

Dynamic Structure of Neural Variability in the Cortical Representation of Speech Sounds

 Benjamin K. Dichter,^{1,2,3} Kristofer E. Bouchard,^{1,2,4,5} and Edward F. Chang^{1,2,3,4,6}

¹Departments of Neurological Surgery and Physiology, University of California–San Francisco, San Francisco, California 94143-0112, ²Center for Integrative Neuroscience, University of California–San Francisco, San Francisco, California 94158, ³University of California–Berkeley and University of California–San Francisco Joint Program in Bioengineering, Berkeley, California 94720-3370, ⁴Center for Neural Engineering and Prosthesis, University of California–San Francisco and University of California–Berkeley, Berkeley, California 94720-3370, ⁵Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, and ⁶University of California–San Francisco Epilepsy Center, San Francisco, California 94143

Accurate sensory discrimination is commonly believed to require precise representations in the nervous system; however, neural stimulus responses can be highly variable, even to identical stimuli. Recent studies suggest that cortical response variability decreases during stimulus processing, but the implications of such effects on stimulus discrimination are unclear. To address this, we examined electrocorticographic cortical field potential recordings from the human nonprimary auditory cortex (superior temporal gyrus) while subjects listened to speech syllables. Compared with a prestimulus baseline, activation variability decreased upon stimulus onset, similar to findings from microelectrode recordings in animal studies. We found that this decrease was simultaneous with encoding and spatially specific for those electrodes that most strongly discriminated speech sounds. We also found that variability was predominantly reduced in a correlated subspace across electrodes. We then compared signal and variability (noise) correlations and found that noise correlations reduce more for electrodes with strong signal correlations. Furthermore, we found that this decrease in variability is strongest in the high gamma band, which correlates with firing rate response. Together, these findings indicate that the structure of single-trial response variability is shaped to enhance discriminability despite non-stimulus-related noise.

Key words: ECoG; encoding; noise correlations; speech; superior temporal gyrus; variability

Significance Statement

Cortical responses can be highly variable to auditory speech sounds. Despite this, sensory perception can be remarkably stable. Here, we recorded from the human superior temporal gyrus, a high-order auditory cortex, and studied the changes in the cortical representation of speech stimuli across multiple repetitions. We found that neural variability is reduced upon stimulus onset across electrodes that encode speech sounds.

Introduction

The human superior temporal gyrus (STG) represents speech sounds with spatiotemporal patterns of neural activity across populations that are tuned to specific acoustic features of the

sounds (Formisano et al., 2008; Chang et al., 2010; Obleser et al., 2010; Steinschneider, 2011; Mesgarani et al., 2014; Nourski et al., 2014). However, neural responses to sensory stimuli are variable, and the brain responds differently to the same stimulus each time it is encountered (Faisal et al., 2008). Speech perception fundamentally involves classifying instances of sounds as members of specific linguistic categories (e.g., phonemes, words, etc.) (Lieberman et al., 1957, 1967; Perkell and Klatt, 1986; Diehl et al., 2004), although the acoustics of these sounds can vary in pitch, location, intensity, etc. The classification problem is compounded by the presence of different neural responses to physically identical sounds (Kisley and Gerstein, 1999). However, despite this variability, human listeners perceive speech effortlessly. This reliability in sensory perception despite variability in the neural response holds across sensory domains: visual (Schiller et al., 1976; Heggelund and Albus, 1978; Rose, 1979; Churchland et al., 2010), somatosensory (Whitsel et al., 1977), and as such, understanding

Received Jan. 10, 2016; revised May 25, 2016; accepted June 2, 2016.

Author contributions: K.E.B. and E.F.C. performed research; B.K.D. and K.E.B. analyzed data; B.K.D., K.E.B., and E.F.C. wrote the paper.

This work was supported by the National Science Foundation Graduate Research Fellowship Grant 1144247, National Institutes of Health Grants R00-NS065120, DP2-OD00862, and R01-DC012379, and the Ester A. and Joseph Klingenstein Foundation. We thank the NVIDIA Corporation for donation of the Tesla K40 GPU used for this research; Connie Cheung for assistance in collecting data; and Matthew Leonard and Joseph Makin for providing helpful feedback on the manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Edward F. Chang, University of California–San Francisco Mission Bay, Sandler Neuroscience Building, 675 Nelson Rising Lane, San Francisco, CA 94143. E-mail: edward.chang@ucsf.edu.

DOI:10.1523/JNEUROSCI.0156-16.2016

Copyright © 2016 the authors 0270-6474/16/367453-11\$15.00/0

neural variability is essential to understanding neural representations in general (Averbeck et al., 2006; Churchland et al., 2011).

Recent studies have identified a variety of factors that modulate trial-to-trial neural response variability in single-neuron firing rate. For example, the variability of neural responses to sensory stimuli is modulated by attentional state (Mitchell et al., 2009; Downer et al., 2015), and the difference in variability accounts for a large change in discriminability of stimuli (Cohen and Maunsell, 2009). Furthermore, neural response variability changes dynamically during the time course of stimulus presentation, with a reduction time-locked to stimulus onset (Cohen and Maunsell, 2009; Churchland et al., 2010). The potential for variability in both single neuron and population activity to hinder perception suggests that its modulation plays an important role in neural signal processing (Shadlen and Newsome, 1998; Abbott and Dayan, 1999; Churchland et al., 2011; Hu et al., 2014; Moreno-Bote et al., 2014).

It is unclear how reduction in neural variability affects stimulus representation and discriminability at the mesoscale of aggregate neural populations (i.e., field potentials). With a few exceptions (He and Zempel, 2013), most of the literature on the dynamics of neural response variability focuses on individual or multiple single-unit recordings, in which at most a few hundred neurons are simultaneously observed, a small subset of those neurons active in the sensory task. Previous studies have found strong encoding of acoustic-phonetic features in the STG using high-density electrocorticography (ECoG) (Mesgarani et al., 2014), where each electrode records from populations several orders of magnitude greater than those observed in multineuron recordings (Chang, 2015). It is unclear how variability dynamics found in single and multiunit recordings extend to these larger neural populations, and how this affects sensory processing.

To address these questions, we recorded cortical field potentials using ECoG from human STG while subjects listened to simple speech sounds. We first determined whether cortical field potentials shared the trends found in firing rate responses: that the variance is correlated with the mean of the response (Tolhurst et al., 1983; Vogels et al., 1989) and that variability decreases after stimulus onset. Furthermore, we explored the relationship between variability and stimulus encoding, testing the hypothesis that the shape of the changes in variability of population neural responses depends on cortical sound representations.

Materials and Methods

The experimental protocol was approved by Human Research Protection Program at the University of California–San Francisco.

Subjects and experimental task. Three native English-speaking human participants (one female) underwent implantation of a high-density, subdural ECoG array as part of their clinically indicated neurosurgical treatment for epilepsy. Participants gave their written informed consent before the day of surgery. Implanted ECoG grids were each 256-channel grids with 2.3-mm-diameter electrodes at 4 mm center-to-center spacing and were placed in the language dominant hemisphere in each patient (as determined with the Wada carotid intra-arterial amobarbital injection), which was left in two subjects and right in one subject. Each participant listened to a recording of consonant-vowel (CV) syllables. Sixteen consonants (/b/, /d/, /g/, /k/, /l/, /m/, /n/, /p/, /r/, /s/, /sh/, /t/, /v/, /w/, /y/, /z/) combined with 3 vowels (/a/, /i/, /u/) were spoken by 6 speakers (3 female), resulting in 288 unique auditory stimuli total. The stimuli had a mean duration of 0.43 s and SD of 0.093 s. The interstimulus interval was jittered across trials, with a mean of 1 s and a SD of 0.15 s. Stimuli were recorded in-house and played with speakers. Each stimulus was presented between 17 and 21 times to each subject.

Anatomical location of STG. We focused our analysis on the STG, a nonprimary auditory area that responds to speech sounds. Visual exam-

ination of coregistered CT and MR scans indicate that the ECoG grid covered the spatial extent of the STG of each patient. STG electrodes were identified through inspection of this coregistration and only those electrodes were used for analysis (see Fig. 1; number of electrodes in STG: S1:51; S2:48; S3:72).

Data acquisition and signal processing. Cortical-surface field potentials were recorded referenced to scalp with a multichannel PZ2 amplifier optically connected to a RZ2 digital signal processor (Tucker-Davis Technologies). ECoG signals were acquired at 3052 Hz. The speaker signal was split and recorded in-line with the ECoG data to ensure synchronization.

