min), and units (μV) as the original ECoG recording. The AA is a formal means of obtaining the envelope of a signal. See Supplementary Information (SI) for details and an illustration of the method. The envelopes of the frequency bands from 70–160 Hz are averaged together to form the γ_{high} AA used in all results shown. Although the full γ_{high} band ranges from ~60–300 Hz, only the range from 70–160 Hz was used in order to avoid electrical artifacts at 60 and 180 Hz, and because this is the frequency range of maximal response. All single-trial data (Figs. 2–5d) are shown in units of *z*-score (calculated by subtracting the baseline mean and dividing by the baseline standard deviation). All other results (Figs. 2–5c, e) are shown in units of percentage (%) changes from the baseline mean. The baseline for all analyses was −250 to 0 ms pre-stimulus.

Event-related averages of γ_{high} AA are taken relative to stimulus and response onsets. In order to show the semantic association interval – which is not strictly time-locked to external events – on one time-line with the sensory-locked and motor-locked responses (Figs. 2c, 3c), the single-trials are realigned to the median RT prior to averaging. The realignment procedure (see Supplementary Methods) only affects the epoch from 500 ms post-stimulus to 250 ms pre-response, which is essentially expanded/contracted to match the median RT. The data before this interval are averaged locked to the stimulus, and the data after are averaged locked to the response.

To assess if a change in AA is significantly different from baseline, we use the resampling method of our prior studies (Edwards et al., 2005, 2009). Briefly, a null distribution of AA is made by resampling single-trial baseline values and averaging 5000 times. The *P* value for a given AA is its percentile location in the null distribution. Raw *P* values are corrected for multiple comparisons using the false-discovery rate (FDR) approach (Benjamini and Hochberg, 1995; Nichols and Hayasaka, 2003). Note that in Figs. 2–5, the same data and *P* values are used in parts c and e, only presented in different form (time course in c, spatial topography in e).

Results

Auditory, motor, and language tasks were performed over the course of 2–3 days in each patient, including the verb generation and picture naming tasks. The SI gives complementary results from additional auditory tasks. TF analysis of the ECoG used a Gaussian filter-bank and the Hilbert transform, yielding a time-series of analytic amplitude (AA, envelope of the filtered signal) for each frequency band. We focus on γ_{high} (averaged over the range 70–160 Hz), since this band shows the most robust tracking of language activations. Significant γ_{high} activations were in the range of ~10–100% greater than baseline levels (Figs. 2–5c), with a sufficient SNR to study single-trial averages with only a 3-trial moving-window average (Figs. 2–5d). Full TF results from 4–250 Hz are presented for comparison in the SI (Fig. S4).

Verb Generation Task

The median RT was 1.25 s for Patient 1 (range: 0.87–3.45 s) and 1.48 s for Patient 2 (range: 0.85–3.1 s). For Patients 3 and 4 (SI), the median RTs were 2.05 s and 1.67 s, respectively. The γ_{high} results show three broad stages of cortical processing for the verb generation task: speech comprehension of the auditorily-presented noun, semantic processing for selecting an associated verb, and a speech production interval surrounding the verbal response. These broad stages, each of ~0.5–1 s duration, follow each other with only partial overlap (serial stages). Within each of these stages, highly overlapping activation time courses are seen (parallel or incremental processing). The results are organized by the three broad intervals of processing.

Speech perception

Several activation foci are seen during the speech perception interval (~50–600 ms following stimulus onset), including Wernicke's area (mid-to-posterior STG and MTG, see Fig. 1 for abbreviations), the precentral gyrus (sPMv), and regions of the central and parietal operculum just dorsal to the Sylvian fissure. All of the activation results are confirmed in detail during a single-word auditory task (Figs. S5, S6).

In Patient 1 (Fig. 2), the activation focus shifts from the pSTG site (dark blue electrode, MNI: −65, −52, 18 mm) at ~100 ms to the pMTG site (turquoise electrode, MNI: −65, −53, 2 mm) at ~300 ms. The sPMv site (orange electrode, MNI: −55, −26, 49 mm) also exhibited significant auditory activations (~75% above baseline, $P<0.02$) peaking at ~310 ms. One electrode on the central operculum (red electrode, MNI: −65, −19, 9 mm) showed a similar activation profile as sPMv, except with a greater weighting for the response (post-stimulus activation ~40%, post-response activation ~120%). Additional auditory-responsive electrodes without any motor responses were found on the most anterior–inferior part of the SMG (parietal operculum, MNI: −64, −39, 17 mm) and on the mid-STG directly lateral from Heschl's gyrus (MNI: −67, −26, 1 mm). It is noted (see Discussion on dual ventral streams) that the mid-STG electrode is located over 2 cm anterior to the pSTG focus, and exhibits no significant activations during the semantic or pre-response intervals of picture naming or verb generation.

In Patient 2 (Fig. 3), the early (~100 ms) auditory activations are focused on the mid-STG, with the peak at the mid-STG electrode directly lateral from Heschl's gyrus (dark blue electrode, MNI: −68, −12, −3 mm). The posterior Sylvian morphology was considerably different in this patient compared to Patient 1, terminating at a relatively anterior level with both ascending and descending terminal rami (Ono et al., 1990, p. 148). Two auditory responsive electrodes were found around these posterior Sylvian branches, but neither seems to be the typical pSTG site (as seen in Patients 1, 3 and probably 4). The first was on the most anterior–inferior SMG or parietal operculum (turquoise electrode, MNI: −67, −30, 17 mm) and it exhibited a purely auditory response profile with peak (55%, $P<0.02$) at ~300 ms. The second was just posterior to the descending terminal ramus (MNI: −70, −44, 11 mm), but its responses were relatively weak (peak ~20%, ~300 ms) and so it is not studied further or shown separately in Fig. 3. It is possible this electrode was near the typical pSTG site located within the descending ramus or another nearby sulcus for this patient. Just superior to this electrode on the TPJ was the important electrode labeled 'Spt/TPJ' (red electrode in Fig. 3, MNI: −69, −41, 22 mm). Although its response was stronger during the motor interval, it also gave significant (30%, $P<0.02$) auditory activations peaking at ~300 ms. A similar activation profile was obtained from the sPMv site (orange electrode, MNI: −54, −6, 52 mm), and from another precentral electrode (dark red, MNI: −62, −1, 29) that gave a weak (~15%) but significant ($P<0.02$) auditory response at ~400 ms. In all electrodes, the peak latency is only a rough indication since a sustained plateau of 100–500 ms duration was the general rule.

