



Published in final edited form as:

Brain Lang. 2011 June ; 117(3): 103–109. doi:10.1016/j.bandl.2010.09.009.

Sub-centimeter language organization in the human temporal lobe

A. Flinker¹, E.F. Chang³, N.M. Barbaro³, M.S. Berger³, and R.T. Knight^{1,2,3}

¹Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA

²Department of Psychology, University of California, Berkeley, CA, USA

³Department of Neurosurgery, University of California, San Francisco, CA, USA

Abstract

The human temporal lobe is well known to be critical for language comprehension. Previous physiological research has focused mainly on noninvasive neuroimaging and electrophysiological techniques with each approach requiring averaging across many trials and subjects. The results of these studies have implicated extended anatomical regions in peri-sylvian cortex in speech perception. These non-invasive studies typically report a spatially homogenous functional pattern of activity across several centimeters of cortex. We examined the spatiotemporal dynamics of word processing using electrophysiological signals acquired from high-density electrode arrays (4 mm spacing) placed directly on the human temporal lobe. Electrographic (ECoG) activity revealed a rich mosaic of language activity, which was functionally distinct at four mm separation. Cortical sites responding specifically to word and not phoneme stimuli were surrounded by sites that responded to both words and phonemes. Other sub-regions of the temporal lobe responded robustly to self-produced speech and minimally to external stimuli while surrounding sites at 4 mm distance exhibited an inverse pattern of activation. These data provide evidence for temporal lobe specificity to words as well as self produced speech. Furthermore, the results provide evidence that cortical processing in the temporal lobe is not spatially homogenous over centimeters of cortex. Rather, language processing is supported by independent and spatially distinct functional sub-regions of cortex at a resolution of at least 4 mm.

Keywords

electrocorticography; ECoG; gamma; auditory processing

Introduction

The role of the human temporal lobe has been vigorously studied since Carl Wernicke's first account of language comprehension deficits in stroke patients (Wernicke, 1874). Numerous neuroimaging studies have employed various paradigms in an attempt to elucidate the neuroanatomical pathways of language perception. Across these studies the superior

© 2010 Elsevier Inc. All rights reserved

Correspondence: Adeen Flinker University of California at Berkeley 132 Barker Hall Berkeley, CA 94720-3190 Telephone: 510 316 8591 Fax: 510 643 3516 adeen.f@gmail.com.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

temporal gyrus and superior temporal sulcus have been consistently implicated in the perception of speech (Binder, et al., 1997; Démonet, Chollet, et al., 1992; Price, et al., 1996; R. J. S. Wise, et al., 1991; Zatorre, Evans, Meyer, & Gjedde, 1992).

Evidence from non-human primates supports a hierarchal organization of functionally distinct subdivisions of auditory cortex (Hackett, Stepniewska, & Kaas, 1998; Rauschecker, 1998). Studies in humans support a similar subdivision of auditory cortex supporting distinct anterior and posterior streams of information processing (Hickok & Poeppel, 2007; Scott, Blank, Rosen, & Wise, 2000; R. J. Wise, et al., 2001). Furthermore, Binder et al. identified different sub-regions of the superior temporal lobe activated in processing speech and non-speech stimuli (Binder, et al., 2000). Nonetheless there remain discrepancies between the studies regarding the exact anatomical pathways as well as the functional significance of the different subdivisions of auditory cortex. Scott et al. reported a posterior and anterior subdivision processing unintelligible and intelligible speech respectively (Scott, et al., 2000). Wise et al. focused on two posterior subdivisions: the posterior STS processing perception and retrieval of words, and the medial temporoparietal junction processing speech production (R. J. Wise, et al., 2001). Lastly, Hickok & Poeppel proposed a dual-stream model of speech processing including a ventral stream (superior middle temporal lobe) processing speech comprehension and a dorsal stream (posterior dorsal temporal lobe, parietal operculum and posterior frontal lobe) processing auditory-motor integration (Hickok & Poeppel, 2007). Neuroimaging studies investigating the neuroanatomical functional organization of the superior temporal gyrus typically report activation sites spanning several centimeters of cortex and represent an average across several subjects (Binder, et al., 1997; Démonet, Chollet, et al., 1992; Price, et al., 1996; R. J. S. Wise, et al., 1991; Zatorre, et al., 1992). Conversely, intraoperative language mapping using electrical cortical stimulation (ESM) report a high degree of inter-subject variability in the location of cortical language sites using 1 cm resolution electrodes (Ojemann, Ojemann, Lettich, & Berger, 1989; Sanai, Mirzadeh, & Berger, 2008). This inter-subject variability suggests that activation maps currently drawn from neuroimaging data are averaging over a distribution of cortical sub-regions involved in speech processing potentially obscuring a finer grain cortical organization of language. While the functional neuroanatomy of language is likely common across subjects, some cortical sites could have a spatially dense organization, which varies across subjects. That is, use of a common coordinate frame across subjects could blur differences in regional cytoarchitecture across individual subjects. In order to address the spatial distribution of cortical activity during word and phoneme processing we recorded intraoperative electrocorticographic (ECoG) activity directly from the surface of the human temporal lobe using high-density (4 mm spacing) electrode arrays. ECoG recordings acquired directly from cortex provide a rich electrophysiological signal with spectral content not readily seen in conventional EEG. Spectral High Gamma band activity ($\gamma_{\text{High}} > 70$ Hz) is an ideal index for cortical activity and has been reported to reliably track neuronal activity in various functional relevant modalities including, auditory- (Crone, Boatman, Gordon, & Hao, 2001; Edwards, Soltani, Deouell, Berger, & Knight, 2005; Trautner, et al., 2006), motor- (Crone, Miglioretti, Gordon, & Lesser, 1998; Miller, et al., 2007) and language- (Brown, et al., 2008; Canolty, et al., 2007; Crone, Hao, et al., 2001; Tanji, Suzuki, Delorme, Shamoto, & Nakasato, 2005; Towle, et al., 2008) related tasks. Furthermore, the γ_{High} response has been linked to neuronal firing rate and is believed to emerge from synchronous firing of neuronal populations (Allen, Pasley, Duong, & Freeman, 2007; Belitski, et al., 2008; Liu & Newsome, 2006; Mukamel, et al., 2005; Ray, Crone, Niebur, Franaszczuk, & Hsiao, 2008).