The time series voltage trace of each channel was visually and quantitatively inspected for artifacts or excessive noise (typically 60 Hz line noise or movement artifacts). Recordings with many artifacts were excluded from analysis, and the signal of the remaining channels were then common average referenced for each 16-channel ECoG strip to remove electrical noise shared across electrodes. The signals of the remaining channels were bandpass filtered using Gaussian bandpass filters with logarithmically increasing center frequencies and semilogarithmically increasing bandwidths from 4 to 200 Hz. The Hilbert transform was then calculated for each band, and the analytic amplitude at 400 Hz tracked the activation in each of the filter bands. The high-gamma power was calculated by averaging the analytic amplitude across the eight bands between 70 and 150 Hz. To mitigate intersession changes in signal strength, high-gamma power was z-scored relative to the mean and SD of the recording session for each channel. Throughout, when we speak of high-gamma power, we refer to this z-scored measure, denoted as $H\gamma$.

Measuring variability. We wish to measure the time course of variability in neural activity. If we were to simply measure variance, the dynamics of the variability would be swamped by the effect of increased activity, which tends to increase following stimulus presentation and is strongly correlated with variance (Tolhurst et al., 1983). To study the effect of a stimulus on response variability, we observe how it changes the relationship between mean and variance of response. Previous studies have accounted for this relationship by using the Fano factor (FF) (Churchland et al., 2010) and varCE (Churchland et al., 2011), both of which are variability measures designed for firing rates that assume a linear relationship between mean activity and variance, and both were used to show a variability reduction following stimulus onset. We establish a similar metric, but cannot use FF or varCE because (z-scored) $H\gamma$ has nonzero variance at zero activation. To account for this difference, we modify previous methods to include a y-intercept and quantify the variability as the slope of affine regression. We quantify the relationship between variance of responses (v) and mean of responses (m) with an affine function as follows:

$$v = Dm + C \quad (1)$$

where D , the slope, is our measure of variability, and C is a y-intercept, which is discarded. Using this method, we are able to account for the effect of mean on variance. D is our analog for FF. It is not the same metric, but it serves the same purpose of quantifying the relationship between the mean and variance of neural activation.

Mean matching. To ensure that the variability reduction was not due to differences in the mean distribution, we use a “mean-matching” procedure adapted from Churchland et al. (2010). For each subject, electrode-stimuli sets were removed so that each time point has the same mean distribution of $H\gamma$ responses. The mean $H\gamma$ responses for each stimulus on each electrode are binned into 30 equally spaced bins for each of the time points in the 400 Hz $H\gamma$ signal spanning 300 ms before to 700 ms after the stimulus onset. A maximum common histogram of mean responses was then constructed, which was where all of the histograms through time overlapped. At each time point, stimulus-electrode responses were removed randomly from bins with mean $H\gamma$ in excess of the maximum common histogram until each histogram was equal to the maximum common histogram. The mean matching procedure resulted in data that had the same mean distribution through time, so differences in the regression coefficient cannot be attributed to differences in mean activity.

Temporal encoding. The relationship between acoustics and H γ was modeled by a token stimulus-encoding model, where the H γ response was modeled for each of the 288 different CV sound stimuli independently. The adjusted R^2 (\bar{R}^2) (Theil, 1961) was used to determine the degree of encoding through time as follows:

$$\bar{R}^2 = 1 - (1 - R^2) \frac{n - 1}{n - p - 1} \quad (2)$$

Where n is the sample size and p is the number of unique stimuli. \bar{R}^2 is similar to R^2 but is altered to adjust for bias in the estimation of variance due to a small sample size. \bar{R}^2 was calculated independently for each H γ measurement in time, then averaged across electrodes to obtain a trend across the STG. Qualitatively similar time courses for all three subjects were obtained with a linear discriminant classifier decoding stimulus identity from H γ across all electrodes in a subject.

Spatial variability and encoding. Electrode-wise encoding and variability changes were analyzed by a comparison of the mean of D during a prepresentation baseline (−300 to 0 ms) and a stimulus-encoding period (100–400 ms). Here, encoding strength was calculated as \bar{R}^2 but not averaged. The variance-mean regression was performed individually for each electrode and not mean-matched.

Factor analysis. Next, we studied how the decrease in variability affected correlated noise across electrodes. Factor analysis (FA) was used to separate the variability of responses into shared variance and private noise. FA is an unsupervised machine learning algorithm that models the data as being generated by a Gaussian distribution on a lower dimensional space (x), corrupted with private (uncorrelated) noise Q into the full dimensionality of the data as follows:

$$\begin{aligned} x &\sim \mathcal{N}(0, I) \\ y &\sim \mathcal{N}(Cx, Q) \end{aligned} \quad (3)$$

where x is a vector of latent variables and y is a vector of the observed H γ , Q is the noise covariance matrix and is constrained to be diagonal, and C is the loadings matrix, which maps from the latent space to the observed space, and $\mathcal{N}(\mu, \Sigma)$ denotes a multivariate normal distribution with mean μ and covariance Σ . The shared component of network variability captured by FA was calculated by $C^T C$, and a private uncorrelated component was the diagonal matrix Q . To have enough trials to robustly fit the FA model, stimuli were combined across the six CV speakers of the stimulus set. To ensure unique and informative stimulus responses, only electrodes with $\bar{R}^2 > 0.5$ during the encoding period were used for each subject, approximately half of the STG electrodes for each subject. H γ was z-scored across trials for each set of stimulus responses on each electrode, stimulus, and time point. We determined the dimensionality of the neural data by conducting principle component analysis on the average response across stimuli and the prestimulus period (Churchland et al., 2010). We found that 5, 7, and 6 dimensions were required to explain 95% of the variance in each subject, respectively. To ensure that we did not overestimate the dimensionality of the data, we modeled the data in a subspace with a dimensionality of 5, the minimum across subjects, which explained 96%, of the variance in the data for S1, 92% for S2, and 93% for S3. FA was conducted independently on each time point. The proportion of variability that was shared was calculated by $\text{mean}(\text{trace}(C^T C))$. This is a proportion because the z-score makes the total variance $\text{mean}(\text{trace}(C^T C + Q)) = 1$ across time. Qualitatively similar results were observed if no common-average reference is performed.

Relationship between signal and noise correlations. To determine whether the reduction in variability was in the directions that benefited stimulus discriminability, we compared signal and noise correlations for each pair of electrodes (Cohen and Maunsell, 2009). “Signal” is the mean H γ during the encoding period, and “noise” is the residual of this activity. Pairwise signal correlations were computed for each pair of electrodes during the encoding period using Pearson correlation. Noise correlations were calculated immediately before and 300 ms after stimulus onset. We calculated the average change in noise correlation. We also calculated the correlation between signal correlation and change in noise correlations

before and after stimulus presentation across electrode pairs for each subject.

Decoding. As the brain responds to stimuli, the mean response changes, distinguishing between different sounds. The correlation between electrodes also changes, which can have additional effects on the discriminability between sounds. To determine the extent to which noise correlations affect discriminability, we developed an analytical method that imposes the noise correlations of the prestimulus period during the stimulus response. The high gamma activity was mean-subtracted for each unique stimulus. Then the response was whitened by premultiplying by $\Sigma_{\text{encoding}}^{-\frac{1}{2}}$ and “colored” by the prestimulus noise covariance by premultiplying by $\Sigma_{\text{pre-stim}}^{\frac{1}{2}}$. The stimulus response is now warped to express the same covariance as the prestimulus activity. The warped activity therefore has prestimulus noise correlations, but also the variances and covariance determinant, which can also affect discriminability. To determine the effects solely of correlations, we premultiply by an additional diagonal matrix, K . This matrix scales the warped responses so that they match the relative variances and the determinant of the original stimulus response covariance matrix. Because K is diagonal, this does not affect the correlations. In summary,

$$\tilde{X}_{\text{warped}} = K \Sigma_{\text{pre-stim}}^{\frac{1}{2}} \Sigma_{\text{encoding}}^{-\frac{1}{2}} \tilde{X}_{\text{encoding}} \quad (4)$$

was used, where \tilde{X} are residual activity matrices, Σ are covariance matrices, and K is the diagonal matrix that scales the result to impose the relative variances and covariance determinant of the encoding responses. Finally, the means were added back, which resulted in the final warped encoding signal. We examined the response both before and after warping and trained linear classifiers on the warped and original data using linear discriminant analysis.

