Sound representation in higher language areas during language generation

Lorenzo Magrassi[a,b,c,1], Giuseppe Aromataris[d], Alessandro Cabrini[d], Valerio Annovazzi-Lodi[d], and Andrea Moro[c]

[a]Neurosurgical Unit, Department of Clinical, Surgical, Diagnostic, and Pediatric Sciences, University of Pavia, 27100 Pavia, Italy; [b]Istituto di Genetica Molecolare - Consiglio Nazionale delle Ricerche, 27100 Pavia, Italy; [c]Neurocognition and Theoretical Syntax Research Center, Institute for Advanced Studies, 27100 Pavia, Italy; and [d]Department of Electrical, Computer, and Biomedical Engineering, University of Pavia, Pavia, Italy

Edited* by Dale Purves, Duke University, Durham, NC, and approved December 17, 2014 (received for review September 21, 2014)

How language is encoded by neural activity in the higher-level language areas of humans is still largely unknown. We investigated whether the electrophysiological activity of Broca's area correlates with the sound of the utterances produced. During speech perception, the electric cortical activity of the auditory areas correlates with the sound envelope of the utterances. In our experiment, we compared the electrocorticogram recorded during awake neurosurgical operations in Broca's area and in the dominant temporal lobe with the sound envelope of single words versus sentences read aloud or mentally by the patients. Our results indicate that the electrocorticogram correlates with the sound envelope of the utterances, starting before any sound is produced and even in the absence of speech, when the patient is reading mentally. No correlations were found when the electrocorticogram was recorded in the superior parietal gyrus, an area not directly involved in language generation, or in Broca's area when the participants were executing a repetitive motor task, which did not include any linguistic content, with their dominant hand. The distribution of suprathreshold correlations across frequencies of cortical activities varied whether the sound envelope derived from words or sentences. Our results suggest the activity of language areas is organized by sound when language is generated before any utterance is produced or heard.

language | sound envelope | Broca's area | morphosyntax | temporal lobe

An important aspect of human language is speech production, although language may be generated independently from sound, as when one writes or thinks. However, introspection seems to suggest that our thoughts resound in our brain, much as if we were listening to an internal speech, yielding the impression/illusion that sound is inseparable from language.

When human subjects listen to utterances, the neural activity in the superior temporal gyrus is modulated to track the envelope of the acoustic stimulus. The correlation between the power envelope of the speech and the neural activity is maximal at low frequencies (2–7 Hz, theta range) corresponding to syllable rates, and becomes less precise at higher frequencies (15–150 Hz, gamma range) corresponding to phoneme rates (1). Entrainment of neural activity to the speech envelope in auditory regions has allowed recognition of the phonetic features heard during speech perception (2–5), and even the reconstruction of simple words (6). This evidence indicates that during listening, speech representation in the auditory cortex and adjacent areas of the superior temporal lobe reflects acoustic features directly related to linguistically defined phonological entities such as phonemes and syllables. The relationship of specific patterns of sound amplitude and frequency to all similar patterns experienced over phylogeny and individual ontogeny is then responsible for sound perceptions (7). Moreover, as for subjects listening to natural speech, spatiotemporal features of the acoustic trace representative of the sound of the listened words have also been detected in the neural activity of cortical areas outside the superior temporal gyrus (8).

Is sound representation essential during the generation of linguistic expression before the implementation of the motor program for speech? Or is the neural activity in higher language areas completely independent of it?

Results

Cortical Activity Correlates with the Sound Envelope of the Words. After obtaining approval from the institutional ethics committee of the Fondazione Policlinico S. Matteo, we retrospectively analyzed the electrocorticographic activity (ECoG) and concomitant sound tracks recorded from the dominant frontal and temporal lobes of native Italian speakers, using high-density surface multielectrode arrays (HDMs), during awake neurosurgical operations. In Fig. 1, we show an example of the typical HDM positioning on the frontal and temporal lobes (Fig. 1*A*) and a map of the positions of all electrodes in all patients involved in the study (Fig. 1*B*). We calculated the cross-correlation between the sound envelope of the words and sentences read aloud by the patients and the corresponding ECoG traces of the prefrontal electrodes (Fig. 1*C*), along with its periodogram, computed by a fast Fourier transform (FFT) algorithm. This correlation cyclically increased with a frequency coherent to the pace of reading (Fig. 2*A*). In contrast, when we compared the sound envelope of the same utterances to the corresponding ECoG activity recorded in the superior parietal gyrus, an area not involved in language generation and showing weak functional MRI signal during discourse comprehension (9), no periodic variations in correlation amplitudes were found (Fig. 2*B*). As a further control, during the same surgical session, in three patients, we also compared the sound envelope of the words and sentences read aloud in one trial with the ECoG activity

