Phonetic Feature Encoding in Human Superior Temporal Gyrus

Nima Mesgarani,[1]* Connie Cheung,[1] Keith Johnson,[2] Edward F. Chang[1]†

During speech perception, linguistic elements such as consonants and vowels are extracted from a complex acoustic speech signal. The superior temporal gyrus (STG) participates in high-order auditory processing of speech, but how it encodes phonetic information is poorly understood. We used high-density direct cortical surface recordings in humans while they listened to natural, continuous speech to reveal the STG representation of the entire English phonetic inventory. At single electrodes, we found response selectivity to distinct phonetic features. Encoding of acoustic properties was mediated by a distributed population response. Phonetic features could be directly related to tuning for spectrotemporal acoustic cues, some of which were encoded in a nonlinear fashion or by integration of multiple cues. These findings demonstrate the acoustic-phonetic representation of speech in human STG.

Phonemes—and the distinctive features composing them—are hypothesized to be the smallest contrastive units that change a word's meaning (e.g., /b/ and /d/ as in bad versus dad) (*1*). The superior temporal gyrus (Brodmann area 22, STG) has a key role in acoustic-phonetic processing because it responds to speech over other sounds (*2*) and focal electrical stimulation there selectively interrupts speech discrimination (*3*). These findings raise fundamental questions about the representation of speech sounds, such as whether local neural encoding is specific for phonemes, acoustic-phonetic features, or low-level spectrotemporal parameters. A major challenge in addressing this in natural speech is that cortical processing of individual speech sounds is extraordinarily spatially discrete and rapid (*4*–*7*).

We recorded direct cortical activity from six human participants implanted with high-density multielectrode arrays as part of their clinical evaluation for epilepsy surgery (*8*). These recordings provide simultaneous high spatial and temporal resolution while sampling population neural activity from temporal lobe auditory speech cortex. We analyzed high gamma (75 to 150 Hz) cortical surface field potentials (*9*, *10*), which correlate with neuronal spiking (*11*, *12*).

Participants listened to natural speech samples featuring a wide range of American English speakers (500 sentences spoken by 400 people) (*13*). Most speech-responsive sites were found in posterior and middle STG (Fig. 1A, 37 to 102 sites per participant, comparing speech versus silence, $P < 0.01$, t test). Neural responses demonstrated a distributed spatiotemporal pattern evoked during listening (Fig. 1, B and C, and figs. S1 and S2).

We segmented the sentences into time-aligned sequences of phonemes to investigate whether STG sites show preferential responses. We estimated the mean neural response at each electrode to every phoneme and found distinct selectiv-

[1]Department of Neurological Surgery, Department of Physiology, and Center for Integrative Neuroscience, University of California, San Francisco, CA 94143, USA. [2]Department of Linguistics, University of California, Berkeley, CA 94720, USA.

*Present address: Department of Electrical Engineering, Columbia University, New York, NY 10027, USA.

