Title

The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening

Permalink

https://escholarship.org/uc/item/0kd7n73h

Authors

Cibelli, Emily S
Leonard, Matthew K
Johnson, Keith
et al.

Publication Date

2015-08-01

DOI

10.1016/j.bandl.2015.05.005

Peer reviewed

The influence of lexical statistics on temporal lobe cortical dynamics during spoken word listening

Emily S. Cibelli[1], **Matthew K. Leonard**[2], **Keith Johnson**[1], and **Edward F. Chang**[2,3,*]

[1] Department of Linguistics, University of California, Berkeley, 1203 Dwinelle Hall, Berkeley, CA 94720, USA

[2] Department of Neurological Surgery, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143, USA

[3] Department of Physiology, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143, USA

Abstract

Neural representations of words are thought to have a complex spatio-temporal cortical basis. It has been suggested that spoken word recognition is not a process of feed-forward computations from phonetic to lexical forms, but rather involves the online integration of bottom-up input with stored lexical knowledge. Using direct neural recordings from the temporal lobe, we examined cortical responses to words and pseudowords. We found that neural populations were not only sensitive to lexical status (real vs. pseudo), but also to cohort size (number of words matching the phonetic input at each time point) and cohort frequency (lexical frequency of those words). These lexical variables modulated neural activity from the posterior to anterior temporal lobe, and also dynamically as the stimuli unfolded on a millisecond time scale. Our findings indicate that word recognition is not purely modular, but relies on rapid and online integration of multiple sources of lexical knowledge.

Keywords

Word comprehension; lexical statistics; pseudowords; temporal lobe; electrocorticography (ECoG)

1. Introduction

Many theoretical accounts of spoken word recognition posit a process of matching acoustic input with stored representations that have rich lexical and semantic structure. These models assume the existence of acoustic-phonetic, phonemic, and lexical targets in the brain that are activated when specific input is received (McClelland & Elman 1986; Norris 1994; Marslen-Wilson 1987, 1989). For example, hearing the word "cat" evokes activity in neural

*Correspondence to: Edward Chang (ChangEd@neurosurg.ucsf.edu), University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143, USA, (415) 353-3489.
(ecibelli@berkeley.edu, kjohnson@berkeley.edu)
(LeonardM@neurosurg.ucsf.edu, ChangEd@neurosurg.ucsf.edu)

populations that are selective to phonetic features like plosives and low front vowels, which in turn activate stored representations of the phonemes /k/, /æ/, and /t/. The activations of these phonemic representations are integrated over time, and serve as inputs to neurons at the lexical level that represent the word "cat". The dynamic nature of this process has led many researchers to suggest that the representations in this hierarchy interact with and influence each other, meaning that word recognition is an iterative process where multiple targets at each level are active until the input is no longer consistent with those targets (e.g., /k-æ-**p**/) (Marslen-Wilson & Welsh, 1978; Heald & Nusbaum 2014).

Several influential models of the neural basis of speech comprehension (Hickok & Poeppel 2007; Scott & Wise 2004) propose a set of cortical regions that perform the transformation from spectrotemporal representations of speech signals to abstracted lexical representations of words. These proposals are based on data from patients with lesions to various cortical areas (Dronkers & Wilkins 2004), and on recent neuroimaging studies that support a distributed and interconnected network of cortical regions thought to be responsible for the representation of words and language (see, e.g., Davis and Gaskell 2009; Turken & Dronkers 2011). Many of these studies observe functional specialization of different regions in the temporal lobe, with acoustic-phonetic and phonemic representations in the posterior superior temporal cortex, and higher-order lexical representations in the middle, anterior, and ventral temporal cortex. This pathway is often referred to as the auditory ventral stream and is argued to link acoustic, phonemic, and lexical processing (see also Dewitt & Rauschecker 2012; Okada, Hickok, & Serences 2009; Lerner, Honey, Silbert, & Hasson, 2011).

However, it remains unclear how this transformation occurs, and specifically how the ventral stream integrates high-level knowledge about the language with bottom-up acoustic input. In particular, the mental lexicon can be characterized by a number of features and statistics that relate the stored representations of individual words with one another, and also with lower-level features like phonemes and phonetic features. As a speech token unfolds, a cohort of forms stored in the lexicon that match the acoustic input is activated (Marslen-Wilson 1987, 1989). This matching set of lexical forms (the cohort) will change over time as more of the target word is heard, thereby changing the lexical competition space on a moment-by-moment basis. It is therefore necessary to capture the temporal dynamics of these changing lexical statistics when describing the processes involved in word comprehension. A primary goal of the present study is to describe the spatiotemporal dynamics of spoken word recognition across the duration of a word, and across the auditory ventral stream.

In the present study, we compare neural responses to real words (e.g. *ceremony, repetition*) and novel forms, or pseudowords (e.g. *moanaserry, piteretion*), to examine how this lexical structure is encoded in the brain. Several studies have found differences in the hemodynamic response to real words and pseudowords (Davis & Gaskell 2009; Mechelli, Gorno-Tempini, & Price 2003; Raettig and Kotz 2008; Tanji, Suzuki, Delorme, Shamoto, & Nakasato 2005; Mainy et al. 2008). However, it is likely that the word/pseudoword difference is not purely binary; in particular, there is behavioral evidence that word-like forms may be processed as potential real words (Meunier & Longtin 2007; De Vaan, Schreuder, & Baayen 2007;

Lindsay, Sedin, & Gaskell 2012). Taken together, these findings suggest that while neural responses to pseudowords can be reliably distinguished from familiar word forms, the processing of novel forms may also rely on information stored in the lexicon, including high-level features like cohort statistics. Therefore, a second goal of this study is to explore how this type of stored lexical information can affect the processing of both pseudowords and real words.

Lexical statistics which capture aspects of lexical competition may be of particular importance for both real word and pseudoword processing. Cohort size (Marslen-Wilson 1987, 1989; Magnuson, Dixon, Tanenhaus, & Aslin, 2007) is defined as the number of words in the lexicon that match the phonemes the listener has heard up to any given point in a word. This provides an incremental metric of potential lexical forms which changes as acoustic input is received. Average cohort frequency, by contrast, is defined as the average lexical frequency of the words in a cohort. Finally, summed cohort frequency sums the lexical frequency of all words in a cohort, thus quantifying the number of words and their relative usage statistics in a single metric. The extent to which neural activity evoked by real words and pseudowords is modulated by these features allows us to explore the specific linguistic processes involved in the acoustic-to-lexical transformation.

To study the on-line processing of lexical forms and cohort statistics, we examined cortical responses to real words and pseudowords using data recorded from high-density electrocorticographic (ECoG) electrodes placed directly on the cortical surface. ECoG provides high spatial and temporal resolution with a relatively high signal-to-noise ratio at the individual electrode level. These properties are critical to our study goals of examining how the lexical status and cohort statistics of specific speech tokens affect neural activity as the input is being processed in real-time. Because the neural representations of words are complex, distributed, and likely represented in a high-dimensional space, these methodological advantages may be necessary to uncover the nature of lexical processing.

2. Methods

2.1. Subjects

Four human subjects underwent surgery placement of a 256-channel subdural electrocorticography (ECoG) array as part of clinical treatment for epilepsy surgery planning. All electrode arrays were placed over the perisylvian region of the language-dominant hemisphere (left hemisphere for all but subject 3; no observable patterns differentiated the right hemisphere data from the left hemisphere data). All subjects gave informed written consent prior to surgery and experimental testing.

2.2. Stimuli

2.2.1. Stimulus design—The stimulus set consisted of ten words and ten pseudowords. Words were four syllables long, and each syllable had consonant-vowel or consonant-vowel-consonant structure. Pseudowords were formed from the real words by scrambling syllables or segments to form phonotactically-legal novel sequences (e.g. "ceremony" became "moanaserry," "repetition" became "piteretion"). Stimuli are listed in Table 1.