Analysis across frequency bands. We explored the relationship between variability decrease and stimulus representation across frequency bands using the canonical frequency bands theta (4–8 Hz), α/μ (8–13 Hz), β (13–30 Hz), gamma (30–70 Hz), and high gamma (70–150 Hz) (Mackay, 1997; Canolty et al., 2006; Crone et al., 2011). Analytic amplitude within these bands was calculated analogously to the methods used for H γ . \bar{R}^2 , D , and decoding accuracy are calculated across time for each frequency band.

Results

To understand the role of neural variability in response to speech sounds, we presented subjects with auditory playback of 288 distinct CV syllables. The stimuli consisted of 16 consonants followed by either /a/, /u/ or /i/ (cardinal vowels), spoken by six different speakers, and were chosen to sample the acoustic and phonological space of American English. We recorded neural activity directly from the surface of the STG in the language dominant hemisphere with high-density ECoG arrays. We examined the structure and dynamics of neural responses across multiple presentations of different syllables (across-syllable variability), and compared this with neural responses across multiple presentations of the same syllable (within-syllable variability) to examine how STG encodes different sounds with distinct representations.

Mean dependence of response variability

Figure 1A shows the STG electrodes for one patient (S1), with a single example electrode highlighted in yellow. Figure 1B (top and middle rows) displays the amplitude waveform and spectrogram for the syllables /di/ and /si/, the consonants of which have very different acoustic structure.

We focused on the cortical response in the high-gamma frequency component of field potentials (H γ , 70–150 Hz), which correlates well with multiunit firing rates (Rasch et al., 2008; Ray et al., 2008; Whittingstall and Logothetis, 2009; Ray and Maun-

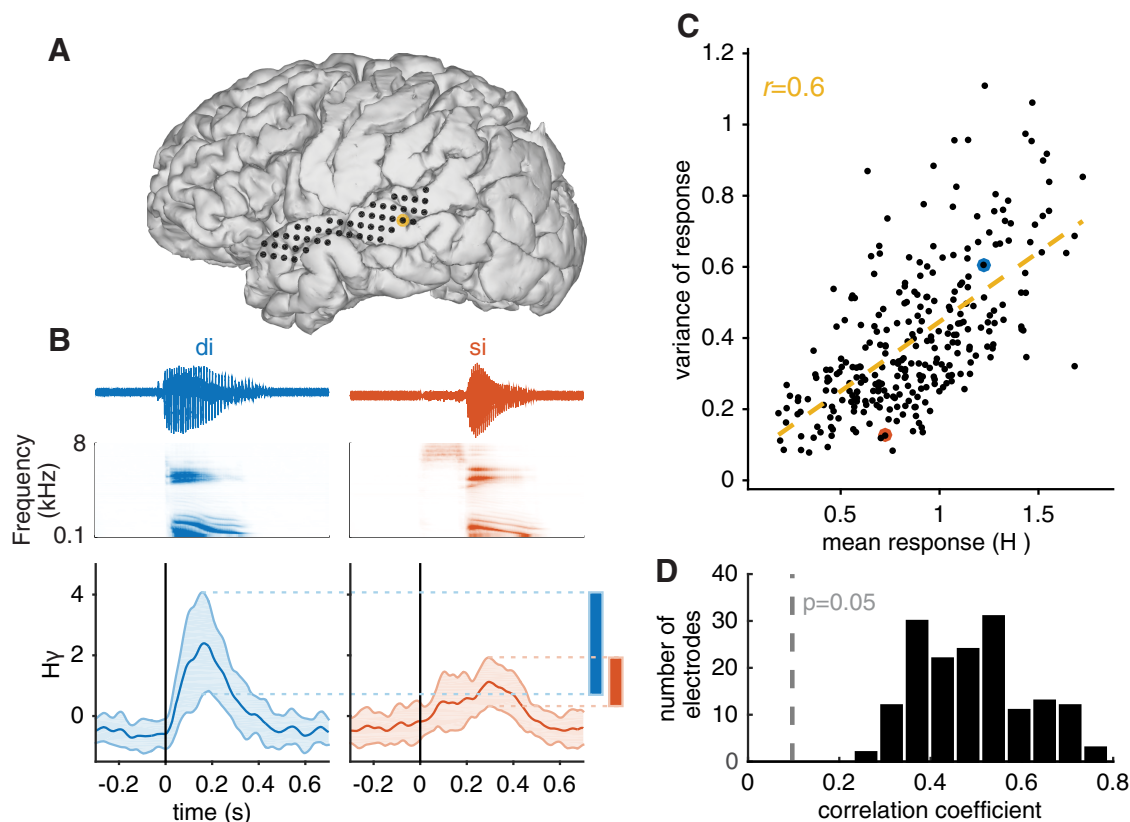


Figure 1. Example response and correlation between mean and variance of response. **A**, Placement of high-density ECoG electrodes on the cortex of an example human subject. The electrode size is anatomically correct. Yellow electrode is the example electrode used for **B** and **C**. **B**, Top, Example consonant–vowel sounds, /di/ and /si/ with acoustic waveform spectrogram. Bottom, Mean and SD of Hγ responses across task time. Rectangles to the right compare the SDs of the two response distributions and illustrate that /di/, the sound that elicits a greater mean response, also elicits a greater variance of responses. **C**, Mean and variance of Hγ over trials of every stimulus for the example electrode (100–400 ms after stimulus onset). This electrode has a positive Pearson correlation of 0.6, indicating that, for this electrode, stimuli that elicit a greater mean response tend to elicit a greater variance of response. **D**, The Pearson correlation coefficient between mean and variance of Hγ responses for each electrode across all three subjects. All electrodes have a correlation greater than chance, which is 0.01 ($p = 0.05$).

sell, 2011; Rey et al., 2014). The Hγ responses from the high-lighted electrode evoked by these syllables are displayed in Figure 1B. Responses increased shortly after the onset of both stimuli and returned to baseline by ~500 ms. In this example, /di/ caused greater mean activity than /si/, with the peak difference occurring at ~200 ms. Additionally, the response to /di/ had greater variability, as is evident from its broader error bars, which show SD and are compared side-by-side to the right. These example stimuli illustrate the general trend that the variance of Hγ responses for a stimulus was positively correlated with the mean. Figure 1C shows the mean and variance of responses across repetitions of each of the 288 unique auditory stimuli for the example electrode. We found a strong positive correlation (Pearson correlation: $r = 0.6$) between mean and variance, as illustrated by the yellow dashed line, which is the best linear fit between the two.

Similar results were observed across all STG electrodes. Figure 1D shows the correlation coefficient calculated in the same manner for all of the STG electrodes across each of the three subjects. We found statistically significant positive correlations between mean and variance of Hγ responses across all of the STG electrodes that we recorded ($r = 0.49 \pm 0.12$, mean \pm SD; $N = 179$; vertical dashed line indicates statistical significance). The positive correlation between mean and variance in Hγ is similar to the relationship between mean and variance that has been well studied in the response statistics of single neurons (Tolhurst et al., 1983), and is consistent with the hypothesis that variability of ECoG Hγ reflects variability in the activity of underlying populations of neurons.

Reduction in variability following stimulus onset

We next investigated how the relationship between mean and variance of neural responses changed through time during the representation of a stimulus. We hypothesized that this relationship would decrease, consistent with experiments (Cohen and Maunsell, 2009; Churchland et al., 2010) and models (Litwin-Kumar and Doiron, 2012) of single neuron firing. Figure 2 illustrates the dynamics of the relationship between the mean and variance of neural responses. We quantified the relationship between the mean to the variance at each time point using linear regression (Eq. 1). The slope of the regression (D) is our metric of relative variability and is similar to FF, except that it is calculated with the inclusion of a y -intercept. To illustrate this procedure and metric, five time points are shown in Figure 2A: -200, 0, 200, 400, and 600 ms. Each point of each panel in Figure 2A shows the mean and variance of Hγ responses for a stimulus on an electrode of Subject S1. In each panel, a regression was performed, and the best-fit line is shown as a red dashed line. We observed the slope of the best-fit line was ~1 before the acoustic onset of the stimulus, decreased following the acoustic onset of the stimulus, and slowly returned to 1. The regression was performed at every sample time and is shown in Figure 2B. Subject S1, the example subject for Figure 2A, is shown in red, and Subjects S2 and S3 are shown in green and blue, respectively. The slope decreased sharply within 100 ms of stimulus onset in each of the three subjects, and this reduction was sustained for 200–300 ms. During the period between 100 and 400 ms after stimulus onset, D decreased by 70%, 38%, and 42% for S1, S2, and S3, respectively.