†Corresponding author. E-mail: changed@neurosurg.ucsf.edu

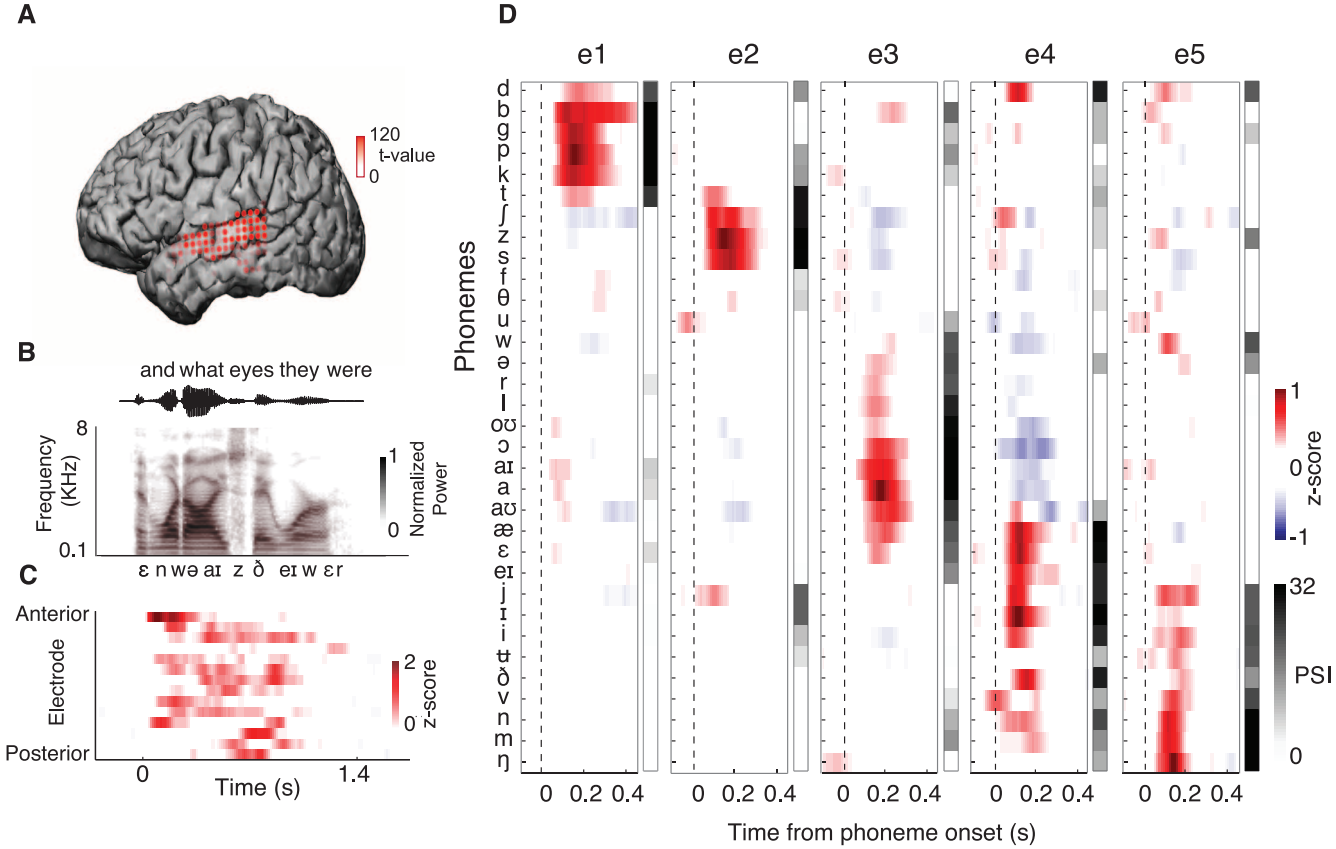

Fig. 1. Human STG cortical selectivity to speech sounds. (**A**) Magnetic resonance image surface reconstruction of one participant's cerebrum. Electrodes (red) are plotted with opacity signifying the t test value when comparing responses to silence and speech ($P < 0.01$, t test). (**B**) Example sentence and its acoustic waveform, spectrogram, and phonetic transcription. (**C**) Neural responses evoked by the sentence at selected electrodes. z score indicates normalized response. (**D**) Average responses at five example electrodes to all English phonemes and their PSI vectors.

ity. For example, electrode e1 (Fig. 1D) showed large evoked responses to plosive phonemes /p/, /t/, /k/, /b/, /d/, and /g/. Electrode e2 showed selective responses to sibilant fricatives: /s/, /ʃ/, and /z/. The next two electrodes showed selective responses to subsets of vowels: low-back (electrode e3, e.g., /a/ and /aʊ/), high-front vowels and glides (electrode e4, e.g., /i/ and /j/). Last, neural activity recorded at electrode e5 was selective for nasals (/n/, /m/, and /ŋ/).

To quantify selectivity at single electrodes, we derived a metric indicating the number of phonemes with cortical responses statistically distinguishable from the response to a particular phoneme. The phoneme selectivity index (PSI) is a dimension of 33 English phonemes; PSI = 0 is nonselective and PSI = 32 is extremely selective (Wilcox rank-sum test, $P < 0.01$, Fig. 1D; methods shown in fig. S3). We determined an optimal analysis time window of 50 ms, centered 150 ms after the phoneme onset by using a phoneme separability analysis (f-statistic, fig. S4A). The average PSI over all phonemes summarizes an electrode's overall selectivity. The average PSI was highly correlated to a site's response magnitude to speech over silence ($r = 0.77$, $P < 0.001$, t test; fig. S5A) and the degree to which the response could be predicted with a linear spectrotemporal receptive field [STRF, $r = 0.88$, $P < 0.001$, t test; fig. S5B (*14*)]. Therefore, the majority of speech-responsive sites in STG are selective to specific phoneme groups.

To investigate the organization of selectivity across the neural population, we constructed an array containing PSI vectors for electrodes across all participants (Fig. 2A). In this array, each column corresponds to a single electrode, and each row corresponds to a single phoneme. Most STG electrodes are selective not to individual but to specific groups of phonemes. To determine selectivity patterns across electrodes and phonemes, we used unsupervised hierarchical clustering analyses. Clustering across rows revealed groupings of phonemes on the basis of similarity of PSI values in the population response (Fig. 2B). Clustering