Result

Intraoperative electrocorticographic (ECoG) activity was recorded from four subjects undergoing neurosurgical treatment. Cortical responses were sampled from a high-density multi-electrode grid (inter-electrode spacing of 4 mm) placed over the posterior superior temporal gyrus (STG). Figure 1 depicts cortical responses indexed by oscillatory high gamma (γ_{High} : 70–150 Hz) activity across a 64 contact grid in subject S1 performing a passive listening task to phonemes (top) and a separate task listening to words (bottom). Significant γ_{High} activity commenced as early as 100 ms post stimulus onset and varied in temporal onset across different electrodes (z-scores corresponding to $p < 0.001$ after FDR correction, see Methods).

Cortical sites responding to phoneme stimuli showed a consistently similar spectral response pattern for word stimuli. γ_{High} activity locked to word stimuli remained active longer than γ_{High} activity locked to phoneme stimuli. Sustained γ_{High} activity locked to words remained active throughout stimulus presentation and in some sites up to 200 ms after stimulus has ended. Several cortical sites, which did not produce a significant response during phoneme listening, produced a robust response during word listening. Furthermore, many of these word specific sites were surrounded by electrodes 4 mm away that responded to both stimuli types.

All three subjects who performed both a phoneme and a word task (see Supplemental Table 1) exhibited cortical sites with word specific responses. While one site showed robust activity for both phoneme and word stimuli (Figure 2, Electrode B), an adjacent electrode revealed significant γ_{High} responses for words and minimal responses for phonemes (Figure 2, Electrode A). The word specific electrodes (Electrode A in each subject, see Supplemental Figure 1 for electrode anatomical positions) showed a significant difference between word and phoneme γ_{High} responses within a time window of the first 200 ms (Wilcoxon rank-sum, $p < 0.001$). Furthermore, we ran an analysis of variance within each subject in order to assess a condition (phoneme or word) across electrode (A or B) interaction. We found a significant condition X electrode interaction within each subject (S1: $F(1,698) = 6.68$, $p < 0.01$; S2: $F(1,598) = 126.76$, $p < 0.001$; S3: $F(1,454) = 9.33$, $p < 0.01$). For each subject a main effect of condition was significant (S1: $F(1,698) = 22.7$, $p < 0.001$; S2: $F(598) = 30.2$, $p < 0.001$; S3: $F(454) = 20.99$, $p < 0.001$) as well as a main effect of electrode (S1: $F(1,698) = 45.99$, $p < 0.001$; S2: $F(598) = 32.27$, $p < 0.001$; S3: $F(454) = 61.61$, $p < 0.001$). A breakdown of the interactions revealed that all subjects showed a significant difference between electrodes in both the phoneme and word conditions as well as a significant difference between conditions within electrode A (a detailed breakdown of the interactions and the main effects are provided in Supplemental Figure 2).

Three of the four subjects performed a phoneme repetition task (see Supplemental Table 1). Figure 3 depicts cortical responses evoked by hearing and then producing phonemes across a 64 contact grid in subject S4. A majority of electrodes exhibited a γ_{High} response locked to hearing phonemes commencing as early as 80 ms post stimulus onset. Responses locked to phoneme production varied across electrodes. A majority of cortical sites exhibited a suppressed response during production as compared with hearing (Wilcoxon rank-sum, $p < 0.001$), while several sites did not show a significant change in response (Wilcoxon rank-sum, $p > 0.05$). One electrode exhibited an enhanced response during production while its eight neighboring electrodes showed a suppressed response as compared with hearing (Figure 3, blue box; arrowhead). A direct comparison of spectral γ_{High} responses locked to hearing and speaking is detailed in Supplemental Figure 3.

