Human speech sounds are produced through a coordinated movement of structures along the vocal tract. Here we show highly structured neuronal encoding of vowel articulation. In medial–frontal neurons, we observe highly specific tuning to individual vowels, whereas superior temporal gyrus neurons have nonspecific, sinusoidally modulated tuning (analogous to motor cortical directional tuning). At the neuronal population level, a decoding analysis reveals that the underlying structure of vowel encoding reflects the anatomical basis of articulatory movements. This structured encoding enables accurate decoding of volitional speech segments and could be applied in the development of brain–machine interfaces for restoring speech in paralysed individuals.
The articulatory features that distinguish different vowel sounds are conventionally described along a two-dimensional coordinate system that naturally represents the position of the highest point of the tongue during articulation, for example, in the International Phonetic Alphabet (IPA) chart. The two axes of this system are height (tongue vertical position relative to the roof of the mouth or the aperture of the jaw) and backness (tongue position relative to the back of the mouth). How is the structured articulatory production encoded and controlled in brain circuitry? The gross functional neuroanatomy of speech production was described by multiple imaging, lesion and stimulation studies, and includes primary, supplementary and pre-motor areas, Broca’s area, superior temporal gyrus (STG), anterior cingulate cortex (ACC) and other medial–frontal regions. The temporal dynamics of collective neuronal activity was studied in Broca’s area using local field potentials. However, the basic encoding of speech features in the firing patterns of neuronal populations remains unknown.

Here, we study the neuronal encoding of vowel articulation in the human cerebrum at both the single-unit and the neuronal population levels. At the single-neuron level, we find signatures of two structured coding strategies: highly specific, sharp tuning to individual vowels (in medial–frontal neurons) and nonspecific, sinusoidally modulated tuning (in the STG). At the neural population level, we find that the encoding of vowels reflects the underlying articulatory movement structure. These findings may have important implications for the development of high-accuracy brain–machine interfaces for the restoration of speech in paralysed individuals.

**Results**

**Speech-related neurons.** Neuronal responses in human temporal and frontal lobes were recorded from 11 patients with intractable epilepsy monitored with intracranial depth electrodes to identify seizure foci for potential surgical treatment (see Methods). Following an auditory cue, subjects uttered one of five vowels (a, e, i, u and o) or simple syllables containing these vowels (consonant + vowel: da, de, di, do and du). We recorded the activity of 716 temporal and frontal lobe units. As this study focuses on speech and owing to the inherent difficulty to distinguish between auditory- and speech-related neuronal activations, we analysed only the 606 units that did not respond to auditory stimuli. A unit was considered speech-related if its firing rate during speech differed significantly from the pre-cue baseline period (see Methods). Overall, 8% of the analysed units (49) were speech-related, of which, more than a half (25) were vowel-tuned, showing significantly different activation for the five vowels (see Supplementary Fig. S1).

**Sharp and broad vowel tuning.** Two areas commonly activated in speech studies, the STG and a medial–frontal region overlying the rostral anterior cingulate and the adjacent medial orbitofrontal cortex (rAC/MOF; Brodmann areas 11 and 12; See Supplementary Fig. S2 for anatomical locations of the electrodes), had the highest proportions of speech-related (75% and 11%, respectively) and vowel-tuned units (58% and 77% of these units). In imaging and electrocorticography studies, the rACC was shown to participate in speech control, the orbitofrontal cortex in speech comprehension and reading, and the STG in speech production at the phoneme level. Involvement of STG neurons in speech production was also observed in earlier single-unit recordings in humans. We analysed neuronal tuning in these two areas and found that it had divergent characteristics: broadly tuned units that responded to all vowels, with a gradual modulation in the firing rate between vowels, comprised 93% of tuned units in STG (13/14) but were not found in rAC/MOF (0/10), whereas sharply tuned units that had significant activation exclusively for one or two vowels comprised 100% of the tuned rAC/MOF units (10/10) but were rare in STG (1/14).

Figure 1 displays responses of five sharply tuned units in rAC/MOF, each exhibiting strong, robust increases in their firing rate specifically for one or two vowel sounds, whereas for the other vowels firing remains at the baseline rate. For example, a single unit in the right rostral anterior cingulate cortex (rACC) (Fig. 1, top row) elevated its firing rate to an average of 97 spikes/s when the patient said ‘a’, compared with 6 spikes/s for ‘i’, ‘e’, ‘o’ and ‘u’ (P < 10^–13, one-sided two-sample t-test). Anecdotally, in the first two trials of this example (red arrow) the firing rate remained at the baseline level, unlike the rest of the ‘a’ trials; in these two trials, the patient wrongly said ‘ah’ rather than ‘a’ (confirmed by the sound recordings).
Figure 2 | Broadly tuned STG units. (a) Raster plots and peri-stimulus time histograms during the utterance of the five vowels a, e, i, u and o. Significant change in firing rate from the baseline occurred for all or most vowels, with modulated firing rate (Methods). Red vertical dashed lines indicate speech onset; vertical bars, 10 spikes/s. (b) Tuning curves of the respective units in a over the vowel space, showing orderly variation in the firing rate of STG units with the articulated vowel.