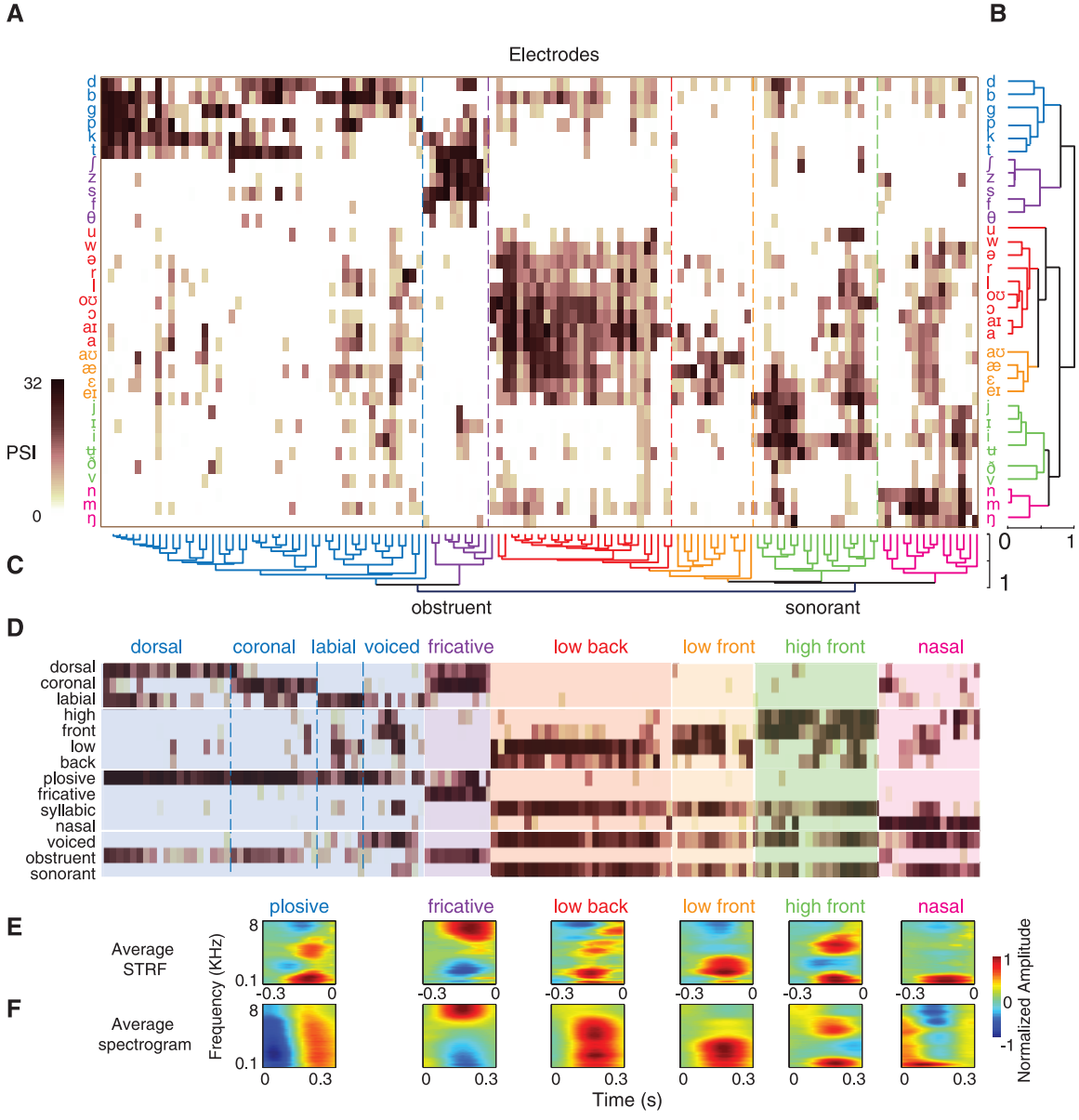

Fig. 2. Hierarchical clustering of single-electrode and population responses. (**A**) PSI vectors of selective electrodes across all participants. Rows correspond to phonemes, and columns correspond to electrodes. (**B**) Clustering across population PSIs (rows). (**C**) Clustering across single electrodes (columns). (**D**) Alternative PSI vectors using rows now corresponding to phonetic features, not phonemes. (**E**) Weighted average STRFs of main electrode clusters. (**F**) Average acoustic spectrograms for phonemes in each population cluster. Correlation between average STRFs and average spectrograms: $r = 0.67$, $P < 0.01$, t test. ($r = 0.50$, 0.78, 0.55, 0.86, 0.86, and 0.47 for plosives, fricatives, vowels, and nasals, respectively; $P < 0.01$, t test).