Two subjects revealed cortical sites that exhibited production specific responses while a neighboring electrode produced an opposite functional response (see Supplemental Figure 1 for electrode anatomical positions). Single trial analysis of the responses showed consistent responses across trials as seen in Figure 4. Stacked γ_{High} power traces for each trial locked to hearing phonemes are shown for two adjacent electrodes in subject S4 (top) and S2 (bottom). A production specific electrode (electrode A) exhibited significantly higher γ_{High} power after production (black line) as compared with hearing external stimuli (Wilcoxon rank-sum, $p < 0.001$). The adjacent electrode (electrode B) at 4 mm distance exhibited a robust response to the external stimuli (relative to baseline) while the response to self-produced speech is severely reduced (relative to hearing in both subjects and relative to baseline in subject S4).

Discussion

We recorded electrocorticographic (ECoG) activity directly from the surface of the temporal cortex of four awake patients undergoing neurosurgery. The aim of the study was to assess the temporal-spatial distribution of cortical activity during processing of phonemes and words. Recording ECoG activity from high-density electrode grids revealed different functional responses of cortical sites separated by 4 mm distance. All three subjects who performed both a word and phoneme related task showed word specific activity adjacent to cortical sites that were non-specific to stimulus type. Two of the three subjects who performed a phoneme repetition task showed production specific activity adjacent to cortical sites that were active during hearing and suppressed during production. These functionally distinct responses within 4 mm of cortex support a dense distribution of independent sub-regions supporting linguistic processing. Evidence of such dense distribution of activity supports a spatially fine graded cortical representation of speech processing in the human temporal lobe. Neuroanatomical differences across subjects in this fine graded representation may obscure current neuroimaging findings, which typically employ averaging techniques over many subjects.

The superior temporal gyrus (STG) and superior temporal sulcus (STS) have long been implicated in both word processing as well as phonological processing (Binder, et al., 1997; Boatman, Lesser, & Gordon, 1995; Démonet, Chollet, et al., 1992; Price, et al., 1996; R. J. S. Wise, et al., 1991; Zatorre, et al., 1992). Studies in non-human primates have shown that superior temporal neurons are broadly tuned and exhibit stronger responses to wideband noise than pure tones. Furthermore, neurons in the lateral surface of the STG are selective to species-specific vocalization (Rauschecker, 1998; Rauschecker, Tian, & Hauser, 1995). Findings in humans have shown superior temporal lobe specificity to intelligible speech (Binder, et al., 2000; Canolty, et al., 2007; Mummery, Ashburner, Scott, & Wise, 1999; Scott, et al., 2000). Our ECoG data, which is sampled from neuronal populations in the lateral surface of the STG, provides similar evidence for complex stimulus dependent local specificity. While one neuronal population responds similarly to both phoneme and word stimuli a neighboring population responds selectively to words. This result can be explained by the attributes of the stimulus presented as well as the specificity of the underlying neuronal populations. One hypothesis is that the neurons in the two neighboring populations have different spectral tuning. The word selective responses could be the result of a large neuronal population responding to wider and more specific frequency bands in the stimuli. An alternate hypothesis is that one population contains more neurons which are responsive to a specific class of complex stimuli. These two hypotheses are not mutually exclusive and could both contribute to the observed neurophysiological signals. Furthermore, we cannot rule out the possibility that the different interstimulus intervals (ISI) in the phoneme (1.5 s ISI) and word (2 s ISI) tasks may partially contribute to the differential functional response (Ceponiene, Cheour, & Näätänen, 1998; Röder, Rösier, & Neville, 1999). Our results, which

complement previous findings in non-human primates (Rauschecker, 1998; Rauschecker, et al., 1995), provide further evidence of a hierarchical organization of the temporal cortex where different neuronal populations preferentially respond to more complex stimuli.