2.2.2. Lexical statistics—Cohort size, average cohort frequency, and summed cohort frequency were calculated for each stimulus at each phoneme. Cohort size was defined as the number of competitors that matched the segment identity (but not necessarily the stress pattern) of a stimulus up to a given point. As an example, after three phonemes, the word *delivery* [**dil**ɪvɚi] has a cohort size of 67, and the cohort includes words such as *delete* [**dil**it] and *delicious* [**dili**ʃɾs]. Average cohort frequency was defined as the average of the lexical frequency of each cohort member (Luce and Large 2001). Summed cohort frequency was calculated by summing the lexical frequency of all members in a cohort. An example of cohort size, average cohort frequency, and summed cohort frequency is shown in Figure 1a.

Cohorts were initially identified by sorting segmental information in the CMU Pronouncing Dictionary (Weide 2008). Because the CMU corpus contains many entries that are proper nouns or extremely unfamiliar forms, the resulting cohorts were filtered through the Irvine Phonotactic Online Dictionary (IPhOD, Vaden, Halpin, & Hickok, 2009), which provides lexical frequency information for more than 54,000 words. We also calculated average cohort frequency and summed cohort frequency at each segment. Frequency counts were drawn from the SUBTLEXus database included in IPhOD (SCDcnt, Brysbaert & New, 2009). Our metrics of cohort frequency allow for a phoneme-by-phoneme accounting of lexical frequency, which was more suited to the non-decomposable pseudowords in our data set. Following Magnuson et al. (2007), cohort statistics were calculated starting at the second phoneme and models described below were likewise fit from the second phoneme onward. Cohort size and summed and average cohort frequency values are plotted as a function of time for all stimuli in Figure 1b.

In preliminary analyses of the stimuli, we examined phonotactic probability as a possible source of sublexical differences between words and pseudowords. We found no differences between the two groups of stimuli in the current study for measures of unstressed biphone probability (unweighted, frequency-weighted, and log-frequency weighted; all $p > 0.05$). Unstressed positional probability marginally differed between the two groups ($t = 1.754$, $p = 0.097$), although the frequency-weighted and log-frequency-weighted metrics of positional probability did not ($p > 0.05$). Because the effect of positional probability was marginal, we included it as a control variable in the initial model, but ultimately removed it, as its inclusion did not improve the model fit. We also investigated the possible effects of acoustic or phonetic properties of the stimuli, but these properties were ruled out as meaningful covariates in the present study (see Supplement for discussion of phonetic controls).

2.3. Experimental procedures

Data from this study were collected from the listening portion of an overt repetition task. Subjects listened to each stimulus and repeated back what they heard. Task design for repetition varied slightly by subject; for subjects 2 and 3, repetition was self-paced, while subjects 1 and 4 were instructed to wait for a cue before repeating the word. Each stimulus was played and repeated back twice per block. Subjects 1 and 2 completed two blocks of the experiment, subject 3 completed five blocks, and subject 4 completed four blocks.

2.4. Data analysis

2.4.1. Data acquisition and pre-processing—ECoG recordings can provide high spatial and temporal resolution of cortical areas that are known to be critical speech regions of the auditory ventral stream. The high-gamma frequency band of the recorded cortical surface potential has several favorable properties, including high signal-to-noise ratio, and is argued to be more spatially and temporally locked to task-related activity than responses at lower frequencies (Crone, Sinai, and Korzeniewska 2006). Because it is thought to be a reliable index for neuronal activity in tasks across a range of functions and modalities (Crone, Boatman, Gordon, & Hao, 2001; Edwards, Soltani, Deouell, Berger, & Knight 2005; Canolty et al. 2007; Flinker, Chang, Barbaro, Berger, & Knight, 2011; Chang et al. 2010), it is well-suited to investigate linguistic processes. For these reasons, high-gamma has become a standard measure in investigations of spoken language processing that use intracranial recordings (e.g. Tanji et al. 2005, Mainy et al. 2008, Towle, Yoon, et al. 2008).

ECoG data was recorded as local field potentials from 256-channel cortical arrays and a multichannel amplifier connected to a digital signal processor (TuckerDavis Technologies). Data was sampled at 3052 Hz and downsampled to 400 Hz for inspection and pre-processing (and ultimately to 100 Hz for analysis with statistical models, following steps described in this section). The data were visually inspected for noisy channels and time points containing artifacts, which were removed prior to analysis; remaining channels were common-average referenced on 16-channel blocks, corresponding to amplifier inputs. The high-gamma band (70-150 Hz) was extracted for analysis from the average of eight logarithmically-spaced band-pass filters using the Hilbert transform (each band logartithmically increasing in center frequency from 75-150 Hz, and bandwidth increasing semi-logarithmically). The analytic amplitude across all eight bands was averaged to calculate the high-gamma power (for more detail, see Bouchard, Mesgarani, Johnson, & Chang, 2013). The high-gamma analytic amplitude was then z-scored relative to a baseline resting period at the end of each recording session for each electrode. Analyses were restricted to electrodes over the temporal lobe (Figure 2a). We further restricted our analyses to only those electrodes that showed responses that were significantly time-locked to the stimulus (relative to a resting baseline period; FDR correction, adjusted $p < 0.05$). A total of 300 electrodes across all subjects were identified for analysis (84 from subject 1; 68 from subject 2; 88 from subject 3; 60 from subject 4).

2.4.2. Audio processing and event time alignment—Acoustic event onsets were identified using automatic phoneme- and word-level alignments generated by the Penn Phonetics Lab Forced Aligner (Yuan & Liberman, 2008). The output of the aligner was hand-corrected in any cases where word boundaries were judged by visual inspection to be inaccurately aligned to the acoustic signal.

2.4.3. Growth curve analysis—We hypothesized that differential neural responses to words and pseudowords, in different electrodes across time, reflect the encoding of various lexical features. Traditional techniques for examining condition-specific effects over time involve performing a series of independent tests on non-independent contiguous time points, or on wide time windows that potentially average out meaningful temporal variation. To

directly examine how lexical processing varies over time, we used growth curve analysis (GCA), a statistical method that explicitly models the encoding of variables of interest (e.g. cohort size, average cohort frequency and summed cohort frequency) as a function of time (Mirman, Dixon, & Magnuson, 2008; Magnuson et al., 2007). GCA is a linear mixed effects regression modeling technique that handles time series data; it uses polynomial time terms to account for the shape and rate changes in a curve, and includes interactions with variables of interest to describe their effect on the trajectory of the dependent variable. Mirman and colleagues (2008) demonstrate that analyses with discrete time bins fail to capture meaningful variability in eye-tracking data, which are temporally fine-grained, and that GCA models are better suited to describing the influences of linguistic predictors at different time points in a trial. Similarly, our data have fine-grained temporal resolution, and require a modeling technique that can capture the dynamics of linguistic processing as speech unfolds over time.

A GCA model was fit to the high-gamma response across all electrodes with a fourth-order (quartic) orthogonal polynomial and fixed effects of lexicality (pseudoword vs. word), cohort size, average cohort frequency, summed cohort frequency, trial order, anterior-posterior grid position (A-P position), and dorsal-ventral grid position (D-V position). Continuous predictors (all but lexicality) were centered around 0. Interactions were included between each of the linguistic main effects (lexicality, cohort size, summed cohort frequency, and average cohort frequency) and each polynomial term and position term (A-P and D-V). Two-way interactions were also included between pairings of the linguistic variables. The random effects structure of the model consisted of random intercepts for subject, stimulus, and electrode, and by-electrode random slopes for the first- and second-order time terms. A full description of the model is presented in Supplementary Table 1. The dependent variable (high-gamma activity) was log-transformed for normality. The window for analysis was from the onset of the second phoneme to 1500 milliseconds after word onset; all stimulus offsets occurred several hundred milliseconds before the end of this window.