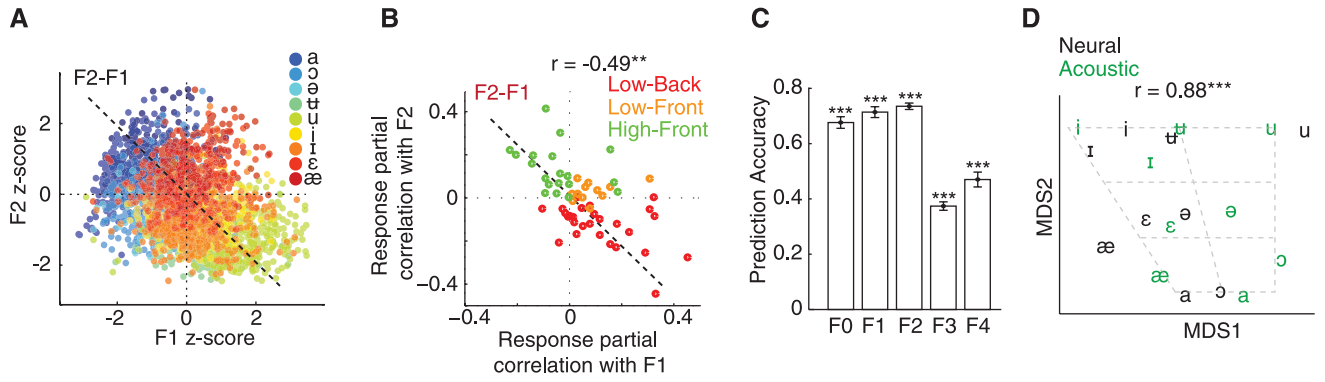

Fig. 3. Neural encoding of vowels. (**A**) Formant frequencies, F1 and F2, for English vowels (F2-F1, dashed line, first principal component). (**B**) F1 and F2 partial correlations for each electrode's response (**$P < 0.01$, *t* test). Dots (electrodes) are color-coded by their cluster membership. (**C**) Neural population decoding of fundamental and formant frequencies. Error bars indicate SEM. (**D**) Multidimensional scaling (MDS) of acoustic and neural space (***$P < 0.001$, *t* test).

across columns revealed single electrodes with similar PSI patterns (Fig. 2C). These two analyses revealed complementary local- and global-level organizational selectivity patterns. We also replotted the array by using 14 phonetic features defined in linguistics to contrast distinctive articulatory and acoustic properties (Fig. 2D; phoneme-feature mapping provided in fig. S7) (*1*, *15*).

The first tier of the single-electrode hierarchy analysis (Fig. 2C) divides STG sites into two distinct groups: obstruent- and sonorant-selective electrodes. The obstruent-selective group is divided into two subgroups: plosive and fricative electrodes (similar to electrodes e1 and e2 in Fig. 1D) (*16*). Among plosive electrodes (blue), some were responsive to all plosives, whereas others were selective to place of articulation (dorsal /g/ and /k/ versus coronal /d/ and /t/ versus labial /p/ and /b/, labeled in Fig. 2D) and voicing (separating voiced /b/, /d/, and /g/ from unvoiced /p/, /t/, and /k/; labeled voiced in Fig. 2D). Fricative-selective electrodes (purple) showed weak, overlapping selectivity to coronal plosives (/d/ and /t/). Sonorant-selective cortical sites, in contrast, were partitioned into four partially overlapping groups: low-back vowels (red), low-front vowels (orange), high-front vowels (green), and nasals (magenta) (labeled in Fig. 2D, similar to e3 to e5 in Fig. 1D).

Both clustering schemes (Fig. 2, B and C) revealed similar phoneme grouping based on shared phonetic features, suggesting that a substantial portion of the population-based organization can be accounted for by local tuning to features at single electrodes (similarity of average PSI values for the local and population subgroups of both clustering analyses is shown in fig. S8; overall $r = 0.73$, $P < 0.001$). Furthermore, selectivity is organized primarily by manner of articulation distinctions and secondarily by place of articulation, corresponding to the degree and the location of constriction in the vocal tract, respectively (*16*). This systematic organization of speech sounds is consistent with auditory perceptual models positing that distinctions are most affected by manner contrasts (*17*, *18*) compared with other feature hierarchies (articulatory or gestural theories) (*19*).

We next determined what spectrotemporal tuning properties accounted for phonetic feature selectivity. We first determined the weighted average STRFs of the six main electrode clusters identified above, weighting them proportionally by their degree of selectivity (average PSI). These STRFs show well-defined spectrotemporal tuning (Fig. 2E) highly similar to average acoustic spectrograms of phonemes in corresponding population clusters (Fig. 2F; average correlation = 0.67, $P < 0.01$, *t* test). For example, the first STRF in Fig. 2E shows tuning for broadband excitation followed by inhibition, similar to the acoustic spectrogram of plosives. The second STRF is tuned to a high frequency, which is a defining feature of sibilant fricatives. STRFs of vowel electrodes show tuning for characteristic formants that define low-back, low-front, and high-front vowels. Last, STRF of nasal-selective electrodes is tuned primarily to low acoustic frequencies generated from heavy voicing and damping of higher frequencies (*16*). The average spectrogram analysis requires a priori phonemic segmentation of speech but is model-independent. The STRF analysis assumes a linear relationship between spectrograms and neural responses but is estimated without segmentation. Despite these differing assumptions, the strong match between these confirms that phonetic feature selectivity results from tuning to signature spectrotemporal cues.

We have thus far focused on local feature selectivity to discrete phonetic feature categories. We next wanted to address the encoding of continuous acoustic parameters that specify phonemes within vowel, plosive, and fricative groups. For vowels, we measured fundamental (F0) and formant (F1 to F4) frequencies (*16*). The first two formants (F1 and F2) play a major perceptual role in distinguishing different English vowels (*16*), despite tremendous variability within and across vowels (Fig. 3A) (*20*). The optimal projection of vowels in formant space was the difference of F2 and F1 (first principal component, dashed line, Fig. 3A), which is consistent with vowel perceptual studies (*21*, *22*). By using partial correlation analysis, we quantified the relationship between electrode response amplitudes and F0 to F4. On average, we observed no correlation between the sensitivity of an electrode to F0 with its sensitivity to F1 or F2. However, sensitivity to F1 and F2 was negatively correlated across all vowel-selective sites (Fig. 3B; $r = -0.49$, $P < 0.01$, *t* test), meaning that single STG sites show an integrated response to both F1 and F2. Furthermore, electrodes selective to low-back and high-front vowels (labeled in Fig. 2D) showed an opposite differential tuning to formants, thereby maximizing vowel discriminability in the neural domain. This complex sound encoding matches the optimal projection in Fig. 3A, suggesting a specialized higher-order encoding of acoustic formant parameters (*23*, *24*) and contrasts with studies of speech sounds in nonhuman species (*25*, *26*).