Significance

The results of our experiments show that a special representation of sound is actually exploited by the brain during language generation, even in the absence of speech. Taking advantage of data collected during neurosurgical operations on awake patients, here we cross-correlated the cortical activity in the frontal and temporal language areas of a person reading aloud or mentally with the envelope of the sound of the corresponding utterances. In both cases, cortical activity and the envelope of the sound of the utterances were significantly correlated. This suggests that in hearing people, sound representation deeply informs generation of linguistic expressions at a much higher level than previously thought. This may help in designing new strategies to help people with language disorders such as aphasia.

Author contributions: L.M. and A.M. designed research; G.A., A.C., and V.A.-L. designed and implemented the codes; A.M. designed the linguistic stimuli; L.M. performed research; L.M., G.A., A.C., and V.A.-L. analyzed data; and L.M., V.A.-L., and A.M. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence should be addressed. Email: lorenzo.magrassi@unipv.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1418162112/-/DCSupplemental.

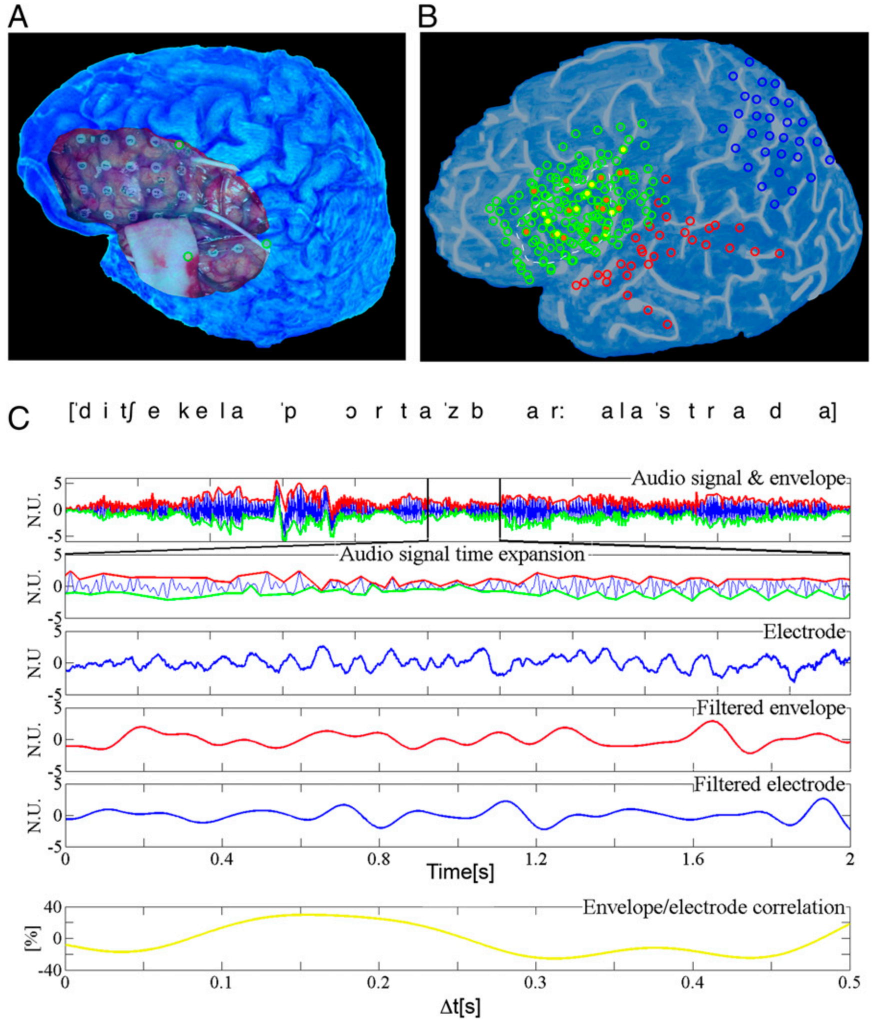

Fig. 1. (*A*) Reconstruction of the surface of the cerebral cortex of one of the patients from magnetic resonance images obtained before surgery. An intraoperative picture of the HDMs was projected over the reconstruction at the corresponding intraoperative position. Electrodes that were partially or completely covered by the dura or a cottonoid used to stabilize the HDM are indicated by green circles. (*B*) The positions of electrodes for all participants recorded from the dominant left hemisphere were mapped on the outline of the surface of the cerebral cortex shown in *A*. Green circles represent electrodes in the dominant frontal lobe, yellow-filled green circles are electrodes positioned over the areas in which stimulation induced speech arrest in all trials, and orange-filled green circles indicate electrodes located in areas in which stimulation induced speech arrest in at least one trial. Red circles represent electrodes in a separate HDM located over the dominant left temporal lobe, and blue circles indicate the position of electrodes located over the superior parietal gyrus. The white dashed line outlines the anatomical Broca's area. (*C*) From top to bottom: phonetic transcription of the sample Italian sentence ("dice che la porta sbarra la strada" [ˈditʃe ke