Semantic association

Activations during the semantic processing interval (~600 ms post-stimulus to ~200 ms pre-response) are distributed over IFG, MFG, and SMG sites, generally considered higher, association areas of neocortex. The IFG and MFG activations are seen most clearly in Patient 1 (Fig. 2) due to more anterior placement of the grid. In Patient 2 (Fig. 3), the more posterior placement allows imaging of posterior SMG sites (yellow electrodes), which exhibit an activity profile similar to the prefrontal pattern. Typical activations in association cortices were 10–40% above baseline during the semantic interval, with the highest peak level at 65% (Patient 2, dark green electrode in MTG). Although these were significant ($P<0.02$), they are generally weaker than sensory and motor activations. In Patient 1, the pMTG site shows significant activation (10–25%, $P<0.02$) during the semantic association interval, in contrast to the STG sites that return completely to baseline or below baseline levels (−10% to 10%) following stimulus processing (~800–1200 ms).

Speech production

Speech production activations begin ~300 ms prior to verbal response in the peri-Rolandic cortices (pre- and post-central gyri) of both patients. In stark contrast to the peri-Rolandic time-courses, STG sites were active in the speech production interval *following* response onset (peak at ~150–500 ms post-response). This is consistent with self-stimulatory auditory activations from the patient's own voice. In both patients, STG and MTG sites were generally less active in the post-response interval than in the post-stimulus interval, consistent with notions that auditory feedback is suppressed during speech production (Houde et al., 2002; Towle et al., 2008). However, in Patient 1 the more anterior pSTG site (blue electrode, MNI: −66, −46, 10 mm) showed equal stimulus-related and response-related activation levels (~100%, $P<0.02$), consistent with an active role in feedback control of speech production (Indefrey and Levelt, 2004; Tourville et al., 2008). This same pSTG site gave the peak auditory activation during a phoneme task (Fig. S5c–d). In Patient 2, both mid-STG sites were less active during production than perception. However, it was the more anterior STG site (dark blue) and the parietal operculum site (turquoise) that showed the strongest activations during the post-response interval, and these were also the sites giving the peak auditory activation during the phoneme task (Fig. S6c–d). This is consistent with participation in an external loop of self-monitoring during speech production, with a special focus at the phonological level of representation (Indefrey and Levelt, 2004).

In Patient 2, a surprising activation was found at a posterior peri-Sylvian site that showed clear *motor*-related activity (red electrode, MNI: −69, −41, 22 mm). This site is located just posterior to the branch point of the Sylvian fissure on the TPJ (Fig. 1b, labeled 'Stp/TPJ'), but its response profile was very similar to premotor sites (Fig. 3c). Thus, this TPJ electrode represents a restricted focus of response-related activation surrounded by completely different activation profiles in the STG and SMG, emphasizing the spatial resolution of γ_{high} effects. This TPJ activation is not understood in terms of classic models of anterior vs. posterior involvement in speech production vs. perception. However, it confirms more recent findings and models that implicate this posterior TPJ area in speech production (called 'Tpt' or 'Spt') (Wise et al., 2001;Hickok and Poeppel, 2007). Our findings suggest it extends onto the free surface of the TPJ in some individuals, particularly those with a short posterior extent of the Sylvian fissure as seen in Patient 2.

Picture Naming Task

Both patients performed this task without difficulty, and neither made any errors or late responses (>3.5 s that would have contaminated the baseline of the following trial). The γ_{high} results are organized according to the stages of visual perception, semantic processing, and speech production.

Visual perception

Visual areas were generally not covered and few visual responses were obtained. In Patient 1 (Fig. 4), only the most inferior–posterior electrode (MNI: −62, −59, 11 mm) showed an early visual processing response at ~100 ms (~40% increase, $P<0.02$). In Patient 2 (Fig. 5), only electrodes from an inferior temporal strip contacting the fusiform gyrus (not shown) exhibited visual responses.

Semantic association

Significant activations with time courses consistent with a role in semantic association (or, in any case, inconsistent with a role in either visual or motor processing) were observed in several of the prefrontal sites implicated in the verb generation task. Prefrontal activations were weaker and less extensive in picture naming relative to verb generation. For example, the MFG electrode of Patient 1 (Fig. 4, dark green electrode) peaked at 25% above baseline, compared to 65% during verb generation. In Patient 2 (Fig. 5), the SMG sites (e.g., the yellow electrode, MNI: −61, −43, 47 mm) again showed an activation time-course similar to prefrontal sites, peaking at ~400–500 ms post-stimulus (~20–40%, $P<0.02$). These were only slightly weaker than during verb generation, but of shorter duration.

In Patient 1 (Fig. 4), the pMTG site (turquoise electrode, MNI: −65, −53, 2 mm) exhibited significant elevation of activity (20–40%, $P<0.02$) during the semantic association interval (~350 ms post-stimulus to ~200 ms pre-response) of the picture naming task. In fact, this was slightly stronger and of longer duration than its activation during the verb generation semantic interval. The pSTG sites (darker blue electrodes), by contrast, show no activation or weak *suppressions* (−10 to 0%) during this interval, supporting the assertion that the pSTG plays a role primarily in acoustic–phonetic processing, whereas the pMTG plays an additional linguistic role at the word-form or semantic level (Binder et al., 2000;Bates et al., 2003;Dronkers et al., 2004;Indefrey and Cutler, 2004;Hickok and Poeppel, 2007).

Speech production

Precentral sites, including sPMv, exhibit robust (~100%, $P<0.02$) γ_{high} increases surrounding the speech production interval in both patients. The pattern is very similar to the speech production interval of the verb generation task. As in the verb generation task, STG sites exhibit robust (40–100%, $P<0.02$) γ_{high} increases beginning *after* speech onset, implicating them only in auditory feedback during speech production.