A completely different encoding of vowels was found in the STG, where the vast majority of tuned units exhibited broad variation of the response over the vowel space, during the articulation of both vowels (Fig. 2a) and simple syllables containing these vowels (Supplementary Fig. S3a). This structured variation is well approximated by sinusoidal tuning curves (Fig. 2b and Supplementary Fig. S3b) analogous to the directional tuning curves commonly observed in motor cortical neurons. Units shown in Fig. 2 had maximal responses (‘preferred vowel; in analogy to ‘preferred direction’) to the vowels ‘i’ and ‘u’, which correspond to a closed articulation where the tongue is maximally raised, and minimal ('anti-preferred') response to ‘a’ and ‘o’ where it is lowered.

Population-level decoding and structure. Unlike directional tuning curves, where angles are naturally ordered, vowels can have different orderings. In the tuning curves of Fig. 2, we ordered the vowels according to their place and manner of articulation as expressed by their location in the IPA chart, but is this ordering natural to the neural representation? Instead of assuming a certain ordering, we could try and deduce the natural organization of speech features represented in the population-level neural code. That is, we can try to infer a neighbourhood structure (or order) of the vowels where similar (neighbouring) neuronal representations are used for neighbouring vowels. We reasoned that this neighbourhood structure could be extracted from the error structure of neuronal classifiers: when a decoder, such as the population vector errors, it is more likely to prefer a value that is a neighbour of the correct value than a more distant one. Thus, classification error rates are expected to be higher between neighbours than between distant features when feature ordering accurately reflects the neural representation. In this case, classification error rates expressed by the classifier’s confusion matrix will have a band-diagonal structure.

To apply this strategy, we decoded the population firing patterns using multivariate linear classifiers with a sparsity constraint to infer the uttered vowel (Methods). The five vowels were decoded with a high average (cross-validated) accuracy of 93% (significantly above the 20% chance level, P < 10^-5, one-sided one-sample t-test, n = 6 cross-validation runs; Supplementary Table S1), and up to 100% when decoding pairs of vowels (Fig. 3a). Next, we selected the vowel ordering that leads to band-diagonal confusion matrices (Fig. 3b). Interestingly, this ordering is consistent across different neuronal subpopulations (Fig. 3b and Supplementary Fig. S4) and exactly matches the organization of vowels according to their place and manner of articulation as reflected by the IPA chart (Fig. 3c). As the vowel chart represents the position of the highest point of the tongue during articulation, the natural organization of speech features by neuronal encoding reflects a functional spatial-anatomical axis in the mouth.

Discussion

These results suggest that speech-related rAC/MOF neurons use sparse coding for vowels, in analogy to the sparse bursts in songbirds’ area HVc and to the sparse, highly selective responses observed in the human medial–temporal lobe. In contradistinction, the gradually modulated speech encoding in STG implies previously unrecognized correlates with a hallmark of motor cortical control—broad, sinusoidal tuning, implying a role in motor control of speech production. Interestingly, speech encoding in these anatomical areas is opposite in nature to that of other modalities: broad tuning for motor control is common in the frontal lobe versus STG in the temporal lobe whereas, sparse tuning to high-level concepts is common in the temporal lobe versus rAC/MOF in the frontal lobe. Analogous to the recently found subpathway between the visual dorsal and ventral streams, our findings may lend support to a speech-related ‘dorsal stream’ where sensorimotor prediction supports speech production by a state feedback control. The sparse rAC/MOF representation may serve as predictor state, in line with anterior cingulate and orbital frontal roles in reward prediction. The broad STG tuning may support evidence that the motor system is capable of modulating the perception system to some degree.

Our finding of sharply tuned neurons in rAC/MOF agrees with the DIVA model of the human speech system, which suggests that higher-level prefrontal cortex regions involved in phonological encoding of an intended utterance sequentially activate speech...
Bandpass-filtered signals (0.3–3 kHz) from these microwires and the sound track computer tomography coregistered to preoperative magnetic resonance imaging. Their locations were verified by magnetic resonance imaging or by weeks, each of which terminated with a set of nine 40-day. Based exclusively on clinical criteria, each patient had 8–12 electrodes for 1–2 ages 19–53) participated in a total of 14 recording sessions, each on a different.