la ˈpɔrta ˈzbarːa la ˈstrada] english translation: "s/he says that the door blocks the way"). Audio signal & envelope: Its acoustic waveform with the corresponding envelopes (positive and negative) outlined in red and green, respectively. Audio signal time expansion: An expanded acoustic waveform and envelopes of the section comprised among the two vertical lines. Electrode: the corresponding ECoG trace recorded from one electrode positioned in the area of speech arrest. Filtered envelope: The corresponding positive envelope (red trace) of the 2,750–3,250-Hz audio band. Filtered electrode: The corresponding ECoG filtered in a selected frequency band (2–8 Hz). The *x* axis of all of the above signals represents time in seconds (*t*[s]). Envelope/electrode correlation: The resulting cross-correlation of the filtered positive sound envelope and filtered ECoG. Δt[s]: Time differential in seconds between the ECoG and the audio trace; in the present example, the correlation maximum entirely contained in the 500-ms window occurs when the ECoG anticipates the start of sound of about 170 ms.

recorded from the dominant prefrontal cortex in a further trial in which the patients were silent. In this trial, instead of reading, they pushed a button with the hand contralateral to the dominant hemisphere each time they saw a slide showing the drawing of a finger pushing a button. Slides with the pushing finger were randomly alternated with black slides at a rate comparable to the rate of the slides displaying linguistic items. Under this experimental condition, we did not find any periodic variations in correlation amplitudes (Fig. 2*C*). Finally, we compared the ECoG activity of patients reading mentally with the sound envelope obtained while they were reading aloud. The words and the sentences the patients were reading mentally were the same, and they were presented at the same pace as when the patients were reading aloud. In this context, the correlation cyclically increased with a frequency coherent to the pace of reading, even in the complete absence of sound (Fig. 2*D*).

To increase the temporal resolution of the study of the cross-correlation of the sound envelope with the ECoG, the ECoG and the sound traces were sliced into periods corresponding to the single utterances of each participant. Periods containing the same utterances in a single trial were put in register and averaged. The average duration of the audio trace of any utterance for each participant was relatively constant, with relative SDs always below 10% (minimum, 5.11%; maximum, 9.64%) of the average