In Patient 1 (Fig. 4), the precentral electrodes attain significance ($P<0.02$) as early as ~550 ms prior to response onset and rise steadily until just after response onset (peak levels at 50–200 ms). The postcentral electrode (dark red, MNI: −63, −26, 33 mm) shows a similar pattern but shifted somewhat later in latency. In the post-response interval, it was again the anterior of the two pSTG sites (blue electrode, MNI: −66, −46, 10 mm) that showed the strongest auditory feedback response (>100%, $P<0.02$). This is the same site showing the peak activation to phoneme stimuli (Fig. S5). These sites showing post-response activations may be involved in an external loop of monitoring during speech production (Levelt, 1999;Indefrey and Levelt, 2004).

In Patient 2 (Fig. 5), activations in ventral peri-Rolandic cortex began ~400 ms prior to verbal response onset, peaking ~50–200 ms after response onset. The same pattern is observed at the Spt/TPJ site (red electrode, MNI: −69, −41, 22 mm) implicated above in speech production. On the other hand, activations in STG and peri-Sylvian auditory-responsive sites begin strictly following response onset (peak ~300–500 ms post-response).

Overall, the results from picture naming confirm in every case the functional roles demonstrated in the verb generation task for speech production, including auditory feedback. The semantic activations exhibited a high degree of overlap, but with less intense and less widespread activation for naming than verb generation.

Discussion

We have used the unique advantages of ECoG γ_{high} mapping (Crone et al., 2006) to study cortical language processing during verb generation and picture naming. Taken together, the results show consistent agreement of γ_{high} activation topographies with results of functional imaging studies, adding further evidence to the strong empirical relation of γ_{high} levels to measures of blood-flow (Darrow and Graf, 1945; Logothetis et al., 2001; Brovelli et al., 2005; Mukamel et al., 2005; Niessing et al., 2005; Lachaux et al., 2007). The activation time-courses revealed by γ_{high} mapping are consistent with known or postulated roles of different cortical areas in language processing.

An important feature of the present study is that a variety of tasks were studied *within* each single subject. Thus, verb generation and single-word processing share common speech perception demands and consistent spatio-temporal activation patterns were observed in the two tasks. Likewise, verb generation and picture naming involve similar speech production demands and show consistent activation patterns during the pre- and post-response intervals. Perhaps most importantly, this study allows comparison of speech perception and production at the identical sites of a single subject. This is in contrast to neuroimaging studies where activations for a single task are averaged over individuals, with the resultant effect of blurring over individual variability in the location of relevant functional regions. Comparisons between speech perception and production sites would typically involve comparison of mean localizations in two data sets obtained with different subjects. We observe distinct activation profiles from sites separated by only 1 cm, and there is large intersubject variability in macroanatomy (sulci, gyri) and in functional localization relative to both macroanatomy and standard coordinate systems (Roland and Zilles, 1998; Rademacher, 2002). Thus, there are distinct advantages to the present approach of studying across tasks *within* a single subject. The major disadvantage of the present approach is the lack of spatial *coverage*, most notably of intrasulcal regions. The other major limitation is the small number of patients studied. Although our main findings are supported by N=4, some other findings here are based only on N=2 or even N=1. These are small numbers compared to most neuroimaging studies in normal subjects.

The discussion is organized according to the three broad stages of activation during the verb generation task: speech perception, semantic association/selection, and speech production. Picture naming is discussed with respect to the latter two.

Speech perception

During speech comprehension (~50–650 ms following stimulus onset), activations are focused in and around Wernicke's area. A large body of evidence supports a role for the pSTG in acoustic–phonetic aspects of speech processing (Boatman et al., 1995; Binder et al., 2000; Crone et al., 2001; Hickok and Poeppel, 2007), whereas more ventral sites such as the pMTG are thought to play a higher linguistic role linking the auditory word form to broadly distributed semantic knowledge (Bates et al., 2003; Dronkers et al., 2004; Hickok and Poeppel, 2007). Accordingly, the activation focus shifts from the pSTG at ~100 ms to the pMTG site at ~300 ms in Patient 1. Also, the pMTG site shows significant activation during the semantic association interval of the verb generation and picture naming tasks, in contrast to the pSTG sites that remain at or below baseline levels during this interval. This is consistent with a greater lexical–semantic role for pMTG relative to a more acoustic–phonetic role for pSTG. The

activation time courses in these areas are highly overlapping, supporting models of speech perception where results of acoustic and phonological processing are made immediately available to higher processing stages as the information accumulates (McClelland and Elman, 1986; Warren and Marslen-Wilson, 1987; Pulvermüller et al., 2006). Traditional serial models, whereby each discrete stage completes its processing before passing its final output to the next stage, are not supported.

Beyond Wernicke's area, speech processing activations are observed in the precentral gyrus of both patients. Although not part of the classic network for speech comprehension, this area is implicated in recent fMRI (Wilson et al., 2004) and other studies (Meister et al., 2007; Brown et al., 2008) of speech perception, and is included in contemporary models (Hickok and Poeppel, 2007). We observe that the premotor area is more responsive to speech sounds compared to simple tone stimuli (Figs. S5, S6), consistent with other evidence that this region is preferentially responsive to speech (Wilson et al., 2004; Hickok and Poeppel, 2007; Meister et al., 2007). We note that speech perception activations are observed in these dorsal stream areas even without sub-segmental processing, suggesting a partial reconsideration of this point in a recent model (Hickok and Poeppel, 2007).

In the ventral stream, we find intersubject variability in the anterior–posterior location of speech activations. Patient 2 shows the main focus at the mid-STG, Patient 3 shows the main focus at the pSTG, and Patients 1 and 4 show both foci (bimodal). The STG can be quite thin (sometimes < 1 cm wide in the dorsal–ventral dimension), so our spacing of 1 cm may have missed one of the foci in some patients. In Patient 2, a weakly responding electrode was found just posterior to a descending ramus of the Sylvian fissure, so the usual pSTG site may have been present within this or another sulcus. Thus, all 4 patients may have had the pSTG focus, but it was not optimally sampled in some patients due to the 1 cm spacing and lack of intrasulcal coverage. The total topographical distribution we observe is bimodal.