Because all of the cohort metrics included in this model are correlated to some extent, we took care to ensure that their inclusion did not lead to model over-fitting. A test between all of our continuous predictors indicated that multicollinearity was not an issue ($\kappa = 8.9$; values below 10 are considered to be safe). We also constructed three similar models, where only two of the cohort metrics were included. The best two-metric model (cohort size and summed cohort frequency) was a worse fit to the data than the three-metric model (Akaike Information Criteron [AIC], two metric: –2396842; AIC, three-metric: –2397393; χ^2 (18) = 947.39, $p < 0.0001$). As a result, we chose to retain all three metrics in the model reported here.

The interpretation of models with many higher-order interactions is complex. To illustrate how the magnitude of model terms affects the predicted curves generated by the model, Figure 1c demonstrates the effect of adjusting time parameters in a sample growth curve model. In this example, we show the effects of adjusting the magnitude of the lexicality * time term interactions in one electrode, holding all other predictors (including the three cohort metrics) constant. The plots show the model prediction for all word stimuli over time

(black lines), with 50% increases and decreases made to each lexicality * time coefficient independent of the others. Each nth-order term affects the curve in a different way. The intercept term shifts the curve up or down; the linear term (lexicality*time) adjusts the steepness or shallowness of the slope; the quadratic term (lexicality*time^2) affects the centering of the peak of the curve; the cubic (lexicality*time^3) and quartic (not shown, as its effects tend to be subtle) terms adjust the curve at its tails (Mirman et al. 2008). Each higher order term introduces an additional inflection point, with a flip in the relative magnitude of the +50% and −50% adjustments at each point. The effect of each predictor on the shape of each electrode's response curve indicates whether a given linguistic variable causes stronger or weaker responses, earlier or later peaks, and sharper or more sustained activity.

3. Results

3.1. Timing and location of responses to real words and pseudowords

Real words and pseudowords evoked different high-gamma neural responses across many temporal lobe electrodes (see Figure 2a for electrode placement in one subject) in all participants, with typically stronger activity for pseudowords (Figure 2b). This lexicality effect was significant between 320-1500 milliseconds (bootstrap $p < 0.05$). In addition to these magnitude differences, there was a clear progression in the timing of the peak of the neural response from posterior to anterior temporal lobe electrodes over time (Figure 2c), that differed between real words ($R^2 = 0.13$) and pseudowords ($R^2 = 0.06$). Across electrodes with significant high-gamma responses, the peak latency for pseudowords was significantly later than for real words (two-sample t-test, $t(497.3) = −5.9248$, $p < 0.001$). Thus, we observed overall magnitude and timing differences between real words and pseudowords throughout the superior temporal lobe during online speech processing.

3.2. Growth curve analysis

The above analyses of the high-gamma responses at each electrode illustrate that differential neural responses to words and pseudowords are dynamic over the course of processing. To examine the fine-scale temporal properties of lexical processing in the temporal lobe, we used growth curve analysis (GCA) to model neural activity on each electrode as a function of lexical variables and time. This technique avoids the problems associated with using discrete and arbitrary time windows, and takes full advantage of the high temporal resolution that ECoG affords. The full table of coefficients for this model is reported in Supplemental Table 1.

3.2.1. Effects of linguistic variables and electrode position—We examined the relationship between high-gamma activity and stimulus lexicality (real word vs. pseudoword), cohort size (the number of words in the lexicon consistent with the phonetic input at a given point in time), average cohort frequency (the mean lexical frequency of all cohort competitors at a given time), and summed cohort frequency (the total lexical frequency of all cohort competitors at a given time). The coefficients for the simple main effects of each of these linguistic variables are not very informative, as each predictor enters into higher-order interactions which are critical to the interpretation their overall effects.

Therefore, we restrict our discussion of the effects of each linguistic variable to their interaction terms, beginning with the effect of position.

We first considered how the effect of lexicality varied across the temporal lobe. The interaction of lexicality and anterior-posterior (A-P) position along the temporal lobe was significant ($\beta = 0.006$, $t = 6.79$). As activity propagated from posterior to anterior, responses to pseudowords increased, while they decreased for words. There was no significant interaction of lexicality and dorsal-ventral (D-V) position ($\beta = -0.0008$, $t = -0.04$).

We also examined whether cohort size varied as a function of electrode position. There was a significant cohort size by A-P position interaction ($\beta = -0.009$, $t = -7.22$). This effect was most apparent at the larger cohorts, which showed smaller responses in posterior sites and larger responses at anterior sites. In smaller cohorts, this effect of A-P position was much less pronounced. There were significant interactions of average cohort frequency with both A-P position ($\beta = -0.0009$, $t = -5.47$) and D-V position ($\beta = -0.0002$, $t = -5.55$). Summed cohort frequency also had significant interactions with both A-P position ($\beta = 0.0001$, $t = 7.64$) and D-V position ($\beta = 0.0004$, $t = 9.07$). These terms suggest that the number of lexical competitors is encoded along the anterior-posterior extent of the temporal lobe, while frequency effects have both anterior-posterior and dorsal-ventral gradients.

Together, the positional effects of lexicality, cohort size, average cohort frequency, and summed cohort frequency demonstrate that hierarchical speech processing in the temporal lobe is influenced not only by the lexical status of a form, but also by high-order statistics that describe learned knowledge about the structure of the mental lexicon, including the complex relationships between words.

3.2.2. Interactions between linguistic variables—A major goal of the study was to understand how the combined influences of several lexical statistics influence the processing of both words and pseudowords. It is possible that the effects of stored lexical knowledge for processing familiar forms differ when compared to the effects for novel forms. Similarly, it is possible that the different cohort metrics interact with one another in different ways as incoming speech is being processed. To test this, we examined all two-way interactions of the linguistic variables (lexicality, cohort size, average cohort frequency, and summed cohort frequency).

Lexicality significantly interacted with cohort size ($\beta = -0.007$, $t = -3.38$), but not with the other cohort metrics. The overall magnitude of the cohort size effect was larger for words than pseudowords. This suggests that there was a larger difference in the neural response at large and small cohorts for words than pseudowords (Figure 3a shows the responses predicted by the model across cohort sizes for this interaction).

There was a significant interaction between cohort size and average cohort frequency ($\beta = 0.0059$, $t = 13.84$; Figure 3b). For visual convenience, average cohort frequency is split into three bins (high, mid, and low). Across all bins, responses are stronger at larger cohort sizes; however, this effect is much more pronounced for the most frequent cohorts, and is attenuated for mid- and low-frequency cohorts.

A different relationship holds for the interaction of cohort size and summed cohort frequency ($\beta = -0.0009$, $t = -6.56$; Figure 3c). In this case, the neural response decreases as cohort size decreases for the highest and mid-range summed frequency bins. However, for the bin with the lowest summed frequency, the opposite effect was observed, where less frequent cohorts have a larger response as cohort size decreases.

Finally, average cohort frequency and summed cohort frequency significantly interacted ($\beta = -0.0019$, $t = -14.61$; Figure 3d), suggesting that the two frequency metrics have dissociable effects. The highest and mid-range summed frequency bins had a negative relationship with average frequency, where decreasing average cohort frequency was associated with an increased neural response. For the lowest summed cohort frequency bin, the response decreased as average cohort frequency decreased.