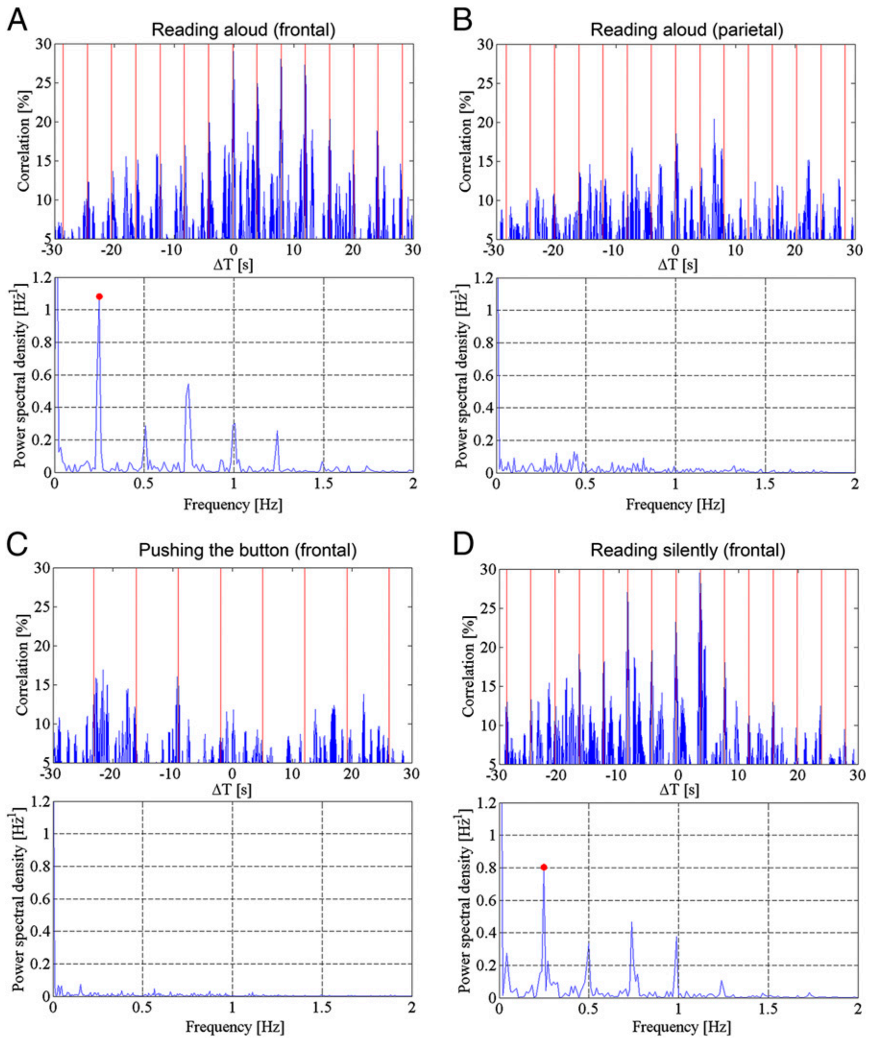

Fig. 2. (*A*) Diagram showing the evolution of cross-correlation between the ECoG recorded by one electrode in Broca's area where speech arrest was obtained during stimulation and the envelope of the corresponding audio signal; both traces were recorded while the patient was reading 18 different utterances aloud. The red vertical lines indicate the time of presentation of the linguistic items (every 4 s). The peaks of cross-correlation follow speech modulations over a period of 72 s. The periodogram, shown in the lower diagram, demonstrates a main frequency of increase of the correlation (red dot) of 0.25 Hz corresponding to a 4-s period, which is equal to the pace of the presentation. The lower peaks correspond to signal harmonics. Audio band, 2,750–3,250 Hz; ECoG band, 2–4 Hz. (*B*) Same as in *A*, but here, the ECoG was recorded by an electrode located in the superior parietal gyrus. The patient was reading aloud the same utterances as in *A*. Cross-correlation peaks do not follow speech modulations over a period of 60 s. As in *A*, the thin vertical lines indicate the time of presentation of the linguistic items (4 s). The periodogram here confirms the absence of any dominant frequency of change of the correlation compatible with the pace of presentation of the linguistics items. (*C*) Same as in *A*: electrode located over Broca's area as in *A*, but the ECoG was recorded while the patient pushed a button with the hand contralateral to the operated dominant hemisphere each time she or he saw a slide showing the drawing of a finger pushing a button. The patient was asked to remain silent and to try to concentrate on the visual image without thinking verbally. The sound track was recorded in a previous trial during the same neurosurgical operation, when the same patient was reading aloud words and sentences. The frequency of the slide presentation in both trials was 7 s and is indicated by the red vertical lines. Changes in the cross-correlation are not following the pace of slide presentation and patient activity for a period of 70 s. As in *B*, the corresponding periodogram represented here does not show any specific or dominant frequency of the correlation compatible with the pace of stimulus presentation. (*D*) Same patient, same pace of presentation of the linguistic items, and same electrode as in *A*, but here the patient reads mentally 12 of the sentences he previously read aloud. The trace represents a period of 50 s read in a completely silent mode. Variations in cross-correlation between the ECoG recorded while the patient was reading mentally and the envelope of the audio recorded in a previous trial when the patient was reading aloud the same linguistic items, follow the presentation time (4 s) of the linguistic items (red vertical lines). The periodogram below shows a dominant frequency of variations of the correlation of 0.25 Hz (red dot), corresponding to a period of 4 s that, as in *A*, equals the rate of presentation of the linguistic items.