Prior evidence for a bimodal representation on the STG also exists, for example in the electrical stimulation results of Miglioretti and Boatman (2003) and in the meta-analysis of Indefrey and Cutler (2004). This may explain some apparent inconsistencies in the literature, and we suggest that the ventral stream traversing the STG can be conceived of as two separate, but overlapping, pathways (overlapping topographically and in terms of function). The more anterior ventral stream corresponds to the "pathway of intelligible speech" of Scott et al. (2000), proceeding laterally from Heschl's gyrus and terminating in the anterior STS. The more posterior ventral stream corresponds to the pathway emphasized in the depiction of Hickok and Poeppel (2007), proceeding from the planum temporale and terminating in the mid-to-posterior STS and MTG. In some subjects, the dual ventral streams are essentially merged at the mid-to-posterior STG, and in others a clear separation is seen. The anterior vs. posterior distinction for temporal lobe speech processing sites follows the findings and models of Scott, Wise and colleagues (Crinion et al., 2003; Scott and Wise, 2004; Jacquemot and Scott, 2006; Spitsyna et al., 2006), where the posterior portion plays a larger role in phonological working memory (also consistent with our data, see below). Thus, the model that is most consistent with our data to date (including a larger cohort of patients not given in this study) is a hybrid of these two leading models. We represent this hybrid by dividing the ventral stream of Hickok and Poeppel (2007) into anterior and posterior components (Fig. 6). Following simultaneously the suggestions of these two models, the anterior ventral stream is more connected with Broca's (1861) area and more involved in sentence-level processing, whereas the posterior ventral stream is more connected to the Spt/TPJ region and more involved in single-word speech production and phonological working memory.

Semantic association

Verb generation and picture naming both require semantic association, but the verb generation task is more difficult and requires more controlled (executive-level) processing in this regard (Thompson-Schill et al., 1998). The semantic activations exhibited a high degree of overlap, but with less intense and less widespread activation for naming than verb generation, consistent with prior studies comparing the two tasks (Herholz et al., 1997; Ojemann et al., 2002). Consistent with evidence for topographical differences in noun vs. verb representation (Damasio and Tranel, 1993), prefrontal sites showed relatively greater activation during verb generation and the pMTG site showed somewhat greater activation in the semantic interval for picture naming. Semantic representation and processing involve widely distributed cortical areas, only some of which find exposure in the current study. This is the major limitation of the ECoG mapping technique – the lack of complete spatial coverage of all relevant cortical areas.

Our results, in combination with prior studies and models, lead to the following model of picture naming. Early visual processing and object recognition take place in occipital and inferior temporal areas ('what' pathway). This process takes ~150 ms, following Levelt et al (1998). The representation of the object in the 'what' pathway and nearby inferior temporal areas itself constitutes a major aspect of the conceptual–semantic representation. Additional semantic and syntactic ("lemma") associations are also activated after ~150 ms (Levelt et al., 1998), and during this interval of highly variable duration (depending on the subject, the difficulty of the current object, etc.), the word to be spoken is selected. This involves some of the same sites – PFC, SMG, and other association cortices –involved in the semantic selection stage of verb generation, but not as strongly given the more constrained set of word choice alternatives. The speech production stage is discussed below.

Speech production

Speech production activations begin ~300 ms prior to verbal response in the peri-Rolandic cortices (pre- and post-central gyri) of both patients, usually peaking at 100–200 ms after response onset. The role of ventral peri-Rolandic cortices in speech motor functions has long been appreciated (Penfield and Roberts, 1959) and is not discussed further. Some of these same sites (e.g., sPMv) also exhibited auditory responses preferential to speech stimuli and are part of the dorsal stream.

Although involvement of Wernicke's area in speech production has been suggested by several authors, this is the first data to document the participation of traditional Wernicke's area (mid-to-posterior STG) only in *post-response* auditory feedback, while demonstrating a clear *pre-response* activation from the nearby Spt/TPJ. One major model of speech production implicates pSTG in pre-response encoding of the phonological word for speech production (Indefrey and Levelt, 2004), but the present data suggest a revision to this aspect of the model. We suggest instead that the common route to speech production is through verbal/phonological working memory (WM) using the same dorsal stream areas (Spt/TPJ, sPMv) implicated in speech perception and phonological WM (Jacquemot and Scott, 2006; Hickok and Poeppel, 2007; Buchsbaum and D'Esposito, 2008). The observed pre-response activations at these dorsal stream sites are suggested to subserve phonological encoding and its translation to the articulatory score for speech. Post-response Wernicke's activations, on the other hand, are involved strictly in auditory self-monitoring.

Several authors (Dell et al., 1997; Levelt et al., 1998; Jacquemot and Scott, 2006) support a model in which the route to speech production runs essentially in reverse of speech perception, i.e., from conceptual level to word form to phonological representation. This framework is generally consistent with our data, with the pMTG activation representing part of the spread

from conceptual to word form representations. The activations may then spread back to the pSTS representing the phonological level (abstract representation of syllable sequence) (Hickok and Poeppel, 2007). However, there is no further reverse spread to the pSTG and we observe response-related activity only from motor and dorsal stream areas (Spt/TPJ, PMv, Broca's area). These same dorsal stream areas are also implicated in verbal/phonological WM (Shallice and Warrington, 1977; Paulesu et al., 1993; Vallar and Papagano, 2002), and we follow the suggestion (Jacquemot and Scott, 2006; Buchsbaum and D'Esposito, 2008) that the common route to single-word speech production is through phonological WM areas. Consistent with our observation in 4/4 patients of absolutely no pre-response activity in the STG, we dispense of any postulated role for the STG in formation of the "phonological word". Rather, the phonological level of representation used in speech production is none other than phonological WM, supported primarily by dorsal stream areas. Ventral stream areas (STG, STS, MTG) also include a phonological level of representation for speech perception and immediate (echoic) WM (Hickok and Poeppel, 2007). The reverse spread of activation in these areas during speech production likely reaches the phonological level (~STS), but no further (STG). The reverse-activated ventral stream areas provide input to the formation of the "phonological word" in dorsal stream areas for phonological WM. Dorsal stream and motor areas translate the phonological code into the articulatory score that guides overt speech.