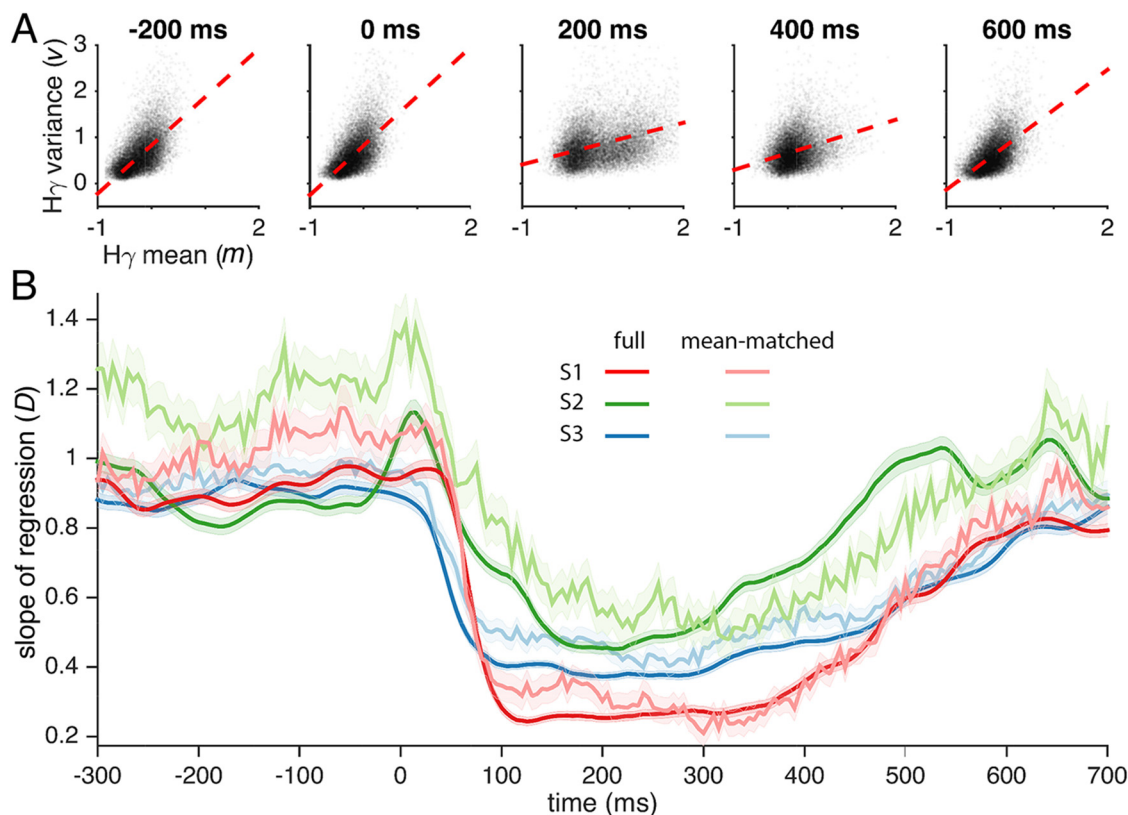


Figure 2. Reduction in mean dependence of variance. **A**, $H\gamma$ for all stimulus-electrode responses for Subject S1. Red lines indicate the result of linear regression with a y -intercept. The regression line has a slope of ~ 1 before and during stimulus onset, decreases 200 and 400 ms after stimulus onset, and returns to ~ 1 by 600 ms. **B**, Slope of regression for each point in time for Subjects S1 (red), S2 (green), and S3 (blue). Lighter shade represents results for mean-matched D , where regression was performed on a subset of stimulus electrodes with a constant mean distribution through time. D decreases sharply after acoustic onset for mean-matched and regular regression. Data are mean \pm SEM.

The decrease in relative variability is similar to the “quenching” in FF of firing rates after stimulus onset previously observed in animal studies (Cohen and Maunsell, 2009; Churchland et al., 2010).

Additionally, to ensure that the change in D was not due to a difference in the mean distribution, a version of “mean-matching,” (Churchland et al., 2007, 2010) modified for ECoG, was also performed. Here, stimulus-electrode pairs were randomly excluded from each time point until only a subset of stimulus-electrode pairs remained that have the same mean distribution through time. Although mean-matching removed approximately half of stimulus-electrode pairs at each time point (S1: 56%, S2: 66%, S3: 50%), it did not qualitatively change the trend of a sharp decrease in D (Fig. 2B, light colors). Together, these results indicate that the relative magnitude of neural variability is dynamically quenched during the presentation of speech stimuli.

Co-occurrence between stimulus encoding and reduction in variability

In order for changes in variability to affect encoding strength, these changes must coincide with the neural response that discriminates between stimuli, and must be present at electrodes that discriminate stimuli. To ascertain the role of reduced neural variability in the representation of stimuli, we examined its relationship with stimulus encoding. Specifically, we tested the hypotheses that the decrease in D temporally coincides with the encoding of a stimulus, and that electrodes which contribute most to the overall decrease in D are the electrodes that most

strongly encode the stimulus. Encoding strength was measured using the coefficient of determination adjusted for degrees of freedom, \bar{R}^2 , which quantifies how much of the variance in the response of an electrode could be explained by which stimulus was presented (Eq. 2).

We found that dynamics of stimulus encoding co-occurred spatially and temporally with the reduction in response variability. The encoding of syllables across the STG for Subject S1 is shown in Figure 3A. To determine the dynamics of encoding strength, \bar{R}^2 was calculated for each STG electrode across the trial, time aligned to the acoustic onset of the stimulus (Fig. 3A, dashed black line). Here, electrodes are ordered by position on the STG, with posterior electrodes on the top and anterior electrodes on the bottom. The ordering of the electrodes was determined by eye. All electrodes had an \bar{R}^2 near 0 before the acoustic onset of the stimulus, which indicates that their variability was not predicted by the upcoming stimulus. Following the stimulus onset, \bar{R}^2 increased first in the posterior electrodes (Hickok and Poeppel, 2000) and peaked in the anterior electrodes over the first 100 ms (DeWitt and Rauschecker, 2012). Similar spatiotemporal structures were observed in the other subjects. Figure 3A (bottom) shows the average \bar{R}^2 across all STG electrodes for each of the three subjects. For each subject, the mean \bar{R}^2 was maximum at ~ 100 ms and slowly decreased over the next 900 ms. Qualitatively, the dynamics of average \bar{R}^2 had a similar time course to the decrease in D shown in Figure 2B. Similar results were obtained with a linear discriminant classifier decoding stimulus identity from $H\gamma$.