duration of the utterance. Furthermore, the sound resulting from the average of the single utterances of a participant was always understandable. Strengthening our findings in longer, unselected periods, when we considered all electrodes, the average value of the cross-correlation of the averaged ECoG, and the envelope of the acoustic trace of words and sentences, was 36.86 (SD, ±15.64), well above the background (patient being silent and not reading mentally) maximal value (i.e., 20.00). Both in the lateral inferior prefrontal cortex containing Broca's area and in the superior temporal gyrus, the amplitude of correlations during reading was maximal (44.72; SD, ±17.12) in the theta range but was also present (29.07; SD, ±8.61) in the gamma range (Fig. 3*A*). These findings are similar to those described in the superior temporal gyrus during listening (1, 6).

Random Shuffling of the ECoG Abrogates the Correlation. When we randomly shuffled segments obtained from the ECoG, the amplitude of the cross-correlation with the corresponding

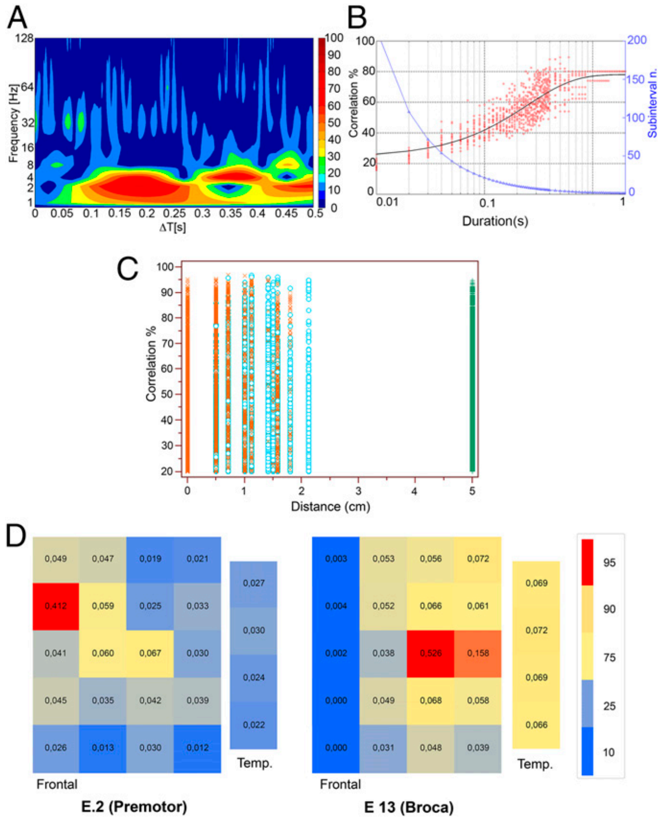

Fig. 3. (*A*) Cross-correlation as a function of frequency of the ECoG and the speech envelope; same sentence as in Fig. 1*C*. Data have been separately normalized on the different frequency bands. Notice that maximal correlation values are within the theta band (2–8 Hz). Frequency (Hz): Frequency of the ECoG in Hz. ΔT[s]: time differential in seconds between the ECoG and the audio trace of overt speech; ECoG response always anticipates sound. (*B*) Plot showing the dependency of the amplitude of cross-correlation (Correlation %) of the ECoG and the speech envelope (left *y* axis) on the correct linear order of the original signal. When we randomly shuffled subintervals of decreasing length obtained by cutting the audio and ECoG trace of the same sentence represented in Fig. 1*C* and *A*, we obtained an exponential decay of the correlation amplitude that halved for an interval duration of ~0.09 s. Red circle: Correlation value obtained of each shuffling; the black line fits the mean values of the correlation amplitude obtained in all independent shuffling for each subinterval length. Subinterval n.: Number of possible subintervals for each interval duration (right *y* axis) is shown by blue triangles. Duration (s): Length of subintervals in seconds. (*C*) Amplitude of the cross-correlation above background maximal value (20%) of the ECoG and the speech envelope as a function of the linear distance (cm) of the recording electrodes from the electrodes overlaying the area in which electrical stimulation induced speech arrest (0). Temporal electrodes, indicated by green "+," were all referred to the average distance of 5 cm. Electrodes located over the anatomical Broca's area are indicated by red "x"; frontal electrodes outside anatomical Broca's area are indicated by cyan circles. All data (90.504 measures) obtained in all patients at all frequencies of audio and ECoG bands are represented. No significant differences in the distribution of the correlation amplitudes are visible between electrodes located over the speech arrest area and the electrode outside it. Please notice that electrodes located over the superior parietal gyrus are not represented here. (*D*) Spatial distribution of all of the electrodes showing a suprathreshold value (20%) for the correlation between the audio envelope and ECoG simultaneous to the correlation peak of one electrode located either over an area of the frontal cortex whose stimulation did not interfered with language [E2 (Premotor)] or over another area of the cortex whose stimulation induced speech arrest [E13 (Broca)]. Each square represents an electrode; colors (red, maximal; blue, minimal) are proportional to the frequency of occurrence of suprathreshold correlation values and are indicated by the number in the middle of the square. Bigger rectangles (Frontal) represent electrodes located over the dominant frontal lobe, and smaller rectangles (Temp.) represent electrodes located over the dominant temporal lobe. A color scale is provided with the indication of percentiles corresponding to each color. The probability distribution of a suprathreshold correlation changes significantly according to the position of the reference electrode considered. The number of electrodes in the dominant temporal lobe that reach suprathreshold correlation values simultaneously with the electrode located in the area of speech arrest is significantly higher than when the electrode was located on a prefrontal area, where speech arrest could not be evoked.