To examine population representation of vowel parameters, we used linear regression to decode F0 to F4 from neural responses. To ensure unbiased estimation, we first removed correlations between F0 to F4 by using linear prediction and decoded the residuals. Relatively high decoding accuracies are shown in Fig. 3C ($P < 0.001$, *t* test), suggesting fundamental and formant variability is well represented in population STG responses (interaction between decoder weights with electrode STRFs shown in fig. S9). By using multidimensional scaling, we found that the relational organization between vowel centroids in the acoustic domain is well preserved in neural space (Fig. 3D; $r = 0.88$, $P < 0.001$).

For plosives, we measured three perceptually important acoustic cues (fig. S10): voice-onset time (VOT), which distinguishes voiced (/b/, /d/, and /g/) from unvoiced plosives (/p/, /t/, and /k/); spectral peak (differentiating labials /p/ and /b/ versus coronal /t/ and /d/ versus dorsal /k/ and /g/); and F2 of the following vowel (*16*). These acoustic parameters could be decoded from population STG responses (Fig. 4A; $P < 0.001$, *t* test). VOTs in particular are temporal cues that are perceived categorically, which suggests a nonlinear encoding (*27*). Figure 4B shows neural responses for three example electrodes plotted for all plosive instances (total of 1200), aligned to their release

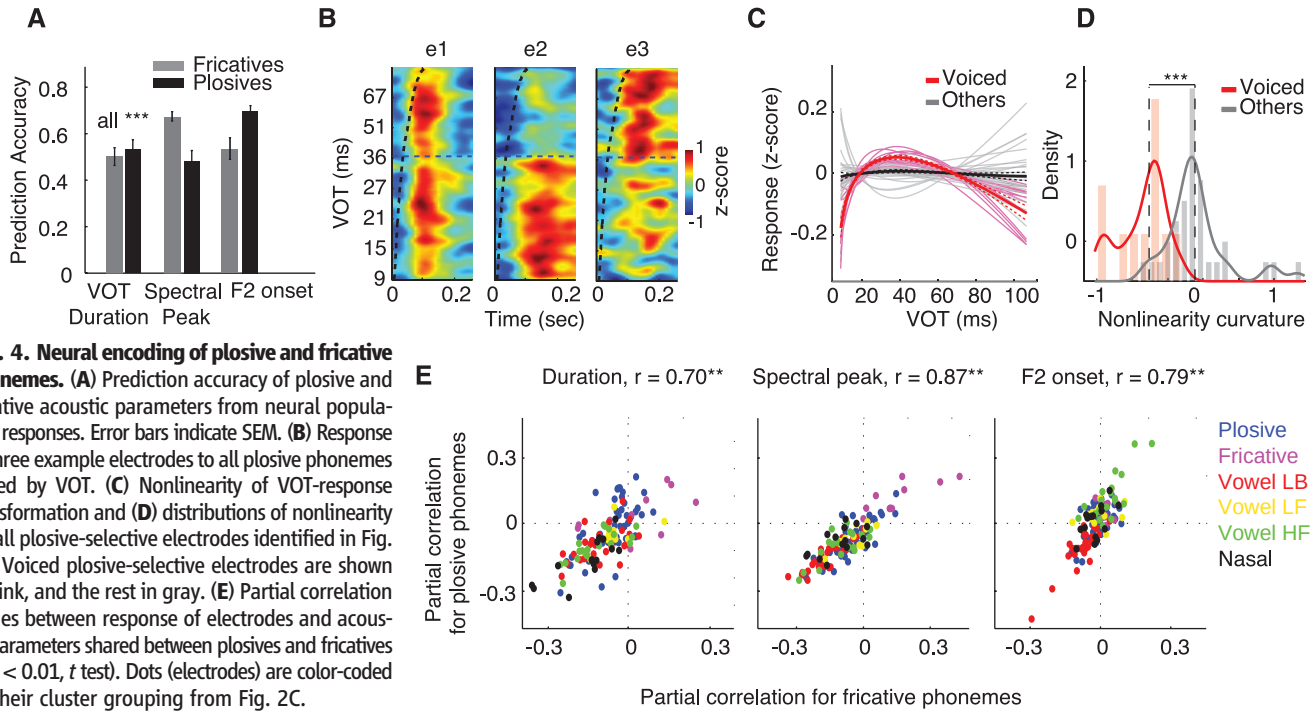

Fig. 4. Neural encoding of plosive and fricative phonemes. (**A**) Prediction accuracy of plosive and fricative acoustic parameters from neural population responses. Error bars indicate SEM. (**B**) Response of three example electrodes to all plosive phonemes sorted by VOT. (**C**) Nonlinearity of VOT-response transformation and (**D**) distributions of nonlinearity for all plosive-selective electrodes identified in Fig. 2D. Voiced plosive-selective electrodes are shown in pink, and the rest in gray. (**E**) Partial correlation values between response of electrodes and acoustic parameters shared between plosives and fricatives (**$P < 0.01$, *t* test). Dots (electrodes) are color-coded by their cluster grouping from Fig. 2C.