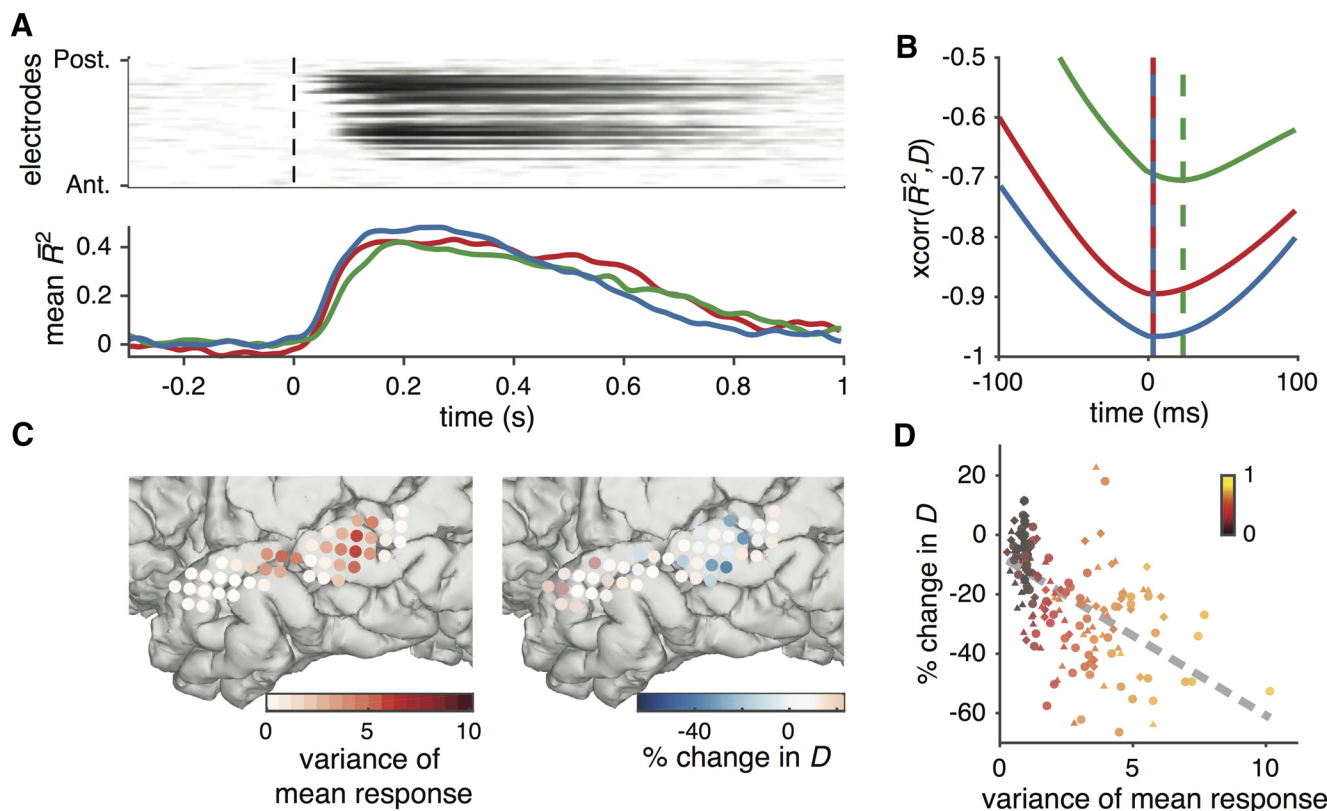


Figure 3. Temporal and spatial correlation between encoding and drop in variability. **A**, \bar{R}^2 through time for each electrode on the superior temporal gyrus (STG) of S1 with electrodes ordered from posterior to anterior region of the STG. Mean \bar{R}^2 across electrodes is shown below for S1 (red), S2 (green), and S3 (blue). **B**, Cross-correlation between variability (D) and \bar{R}^2 for the three subjects. Minimum cross-correlation is marked as dashed lines colored by subject and was close to 0 in all three subjects. **C**, Left, For S1, the variance of the mean response during the encoding period (100–400 ms) is shown for each electrode. Darker electrodes are those that have more different responses for different stimuli. Right, D was calculated for each electrode independently. Electrodes are colored according to the change in D from the average of the pre-stim period (−300 to 0 ms) to the average of the encoding period (100–400 ms). **D**, Mean response variance and change in D are shown for each electrode across all three subjects. Individual electrodes from S1 (triangles), S2 (diamonds), and S3 (circles). Electrodes are also colored according to \bar{R}^2 . Gray dashed line indicates the line of best fit. The negative correlation shows that electrodes that respond differently to different stimuli tend to have a reduced D during stimulus encoding. Data are mean \pm SEM.

To quantitatively compare the dynamics of \bar{R}^2 to D , we used a cross-correlation analysis for each subject (Fig. 3B). The time of the minimum cross-correlation represents the overall lag between decreased variability and increased encoding strength. We found that the time lags were close to 0 for each subject, demonstrating that the decrease in variability was nearly synchronous with stimulus encoding, preceding it only slightly for each subject. The large negative correlations close to -1 indicate that encoding strength and variability reduction have similar overall dynamics (S1: $r = -0.89$ at 5 ms, S2: $r = -0.70$ at 25 ms, S3: $r = -0.97$ at 5 ms). These results show that the representation of stimuli occurs simultaneously with the decrease in variability, which is necessary for the shaping of variability to enhance discriminability.

Colocalization of stimulus encoding and change in variability

To determine the degree to which different speech sounds gave rise to different responses, we calculated the variance of the mean responses to the different stimuli (across-stimulus variance), which would be high if the activity of that electrode had different responses for different acoustic stimuli. The across-stimulus variance of each electrode was calculated during the “encoding period,” 100–400 ms after stimulus onset. Figure 3C (left) shows each electrode of Subject S1 with the color of the electrode indicating the variance of the mean response (darker red electrodes had a greater across-stimulus variance). D was also calculated separately for each electrode across time, and a change in D was

determined by comparing the encoding period with the baseline period (300 ms before stimulus onset to the moment of stimulus onset). The change in variability is shown in Figure 3C (right). Electrodes that are dark blue had a strong decrease in the linear relationship between mean and variance, and electrodes that are red had a slight increase. Figure 3D summarizes the relationship between stimulus variance and variability for all three subjects. We found a robust negative correlation between interstimulus variance and D across subjects ($r = -0.56 \pm 0.09$; slope of best fit regression line = -5.4 ± 1.2 ; both are 95% CI), indicating that the electrodes with a high interstimulus variance tended to also have a greater decrease in variability. The \bar{R}^2 of H γ of each electrode during the encoding period is indicated by color. Electrodes with a high across-stimulus variance and a strong decrease in variability had the strongest encoding strength. Together with the previous analyses, these results demonstrate that the drop in variability and increase in stimulus encoding were nearly synchronous in time, and colocalized to the same electrodes.

Reduction in population variability improves discriminability

Phonetic features are encoded by a spatially distributed network of activity along the STG (Mesgarani et al., 2014). In distributed representations, noise in the activity of individual sites that is uncorrelated across the sites can be removed through averaging. However, if the variability is correlated across the different sites, averaging will not remove it, and this correlated noise can have a

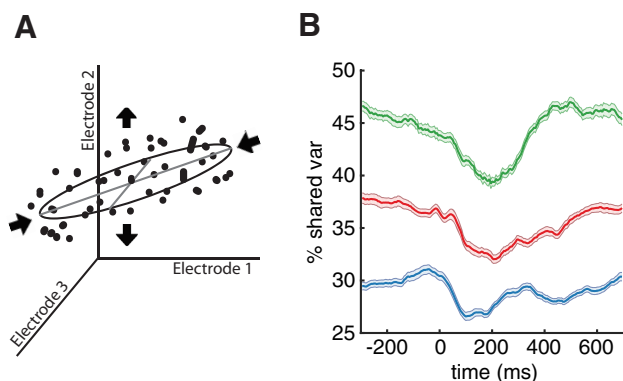


Figure 4. Factor analysis. **A**, Factor analysis schematic illustrating a change in variability from more correlated to less correlated. Shown is the activity on three example electrodes for a single stimulus. Oval represents correlated variability in a 2D subspace of the three electrodes. Dots represent individual data points after the addition of private uncorrelated noise. Arrows indicate a reduction in shared variability and an increase in private noise. The schematic represents a 2D subspace for illustrative purposes, but the data were actually modeled on a 5-dimensional subspace. **B**, Factor analysis results. The proportion of variability that was in the shared subspace is shown over time for Subjects S1 (red), S2 (green), and S3 (blue). Although the portion of variability that was shared across electrodes differs between subjects, it decreases upon stimulus onset in all three subjects. Data are mean \pm SEM.

large impact on discriminability (Zohary et al., 1994). We have observed a change in variability coincident with stimulus representation, but the effect of this variability on strength of encoding depends on the shape of the variability. To understand how the decrease in variability was distributed across all of the electrodes, we used factor analysis (Eq. 3) (Churchland et al., 2010). Factor analysis distinguishes between correlated and uncorrelated variability by modeling the data with a generative process that contains a correlated subspace that is “shared” across observations, and corrupted by uncorrelated “private” noise. We hypothesized that, during speech perception, shared variability in the STG neural population would decrease more than the uncorrelated variability, and that the dynamics would be closely matched to the dynamics of variability reduction at individual sites. This hypothesis is illustrated as a schematic in Figure 4A. Here, factor analysis finds a subspace, depicted by the ellipse that spans mostly Electrode 1 and Electrode 2. The activity is then corrupted by Gaussian noise, which extends activity mostly into the direction of Electrode 3. The arrows illustrated a reduction in shared variability, marked by a shrinking of the ellipse and a relative increase in variability in the direction of Electrode 3, the uncorrelated electrode. The resulting noise structure with less shared noise will be more spherical.