unshuffled sound envelope decayed exponentially with the inverse of segment duration (Fig. 3*B*). On average, the amplitude of the cross-correlation was reduced to half of the original when 90-ms segments were randomly shuffled. This, together with the negative results of the experiments on long correlations obtained from the superior parietal gyrus or from the frontal lobe when the patient was involved in a motor task, confirms that the cross-correlation between the sound envelope and the corresponding ECoG is not a result of interference or random fluctuations of the signals.

Cortical Activity of Areas Close to Broca's Area Also Correlate with the Sound Envelope of the Linguistic Expressions Generated. Comparing the mean value of cross-correlation among ECoG activities recorded in each explored site of the dominant frontal and temporal lobes with the envelope of the sound of the words and sentences pronounced, we found correlation values above background in many electrodes located in the dominant frontal and temporal lobes close to, but outside, the anatomically defined Broca's and Wernicke areas (Fig. 3*C*). Our results are in line with previous results derived from stimulation mapping of language areas during neurosurgery (10, 11), showing that the location of sites whose stimulation could interfere with language varied widely from patient to patient and were not limited to the anatomically defined Broca's and Wernicke areas. We did not find statistically significant differences (t test unpaired $P = 0.158$) between electrodes in the frontal lobe, where cortical stimulation induced speech arrest during naming (average, 36.52; SD, ±15.42) with respect to those where the stimulation did not interfere with language activity (average, 36.85; SD, ±15.68). When we compared electrodes located over the anatomically defined Broca's area (average, 36.99; SD, ±15.75) independent from the result of cortical stimulation versus electrodes in the frontal lobe that were located elsewhere (average, 36.21; SD, ±15.52), we found a statistically significant difference in amplitude (t test unpaired $P < 0.0001$); however, this difference is so small compared with the average fluctuations of background correlation that we do not consider it biologically relevant. The only robust difference between electrodes located above the areas of speech arrest versus those that were located elsewhere was that electrodes that had a correlation value above background when those in the speech arrest had maximal correlation values that were significantly different [χ^2, 746,429; degrees of freedom (DF), 529; $P < 0.0001$] from those that had maximal correlation values when the electrodes not located in speech arrest areas showed maximal correlation values (Fig. 3*D*).

The Average Anticipation of the ECoG over the Sound Envelope Precedes Phonological Processing. Language-related neural activity in Broca's area, as shown by large field potentials, starts ~400 ms before sound is produced and is characterized by three further components at 280, 150, and 50 ms (12). In our experiment, the average anticipation of the ECoG over the sound envelope was 245 ms (SD, ±11.27). This value that maximized the correlation between the two signals is consistent with the second component of large field potentials that is correlated with grammatical processing and precedes later components related to both the phonological and phonetic processing and the articulatory motor commands (12, 13). This value was similar when patients were reading aloud, as well as when they were reading the same sentences and words mentally. Further indication that the correlation between sound envelope and ECoG has biological significance comes from the observation that the length of the segments that significantly start to reduce correlation after shuffling falls into the same range of length of the segments of an utterance that, when locally time-reversed, make it partially unintelligible (14).