3.2.3. Lexicality effects over time—A second major goal of the present study was to understand how lexical variables affect neural activity over time, as information about the stimulus identity is changing. While the main effects of each of these variables is informative for understanding the data set at an average time and position, language comprehension is a dynamic process that unfolds throughout the entire timecourse of acoustic input. To understand how real words and pseudowords differentially modulate activity over time, we examined the interaction between lexicality and the polynomial time terms included in the GCA model. Each polynomial term in the model captures a different way in which the neural signal can change over time. The interaction of each predictor with the linear term accounts for changes in the slope of the curve, while the quadratic term describes the steepness of the curve and how symmetrical or centered the curve is. Effects in the tails of the curves are captured by the cubic and quartic terms (Mirman et al. 2008), which contributed significantly but relatively subtly to many of the models, and are thus not discussed in detail below.

In the raw data, we observed that the latency to the peak of a response was later for pseudowords than for real words (Figure 4a, upper panel). As the results for individual electrodes across the A-P axis of the temporal lobe illustrate (Figure 4a, lower panel), GCA successfully captured the fine-scale dynamics of magnitude and latency differences between real words and pseudowords.

To understand specifically how these effects vary as a function of time, we examined the interaction of lexicality with the polynomial time terms. There was a significant interaction with all four time terms, including the linear term ($\beta = 3.512$, $t = 13.53$) and quadratic term ($\beta = 3.261$, $t = 17.48$). Positive interactions with the first- and second-order time terms, indicates that compared to pseudowords, words had a shallower slope (linear term) and an earlier or less centered peak than pseudowords (quadratic term). Thus, responses to words often reached the peak of the neural response more quickly than responses to pseudowords, and typically also decreased more rapidly over the course of a trial (Figure 4a, lower panels).

Given that neural activity in the auditory ventral stream flows primarily in the posterodorsal-to-anteroventral direction over time, we included a three-way interaction between lexicality, each position term (A-P and D-V), and each time term in the model. A-P position had

significant interaction effects with all four time terms (linear term: $\beta = 0.111$, $t = 7.83$; quadratic term: $\beta = 0.010$, $t = 9.48$; quartic term: $\beta = 0.005$, $t = 11.18$; cubic term: $\beta = 0.0012$, $t = 12.33$), indicating that the steepness, peak timing, and tails of the response curves to words and pseudowords varied across the anterior-posterior axis (see relative curve shapes in Figure 4a). The interaction with dorsal-ventral position was not significant for any time term, indicating that distance from the Sylvian fissure did not impact the difference between words and pseudowords at any point in time.

In the linear term, both words and pseudowords at early time points have larger responses in posterior than anterior sites. In contrast, at late time points, both stimulus types have larger responses in anterior than posterior sites. However, at time points near the middle of the trial (centered around 750 ms after stimulus onset), words have greater responses in posterior sites, while pseudowords have greater responses at anterior sites. This indicates that the greatest distinction between the steepness of the curve in response to words and pseudowords comes at the midpoint of the trial.

These results suggest that while the overall responses in STG neural populations reflect word-pseudoword differences, the underlying computations associated with lexicality are temporally and spatially distributed. The interactions across electrodes between lexicality and time indicate that lexical status influences neural processing well before and well after the point at which it is possible, through a search of the lexical cohort, to determine whether a familiar or unfamiliar word was heard.

3.2.4. Cohort statistics over time—Cohort size, average cohort frequency, and summed cohort frequency are all dynamic properties of the mental lexicon; they change as acoustic information arrives and as predictions about a word form's identity are updated (Figure 1a-b). Therefore, we examined the interactions between these lexical statistics and the polynomial time terms to understand their varying influences over time. We restrict our discussion here to the linear and quadratic terms, as the effects of the cubic and quartic terms are subtle and are reflected primarily in the tails of the curves. Full model coefficients are reported in Supplemental Table 1.

The interaction between cohort size and the linear time term was negative ($\beta = -5.254$, $t = -14.07$). A negative cohort size by linear time term interaction indicates a steeper slope at larger cohort sizes. This pattern is expected, as larger cohort sizes are found at the beginning of a stimulus, when more lexical items fit the acoustic input. We observed the same pattern for the quadratic term ($\beta = -4.222$, $t = -15.69$). A negative cohort size by quadratic term indicates that as cohort size decreased, the neural response reached its peak more slowly. This effect can be seen in Figure 4b, where, particularly in the posterior and middle STG electrodes, responses to larger cohorts tended to peak, and then decline, more quickly than smaller cohorts, which have more centered curves.

Next, we examined the effects of the two cohort frequency metrics. Average cohort frequency had a positive interaction with the linear term ($\beta = 0.0766$, $t = 9.4$), and a negative quadratic term interaction ($\beta = -0.0467$, $t = -6.46$). These results indicate that cohorts that had lower lexical frequencies had steeper slopes, but reached their peak response more

slowly (Figure 4c). Summed cohort frequency had a significant negative interaction with the linear term ($\beta = -0.0587$, $t = -7.07$), but not the quadratic term, indicating steeper slopes at cohorts with a larger number of frequent members (Figure 4d).

As with the main effects of cohort size and cohort frequency, the direction of the interaction terms varied by location along the anterior-posterior and dorsal-ventral axes of the temporal lobe. For cohort size, these effects were primarily present along the anterior-posterior axis (linear term: $\beta = -0.191$, $t = -9.49$; quadratic term: $\beta = -0.157$, $t = -10.87$; Figure 4b). There were no significant effects of cohort size along the dorsal-ventral axis for any time terms.

For average cohort frequency, the linear ($\beta = 0.0013$, $t = 4.42$) and quadratic ($\beta = -0.002$, $t = -7.78$) terms were significant along the anterior-posterior axis (Figure 4c). Unlike cohort size, average cohort frequency did interact along the dorsal-ventral axis (linear: $\beta = -.0070$, $t = 8.89$; quadratic: $\beta = -0.0064$, $t = -8.85$). Summed cohort frequency also interacted along the A-P (linear: $\beta = -0.0015$, $t = -5.02$; quadratic: $\beta = 0.0016$, $t = 4.20$; Figure 4d) and D-V (linear: $\beta = -0.0056$, $t = -7.18$; quadratic: $\beta = 0.0041$, $t = 5.75$) axes.

We found that the signs of the anterior-posterior interaction terms were opposite for average cohort frequency and summed cohort frequency, indicating a dissociation between the effects of the two metrics. Figures 4c and 4d illustrate the difference between these effects in three example electrodes along the A-P axis. For average cohort frequency, the least frequent cohorts had the strongest responses, latest peaks, and least steep slopes. By contrast, cohorts with a mid-range summed cohort frequency are the ones with the strongest, latest peak responses. These differences between the two metrics are particularly clear in the anterior electrode.

Taken together, these results suggest that the integration of lexical information unfolds across the temporal lobe as a word form is heard and processed. The number of lexical competitors (cohort size) appears to be exclusively encoded in the anterior-posterior dimension. Both average and summed cohort frequency information vary along both the anterior-posterior and dorsal-ventral axes. The anterior-posterior dimension also appears to encode differences between summed and average cohort frequency across time.

4. Discussion

We used high-resolution direct intracranial recordings to examine how the ventral stream for speech processes lexical information in both familiar and novel word forms. We found that neural activity recorded from the temporal lobe while participants listened to words and pseudowords reflects differences between these broad categories with relatively complex temporal dynamics. Further, responses to spoken stimuli were modulated by language-level features – cohort size, average cohort frequency, and summed cohort frequency - that reflect the structure and organization of the mental lexicon. Finally, all of these variables modulated neural activity on specific time scales, suggesting that multiple lexical features are integrated with bottom-up acoustic-phonetic input throughout the course of hearing a spoken stimulus.