How can we account for the results from two studies (Levelt et al., 1998; Indefrey and Levelt, 2004) implicating the Wernicke's area (pSTG) in pre-response phonological encoding? The most likely answer to this question is that both of these studies may have lacked sufficient spatial resolution. The pSTG is flanked on both sides by sites that are activated during the pre-response interval (Spt/TPJ above and STS/MTG below). Levelt et al. (1998) implicated the pSTG in pre-response activity based on dipole fitting of evoked MEG field levels. This method may lack the spatial resolution to avoid mixing from nearby Spt/TPJ and STS/MTG sources, especially given that ERPs exhibit a different and more widespread topography than γ_{high} (Crone et al., 2001; Edwards et al., 2005, 2009). Also, most of their dipoles were actually localized to the region of the Spt/TPJ, not pSTG (see their Fig. 5). The meta-analysis of Indefrey and Levelt, (2004) considered ROIs on a Talairach-based parcellation of cortex that did not distinguish intrasulcal regions. Activations "near the border" of two ROIs were counted towards both ROIs, so both STS/MTG and Spt/TPJ activations would likely have been included in the pSTG ROI. The effects of smoothing and averaging discussed above also apply to this meta-analysis. Also, given the low temporal resolution of PET and fMRI, *post*-response pSTG activations may not have been excluded.

One qualification must be placed on the above critique – the superior bank of the STS can be included in the pSTG. For example, the fMRI study of Buchsbaum et al. (2001) found two "pSTG" sites involved in speech production. One was centered on posterior intra-Sylvian cortex and parietal operculum (i.e. area Spt/TPJ), and the other was the posterior STS. Taking this larger definition of pSTG that includes several dorsal and ventral stream areas, a role for pSTG in phonological representation for speech production can be claimed. However, Indefrey and Levelt, (2004) imply that this representation is centered on the free surface of the pSTG, and no distinctions were made for intra-Sylvian or STS sources. The present data provide greater anatomical specificity to the claim of Wer-nicke's and "pSTG" involvement in speech production – the early ~half of the ventral stream, including the free surface of the pSTG in most subjects, is only involved in post-response auditory feedback. The second ~half of the ventral stream (MTG, STS) is reverse activated in going from conceptual to lexical to phonological levels of representation, and these link to dorsal stream WM areas including Spt/TPJ for eventual articulation. Previous authors have noted that the Spt/TPJ region displays a heterogeneous and finely subdivided set of functional and cytoarchitectonic fields (Scott and Johnsrude, 2003), and we observe completely distinct activation profiles separated by 1 cm in this region. Thus, we cannot fault previous authors for overlooking such finely spaced

distinctions, but further progress in this area will require the type of anatomical specificity discussed here.

We have studied speech production using two common language tasks, verb generation and picture naming. We observe robust pre-response activations in dorsal stream areas also implicated in phonological WM and suggest these sites as the common route to speech production, including phonological representation of the to-be-spoken word. Our results provide the first unambiguous demonstration of post-response (feedback) activation of STG, with a clear absence of pre-response activations. Such evidence is only obtainable with a method combining high spatial and temporal resolution, such as ECoG mapping with γ_{high}. A working model for the physiological origins of the γ_{high} signal is presented in the SI.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF. Voxel-based lesion-symptom mapping. Nat Neurosci 2003;6:448–450. [PubMed: 12704393]

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate – a practical and powerful approach to multiple testing. J R Stat Soc, Series B (Methodological) 1995;57:289–300.

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 2000;10:512–528. [PubMed: 10847601]

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

Broca P. Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau. Bull Soc Anthropol Paris 1861;2:235–238.

Brovelli A, Lachaux JP, Kahane P, Boussaoud D. High gamma frequency oscillatory activity dissociates attention from intention in the human premotor cortex. NeuroImage 2005;28:154–164. [PubMed: 16023374]

Brown EC, Rothermel R, Nishida M, Juhász C, Muzik O, Hoechstetter K, Sood S, Chugani HT, Asano E. In vivo animation of auditory-language-induced gamma-oscillations in children with intractable focal epilepsy. NeuroImage 2008;41:1120–1131. [PubMed: 18455440]

Buchsbaum BR, D'Esposito M. The search for the phonological store: from loop to convolution. J Cogn Neurosci 2008;20:762–778. [PubMed: 18201133]

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

Buchsbaum BR, Olsen RK, Koch PF, Kohn P, Kippenhan JS, Berman KF. Reading, hearing, and the planum temporale. NeuroImage 2005;24:444–454. [PubMed: 15627586]

Canolty RT, Soltani M, Dalal SS, Edwards E, Dronkers NF, Nagarajan SS, Kirsch HE, Barbaro NM, Knight RT. Spatiotemporal dynamics of word processing in the human brain. Front Neurosci 2007;1:185–196. [PubMed: 18982128]

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

Crone NE, Miglioretti DL, Gordon B, Lesser RP. Functional mapping of human sensorimotor cortex with electrocorticographic spectral analysis. II. Event-related synchronization in the gamma band. Brain 1998;121:2301–2315. [PubMed: 9874481]

Crone NE, Boatman D, Gordon B, Hao L. Induced electrocorticographic gamma activity during auditory perception. Brazier Award-winning article. Clin Neurophysiol 2001;112:565–582. [PubMed: 11275528]

Crone NE, Sinai A, Korzeniewska A. High-frequency gamma oscillations and human brain mapping with electrocorticography. Progr Brain Res 2006;159:275–295.