The human temporal cortex has consistently shown a suppressed pattern of activity during self-produced speech as compared with external stimuli (Christoffels, Formisano, & Schiller, 2007; Curio, Neuloh, Numminen, Jousmäki, & Hari, 2000; Ford, et al., 2001; Houde, Nagarajan, Sekihara, & Merzenich, 2002). Single unit studies in both non-human and human primates have provided evidence of STG neurons that are suppressed during vocalization as well as a subset of neurons which do not show suppression (Müller-Preuss 1981; Eliades 2003; Eliades 2008; Creutzfeldt 1989). Our results showing suppressed auditory responses during production replicate previous ECoG findings (Crone 2001; Towle 2008) as well as non-invasive electrophysiological results (Ford 2001; Numminen 1999; Houde 2002). The fact that we find cortical sites exhibiting both suppressed, non-suppressed and excited responses could be due to variability in the proportion of suppressed neurons in the sampled region of cortex. Such an interpretation could explain the differing proportions of suppressed neurons reported in single unit studies (Creutzfeldt, Ojemann, & Lettich, 1989; S. J. Eliades & Wang, 2003; S. Eliades & Wang, 2008; Müller-Preuss & Ploog, 1981). The speech-specific cortical sites we have found in two patients provide evidence for monitoring of self-speech. These electrophysiological responses could represent a population of neurons with specificity to self-produced vocalization. We currently cannot rule out the possibility that these responses are due to neuronal tuning to the spectral attributes of self-speech such as the acoustic masking of the auditory signal in the middle ear due to bone conduction (v. Békésy, 1949). Nevertheless, our results provide further evidence for a rich dynamic of spatial cortical responses in the STG, which are functionally distinct.

Evidence from both non-human primates (Hackett, et al., 1998; Rauschecker, 1998; Romanski, et al., 1999) as well as humans (Hickok & Poeppel, 2007; Scott, et al., 2000; R. J. Wise, et al., 2001) supports a hierarchical organization of functionally distinct subdivisions of the auditory cortex. The current emphasis is on anterior and posterior functional streams of processing similar to those in the visual domain (Hickok & Poeppel, 2007; Rauschecker & Tian, 2000). Our current data supports a richer spatial organization within the temporal cortex suggesting a more complex and finer functional subdivision.

Methods

Subjects

Four subjects participated in the study at University of California San Francisco (UCSF) while undergoing intraoperative neurosurgical treatment for refractory epilepsy (Subjects S1 and S3) or tumor resection (Subjects S2 and S4). Treatment involved a surgical procedure including intraoperative awake language and motor mapping followed by tailored resection of the damaged tissue under ECoG guidance. After all clinical mapping was performed, the surgeon (EC) placed a high density electrode array with inter-electrode spacing of 4 mm. The subjects performed either passive listening or repetition tasks (Supplemental Table 1) for several minutes after which time the grid was removed and the surgeon continued clinical treatment. The subjects provided written and oral consent prior to the surgery and were informed that the task was for research purposes. All subjects were fluent in English and had no language deficits. During surgery, the subjects were informed by the surgeon when the clinical mapping was over and the research task was completed under the discretion of the surgeon. All medical treatment including the size and location of the craniotomy site were dictated solely by the clinical needs of the patient under the discretion of the surgeon. The study protocol was approved by the UC San Francisco and UC Berkeley Committees on Human Research.

Tasks and Stimuli

Four subjects (S1–4) participated in short five-minute tasks consisting of either passive listening (S1–2) or active repetition (S2–4) of phonemes or words (see Supplemental Table 1 for breakdown per subject). The phoneme passive listening task consisted of the synthesized syllables /ba/, /da/ and /ga/ digitized with a sampling rate of 20 KHz and 16 bit precision. Phoneme stimuli were presented with a jittered inter stimulus interval (ISI) of 1.5 seconds \pm 100 ms (random jitter), each stimulus lasted approximately 150 ms. Subjects were instructed that they were going to hear several speech syllables and they were asked to concentrate on whether they heard /ba/, /da/ or /ga/. The phoneme repetition task consisted of the synthesized syllables /ba/ and /pa/ digitized with a sampling rate of 20 KHz and 16 bit precision. Phoneme stimuli were presented with a jittered inter stimulus interval (ISI) of 3 seconds \pm 100 ms (random jitter), each stimulus lasted approximately 150 ms. Subjects were instructed to listen and then repeat aloud each speech syllable as best they could. The word passive listening task consisted of word stimuli, digitally prerecorded from a female native speaker of English, acquired at a sampling rate of 44 KHz and 16 bit precision. Word stimuli consisted of 23 pseudo-words (3 phonemes in length), 23 real words (3 phonemes in length) and 4 proper names (5 phonemes in length). Real words were all high frequency (Kucera Francis log scale 2–2.4), both real and pseudo-words were controlled for phonotactic probabilities using the Irvine Phonotactic Online Dictionary (Vaden, Hickok, & Halpin, 2009).