For each subject, we modeled the shared component of H γ variability with a 5-dimensional subspace (see Materials and Methods). Factor analysis was performed separately on z-scored H γ for each sample time of each consonant-vowel stimulus. The overall difference between subjects in the shared variance is explained by the number of electrodes on the STG of each subject. Subject S2 has the fewest electrodes and the most shared variance, and Subject S3 has the most electrodes and the least shared variance. Following stimulus onset, there was a pronounced decrease in the proportion of the population variance that was correlated across electrodes, and the private noise makes up a larger proportion of the variance in each of the three subjects (Fig. 4B). This shows that neural variability is primarily constrained in the dimensions that are most variable during multiple presentations of a stimulus.

Figure 5A shows a schematic of how noise correlations might affect discriminability at the example electrode in Figure 1. In the schematic, “di” causes a higher response than “si” in both electrode 1 and 2 (both 1 and 2 respond like the example electrode in Fig. 1). Because the two electrodes respond similarly to stimuli, they have positive signal correlation. The signal is corrupted by noise, indicated by the ellipses around the mean responses. The orange ellipses show noise that is positively correlated, which lowers the discriminability of the sounds. By reducing the magnitude of positive correlations, the distance between the ellipses increases. Thus, discriminability improves if the positive correlations are reduced, as shown by the blue ellipses.

To determine how changes in noise correlations affect stimulus discriminability, we examined the relationship between signal and noise correlations for each pair of electrodes before the stimulus and during representation (Cohen and Maunsell, 2009) (Fig. 5B). We find that noise correlations between electrode pairs are uniformly positive, and larger for electrodes with higher signal correlation, both of which are in agreement with neuronal firing rate correlations (Kohn and Smith, 2005; Cohen and Kohn, 2011). The majority of pairs (90%) of electrodes have a positive signal correlation. Noise correlations decreased on average during stimulus encoding for each subject (S1: $\Delta r_{\text{noise}} = -0.018 \pm 1.7e-3$; S2: $\Delta r_{\text{noise}} = -0.034 \pm 1.1e-3$; S3: $\Delta r_{\text{noise}} = -0.014 \pm 1.6e-3$; 95% CIs). Moreover, electrode pairs with positive correlations have more strongly reduced noise correlations, although this was only statistically significant in S1 and S3 (S1: $r = 0.28 \pm 0.04$; S2: $r = 0.05 \pm 0.058$; S3: $r = 0.24 \pm 0.04$; 95% CIs, $N = 1711, 1128$, and 2556). This is shown as an increasing difference between noise correlations before and during representation for higher signal correlations. These results provide another key insight into the shape of variability: the reduction in variability during stimulus representation is stronger in the directions of neural space that benefits discriminability.

Recent theoretical work has demonstrated that noise correlations alone can give a misleading impression of change in strength of representation (Moreno-Bote et al., 2014). To determine the extent to which the observed correlations change the discriminability of speech sounds, we imposed the prestimulus correlations to the neural activity during the encoding period and quantified the difference in the accuracy of a decoder trained to distinguish consonants across time. To evaluate the effect of the change in correlations on discriminability across the STG, we evaluated the change in decoding performance at 200 ms for all pairs of electrodes with $\bar{R}^2 > 0.5$ (Fig. 5C). Pairs that had a greater signal correlation tended to discriminate better with the original signal than with the warped signal, whereas electrodes that had anticorrelated responses performed better with the warped signal ($r = 0.27$; $p < 1e-4$). However, warping the signal correlations had no effect on the decode accuracy of a classifier using all electrodes. Although the noise correlations were shaped by the similarity in stimulus response, these changes in the noise did not affect decoding performance across the STG.

Variability reduction was correlated with discriminability across frequency bands

Although we have emphasized H γ , other frequency bands of the ECoG field potential may play important roles in speech processing. To explore these other bands, we compute the variability and encoding strength for the canonical frequency bands: theta (4–8 Hz), α/μ (8–13 Hz), β (13–30 Hz), gamma (30–70 Hz), and high gamma (70–150 Hz) (Mackay, 1997; Canolty et al., 2006; Crone et al., 2011). The time course

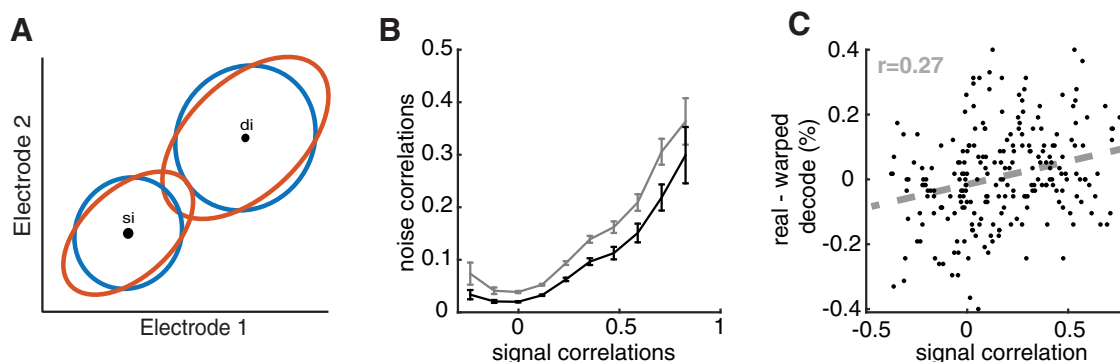


Figure 5. Noise correlations and decoding. **A**, Schematic of signal and noise correlations. The responses of two hypothetical electrodes are shown for two stimuli, “si” and “di.” Circles in the center of the ellipses represent mean responses for these stimuli. Both of these electrodes have a stronger response to “di” than to “si,” like the real example electrode in Figure 1. Orange ellipses represent response distribution with a positive noise correlation because the major axis of the ellipse has a positive slope. Blue ellipses represent response distributions with reduced noise correlations, which would improve discriminability by increasing the separation between the ellipses. **B**, Noise correlation as a function of signal correlation across all electrode pairs, for example, Subject S1: gray represents noise correlations just before stimulus onset; black represents noise correlations at 200 ms after stimulus onset. Data are mean \pm SEM. **C**, The effect of prestimulus noise correlations on stimulus discriminability of speech sounds is illustrated on two example electrodes. Leave-one-out linear discriminant analysis is used to test the separability of consonants. The true stimulus response is compared with a warped response that has noise correlations of the prestimulus period at 200 ms. The difference between decoding performance for original and warped activity for each electrode pair for Subject S1. Electrode pairs that had a greater signal correlation had a greater improvement of decoding compared with the warped signal decode (Pearson $r = 0.27$; $p < 1e-4$).

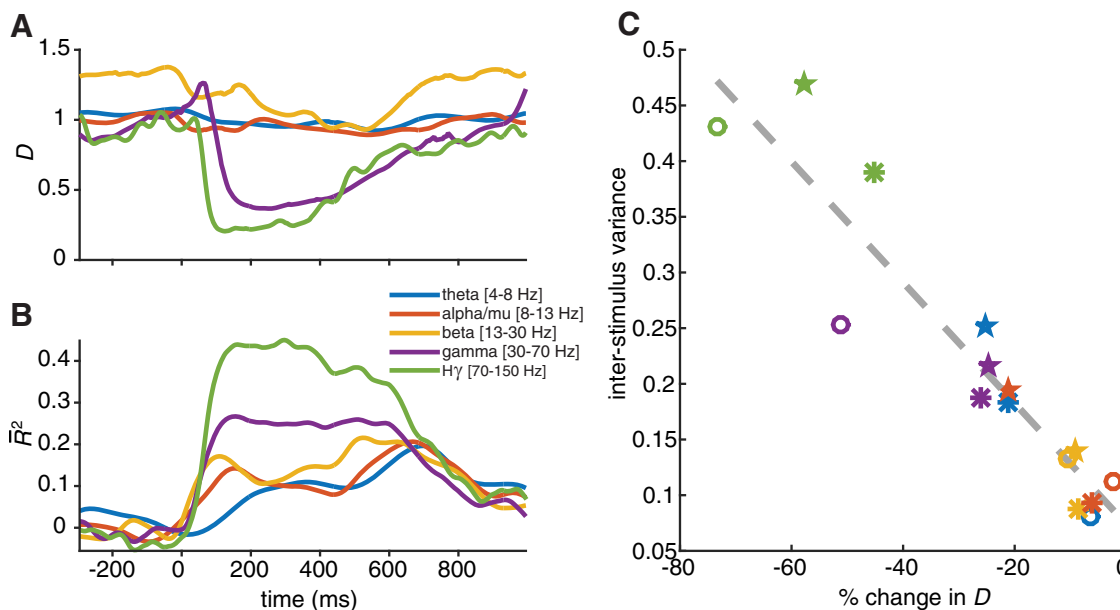


Figure 6. Multiband change in variability and encoding. **A**, The linear relationship between mean and variance of response, for each of the functional bands through time. H γ and gamma have the strongest modulation in variability (D), and the lower frequency bands show little or no modulation. **B**, Encoding strength shown as \bar{R}^2 across time for each of the functional bands. H γ shows the greatest increase in \bar{R}^2 . **C**, Encoding strength versus change in D for all subjects (S1: circles S2: asterisks S3: 5 stars) during the encoding period (100–400 ms after stimulus onset). Gray dashed line indicates line of best fit.