Dalal SS, Edwards E, Kirsch HE, Barbaro NM, Knight RT, Nagarajan SS. Localization of neurosurgically implanted electrodes via photograph–MRI–radiograph coregistration. J Neurosci Methods 2008;174:106–115. [PubMed: 18657573]

Damasio AR, Tranel D. Nouns and verbs are retrieved with differently distributed neural systems. Proc Natl Acad Sci U S A 1993;90:4957–4960. [PubMed: 8506341]

Darrow CW, Graf CG. Relation of electroencephalogram to photometrically observed vasomotor changes in the brain. J Neurophysiol 1945;8:449–461.

Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. Psychol Rev 1997;104:801–838. [PubMed: 9337631]

Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition 2004;92:145–177. [PubMed: 15037129]

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J Neurophysiol 2005;94:4269–4280. [PubMed: 16093343]

Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, Knight RT. Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex. J Neurophysiol 2009;102:377–386. [PubMed: 19439673]

Goodglass, H.; Kaplan, E.; Barresi, B. The assessment of aphasia and related disorders. Lippincott Williams & Wilkins; Philadelphia: 2000.

Griffiths TD, Warren JD. The planum temporale as a computational hub. Trends Neurosci 2002;25:348–353. [PubMed: 12079762]

Herholz K, Reulen HJ, von Stockhausen HM, Thiel A, Ilmberger J, Kessler J, Eisner W, Yousry TA, Heiss WD. Preoperative activation and intraoperative stimulation of language-related areas in patients with glioma. Neurosurgery 1997;41:1253–1260. [PubMed: 9402576]

Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev, Neurosci 2007;8:393–402. [PubMed: 17431404]

Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: an MEG study. J Cogn Neurosci 2002;14:1125–1138. [PubMed: 12495520]

Indefrey, P.; Cutler, A. Prelexical and lexical processing in listening. In: Gazzaniga, MS., editor. The cognitive neurosciences. MIT Press; Cambridge, MA: 2004. p. 759-774.

Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. Cognition 2004;92:101–144. [PubMed: 15037128]

Jacquemot C, Scott SK. What is the relationship between phonological short-term memory and speech processing? Trends Cogn Sci 2006;10:480–486. [PubMed: 16997610]

Kaplan, E.; Goodglass, H.; Weintraub, S. Boston naming test. Lea & Febiger; Philadelphia: 1983.

Lachaux JP, Fonlupt P, Kahane P, Minotti L, Hoffmann D, Bertrand O, Baciu M. Relationship between task-related gamma oscillations and BOLD signal: New insights from combined fMRI and intracranial EEG. Hum Brain Mapp 2007;28:1368–1375. [PubMed: 17274021]

Levelt, WJM. Producing spoken language: a blueprint of the speaker. In: Brown, CM.; Hagoort, P., editors. The neurocognition of language. Oxford University Press; Oxford; New York: 1999. p. 83-122.

Levelt WJM, Praamstra P, Meyer AS, Helenius P, Salmelin R. An MEG study of picture naming. J Cogn Neurosci 1998;10:553–567. [PubMed: 9802989]

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. Psychol Rev 1967;74:431–461. [PubMed: 4170865]

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysio-logical investigation of the basis of the fMRI signal. Nature 2001;412:150–157. [PubMed: 11449264]

Loring DW, Hermann BP, Perrine K, Plenger PM, Lee GP, Meador KJ. Effect of Wada memory stimulus type in discriminating lateralized temporal lobe impairment. Epilepsia 1997;38:219–224. [PubMed: 9048675]

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

Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. The essential role of premotor cortex in speech perception. Curr Biol 2007;17:1692–1696. [PubMed: 17900904]

Miglioretti DL, Boatman D. Modeling variability in cortical representations of human complex sound perception. Exp Brain Res 2003;153:382–387. [PubMed: 14534769]

Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. Science 2005;309:951–954. [PubMed: 16081741]

Nichols T, Hayasaka S. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res 2003;12:419–446. [PubMed: 14599004]

Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA. Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 2005;309:948–951. [PubMed: 16081740]

Ojemann GA, Ojemann JG, Lettich E, Berger MS. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. J Neurosurg 1989;71:316–326. [PubMed: 2769383]

Ojemann JG, Ojemann GA, Lettich E. Cortical stimulation mapping of language cortex by using a verb generation task: effects of learning and comparison to mapping based on object naming. J Neurosurg 2002;97:33–38. [PubMed: 12134930]

Ono, M.; Kubik, S.; Abernathey, CD. Atlas of the cerebral sulci. G. Thieme Verlag; Thieme Medical Publishers; Stuttgart; New York: 1990.

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

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

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 1988;331:585–589. [PubMed: 3277066]

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron emission tomographic studies of the processing of single words. J Cogn Neurosci 1989;1:153–170.

Pulvermüller F, Shtyrov Y, Ilmoniemi RJ, Marslen-Wilson WD. Tracking speech comprehension in space and time. NeuroImage 2006;31:1297–1305. [PubMed: 16556504]

Rademacher, J. Topographical variability of cytoarchitectonic areas. In: Schüz, A.; Miller, R., editors. Cortical areas: unity and diversity. Taylor & Francis; London; New York: 2002. p. 53-78.

Roland PE, Zilles K. Structural divisions and functional fields in the human cerebral cortex. Brain Res Brain Res Rev 1998;26:87–105. [PubMed: 9651489]

Scott SK, Johnsrude IS. The neuroanatomical and functional organization of speech perception. Trends Neurosci 2003;26:100–107. [PubMed: 12536133]

Scott SK, Wise RJS. The functional neuroanatomy of prelexical processing in speech perception. Cognition 2004;92:13–45. [PubMed: 15037125]

Scott SK, Blank SC, Rosen S, Wise RJS. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 2000;123:2400–2406. [PubMed: 11099443]

Shallice T, Warrington EK. Auditory–verbal short-term memory impairment and conduction aphasia. Brain Lang 1977;4:479–491. [PubMed: 922463]

Sinai A, Bowers CW, Crainiceanu CM, Boatman D, Gordon B, Lesser RP, Lenz FA, Crone NE. Electrocorticographic high gamma activity versus electrical cortical stimulation mapping of naming. Brain 2005;128:1556–1570. [PubMed: 15817517]