Word stimuli were presented with a jittered inter stimulus interval (ISI) of 2 seconds \pm 250 ms (random jitter), stimuli varied in length (400–700 ms) with a mean of 525 ms and standard deviation of 100 ms. Subjects were instructed that they were going to hear several words, some with meaning and some without, and they were asked to listen to what they heard. The word repetition task consisted of the same word stimuli described above. Word stimuli were presented with a jittered inter stimulus interval (ISI) of 4 seconds \pm 250 ms (random jitter). Subjects were instructed that they were going to hear several words, some with meaning and some without, and they were asked to repeat aloud each word they heard as best they could. All subjects were presented with the digital audio stimuli vowels via two speakers placed in front of them. During the repetition tasks vocalization responses were recorded by using a microphone placed in front of the patient. All peripheral equipment was placed in the same locations traditionally used by the neurology team who administer language mapping tasks. Both the speaker and microphone channels were fed directly to the recording system in order to record stimulus presentation and responses simultaneously with the electrophysiological signals.

Electrode Placement and Selection

High density grids consisted of electrode arrays with 4 mm inter-electrode spacing manufactured by Ad-Tech Medical Instrument Corporation (<http://www.adtechmedical.com/>) assuring hospital-grade, FDA-approved standards. All electrode arrays were placed by the surgeon and covered the posterior lateral surface of the superior temporal gyrus. Supplemental Figure 4 shows an intraoperative image with and without the electrode grid alongside a reconstruction of the MRI. Grid locations were ascertained by the performing neurosurgeon using the intraoperative images as well as stereotactic coordinates. Word specific sites reported in Figure 2 were all inferior to the sylvian fissure on the lateral surface of the superior temporal gyrus (S1: electrodes 23 and 14, 5.66 mm center-to-center; S2: electrodes 56 and 63, 5.66 mm center-to-center; S3: electrodes 31 and 39, 4 mm center-to-center). During the recording of subject S2 the grid had to be repositioned by the surgeon, as a result the phoneme repetition task was recorded in a different position than the word repetition and passive phoneme task. The word-specific results for subject S2 compared passive phoneme listening to active word repetition. Word-

specific sites were still found when comparing two nearby electrodes (< 1 cm) in the phoneme repetition and word repetition tasks (Wilcoxon rank-sum, $p < 0.001$). Speech specific sites reported in Figure 4 were inferior to the sylvian fissure on the lateral surface of the temporal gyrus (S3: electrodes 40 and 48, 4 mm center-to-center; S4: electrodes 19 and 27, 4 mm center-to-center).

Data Acquisition

Electrophysiological and peripheral auditory channels were acquired using a custom made Tucker Davis Technologies (<http://www.tdt.com/>) recording system (256 channel amplifier and Z-series digital signal processor board). Channels were sampled at 3051 Hz and the peripheral auditory channels were sampled at 24.4 KHz. Electrophysiological data was recorded using a subdural electrode as reference (corner grid electrode) and an additional electrode as ground.

Data Preprocessing

All EEG channels were manually inspected by a neurologist in order to identify epileptic channels as well as epochs of ictal activity which spread to other channels. Channels contaminated by sustained epileptic activity, electrical line noise (60 Hz) or abnormal signal were removed from further analysis. The raw time series, voltage histograms, and power spectra were used to identify noisy channels. All channels were re-referenced to a common averaged reference defined as the mean of all the remaining channels. Epochs in which ictal activity spread to non-epileptic channels were also removed from further analysis. Speaker and microphone channels recorded simultaneously with EEG activity were manually inspected in order to mark onset of stimulus and subsequent response. The audio channels were inspected using both the raw time-series as well as a time-frequency representation (spectrogram) to ensure accurate onset estimation. Trials in which the subject did not respond were removed from analysis, similarly trials overlapping with ictal activity were discarded. All analyses were done using custom scripts written for MATLAB, The Mathworks Inc. (<http://www.mathworks.com/>).

Data Analysis

Event Related Spectral Perturbations (ERSP) were created by first computing power series across the entire task for several spectral bands, producing a time-frequency representation of the entire task. Subsequently event related power averages locked to stimuli onset were computed and baseline corrected ($-300 \rightarrow -100$ ms) for each spectral band using a temporal window of $-100 \rightarrow 750$ ms in the passive tasks and $-100 \rightarrow 2200$ in the repetition task. Spectral bands were defined using center frequencies logarithmically spaced from 1 through 250 Hz and a fractional bandwidth of 20% the center frequency (i.e. a band centered at 4 Hz will be 3.6–4.4 Hz and a band centered at 100 Hz will be 90–110 Hz). Statistical significance was assessed using a bootstrapping approach similar to the method used by Canolty et al. (Canolty, et al., 2007). In brief, an ERSP is created using N randomly generated time stamps (time windows do not overlap with ictal activity) instead of the N real stimuli onsets (N =number of real stimuli onsets). This process is then repeated over 1000 times producing a surrogate distribution of ERSPs. Each time-frequency point in the real ERSP can now be transformed using a z-score based on the mean and standard deviation of the surrogate distribution (which is normally distributed). All time-frequency plots are corrected for multiple comparisons (number of frequencies, time points and channels) using False Discovery Rate (FDR) correction of $q=0.05$ (Benjamini & Hochberg, 1995).