Patients and electrophysiology. Eleven patients with pharmacologically resistant epilepsy undergoing invasive monitoring with intracranial depth electrodes to identify the seizure focus for potential surgical treatment (9 right handed, 7 females, ages 19–53) participated in a total of 14 recording sessions, each on a different day. Based exclusively on clinical criteria, each patient had 12 electrodes for 1–2 weeks, each of which terminated with a set of nine 40-μm platinum–iridium microwires. Their locations were verified by magnetic resonance imaging or by computer tomography coregistered to preoperative magnetic resonance imaging. Bandpass-filtered signals (0.3–3 kHz) from these microwires and the sound track were synchronously recorded at 30 kHz using a 128-channel acquisition system (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Methods

Patients and electrophysiology. Eleven patients with pharmacologically resistant epilepsy undergoing invasive monitoring with intracranial depth electrodes to identify the seizure focus for potential surgical treatment (9 right handed, 7 females, ages 19–53) participated in a total of 14 recording sessions, each on a different day. Based exclusively on clinical criteria, each patient had 12 electrodes for 1–2 weeks, each of which terminated with a set of nine 40-μm platinum–iridium microwires. Their locations were verified by magnetic resonance imaging or by computer tomography coregistered to preoperative magnetic resonance imaging. Bandpass-filtered signals (0.3–3 kHz) from these microwires and the sound track were synchronously recorded at 30 kHz using a 128-channel acquisition system (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.

Experimental paradigms. Patients first listened to isolated auditory cues (beeps) and to another individual uttering the vowel sounds and three syllables (me/hu/ha) following beeps (auditory controls). Then, following an oral instruction, patients uttered the instructed syllable multiple times, each following a randomly spaced (2–3 s) beep. Syllables consisted of either monophthongal vowels (u/i/o/ and u/i) or a consonant (of: d/di, g/gi, h/hi, t/dit, v/vi, m/min, n/nni, s/si, t/diti, s/n/ and v/vi), and one of the aforementioned vowels (for example, d/di, d/dit, d/diti, d/di, d/dit, and d/di).

Our results further demonstrate that the neuronal population encoding of vowel generation appears to be organized according to a functional representation of a spatial-anatomical axis: tongue height. This axis was shown to have a significant main effect on decoding from speech motor cortex units. Whether these structures multi-level encoding schemes also exist in other speech areas (Neuroport, Blackrock). Sorted units (WaveClus and SUMU) recorded in different sessions are treated as different in this study. All studies conformed with the guidelines of the Medical Institutional Review Board at the University of California Los Angeles.
lowered speech onset; 6: 5-element binary vectors coding the individual vowels uniquely. All decoding results were sixfold cross-validated using trials that were not used for decoder training. The decoder was trained on all of the aforementioned features from the training set only, with no selection of the input neurons or their features. Instead, the sparse decoder described above automatically selects task-relevant features by higher weights it allocates to them using the minimal ||w||1 ||x||1 constraint; task-unrelated features are thus diminished by low weights. Owing to the high accuracy in decoding, we randomly dropped 20% of the units (in each cross-validation training) when computing confusion matrices, to increase the amount of confusions and allow the extraction of a meaningful band-diagonal structure (except for the STG-only training, Supplementary Fig S4).

The vowels in Fig 3c were placed on the IPA chart according to the locations previously calculated for American speakers (ref. 1, page 42), and the overlaid connections (blue lines) were inferred by the maximal non-diagonal element previously calculated for American speakers (ref. 28, page 42). The overlaid connections (blue lines) were inferred by the maximal non-diagonal element previously calculated for American speakers (ref. 28, page 42).