time and sorted by VOT. The first electrode responds to all plosives with same approximate latency and amplitude, irrespective of VOT. The second electrode responds only to plosive phonemes with short VOT (voiced), and the third electrode responds primarily to plosives with long VOT (unvoiced).

To examine the nonlinear relationship between VOT and response amplitude for voiced-plosive electrodes (labeled voiced in Fig. 2D) compared with plosive electrodes with no sensitivity to voicing feature (labeled coronal, labial and dorsal in Fig. 2D), we fitted a linear and exponential function to VOT-response pairs (fig. S11B). The difference between these two fits specifies the nonlinearity of this transformation, shown for all plosive electrodes in Fig. 4C. Voiced-plosive electrodes (pink) all show strong nonlinear bias for short VOTs compared with all other plosive electrodes (gray). We quantified the degree and direction of this nonlinear bias for these two groups of plosive electrodes by measuring the average second-derivative of the curves in Fig. 4C. This measure maps electrodes with nonlinear preference for short VOTs (e.g., electrode e2 in Fig. 4B) to negative values and electrodes with nonlinear preference for long VOTs (e.g., electrode e3 in Fig. 4B) to positive values. The distribution of this measure for voiced-plosive electrodes (Fig. 4D, red distribution) shows significantly greater nonlinear bias compared with the remaining plosive electrodes (Fig. 4D, gray distribution) ($P < 0.001$, Wilcox rank-sum test). This suggests a specialized mechanism for spatially distributed, nonlinear rate encoding of VOT and contrasts with previously described temporal encoding mechanisms (*26*, *28*).

We performed a similar analysis for fricatives, measuring duration, which aids the distinction between voiced (/z/ and /v/) and unvoiced fricatives (/s/, /ʃ/, /θ/, /f/); spectral peak, which differentiates /f/ and /v/ versus coronal /s/ and /z/ versus dorsal /ʃ/; and F2 of the following vowel (*16*) (fig. S12). These parameters can be decoded reliably from population responses (Fig. 4A; $P < 0.001$, *t* test).

Because plosives and fricatives can be subspecified by using similar acoustic parameters, we determined whether the response of electrodes to these parameters depends on their phonetic category (i.e., fricative or plosive). We compared the partial correlation values of neural responses with spectral peak, duration, and F2 onset of fricative and plosive phonemes (Fig. 4E), where each point corresponds to an electrode color-coded by its cluster grouping in Fig. 2D. High correlation values ($r = 0.70$, 0.87, and 0.79; $P < 0.001$; *t* test) suggest that electrodes respond to these acoustic parameters independent of their phonetic context. The similarity of responses to these isolated acoustic parameters suggests that electrode selectivity to a specific phonetic features (shown with colors in Fig. 4E) emerges from combined tuning to multiple acoustic parameters that define phonetic contrasts (*24*, *25*).

We have characterized the STG representation of the entire American English phonetic inventory. We used direct cortical recordings with high spatial and temporal resolution to determine how selectivity for phonetic features is correlated to acoustic spectrotemporal receptive field properties in STG. We found evidence for both spatially local and distributed selectivity to perceptually relevant aspects of speech sounds, which together appear to give rise to our internal representation of a phoneme.

We found selectivity for some higher-order acoustic parameters, such as examples of nonlinear, spatial encoding of VOT, which could have important implications for the categorical representation of this temporal cue. Furthermore, we observed a joint differential encoding of F1 and F2 at single cortical sites, suggesting evidence of spectral integration previously speculated in theories of combination-sensitive neurons for vowels (*23*–*25*, *29*).

Our results are consistent with previous single-unit recordings in human STG, which have not demonstrated invariant, local selectivity to single phonemes (*30*, *31*). Instead, our findings suggest a multidimensional feature space for encoding the acoustic parameters of speech sounds (*25*). Phonetic features defined by distinct acoustic cues for manner of articulation were the strongest determinants of selectivity, whereas place-of-articulation cues were less discriminable. This might explain some patterns of perceptual confusability between phonemes (*32*) and is consistent with feature hierarchies organized around acoustic cues (*17*), where phoneme similarity space in STG is driven more by auditory-acoustic properties than articulatory ones (*33*). A featural representation has greater universality across languages, minimizes the need for precise unit boundaries, and can account for coarticulation and temporal overlap over phoneme-based models for speech perception (*17*).