Generally, our results are consistent with the dominant theoretical accounts of speech processing, which posit a hierarchical organization of spoken language representations along

the temporal lobe (Hickok and Poeppel, 2007; Scott & Wise, 2004). The temporal progression of peak latency high-gamma activity in our data suggests a flow of information processing from more posterior-dorsal to more anterior-ventral temporal lobe sites. Existing theories of the neural basis of speech processing typically argue that this progression reflects distinct, but partially overlapping representations, including acoustic, phonetic, phonemic, and lexical sub-processes (Scott et al. 2000; DeWitt & Rauschecker 2012; Gow, Keller, Eskander, Meng, & Cash, 2009). Our data add a new level of specificity to describe the flow of information as not simply being a lexical search or phonemic concatenation process, but rather a temporally non-linear lexical retrieval process. These results suggest that words are not stored and accessed directly as information propagates through the temporal lobe, but rather are integrated with multiple levels of lexical and sub-lexical knowledge.

We hypothesized that activity along the auditory ventral stream additionally reflects the online integration of high-order lexical knowledge, in this case in the form of cohort statistics. Our results provide novel evidence that the brain uses information about the other words in the mental lexicon that are consistent with the input at a given time point to constrain and focus the lexical recognition process. Established theories of spoken word recognition generally claim that areas downstream of primary and secondary auditory cortex are primarily responsible for lexical processing, and in fact store word form and associated information (Gow, 2012; Davis & Gaskell, 2009; Hickok & Poeppel, 2007). Our results, by contrast, suggest a different process for how lexical information is represented during online speech processing. Incoming acoustic input is not simply sequentially matched with stored templates, but rather is processed in the context of deeply learned knowledge about other words in the language that share important lexical features (Toscano, Anderson, & McMurray, 2013). This provides empirical information about how spatially and temporally distributed lexical representations (Marslen-Wilson, 1997; McClelland & Elman, 1986) are stored and accessed rapidly during speech perception, often dependent on the context in which they are heard (Leonard & Chang, 2014).

Our data suggest that these same processes of lexical integration are at work when novel forms are encountered. Both familiar words and pseudowords showed interactions with temporal features, although not always in the same manner or direction, supporting the idea that the language comprehension system treats pseudowords as potentially valid word forms (De Vaan et al. 2007). Given our findings, it is unlikely that the brain is primarily concerned with binary distinctions such as real words versus pseudowords. It is possible that in many contexts, pseudowords may be treated simply as very low frequency real words, suggesting that a frequency gradient drives most lexicality effects reported in the literature (Prabhakaran, Blumstein, Myers, Hutchinson, & Britton, 2006). Thus, the salient distinctions between real words and pseudowords may be based on language-level statistics and distributions of lexical information.

We hypothesized that three critical lexical features are cohort size (Marslen-Wilson 1987, 1989; Magnuson et al. 2007), average cohort frequency (Luce & Large 2001), and summed cohort frequency, which describe target words in relation to other words that share similar features. We found these lexical statistics modulate neural activity, suggesting that temporal lobe responses are being mediated on-line by information from learned lexical information.

These statistics interacted with both lexicality and time, consistent with unfamiliar word forms being processed as potential known words using information about their cohorts. Thus, the process of spoken word recognition is not simply a bottom-up analysis of acoustic signals, and lexical processing is not spatially or temporally restricted to occur only after lower-level processing is complete. Rather, lexical recognition is also influenced by statistical information from the lexicon, helping to constrain the space of possible forms matching the acoustic input. This occurs even at posterior temporal lobe sites typically attributed to phonetic or phonological processing, suggesting that lexical competition and selection are encoded along the pathway of the auditory ventral stream, and do not exclusively rely on connections from other brain regions.

The exact nature of the relationship between cohort size, summed cohort frequency, and average cohort frequency remains an open question. We found evidence for dissociation between these metrics, but we do not have a complete understanding of how these statistics contribute to the process of lexical competition and selection. One possibility is that the processing of spoken words along the ventral stream involves multiple simultaneous quantifications of lexical competitors. One process may be concerned with the number of competitors and their average lexical frequency, while another may track a summation of the number and frequency of competitors. Future work is needed to understand the independent contribution of each metric in more detail, and how each of these lexical statistics contributes to the rapid and largely automatic process of spoken word recognition.

This investigation of lexical comprehension did not exhaustively examine every possible lexical statistic, and there is more to be learned about the ways that statistical properties of language shape ventral stream representations during spoken word recognition. In particular, phonotactic probability and neighborhood density (Vitevitch & Luce, 1999; Pylkkänen, Stringfellow, & Marantz, 2002) are two properties that are known to influence lexical processing. While we did not find an effect of phonotactics in this data set, which utilized a selective set of stimuli, it is known that phonotactic probability mediates the neural processing of acoustic input and stored lexical knowledge (Leonard et al., 2015). Stimulus sets that carefully control for phonotactics could shed light on how sub-lexical statistics interact with the specific types of lexical statistics we describe here. Similarly, stimuli that systematically vary neighborhood density would provide more information about the influence of lexical competitors that have a very similar form to the target word. Our four-syllable stimuli had very few neighbors, and they also had relatively low lexical frequencies. Future work that explicitly leverages stimulus variety (high and low phonotactics, long and short forms, highly frequent and infrequent words, in dense and sparse neighborhoods) would provide a valuable test of how the cohort statistics described here interact with other metrics of lexical and sub-lexical competition. It will be particularly useful to learn if more common forms engage a different processing pathway than the one we describe here, or whether differences are primarily reflected in the magnitude of the response within overlapping ventral stream regions. Finely-controlled stimulus sets could also take advantage of decoding analyses (King and Dehaene 2014) by explicitly quantifying the degree of confusion between competing lexical forms which share particular lexical or sub-lexical properties at a given point in time.

Human intracranial studies necessarily involve subject populations that have abnormal neurological conditions (intractable epilepsy). Furthermore, due to the rarity of these recordings, most ECoG studies have small numbers of participants. For these reasons, the data must be interpreted in the context of other non-invasive recording modalities that potentially provide a broader and more general picture of neural activity. However, the unique qualities of intracranial recordings make these data an important complement to more traditional neurolinguistic studies, in that they provide simultaneously fine spatial and temporal resolution that is unattainable with non-invasive recording techniques. In general, intracranial methods agree with non-invasive data (e.g. McDonald et al., 2010), and therefore represent an important contribution to studies of neural processing of language.

We also used an analysis technique that overcomes several limitations of traditional methods for time-series data. The integration of spatial and temporal information was critical to describing response patterns to these stimuli. In the temporal dimension, a single electrode may show a stronger response to one stimulus type early in a trial, and the other later in a trial. In terms of space, two electrodes at different sites can show a preference for either words or pseudowords, even if their relative timing is similar. Collapsing across either the spatial or temporal dimension could potentially mask these effects. Here, GCA (Mirman et al., 2008; Magnuson et al., 2007) allowed us to model neural data and lexical variables explicitly in terms of how they change over time on a millisecond level. The fine-scale dynamics we observed suggest that as speech input is processed in real time, numerous sub-lexical and lexical features are contributing to the activity in the temporal lobe speech network.

Taken together, our results suggest that the language comprehension system is a complex pathway in which the response to acoustic input is influenced by the statistical regularities of the listener's prior linguistic experience. This pathway is flexible enough to engage in similar processing strategies for familiar words and novel forms; the effect of prior experience with word forms influences the processing of both types of stimuli. The fact that both types of stimuli are processed in the context of learned lexical statistics suggests that word understanding is the result of a continuous and highly interconnected network of neural computations, rather than a modular system with discrete spatial and temporal elements.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

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

Brysbaert M, New B. Moving beyond Ku era and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behavior Research Methods. 2009; 41(4):977–990. [PubMed: 19897807]

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. Nature Neuroscience. 2010; 13(11):1428–1432. [PubMed: 20890293]

Crone NE, Boatman D, Gordon B, Hao L. Induced electrocorticographic gamma activity during auditory perception. Clinical Neurophysiology. 2001; 112(4):565–582. [PubMed: 11275528]

Crone NE, Sinai A, Korzeniewska A. High-frequency gamma oscillations and human brain mapping with electrocorticography. Progress in Brain Research. 2006; 159:275–295. [PubMed: 17071238]

Davis MH, Gaskell MG. A complementary systems account of word learning: neural and behavioural evidence. Philosophical Transactions of the Royal Society B: Biological Sciences. 2009; 364(1536): 3773–3800.