In order to assess statistical differences between ERSPs locked to phonemes and words the raw power values were directly tested using a non-parametric Wilcoxon rank sum test. For each stimuli onset an average was computed across a temporal window of $0 \rightarrow 200$ ms and a

spectral window of 75.9 → 144.5 Hz (all center frequency bands within range). The average power values for all phoneme stimuli were tested against those of all word stimuli. Furthermore, an analysis of variance was run on raw power values across a temporal window of 0→400 ms and a spectral window of 75.9 → 144.5 Hz to assess condition × electrode interactions. Interactions were statistically assessed using a post-hoc pair-wise t-test.

The same approach was used in order to compare responses locked to hearing and speaking phonemes (a temporal window of 0→200 ms and a spectral window of 75.9 → 144.5 Hz)

Onset and duration times of γ_{High} responses locked to phonemes and words were assessed during the first 750 ms post-stimulus using a spectral window of 75.9 → 144.5 Hz averaged across trials.

Single trial γ_{High} traces were computed by first calculating the spectral power time series (70–150 Hz) for the entire block of data. Event related windows of the time series were extracted and transferred to units of percent change compared with baseline (averaged spectral power within -300 → -100 ms pre-stimulus). Statistical differences between hearing and speaking were assessed by comparing average γ_{High} responses within a temporal window of 0→200 ms locked to stimuli.

Spectral Decomposition

Spectral signal analysis was implemented using a frequency domain Gaussian filter (similarly to (Canolty, et al., 2007)). An input signal X was transformed to the frequency domain signal X_f using an N -point fft (where N is defined by the number of points in the time-series X). In the frequency domain a Gaussian filter was constructed (for both the positive and negative frequencies) and multiplied with the signal X_f . The subsequent filtered signal was transformed back to the time-domain using an inverse fft. Power estimates were calculated by taking the Hilbert transform of the frequency filtered signal and squaring the absolute value. All frequency domain filtering and power estimations are comparable to other filtering techniques, such as the wavelet approach (Bruns, 2004).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to convey our warm gratitude to Beth Mormino who provided assistance in stimulus recording.

This research was supported by the National Institute of Health grants NS059804 (to R.T.K.), NS21135 (to R.T.K.), PO40813 (to R.T.K.) as well as NIH grants F32NS061552 (to E.C.) and K99NS065120 (to E.C.).

References

- Allen EA, Pasley BN, Duong T, Freeman RD. Transcranial magnetic stimulation elicits coupled neural and hemodynamic consequences. *Science*. 2007; 317(5846):1918–1921. [PubMed: 17901333]
- Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis N, et al. Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *Journal of Neuroscience*. 2008; 28(22):5696–5709. [PubMed: 18509031]
- Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*. 1995; 57(1):289–300. Series B (Methodological)

- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN. Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex*. 2000; 10(5):512–528. [PubMed: 10847601]
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *Neurosci*. 1997; 17(1):353–362.
- 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]
- Brown, Erik C.; Rothermel, R.; Nishida, M.; Juhász, C.; Muzik, Otto; Hoehstetter, Karsten, et al. In-Vivo Animation of Auditory-Language-Induced Gamma-Oscillations in Children with Intractable Focal Epilepsy. *NeuroImage*. 2008; 41(3):1120.
- Bruns A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *Journal of Neuroscience Methods*. 2004; 137(2):321–332. [PubMed: 15262077]
- Canolty R, Soltani M, Dalai SS, Edwards E, Dronkers NF, Nagarajan SS, et al. Spatiotemporal dynamics of word processing in the human brain. *Front Neurosci*. 2007; 1(1):185–196. [PubMed: 18982128]
- Ceponiene R, Cheour M, Näätänen R. Interstimulus interval and auditory event-related potentials in children: evidence for multiple generators. *Electroencephalogr Clin Neurophysiol*. 1998; 108(4):345–354. [PubMed: 9714376]
- Christoffels IK, Formisano E, Schiller NO. Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Human Brain Mapping*. 2007; 28(9):868–879. [PubMed: 17266104]
- Creutzfeldt O, Ojemann, Lettich. Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Exp Brain Res*. 1989; 77(3):476–489. [PubMed: 2806442]
- Crone NE, Boatman D, Gordon B, Hao L. Induced electrocorticographic gamma activity during auditory perception. Brazier Award-winning article, 2001. *Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology*. 2001; 112(4):565–582. [PubMed: 11275528]
- Crone NE, Hao L, Hart J, Boatman D, Lesser RP, Irizarry R, et al. Electrocorticographic gamma activity during word production in spoken and sign language. *Neurology*. 2001; 57(11):2045–2053. [PubMed: 11739824]
- 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(Pt 12):2301–2315. [PubMed: 9874481]
- Curio G, Neuloh G, Numminen J, Jousmäki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*. 2000; 9(4):183–191. [PubMed: 10770228]
- Démonet JF, Chollet F, Ramsay S, Cardebat D, Nespoulous JL, Wise RJS, et al. The anatomy of phonological and semantic processing in normal subjects. *Brain*. 1992; 115(Pt 6):1753–1768. [PubMed: 1486459]
- Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. *Journal of Neurophysiology*. 2005; 94(6):4269–4280. [PubMed: 16093343]
- Eliades SJ, Wang X. Sensory-Motor Interaction in the Primate Auditory Cortex During Self-Initiated Vocalizations. *Journal of Neurophysiology*. 2003; 89(4):2194. [PubMed: 12612021]
- Eliades S, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*. 2008; 453(7198):1102–1106. [PubMed: 18454135]
- Ford JM, Mathalon DH, Heinks T, Kalba S, Faustman WO, Roth WT. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *The American journal of psychiatry*. 2001; 158(12):2069–2071. [PubMed: 11729029]
- Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *THE JOURNAL OF COMPARATIVE NEUROLOGY*. 1998; 394(4):475–495. [PubMed: 9590556]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci*. 2007; 8(5):393–402. [PubMed: 17431404]

- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: an MEG study. *Journal of cognitive neuroscience*. 2002; 14(8):1125–1138. [PubMed: 12495520]
- Liu J, Newsome WT. Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *Journal of Neuroscience*. 2006; 26(30):7779–7790. [PubMed: 16870724]
- Miller KJ, Leuthardt EC, Schalk G, Rao RP, Anderson NR, Moran DW. Spectral changes in cortical surface potentials during motor movement. *Journal of Neuroscience*. 2007; 27(9):2424–2432. [PubMed: 17329441]
- 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(5736):951–954. [PubMed: 16081741]
- Müller-Preuss P, Ploog D. Inhibition of auditory cortical neurons during phonation. *Brain Research*. 1981; 215(1–2):61–76. [PubMed: 7260601]
- Mummery CJ, Ashburner J, Scott SK, Wise RJ. Functional neuroimaging of speech perception in six normal and two aphasic subjects. *The Journal of the Acoustical Society of America*. 1999; 106(1):449–457. [PubMed: 10420635]
- Ojemann, George; Ojemann, Jeff; Lettich, E.; Berger, M. Cortical language localization in left, dominant hemisphere. *Journal of Neurosurgery*. 1989; 71(3):316–326. [PubMed: 2769383]
- Price CJ, Wise RJ, Warburton EA, Moore CJ, Howard D, Patterson K, et al. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain*. 1996; 119(Pt 3):919–931. [PubMed: 8673502]
- Rauschecker JP. Parallel processing in the auditory cortex of primates. *Audiol Neurootol*. 1998; 3(2–3):86–103. [PubMed: 9575379]
- Rauschecker JP, Tian B. Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proc Natl Acad Sci USA*. 2000; 97(22):11800–11806. [PubMed: 11050212]
- Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*. 1995; 268(5207):111–114. [PubMed: 7701330]
- Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS. Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *Journal of Neuroscience*. 2008; 28(45):11526–11536. [PubMed: 18987189]
- Röder B, Rösler F, Neville HJ. Effects of interstimulus interval on auditory event-related potentials in congenitally blind and normally sighted humans. *Neuroscience Letters*. 1999; 264(1–3):53–56. [PubMed: 10320012]
- Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature Neuroscience*. 1999; 2(12):1131–1136.
- Sanai N, Mirzadeh Z, Berger MS. Functional outcome after language mapping for glioma resection. *The new england journal of medicine*. 2008; 358(1):18–27. [PubMed: 18172171]
- Scott SK, Blank CC, Rosen S, Wise RJ. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*. 2000; 123(Pt 12):2400–2406. [PubMed: 11099443]
- Tanji K, Suzuki K, Delorme A, Shamoto H, Nakasato N. High-frequency gamma-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *Journal of Neuroscience*. 2005; 25(13):3287–3293. [PubMed: 15800183]
- Towle VL, Yoon H, Castelle M, Edgar JC, Biassou NM, Frim DM, et al. ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. *Brain*. 2008; 131(8):2013. [PubMed: 18669510]
- Trautner P, Rosburg T, Dietl T, Fell J, Korzyukov OA, Kurthen M, et al. Sensory gating of auditory evoked and induced gamma band activity in intracranial recordings. *NeuroImage*. 2006; 32(2):790–798. [PubMed: 16809054]
- v. Békésy G. The Structure of the Middle Ear and the Hearing of One's Own Voice by Bone Conduction. *The Journal of the Acoustical Society of America*. 1949; 21:217.
- Vaden, KI.; Hickok, GS.; Halpin, HR. Irvine Phonotactic Online Dictionary. Version 1.4*2009. Data file Available from <http://www.iphod.com>