of encoding and variability across electrodes for Subject S1 are shown in Figure 6, which is representative of the three subjects. Each of the standard frequency bands is represented by a different color. Variability primarily decreases in gamma (purple) and H γ (green) (Fig. 6A). In all bands, mean \bar{R}^2 was ~ 0 before stimulus onset, rose after the acoustic onset of the stimulus, and decreased over the next 1 s (Fig. 6B). H γ (green) is shown for comparison, and has the strongest encoding. Figure 6C summarizes the across-frequency analysis of the three subjects. In each subject, H γ has the greatest drop in variability and the greatest encoding strength. Moreover, there was a strong negative correlation (Pearson $r = -0.93$, slope of line of best fit = $-5.4e-3 \pm 1.2e-3$, 95% CI) between the encoding strength of a band and the drop in variability. This suggests

that variability shaping is a primary encoding mechanism that distinguishes high-gamma from other frequency bands.

Discussion

Speech perception relies on the ability to discriminate all of the different sounds of a given language. This requires that the representations within a sound category are similar to each other, and that the representations across sound categories are sufficiently distinct. Here, we examined the human nonprimary auditory cortex (STG) for encodings of speech sounds that underlie their discrimination. High-density ECoG permitted the study of variability on the mesoscopic level, with the spatiotemporal resolution necessary to observe neural dynamics associated with phonetic perception. We found a strong positive correlation be-

tween cortical response magnitude and variance that decreased dramatically following onset of an auditory stimulus, similar to neuronal action potentials across several species and brain areas. These results are evidence that H γ trial-by-trial fluctuations are not merely measurement noise but reflect trial-by-trial differences in firing rates.

H γ on a single electrode reflects neural activity across an estimated several thousand neurons under that electrode (Miller et al., 2009). Previous studies of firing rate responses have found that neuronal variability decreases following stimulus onset in a shared subspace across neurons (Cohen and Maunsell, 2009; Churchland et al., 2010). However, these previous studies recorded from at most a few hundred neurons at a time. Our study expands upon previous findings, showing that the decrease in shared variability extends to a much larger neural population. The brain is thought to process sensory stimuli with large populations of neurons (Hinton, 1984). Therefore, variability of H γ on individual electrodes may reflect network properties that are important for sensory representation and computational processing.

The change in variability was monitored through the slope of the regression relating neural response mean to variance, which we called D . D is an analog of FF for ECoG data analysis; however, care must be taken in its interpretation. A FF of 1 is meaningful because it matches a Poisson process, but D should not be interpreted this way; only relative changes in D are meaningful. Nevertheless, the decrease in D observed here is consistent with the hypothesis that variability in H γ reflects variability in the activity of the underlying neuronal population. One concern could be that the relationship between mean and variance was not changing, but instead it was simply a change in the mean distribution that was causing this change in D . If the variance of response does not change, D could decrease simply due to normalization by greater mean responses. Another possible concern is that there is an upper bound of activation for each electrode due to physiological limitations and that this upper bound might lower variability of high responses due to a ceiling effect. The mean-matching control analysis addresses both of these concerns. Furthermore, our model (Eq. 1) includes a linear offset and a slope, capturing the additive and multiplicative noise components, respectively. Additive noise was captured in the offset term, so the slope, D , reflects specifically multiplicative noise. If the mean changed but the correlation with variance did not, this would be reflected in the offset term, not D .

Computational models have provided some explanations for how a decrease in variability could be achieved mechanistically. In models of populations of neurons, networks that are clustered into subgroups exhibit a more stable response to stimuli, lowering the variability of activity (Litwin-Kumar and Doiron, 2012). Such subgrouping organization may also account for the decrease in response variability we observed in field potentials. Alternatively, reduction in ECoG H γ variability could reflect an increase in the synchrony of neural activation of the underlying population (de la Rocha et al., 2007).

To determine the effect of variability on stimulus representation, we compared it with a stimulus-encoding model. We found that the reduction in variability occurred nearly simultaneously with peak stimulus discriminability and was spatially specific to those electrodes that discriminate between different sounds. If variability reduction had been primarily on electrodes that did not have distinct responses for different sounds, or if it had a very different time course from stimulus representation, the decrease in variability would have had no influence discriminability. Yet

we show that variability decrease does occur together with stimulus representation, which is necessary for changes in variability to influence discriminability.

The effect of changes in variability on the discriminability of stimuli depends on how these changes are correlated across electrodes. We used factor analysis to determine the extent to which changes in variability were correlated across electrodes. The use of factor analysis is similar to a previous work from He and Zempel (2013), where variability was examined using principal components analysis. Here we use factor analysis to model variability explicitly, and apply it to H γ rather than broadband voltage traces. We found that the reduction in variability was primarily shared in the subspace of noise correlations, rather than private to individual electrodes, similar to previous work from Churchland et al. (2010), who studied noise correlations between single neurons. Combined, our findings suggest that a reduction in noise correlations is a multiscale neural phenomenon that may play an important computational role in representation on several spatial scales.

The decorrelation of noise we observe during stimulus representation may be a consequence of feedforward inhibitory input. Computational modeling has demonstrated that variability shaping can be caused by nonlinearities that alter the distribution of neural activity. Electrodes with positive signal correlations are decorrelated during stimulus representation because they receive a common feedforward input that changes the response distribution (Middleton et al., 2012). However, it is unclear how this change in the shape of variance on the scale of neurons would generalize to aggregate neural populations in ECoG.

Noise correlations can either decrease or increase discriminability of stimuli, depending on their relationship with the signal correlations. The “sign rule” (Hu et al., 2014) indicates that noise correlations that are correlated with signal correlations decrease stimulus discriminability. On the other hand, noise correlations that are anticorrelated with signal correlations facilitate discriminability (Abbott and Dayan, 1999). We found the largest decrease in noise correlations between electrodes that are similarly tuned. According to the sign rule, these noise correlations lead to less discriminable representations. Therefore, the noise correlations that are mostly strongly diminished during representation are those that most hinder the discrimination of sounds. We used a linear decoder to quantify the change in information due to noise correlations (Moreno-Bote et al., 2014) and found a small but consistent effect for electrodes with high signal correlations.

Our results show changes in the correlation structure that improves the discriminability of stimuli for electrodes with high signal correlations, but the improvement is very slight and only on a subset of electrode pairs. Previous studies have found a larger effect on discriminability of neuron firing rates from multiunit recordings (Cohen and Maunsell, 2009). One possible explanation is that the effect of noise correlations on discriminability is indeed very small for this task, but we think a more likely explanation is that the improvement in discriminability is stronger on the single-neuron level when recording densely from a relatively small volume of cortex. In contrast, ECoG samples over a very large area of cortex, and the average pairwise signal correlation between ECoG electrodes is typically less than single-unit recordings (Downer et al., 2015). Our results imply that decoding activity from ECoG may be less sensitive to the noise structure of neural activity than from recordings of single units.