References and Notes

1. N. Chomsky, M. Halle, *The Sound Pattern of English* (Harper and Row, New York, 1968).
2. J. R. Binder *et al.*, *Cereb. Cortex* **10**, 512–528 (2000).
3. D. Boatman, C. Hall, M. H. Goldstein, R. Lesser, B. Gordon, *Cortex* **33**, 83–98 (1997).
4. E. F. Chang *et al.*, *Nat. Neurosci.* **13**, 1428–1432 (2010).

5. E. Formisano, F. De Martino, M. Bonte, R. Goebel, *Science* **322**, 970–973 (2008).
6. J. Obleser, A. M. Leaver, J. Vanmeter, J. P. Rauschecker, *Front. Psychol.* **1**, 232 (2010).
7. M. Steinschneider *et al.*, *Cereb. Cortex* **21**, 2332–2347 (2011).
8. Materials and methods are available as supplementary materials on *Science* Online.
9. N. E. Crone, D. Boatman, B. Gordon, L. Hao, *Clin. Neurophysiol.* **112**, 565–582 (2001).
10. E. Edwards *et al.*, *J. Neurophysiol.* **102**, 377–386 (2009).
11. M. Steinschneider, Y. I. Fishman, J. C. Arezzo, *Cereb. Cortex* **18**, 610–625 (2008).
12. S. Ray, J. H. R. Maunsell, *PLOS Biol.* **9**, e1000610 (2011).
13. J. S. Garofolo, TIMIT: Acoustic-Phonetic Continuous Speech Corpus (Linguistic Data Consortium, Philadelphia, 1993).
14. F. E. Theunissen *et al.*, *Network* **12**, 289–316 (2001).
15. M. Halle, K. Stevens, in *Music, Language, Speech, and Brain*, J. Sundberg, L. Nord, R. Carlson, Eds. (Wenner-Gren International Symposium Series vol. 59, Macmillan, Basingstoke, UK, 1991).
16. P. Ladefoged, K. Johnson, *A Course in Phonetics* (Cengage Learning, Stamford, CT, 2010).
17. K. N. Stevens, *J. Acoust. Soc. Am.* **111**, 1872–1891 (2002).
18. G. Clements, *Phonol. Yearb.* **2**, 225–252 (1985).
19. C. A. Fowler, *J. Phonetics* **14**, 3–28 (1986).
20. G. E. Peterson, H. L. Barney, *J. Acoust. Soc. Am.* **24**, 175 (1952).
21. J. D. Miller, *J. Acoust. Soc. Am.* **85**, 2114–2134 (1989).
22. A. K. Syrdal, H. S. Gopal, *J. Acoust. Soc. Am.* **79**, 1086–1100 (1986).
23. H. M. Sussman, *Brain Lang.* **28**, 12–23 (1986).
24. I. Nelken, *Curr. Opin. Neurobiol.* **18**, 413–417 (2008).
25. N. Mesgarani, S. V. David, J. B. Fritz, S. A. Shamma, *J. Acoust. Soc. Am.* **123**, 899–909 (2008).
26. C. T. Engineer *et al.*, *Nat. Neurosci.* **11**, 603–608 (2008).
27. L. Lisker, A. S. Abramson, *Lang. Speech* **10**, 1–28 (1967).
28. M. Steinschneider *et al.*, *Cereb. Cortex* **15**, 170–186 (2005).
29. G. Chechik, I. Nelken, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18968–18973 (2012).
30. A. M. Chan *et al.*, *Cereb. Cortex*, published online 16 May 2013 (10.1093/cercor/bht127).
31. O. Creutzfeldt, G. Ojemann, E. Lettich, *Exp. Brain Res.* **77**, 451–475 (1989).
32. G. A. Miller, P. E. Nicely, *J. Acoust. Soc. Am.* **27**, 338 (1955).
33. A. M. Liberman, *Speech: A Special Code* (MIT Press, Cambridge, MA, 1996).

Acknowledgments: We thank A. Ren for technical help with data collection and preprocessing. S. Shamma, C. Espy-Wilson, E. Cibelli, K. Bouchard, and I. Garner provided helpful comments on the manuscript. E.F.C. was funded by NIH grants R01-DC012379, R00-NS065120, and DP2-OD00862 and the Ester A. and Joseph Klingenstein Foundation. E.F.C., C.C., and N.M. collected the data. N.M. and C.C. performed the analysis. N.M. and E.F.C. wrote the manuscript. K.J. provided phonetic consultation. E.F.C. supervised the project.