- Wernicke, C. *Boston Studies in the Philosophy of Science: Proceedings of the Boston Colloquium for the Philosophy of Science*. Vol. Volume IV. D. Reidel Publishing Co; Dordrecht-Holland: 1874. 1966/1968. The symptom complex of aphasia. A psychological study of an anatomical basis.
- Wise RJS, Chollet F, Hadar U, FRISTON K, Hoffner E, Frackowiak RSJ. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*. 1991; 114(Pt 4): 1803–1817. [PubMed: 1884179]
- Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural subsystems within 'Wernicke's area'. *Brain*. 2001; 124(Pt 1):83–95. [PubMed: 11133789]
- Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science*. 1992; 256(5058):846–849. [PubMed: 1589767]

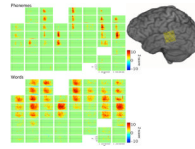


Figure 1. Spatiotemporal responses to phonemes (top) and words (bottoms) across a 64 contact 8×8 electrode grid in subject S1. Event related spectral perturbations are shown for each electrode locked to the onset of stimuli. Color scale represents statistically significant changes in power compared to a bootstrapped surrogate distribution (only significant data is shown, $p < 0.05$ after FDR multiple comparison correction). Electrodes with no contact or abnormal signal are not shown.

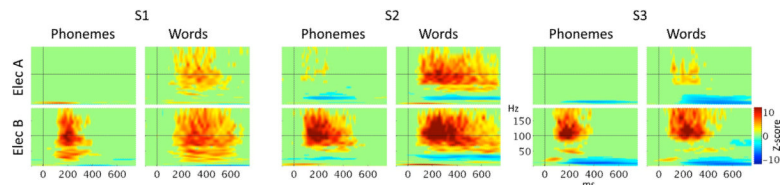


Figure 2.

Event related spectral perturbations of two adjacent electrodes in three subjects locked to phonemes and words. Electrode A (top row) responds selectively to word stimuli while Electrode B (bottom row) 4 mm away responds to both stimuli types. Vertical line marks stimuli onset, horizontal line marks 100 Hz, the color scale represents statistical significant changes in power compared to a bootstrapped surrogate distribution. Only statistically significant values are shown ($p < 0.05$, after FDR multiple comparison correction). See Supplemental Figure 1 for electrode anatomical positions.

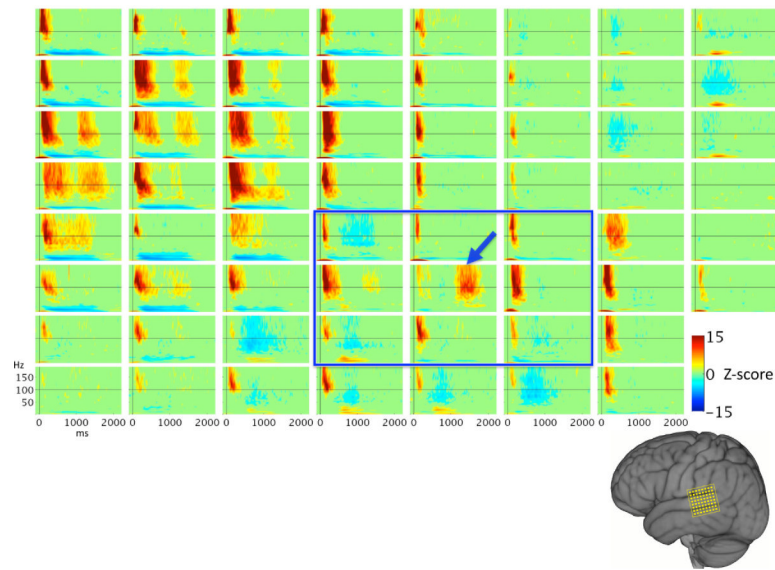


Figure 3. Spatiotemporal responses to hearing and then producing phonemes across a 64 contact 8×8 electrode grid in subject S4. Event related spectral perturbations are shown for each electrode locked to the onset of phonemes. Color scale represents statistically significant changes in power compared to a bootstrapped surrogate distribution (only significant data is shown, $p < 0.05$ after FDR correction). Blue box marks an area of approximately 1 cm² of cortex; blue arrowhead marks an electrode exhibiting self-speech specificity. Electrodes with no contact or abnormal signal are not shown. See Supplemental Figure 3 for a direct comparison of spectral γ_{High} responses locked to hearing and speaking.

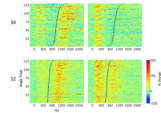


Figure 4. Single trial γ_{High} power traces vertically stacked for two adjacent electrodes in subjects S2 and S4. Zero ms marks phoneme onset and black lines mark production onset (trials are sorted for display purposes). Color-scale represents percent change from the pre-stimulus baseline in each trial. See Supplemental Figure 1 for electrode anatomical positions.