Sinai A, Crone NE, Wied HM, Franaszczuk PJ, Miglioretti DL, Boatman-Reich D. Intracranial mapping of auditory perception: event-related responses and electrocortical stimulation. Clin Neurophysiol 2009;120:140–149. [PubMed: 19070540]

Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS. Converging language streams in the human temporal lobe. J Neurosci 2006;26:7328–7336. [PubMed: 16837579]

Thiel A, Herholz K, von Stockhausen HM, van Leyen-Pilgram K, Pietrzyk U, Kessler J, Wienhard K, Klug N, Heiss WD. Localization of language-related cortex with ^{15}O-labeled water PET in patients with gliomas. NeuroImage 1998;7:284–295. [PubMed: 9626669]

Thompson-Schill SL, Swick D, Farah MJ, D'Esposito M, Kan IP, Knight RT. Verb generation in patients with focal frontal lesions: a neuropsycholog-ical test of neuroimaging findings. Proc Natl Acad Sci U S A 1998;95:15855–15860. [PubMed: 9861060]

Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. NeuroImage 2008;39:1429–1443. [PubMed: 18035557]

Towle VL, Yoon HA, Castelle M, Edgar JC, Biassou NM, Frim DM, Spire JP, Kohrman MH. ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. Brain 2008;131:2013–2027. [PubMed: 18669510]

Vallar, G.; Papagano, C. Neuropsychological impairments of verbal short-term memory. In: Baddeley, AD., et al., editors. The handbook of memory disorders. J. Wiley & Sons; New York: 2002. p. 249-270.

Warren P, Marslen-Wilson WD. Continuous uptake of acoustic cues in spoken word recognition. Percept Psychophys 1987;41:262–275. [PubMed: 3575084]

Wernicke, C. Der aphasische Symptomenkomplex: eine psychologische Studie auf anatomischer. Basis Cohn & Weigert; Breslau: 1874.

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

Wise RJS, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural subsystems within 'Wernicke's area'. Brain 2001;124:83–95. [PubMed: 11133789]