De Vaan L, Schreuder R, Baayen RH. Regular morphologically complex neologisms leave detectable traces in the mental lexicon. The Mental Lexicon. 2007; 2(1):1–24.

DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream. Proceedings of the National Academy of Sciences. 2012; 109(8):E505–E514.

Dronkers NF, Wilkins DP. Lesion analysis of the brain areas involved in language comprehension. Cognition. 2004; 92(1):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. Journal of Neurophysiology. 2005; 94(6):4269–4280. [PubMed: 16093343]

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

Garofalo, JS.; Lamel, LF.; Fisher, WM.; Fiscus, JG.; Pallett, DS.; Dahlgren, NL. The DARPA TIMIT acoustic-phonetic continuous speech corpus LDC93S1. Web Download. Linguistic Data Consortium; Philadelphia: 1993.

Gow DW Jr, Keller CJ, Eskandar E, Meng N, Cash SS. Parallel versus serial processing dependencies in the perisylvian speech network: a Granger analysis of intracranial EEG data. Brain and Language. 2009; 110(1):43–48. [PubMed: 19356793]

Gow DW Jr. The cortical organization of lexical knowledge: a dual lexicon model of spoken language processing. Brain and Language. 2012; 121(3):273–288. [PubMed: 22498237]

Heald SL, Nusbaum HC. Speech perception as an active cognitive process. Frontiers in Systems Neuroscience. 2014; 8:1–5. [PubMed: 24478639]

Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews Neuroscience. 2007; 8(5):393–402. [PubMed: 17431404]

Kemps RJ, Wurm LH, Ernestus M, Schreuder R, Baayen H. Prosodic cues for morphological complexity in Dutch and English. Language and Cognitive Processes. 2005; 20(1-2):43–73.

King JR, Dehaene S. Characterizing the dynamics of mental representations: the temporal generalization method. Trends in Cognitive Sciences. 2014; 18(4):203–210. [PubMed: 24593982]

Leonard MK, K.E. Tang C, Chang EF. Dynamic encoding of speech sequence probability in human temporal cortex. The Journal of Neuroscience. 2015 doi:10.1523/jneurosci.4100-14.2015.

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

Lerner, Yulia; Honey, Christopher J.; Silbert, Lauren J.; Hasson, Uri. Topographic Mapping of a Hierarchy of Temporal Receptive Windows Using a Narrated Story. The Journal of Neuroscience. 31(2011):2906–2915. [PubMed: 21414912]

Lindsay S, Sedin LM, Gaskell MG. Acquiring novel words and their past tenses: Evidence from lexical effects on phonetic categorisation. Journal of Memory and Language. 2012; 66(1):210–225.

Luce PA, Large NR. Phonotactics, density, and entropy in spoken word recognition. Language and Cognitive Processes. 2001; 16(5-6):565–581.

Magnuson JS, Dixon JA, Tanenhaus MK, Aslin RN. The dynamics of lexical competition during spoken word recognition. Cognitive Science. 2007; 31(1):133–156. [PubMed: 21635290]

Mainy N, Jung J, Baciu M, Kahane P, Schoendorff B, Minotti L, Hoffmann D, Bertrand O, Lachaux JP. Cortical dynamics of word recognition. Human Brain Mapping. 2008; 29(11):1215–1230. [PubMed: 17712785]

Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition. 1987; 25(1):71–102. [PubMed: 3581730]

Marslen-Wilson, WD. Access and integration: Projecting sound onto meaning. In: Marslen-Wilson, WD., editor. Lexical representation and process. MIT Press; Cambridge, MA: 1989.

Marslen-Wilson W, Welsh A. Processing interactions and lexical access during word recognition in continuous speech. Cognitive Psychology. 1978; 10:29–63.

McClelland JL, Elman JL. The TRACE model of speech perception. CognitivePsychology. 1986; 18(1):1–86.

McDonald CR, Thesen T, Carlson C, Blumberg M, Girard HM, Trongnetrpunya A, Sherfey JS, Devinsky O, Kuzniecky R, Dolye WK, Cash SS, Leonard MK, Halger DJ Jr. Dale AM, Halgren E. Multimodal imaging of repetition priming: using fMRI, MEG, and intracranial EEG to reveal spatiotemporal profiles of word processing. Neuroimage. 2010; 53(2):707–717. [PubMed: 20620212]

Mechelli A, Gorno-Tempini M, Price C. Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. Journal of Cognitive Neuroscience. 2003; 15(2): 260–271. [PubMed: 12676063]

Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. Nature. 2012; 485(7397):233–236. [PubMed: 22522927]

Meunier F, Longtin CM. Morphological decomposition and semantic integration in word processing. Journal of Memory and Language. 2007; 56(4):457–471.

Mirman D, Dixon JA, Magnuson JS. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. Journal of Memory and Language. 2008; 59(4):475–494. [PubMed: 19060958]

Norris D. Shortlist: A connectionist model of continuous speech recognition. Cognition. 1994; 52(3): 189–234.

Prabhakaran R, Blumstein SE, Myers EB, Hutchison E, Britton B. An event-related fMRI investigation of phonological–lexical competition. Neuropsychologia. 2006; 44(12):2209–2221. [PubMed: 16842827]

Pylkkänen L, Stringfellow A, Marantz A. Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. Brain and Language. 2002; 81(1):666–678. [PubMed: 12081430]

Raettig T, Kotz SA. Auditory processing of different types of pseudo-words: an event-related fMRI study. Neuroimage. 2008; 39(3):1420–1428. [PubMed: 17981055]

Scott SK, Blank CC, Rosen S, Wise RJ. Identification of a pathway for intelligible speech in the left temporal lobe. Brain. 2000; 123(12):2400–2406. [PubMed: 11099443]

Scott SK, Wise RJ. The functional neuroanatomy of prelexical processing in speech perception. Cognition. 2004; 92(1):13–45. [PubMed: 15037125]

Tanji K, Suzuki K, Delorme A, Shamoto H, Nakasato N. High-frequency γ-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. The Journal of Neuroscience. 2005; 25(13):3287–3293. [PubMed: 15800183]

Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network: Computation in Neural Systems. 2001; 12(3):289–316.

Toscano JC, Anderson ND, McMurray B. Reconsidering the role of temporal order in spoken word recognition. Psychonomic Bulletin & Review. 2013; 20(5):981–987. [PubMed: 23456328]

Towle VL, Yoon HA, Castelle M, Edgar JC, Biassou NM, Frim DM, Spire J-P, Kohrman MH. ECoG gamma activity during a language task: differentiating expressive and receptive speech areas. Brain. 2008; 131(8):2013–2027. [PubMed: 18669510]

Turken U, Dronkers NF. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. Frontiers in Systems Neuroscience. 2011; 5:1. [PubMed: 21347218]

Vaden KI, Halpin HR, Hickok GS. Irvine Phonotactic Online Dictionary, Version 2.0 [Data file]. 2009

Weide, R. The CMU Pronouncing Dictionary, release 0.7a. Carnegie Mellon University [Data file]; 2007.

Yuan J, Liberman M. Speaker identification on the SCOTUS corpus. Journal of the Acoustical Society of America. 2008; 123(5):3878.

Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. Journal of Memory and Language. 1999; 40(3):374–408.
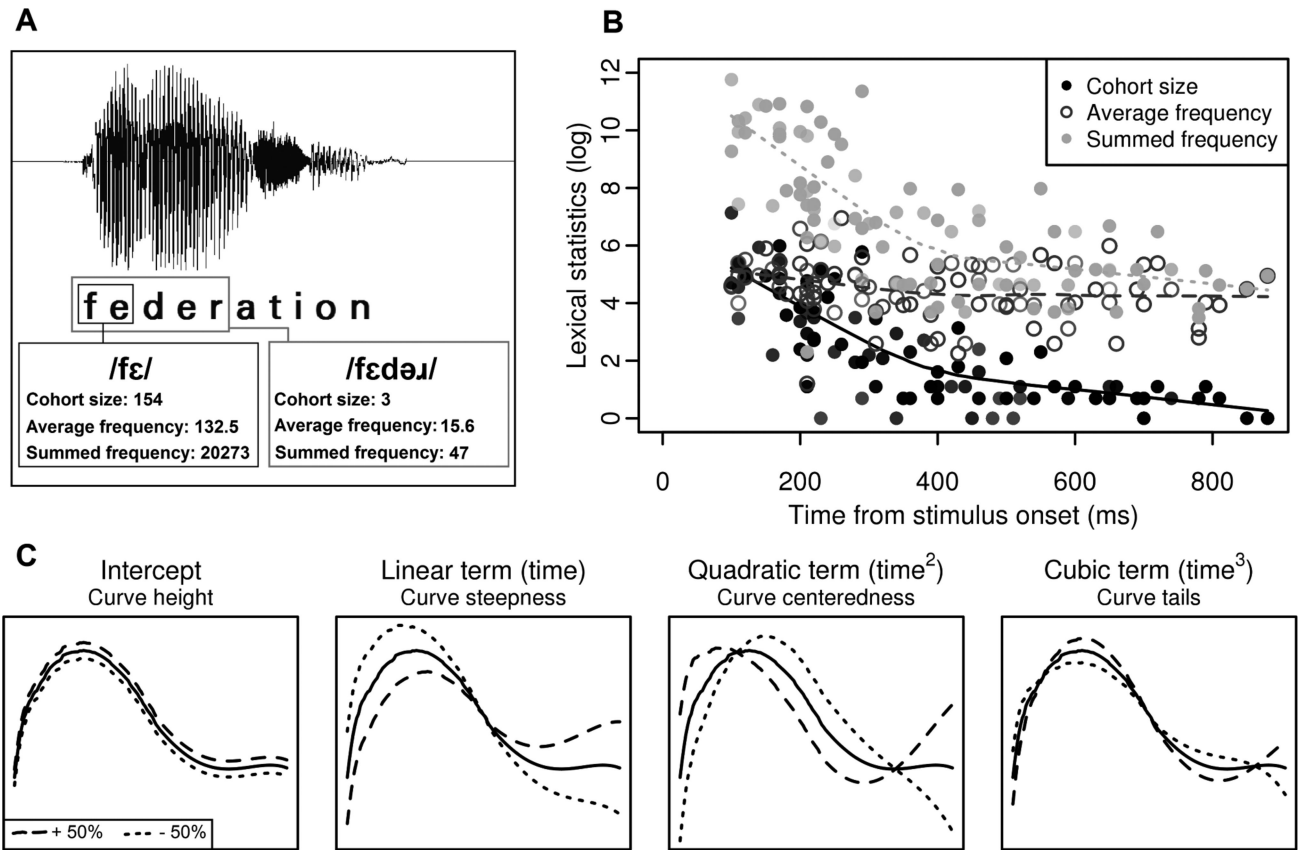
Figure 1. Lexical statistics and neural data analysis approach

(a) Cohort size and cohort frequency for the stimulus "federation" change as more of the acoustic waveform is heard. At the offset of the second phoneme, many word forms (e.g. "feral", "felt", "federation", "fend", "felon") are possible matches to the acoustic input. As a listener hears each successive segment in the stimulus, the number of possible English words shrinks, and the average lexical frequency of that set of words changes. (b) Lexical statistics for all stimuli across their duration. Each point represents a segment in a stimulus and the corresponding lexical statistics associated with that stimulus at that time in the duration of that stimulus. (c) To examine how these lexical statistics are encoded in the temporal lobe, growth curve analysis (GCA) was used to model interactions between lexical statistics and time. The black curve on each plot shows an example of the predicted neural response in a single electrode, as generated by a GCA model. To illustrate the effect that each polynomial time term has on the predicted curve generated by the model, the dashed and dotted lines show simulated 50% positive and negative adjustments of each interaction term, holding all others constant. Each term's independent contribution to the prediction affects either the height of the curve, steepness of the curve, centeredness of the curve, or shape of the curve at its tails.
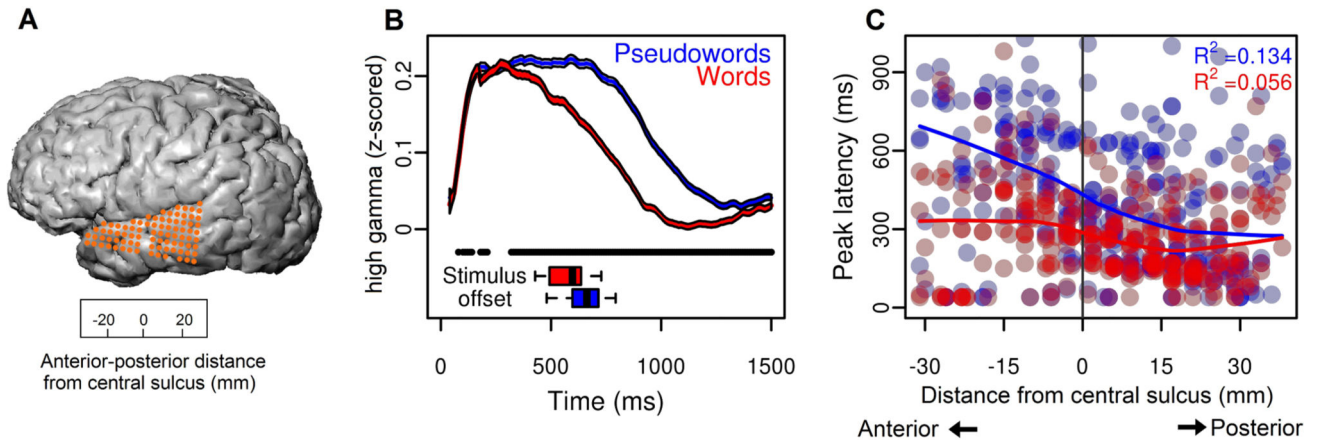
Figure 2. Spatiotemporal patterns of neural activity across electrodes
(a) Electrode coverage for subject 1. The spatial scale of the grid is indicated in the lower box. Distance is measured along an anterior-posterior line, along the Sylvian fissure, with anatomical reference point of the ventral-most aspect of the central sulcus (position=0). (b) Average high-gamma analytic amplitude (AA) response to words and pseudowords across all electrodes (combined across all subjects). Neural responses to words (red) and pseudowords (blue; shading indicates 95% confidence intervals around the mean at each time point) show prolonged differences, with generally greater responses to pseudowords (black bar; bootstrap $p < 0.05$. Lower box plots show stimulus length for words and pseudowords, which are not significantly different ($p = 0.148$). (c) Timing of the peak of neural responses to words and pseudowords in all electrodes, in all subjects, across the anterior-posterior span of the temporal lobe.
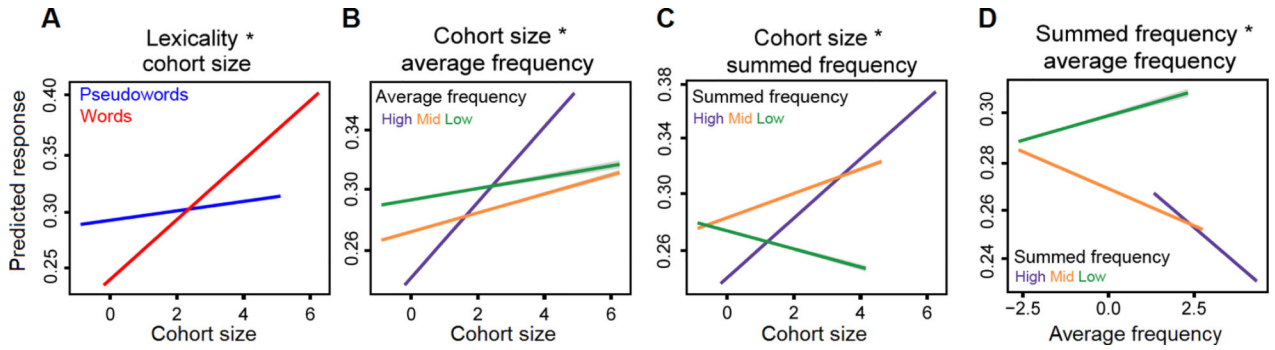
Figure 3. Interactions of linguistic predictors