References

1. International Phonetic Association. Handbook of the International Phonetic Association (Cambridge University Press, 1999).
2. Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D. & Halgren, E. Sequential processing of lexical, grammatical, and phonological information within Broca’s area. Science 326, 449–449 (2009).
3. Hickok, G., Hauk, J. & Rong, F. Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69, 407–422 (2011).
4. Ghosh, S. S., Tourville, J. A. & Guenther, F. H. A neuroimaging study of premotor lateralization and cerebellar involvement in the production of phonemes and syllables. J. Speech Lang. Hear. Res. 51, 1183–1202 (2008).
5. Halgren, E. et al. Spatiotemporal stages in face and word-processing. J. Physiol. Paris 88, 51–80 (1994).
6. Pass, T., Petrides, M., Evans, A. C. & Meyer, E. Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J. Neurophysiol. 70, 453–469 (1993).
7. Sabri, M. et al. Attentional and linguistic interactions in speech perception. NeuroImage 39, 1444–1456 (2008).
8. Buchsbaum, B. R., Hickok, G. & Humphries, C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. Cognitive Sci. 25, 663–678 (2001).
9. Ojemann, G. A., Creutzfeldt, O., Lettich, E. & Haglund, M. Neuronal activity in human lateral temporal cortex related to short-term verbal memory, naming and reading. Brain 111 (Pt 6), 1383–1403 (1988).
10. Georgopoulos, A. P., Kalaska, J. F., Caminiti, R. & Massey, J. T. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J. Neurosci. 2, 1527–1537 (1982).
11. Georgopoulos, A. P., Schwartz, A. B. & Kettner, R. E. Neuronal population coding of movement direction. Science 233, 1416–1419 (1986).
12. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature 419, 65–70 (2002).
13. Quian-Quiroga, R., Reddy, L., Kreiman, G., Koch, C. & Fried, I. Invariant visual representation by single neurons in the human brain. Nature 435, 1102–1107 (2005).
14. Tankus, A. & Fried, I. Visuomotor coordination and motor representation by human temporal lobe neurons. J. Cogn. Neurosci. 24, 600–610 (2011).
15. Haylen, R. V., Heilbronner, S. R., Pearson, J. M. & Platt, M. L. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. J. Neurosci. 31, 4178–4187 (2011).
16. Wallis, J. D. Cross-species studies of orbitofrontal cortex and value-based decision making. Nat. Neurosci. 15, 13–19 (2012).
17. Hickok, G. & Poeppel, D. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402 (2007).
18. Hickok, G. Computational neuroanatomy of speech production. Nat. Rev. Neurosci. 13, 133–145 (2012).
19. Guenther, F. H. & Vladusich, T. A neural theory of speech acquisition and production. J. Neuroeconomics 25, 408–422 (2012).
20. Cavada, C., Coïa, T., Tejedor, J., Cruz-Rizzolo, R. J. & Reinoso-Suárez, F. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. Cereb. Cortex 10, 220–242 (2000).
21. Carmichael, S. T. & Price, J. L. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. 363, 642–664 (1995).
22. Brumberg, J. S., Wright, E. J., Andreasen, D. S., Guenther, F. H. & Kennedy, P. R. Classification of intended phoneme production from chronic intracortical microelectrode recordings in speech-motor cortex. Front. Neurosci. 5, 65 (2011).
23. Guenther, F. H. et al. A wireless brain-machine interface for real-time speech synthesis. PLoS ONE 4, e8218 (2009).
24. Brumberg, J. S., Nieto-Castanon, A., Kennedy, P. R. & Guenther, F. H. Brain-computer interfaces for speech communication. Speech Commun. 52, 367–379 (2010).
25. Pasley, B. N. et al. Reconstructing speech from human auditory cortex. PLoS Biol. 10, e1001251 (2012).
26. Pei, X., Barbour, D. L., Leuthardt, E. C. & Schall, G. Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans. J. Neural Eng. 8, 046028 (2011).
27. Leuthardt, E. C. et al. Using the electrocorticographic speech network to control a brain-computer interface in humans. J. Neural Eng. 8, 036004 (2011).
28. Formisano, E., De Martino, F., Bonte, M. & Goebel, R. ‘Who’ is saying ’what’? Brain-based decoding of human voice and speech. Science 322, 970–973 (2008).
29. Fried, I. et al. Cerebral microdialysis combined with single-neuron and electroencephalographic recording in neurological patients. J. Neurosurg. 91, 697–705 (1999).
30. Quian Quiroga, R., Nadasy, Z. & Ben-Shaul, Y. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. Neural Comput. 16, 1661–1687 (2004).
31. Tankus, A., Yeshurun, Y. & Fried, I. An automatic measure for classifying clusters of suspected spikes into single cells versus multunit. J. Neural Eng. 6, 056001 (2009).
32. McCormick, L. M. et al. Anterior cingulate cortex: an MRI-based parcellation method. NeuroImage 32, 1167–1175 (2006).
33. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57, 289–300 (1995).
34. van den Berg, E. & Friedlander, M. P. Probing the pareto frontier for basis pursuit solutions. SIAM J. Sci. Comput. 31, 890–912 (2008).
35. Tankus, A., Fried, I. & Shoham, S. Sparse decoding of multiple spike trains for brain–machine interfaces. J. Neural Eng. 9 (in press) (2012).

Acknowledgements

We thank D. Pourshayan, E. Behnke, T. Fields and Prof. P. Keating of UCLA, and Prof. A. Cohen and A. Alassy of the Technion for assistance, and the European Research Council (STG 211055), NINDS, Dana Foundation, Lady Davis and L. and L. Richmond research funds for financial support.

Author contributions

A.T., I.F. and S.S. designed the study and wrote the manuscript; I.F. performed neuro-surgeries; A.T. prepared experimental setup and performed the experiments; A.T. and S.S. analysed the data.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests

The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Tankus, A. et al. Structured neuronal encoding and decoding of human speech features. Nat. Commun. 3:1015 doi: 10.1038/ncomms1995 (2012).