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi: 10.1016/j.neuroimage.2009.12.035.
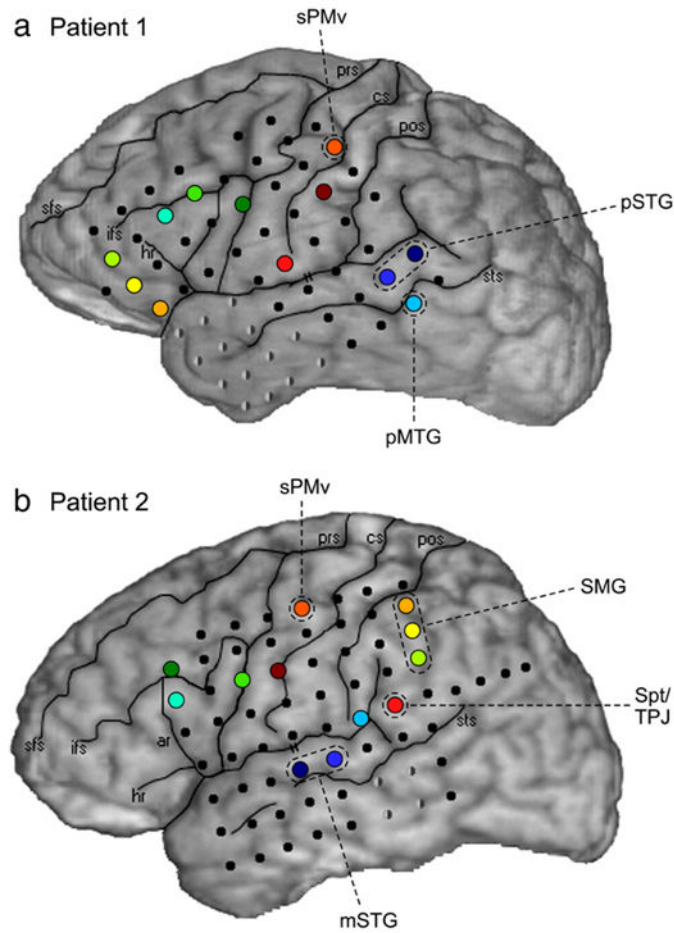
Fig. 1.
Electrode locations on 3D MRI reconstructions. Electrodes are embedded in silastic in an 8×8 array with 1 cm spacing and implanted on the cortical surface. Additional strips of electrodes contact the orbitofronal cortex (Patient 1) or parietal cortex (Patient 2). Electrodes of special interest in this study (shown in Figs. 2–5) are illustrated in color. Electrodes in grey–white exhibited interictal epileptiform activity and were rejected for analysis. The seizure focus was in the medial temporal lobe in both patients. Major sulci are drawn over the reconstruction to permit faster anatomical interpretations. sts: superior temporal sulcus. pos: postcentral sulcus. cs: central sulcus. prs: precentral sulcus. sfs: superior frontal sulcus. ifs: inferior frontal sulcus. hr: horizontal ramus of the Sylvian fissure (defines border between pars triangularis and pars orbitalis of the IFG). ar: ascending ramus (defines the border between pars triangularis and pars opercularis). The small "H" on the Sylvian fissure indicates the approximate location of Heschl's gyrus within the fissure. Electrodes given special mention in the text are labeled with dashed lines for quick reference. mSTG: middle superior temporal gyrus. pSTG: posterior superior temporal gyrus. pMTG: posterior middle temporal gyrus. sPMv: superior portion of ventral premotor cortex (Wilson et al., 2004). SMG: supramarginal gyrus. TPJ: temporal–parietal junction (this site is suggested to correspond functionally to the speech production area called 'Tpt' or 'Spt').
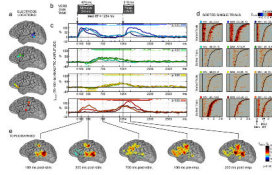
Fig. 2.
Verb generation results for Patient 1. (a) Locations for 12 selected electrodes (refer to Fig. 1 for anatomical detail), grouped according to similar response profiles. (b) Verb generation task. The patient heard auditorily-presented nouns (median duration 470 ms) and responded overtly with an associated verb (median response duration 310 ms). (c) γ_{high} (70–160 Hz) analytic amplitude (AA) averaged across trials after realigning to the median response time (Med.RT=1254 ms). Straight lines above (below) the traces indicate significant increases (decreases) of γ_{high} AA relative to pre-stimulus baseline ($P < 0.02$; FDR corrected for multiple-comparisons) (d) Single-trial γ_{high} AA is shown in z-score units and sorted according to RT. Response onset is indicated by the curved black line in each plot (range 0.87–3.45 s). Single-trial activity was smoothed with a 3-trial moving window average. (e) Topographies of γ_{high} AA are shown for several latencies taken from (c). Activations are masked for significance at $P<0.02$.
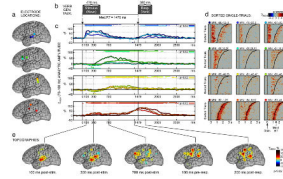
Fig. 3.
Verb generation results for Patient 2. Plotting conventions same as Fig. 2. The second electrode group (green colors) includes sites on the inferior and middle frontal gyri (IFG, MFG). The grid placement was more posterior in this patient, so no coverage was obtained anterior to *pars opercularis* of the IFG, but coverage of the inferior parietal lobe (IPL) was gained. The third electrode group includes three IPL electrodes on the posterior SMG. The fourth group (red colors) includes one electrode from the TPJ cortex just posterior to the branch point of the Sylvian fissure. Although it is a surprising location from which to observe response-related activity, this electrode responded similarly to the two premotor electrodes and is therefore grouped with them.
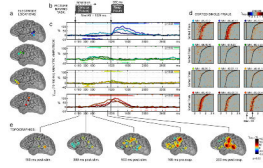
Fig. 4.
Picture naming results for Patient 1. Plotting conventions are the same as Fig. 2. Picture stimuli are black-and-white line drawings and remain displayed throughout each trial. The median RT (1026 ms) is faster for this task than for verb generation. Note that in the absence of visual cortex exposure, only the most posterior–inferior electrode shows strong (~50%) stimulus-locked activation (topography at 100 ms post-stim.). Other activations are discussed in the text and confirm the pattern seen for speech production in the verb generation task. In particular, pSTG electrodes (top row) exhibit only *post-response* activations.
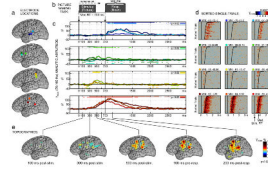
Fig. 5.
Picture naming results for Patient 2. This patient had a fast response time (median RT 733 ms) in this task. The pattern seen for speech production during verb generation is confirmed, including clear response-related activation of the Spt/TPJ site (bottom row). Wernicke's area (top row) showed only post-response activations.
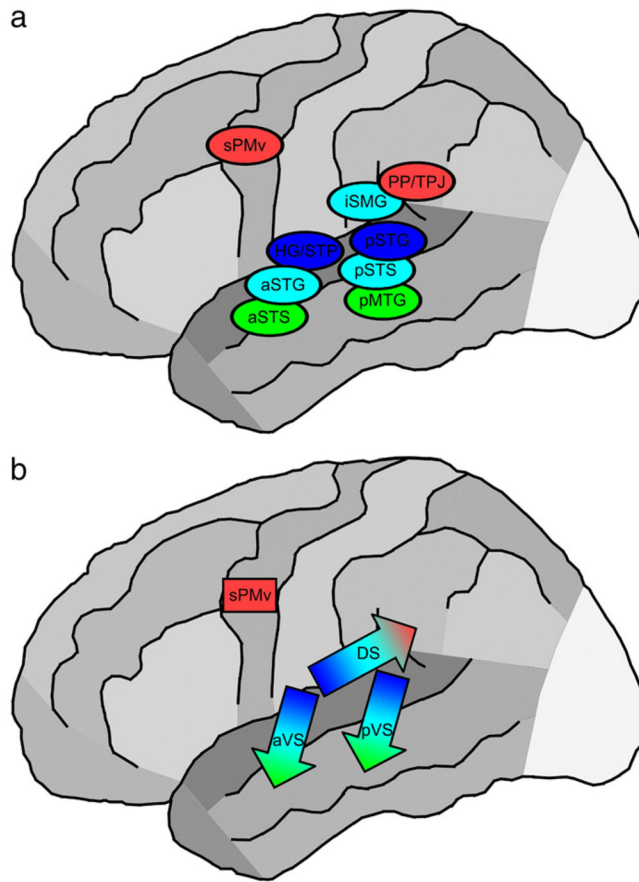
Fig. 6.
Illustration of a "dual ventral stream" model for auditory speech processing, consistent with our results and prior studies (Crinion et al., 2003; Miglioretti and Boatman, 2003; Indefrey and Cutler, 2004; Scott and Wise, 2004). The model is identical to that of Hickok and Poeppel (2007) in the overall division between a dorsal stream (DS) and a ventral stream (VS). It additionally specifies that the ventral stream is often split topographically into two streams traversing the STG to the STS and MTG. The anterior ventral stream (aVS) proceeds ventro-laterally from Heschl's gyrus (HG) and belt regions (Scott et al., 2000). The posterior VS (pVS) and the dorsal stream arise from the planum temporale (Griffiths and Warren, 2002). In some subjects the anterior and posterior ventral streams may overlap at the mid-STG, but the most common phenotype, and the pattern most likely obtained in a meta-analysis across patients, is a bimodal (dual stream) distribution. The colors in the figure roughly indicate function as acoustic–phonetic (blue), phonological (turquoise), and word-form/lexico-semantic (green). For the dorsal stream, red indicates involvement in speech production. Note that not all dorsal stream areas are labeled here (operculum, insula, etc.). The arrows with color gradients in b indicate that there are no sharp divisions between regions supporting these functional roles. Rather, the auditory processing becomes increasingly linguistic with progression along the ventral stream and increasingly WM- and response-related along the dorsal stream.