Predicted values for significant ($t > 2$) interactions of linguistic predictors. In all figures, grey error bars indicate 99% confidence intervals. (a) Cohort size significant interacted with lexicality. There is a strong decrease in the predicted neural response to words as cohort size shrinks. This effect is relatively attenuated for pseudowords. (b) Cohort size and average cohort frequency also interacted. The effect of decreasing cohort size on the neural response was strongest in the most frequent cohorts. (c) Cohort size and summed cohort frequency had a different interaction pattern compared to average cohort frequency. For the highest and mid-range summed frequencies, the predicted response decreased as cohort size decreased. This effect was reversed for cohorts with the smallest summed cohort frequencies. (d) Average and summed cohort frequency interacted, suggesting a dissociation between these metrics. Responses in high and mid summed cohorts increased as average cohort frequency decreased. This effect was reversed for low summed cohorts.
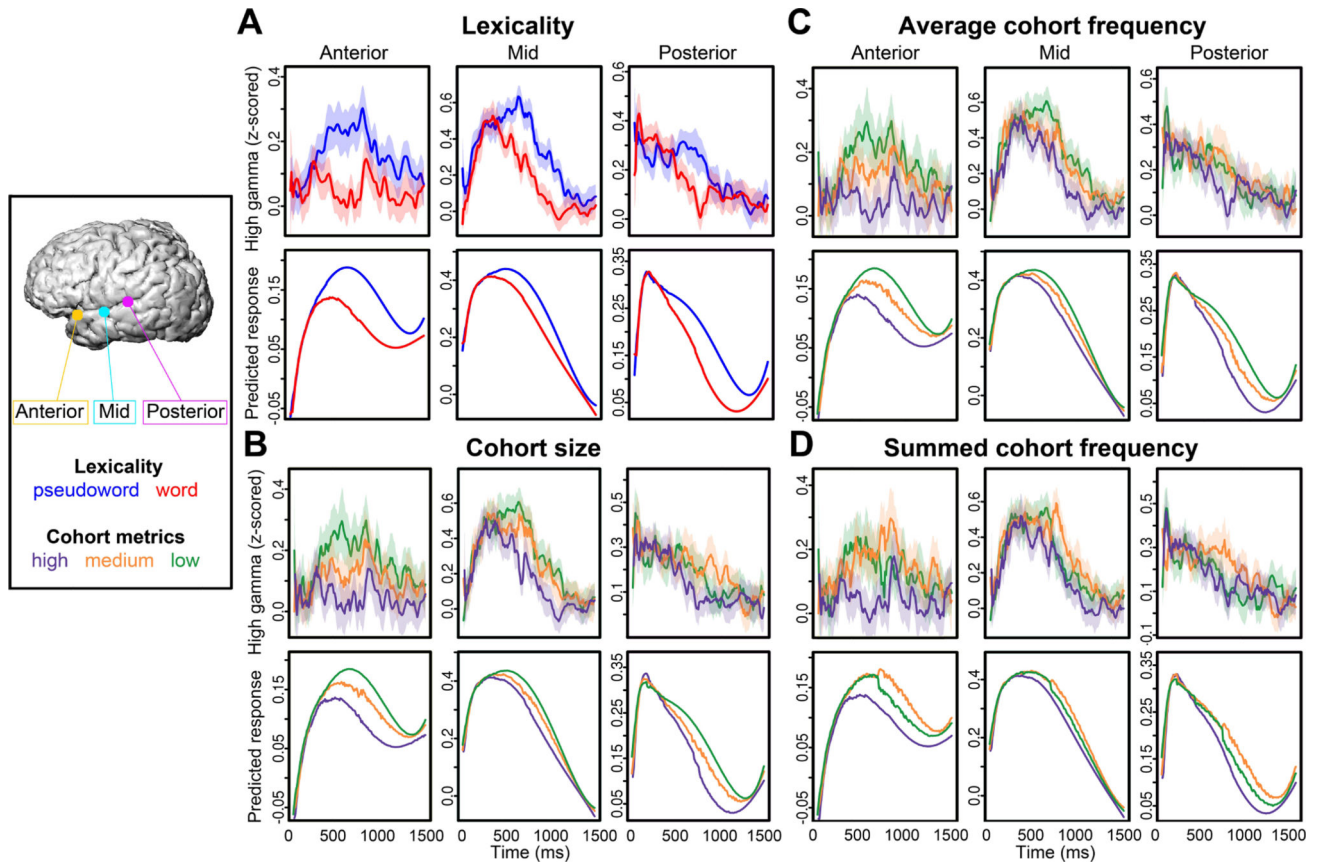
Figure 4. Linguistic variables show different effects on neural activity across time

(a) Actual neural responses (upper plot) to words (red) and pseudowords (blue) and GCA predicted responses (lower plot) for three single electrodes show that the GCA model captures the temporal dynamics of neural activity. These responses to words and pseudowords (with 95% confidence intervals) over time from anterior, middle, and posterior temporal lobe electrodes illustrate different lexicality differences across the anterior-posterior axis of the temporal lobe. Critically, the difference between words and pseudowords varies as a function of both time and location. GCA model predicted responses for the same electrodes show similar effects, demonstrating the efficacy of this statistical technique for ECoG data. (b) Actual (upper) and predicted (lower) responses to cohort size in the same three electrodes. Cohort size is binned into high, mid, and low within each time point to show the predictor's effect on the shape of the response. The smallest cohorts had the biggest responses, particularly at anterior sites. (c) Responses and predictions for three levels of average cohort frequency in the same electrodes. Responses to cohorts with low average frequency had the strongest responses. (d) Responses and predictions for summed cohort frequency; in contrast to average cohort frequency, cohorts with mid-range summed frequency had the highest responses.

Table 1

Stimuli used in the study.

Words	Pseudowords	
ceremony	moanaserry	[moʊnəsɛɹi]
delivery	lerivedy	[ləɾɪvədi]
federation	reifadetion	[ɹeɪfədɛʃən]
majority	jatoremee	[d͡ʒətɑɹəmi]
minority	tomeereneye	[təmiɹənai]
motivation	veimatoshen	[veɪmətoʊʃən]
repetition	piteretion	[pɪtɚɛʃən]
solitary	letarossy	[lɛɾɚɑsi]
velocity	tesolivy	[təsɑləvi]
voluntary	reventally	[ɹɛvəntɑli]