We compared variability reduction across frequency bands and found that γ activity has the highest encoding strength and that it exhibits the strongest reduction in variability upon stimulus onset. The gamma band, which also has been shown to modulate with firing rate (Cui et al., 2016), shows a similar but weaker reduction in variability. These results suggest not only that higher bands in particular hold the most information about the acoustic stimuli, but also that they uniquely exhibit variability changes similar to those found in neurons. These results provide further evidence that spectral power modulation in the higher range is distinct from lower bands, reflecting excitation of neuron populations (Rasch et al., 2008; Ray et al., 2008; Crone et al., 2011; Buzsáki et al., 2012).

A complete understanding of neural representation requires not only the mean response to sensory stimuli, but the entire distribution of neural responses. Variability is an integral part of stimulus representation, and changes dynamically as the stimulus is represented. Indeed, we find that neural activity is dynamically shaped in a manner that enhances the discriminability of sounds.

References

- Abbott LF, Dayan P (1999) The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11:91–101. [CrossRef Medline](#)
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7:358–366. [CrossRef Medline](#)
- Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents: EEG, ECoG, LFP and spikes. *Nat Rev Neurosci* 13:407–420. [CrossRef Medline](#)
- Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS, Barbaro NM, Knight RT (2006) High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313:1626–1628. [CrossRef Medline](#)
- Chang EF (2015) Towards large-scale, human-based, mesoscopic neurotechnologies. *Neuron* 86:68–78. [CrossRef Medline](#)
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT (2010) Categorical speech representation in human superior temporal gyrus. *Nat Neurosci* 13:1428–1432. [CrossRef Medline](#)
- Churchland AK, Kiani R, Chaudhuri R, Wang XJ, Pouget A, Shadlen MN (2011) Variance as a signature of neural computations during decision making. *Neuron* 69:818–831. [CrossRef Medline](#)
- Churchland MM, Yu BM, Sahani M, Shenoy KV (2007) Techniques for extracting single-trial activity patterns from large-scale neural recordings. *Curr Opin Neurobiol* 17:609–618. [CrossRef Medline](#)
- Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci* 13:369–378. [CrossRef Medline](#)
- Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci* 14:811–819. [CrossRef Medline](#)
- Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12:1594–1600. [CrossRef Medline](#)
- Crone NE, Korzeniewska A, Franaszczuk PJ (2011) Cortical gamma responses: searching high and low. *Int J Psychophysiol* 79:9–15. [CrossRef Medline](#)
- Cui Y, Liu LD, McFarland JM, Pack CC, Butts DA (2016) Inferring cortical variability from local field potentials. *J Neurosci* 36:4121–4135. [CrossRef Medline](#)
- de la Rocha J, Doiron B, Shea-Brown E, Josiæ K, Reyes A (2007) Correlation between neural spike trains increases with firing rate. *Nature* 448:802–806. [CrossRef Medline](#)
- DeWitt I, Rauschecker JP (2012) Phoneme and word recognition in the auditory ventral stream. *Proc Natl Acad Sci USA* 109:E505–E514. [CrossRef Medline](#)
- Diehl RL, Lotto AJ, Holt LL (2004) Speech perception. *Annu Rev Psychol* 55:149–179. [CrossRef Medline](#)
- Downer JD, Niwa M, Sutter ML (2015) Task engagement selectively modulates neural correlations in primary auditory cortex. *J Neurosci* 35:7565–7574. [CrossRef Medline](#)
- Faisal AA, Selen LP, Wolpert DM (2008) Noise in the nervous system. *Nat Rev Neurosci* 9:292–303. [CrossRef Medline](#)
- Formisano E, De Martino F, Bonte M, Goebel R (2008) “Who” decoding “what”? Brain-based and speech of human voice and speech. *Science* 322:970–973. [CrossRef Medline](#)
- Heggelund P, Albus K (1978) Response variability and orientation discrimination of single cells in striate cortex of cat. *Exp Brain Res* 32:197–211. [CrossRef Medline](#)
- He BJ, Zempel JM (2013) Average is optimal: an inverted-U relationship between trial-to-trial brain activity and behavioral performance. *PLoS Comput Biol* 9:e1003348. [CrossRef Medline](#)
- Hickok G, Poeppel D (2000) Towards a functional neuroanatomy of speech perception. *Trends Cogn Sci* 4:131–138. [CrossRef Medline](#)
- Hinton GE (1984) Distributed Representations. Technical Report CMU-CS-84-157. Carnegie Mellon University, Pittsburgh, PA.
- Hu Y, Zylberberg J, Shea-Brown E (2014) The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. *PLoS Comput Biol* 10:e1003469. [CrossRef Medline](#)
- Kisley MA, Gerstein GL (1999) Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. *J Neurosci* 19:10451–10460. [CrossRef Medline](#)
- Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25:3661–3673.
- Lieberman AM, Harris KS, Hoffman HS, Griffith BC (1957) The discrimination of speech sounds within and across phoneme boundaries. *J Exp Psychol* 54:358–368. [CrossRef Medline](#)
- Lieberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. *Psychol Rev* 74:431–461. [CrossRef Medline](#)
- Litwin-Kumar A, Doiron B (2012) Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci* 15:1498–1505. [CrossRef Medline](#)
- Mackay WA (1997) Synchronized neuronal oscillations and their role in motor processes. *Trends Cogn Sci* 1:176–183. [CrossRef Medline](#)
- Mesgarani N, Cheung C, Johnson K, Chang EF (2014) Phonetic feature encoding in human superior temporal gyrus. *Science* 343:1006–1010. [CrossRef Medline](#)
- Middleton JW, Omar C, Doiron B, Simons DJ (2012) Neural correlation is stimulus modulated by feedforward inhibitory circuitry. *J Neurosci* 32:506–518. [CrossRef Medline](#)
- Miller KJ, Sorensen LB, Ojemann JG, den Nijs M (2009) Power-law scaling in the brain surface electric potential. *PLoS Comput Biol* 5:e1000609. [CrossRef Medline](#)
- Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in Macaque area V4. *Neuron* 63:879–888. [CrossRef Medline](#)
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A (2014) Information-limiting correlations. *Nat Neurosci* 17:1410–1417. [CrossRef Medline](#)
- Nourski KV, Steinschneider M, Oya H, Kawasaki H, Jones RD, Howard MA (2014) Spectral organization of the human lateral superior temporal gyrus revealed by intracranial recordings. *Cereb Cortex* 24:340–352. [CrossRef Medline](#)
- Obleser J, Leaver AM, Vanmeter J, Rauschecker JP (2010) Segregation of vowels and consonants in human auditory cortex: evidence for distributed hierarchical organization. *Front Psychol* 1:232. [CrossRef Medline](#)
- Perkell JS, Klatt DH (1986) Invariance and variability in speech processes. Hillsdale, NJ: Lawrence Erlbaum.
- Rasch MJ, Gretton A, Murayama Y, Maass W, Logothetis NK (2008) Inferring spike trains from local field potentials. *J Neurophysiol* 99:1461–1476. [CrossRef Medline](#)
- Ray S, Crone NE, Niebur E, Franaszczuk PJ, Hsiao SS (2008) Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J Neurosci* 28:11526–11536. [CrossRef Medline](#)
- Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol* 9:e1000610. [CrossRef Medline](#)

- Rey HG, Fried I, Quiroga R (2014) Timing of single-neuron and local field potential responses in the human medial temporal lobe. *Curr Biol* 24:299–304. [CrossRef Medline](#)
- Rose D (1979) An analysis of the variability of unit activity in the cat's visual cortex. *Exp Brain Res* 37:595–604. [Medline](#)
- Schiller PH, Finlay BL, Volman SF (1976) Short-term response variability of monkey striate neurons. *Brain Res* 105:347–349. [CrossRef Medline](#)
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18:3870–3896. [Medline](#)
- Steinschneider M (2011) Unlocking the role of the superior temporal gyrus for speech sound categorization. *J Neurophysiol* 105:2631–2633. [CrossRef Medline](#)
- Theil H (1961) *Economic forecasts and policy*. Amsterdam: North-Holland.
- Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23:775–785. [CrossRef Medline](#)
- Vogels R, Spileers W, Orban GA (1989) The response variability of striate cortical neurons in the behaving monkey. *Exp Brain Res* 77:432–436. [CrossRef Medline](#)
- Whitsel BL, Schreiner RC, Essick GK (1977) An analysis of variability in somatosensory cortical neuron discharge. *J Neurophysiol* 40:589–607. [Medline](#)
- Whittingstall K, Logothetis NK (2009) Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron* 64:281–289. [CrossRef Medline](#)
- Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–143. [CrossRef Medline](#)