Supplementary Materials

www.sciencemag.org/content/343/6174/1006/suppl/DC1
Materials and Methods
Figs. S1 to S12
Reference (*34*)

16 September 2013; accepted 17 January 2014
Published online 30 January 2014;
10.1126/science.1245994

Detection of a Recurrent *DNAJB1-PRKACA* Chimeric Transcript in Fibrolamellar Hepatocellular Carcinoma

Joshua N. Honeyman,[1,2]* Elana P. Simon,[1,3]* Nicolas Robine,[4]* Rachel Chiaroni-Clarke,[1] David G. Darcy,[1,2] Irene Isabel P. Lim,[1,2] Caroline E. Gleason,[1] Jennifer M. Murphy,[1,2] Brad R. Rosenberg,[5] Lydia Teegan,[1] Constantin N. Takacs,[1] Sergio Botero,[1] Rachel Belote,[1] Soren Germer,[4] Anne-Katrin Emde,[4] Vladimir Vacic,[4] Umesh Bhanot,[6] Michael P. LaQuaglia,[2] Sanford M. Simon[1]†

Fibrolamellar hepatocellular carcinoma (FL-HCC) is a rare liver tumor affecting adolescents and young adults with no history of primary liver disease or cirrhosis. We identified a chimeric transcript that is expressed in FL-HCC but not in adjacent normal liver and that arises as the result of a ~400-kilobase deletion on chromosome 19. The chimeric RNA is predicted to code for a protein containing the amino-terminal domain of DNAJB1, a homolog of the molecular chaperone DNAJ, fused in frame with PRKACA, the catalytic domain of protein kinase A. Immunoprecipitation and Western blot analyses confirmed that the chimeric protein is expressed in tumor tissue, and a cell culture assay indicated that it retains kinase activity. Evidence supporting the presence of the *DNAJB1-PRKACA* chimeric transcript in 100% of the FL-HCCs examined (15/15) suggests that this genetic alteration contributes to tumor pathogenesis.

Fibrolamellar hepatocellular carcinoma (FL-HCC) is a rare liver tumor that was first described in 1956 and that historically has been considered a variant of hepatocellular carcinoma (*1*, *2*). It is histologically characterized by well-differentiated neoplastic hepatocytes and thick fibrous bands in a noncirrhotic background (*3*, *4*). FL-HCC has a clinical phenotype distinct from conventional hepatocellular carcinoma and usually occurs in adolescents and young adults. Patients have normal levels of alpha fetoprotein without underlying liver disease or history of viral hepatitis (*3*–*6*). Little is known of its molecular pathogenesis. FL-HCC tumors do not respond well to chemotherapy (*7*, *8*), and surgical resection remains the mainstay of therapy, with overall survival reported to be 30 to 45% at 5 years (*1*, *6*, *8*, *9*).

To investigate the molecular basis of FL-HCC, we performed whole-transcriptome and whole-genome sequencing of paired tumor and adjacent normal liver samples. To determine whether there were tumor-specific fusion transcripts among the coding RNA, we ran the program FusionCatcher (*10*) on RNA sequencing (RNA-Seq) data from 29 samples, including primary tumors, metastases, recurrences, and matched normal tissue samples, derived from a total of 11 patients (table S1). There was only one recurrent candidate chimeric transcript detected in every tumor sample. This candidate transcript is predicted to result from the in-frame fusion of exon 1 from the *DNAJB1* gene, which encodes a member of the heat shock 40 protein family, with exons 2 to 10 from *PRKACA*, the gene encoding the adenosine 3′,5′-monophosphate (cAMP)–dependent protein kinase A (PKA) catalytic subunit alpha. This fusion transcript was not detected in any of the available paired normal tissue samples ($n = 9$). This fusion is not found in the COSMIC database (*11*) and has not previously been reported in the literature.

To further characterize the candidate fusion transcript, we directly examined those RNA-Seq reads that mapped to *PRKACA* and *DNAJB1*. We examined *PRKACA* transcript levels with DESeq2 (*12*) and found that they were increased relative to normal in tumors from all nine patients tested [P value adjusted for multiple testing (pAdj) $< 10^{-12}$, range three- to eightfold]. To determine whether the increased expression was attributable to a specific isoform of PRKACA, we quantified reads mapping to different exons and evaluated differential expression using DEXSeq (*13*). In all nine patients, there was an increase in the expression of exons 2 to 10 of *PRKACA* in the tumor relative to exon 1 and relative to the expression in normal tissue (Fig. 1A, left). This exon expression pattern does not correspond to a known isoform of PRKACA. Rather, it reflects an increase in *PRKACA* transcripts lacking the first exon, which encodes the domain that engages the regulatory subunits of PKA. All reads mapping to *PRKACA* in normal tissue were either contained within exons or bridged the junctions between adjacent exons at annotated splicing sites (Fig. 1B, left, blue). All tumor samples additionally had reads mapping from the start of the second exon of

[1]Laboratory of Cellular Biophysics, Rockefeller University, 1230 York Avenue, New York, NY 10065, USA. [2]Division of Pediatric Surgery, Department of Surgery, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. [3]The Dalton School, 108 East 89th Street, New York, NY 10128, USA. [4]New York Genome Center, 101 Avenue of the Americas, New York, NY 10013, USA. [5]Whitehead Presidential Fellows Program, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA. [6]Pathology Core Facility Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA.

*These authors contributed equally to this work.
†Corresponding author. E-mail: simon@rockefeller.edu

Phonetic Feature Encoding in Human Superior Temporal Gyrus

Nima Mesgarani *et al.*

Science **343**, 1006 (2014);
DOI: 10.1126/science.1245994

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.