Differences in Spectral Distributions of the Correlation Amplitude Distinguish Words Versus Sentences. Utterances may be distinguished by their phonological and grammatical contents. We found that values above background of the cross-correlation between the sound envelope and the ECoG are differentially distributed in the frequency bands, according to the presence of a specific morphosyntactic structure (i.e., words vs. sentences) contained in the text read by the patients (Fig. 4). Differences in spectral distributions are statistically significant and allow a distinction between words pronounced in isolation versus sentences (χ^2, 1,112.469; DF, 7; $P < 0.0001$). This difference remains significant, even when sentences and words of similar time duration are compared (χ^2, 23.717; DF, 7; $P < 0.0013$). Interestingly, most differences are in the high gamma range of the ECoG (Fig. 4), where individual suprasegmental features, including prosody effects, should be less relevant (1).

Discussion

Our experiment shows that neural activity in Broca's area, a prototypical high-level language area governing morphosyntactic structures across languages (15–19), and in the high-level language areas of the dominant temporal lobe is informed by the envelope of the sound of the linguistic expressions that that same activity is encoding. This is true whether the encoded linguistic items will then be uttered or not (as when we think). Our findings concerning the correlation of cortical activity in high-level language areas to the sound envelope of the encoded linguistic items are symmetrical to those emerging by studying cortical activity during speech perception (3, 5, 6). Even if we assume that subvocalization with subthreshold activation of the phonatory apparatus is always active when we think (20), the correlation between ECoG and the envelope of the sound of the linguistic items involved is not limited to the late phonological and phonetic processing and the encoding of the articulatory motor commands that immediately precede sound production. The value of 245 ms (SD, ±11.27) that we found to maximize the correlation between the sound envelope and ECoG signals precedes by about 90 ms the start of those later activities and is consistent with earlier activities in high-level language areas that are correlated with more abstract linguistic features such as grammatical processing (12, 13). Our ability to find significant differences in the spectral distributions of the correlation amplitude between words and sentences of equal duration also support the notion that sound encoding is involved in the early steps of linguistic production and, at least in normal subjects, is not limited to speech production. Prelingually deaf schizophrenic patients who depended on the same left hemispheric areas that control language production in hearing people for sign language (21) report hearing inner voices with frequencies comparable to normally hearing people (22). Although the current interpretation of hearing voices in congenitally deaf people is controversial (23), it is possible that some rudimentary representation of the auditory consequence of articulation is present in some deaf people (23).

Our results suggest that in normal hearing people, sound representation is at the heart of language and not simply a vehicle for expressing some otherwise mysterious symbolic activity of our brain.

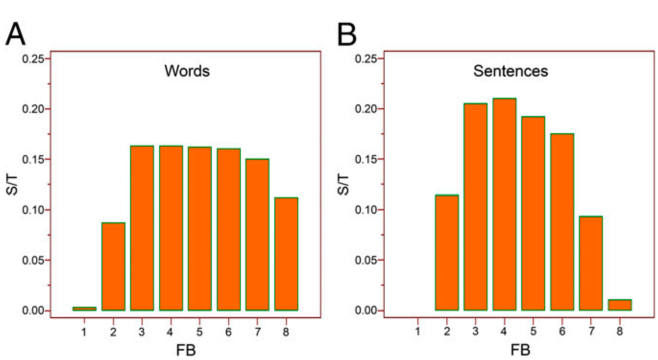

Fig. 4. The cross-correlation of the envelope of the speech sounds with the ECoG above background is differentially distributed in the frequency bands according to the presence of specific morphosyntactic structures (words vs. sentences) contained in the text read by the patients. (*A*) Histogram showing the number of electrodes with suprathreshold correlation in each ECoG frequency band (FB) when patients were reading words. S/T: Ratio between electrodes showing suprathreshold correlation in each ECoG frequency band and the total number of electrodes in all ECoG frequency bands. (*B*) Same as above, but values referred to patients reading sentences.

Materials and Methods

For an extended description of the techniques used, see *SI Materials and Methods*.

Participants and Cortical Mapping. The present study [named Language Induced Potential Study 1 (LIPS1)] was approved by the Fondazione Istituto di Ricerca e Cura a Carattere Scientifico Policlinico S. Matteo institutional review board and ethics committee on human research. All analyses, the results of which are presented in our study, were performed offline and did not interfere with the clinical management of the patients. Acute intraoperative recordings obtained from 16 Italian native speakers (12 males and 4 females) affected by primary or secondary malignant tumors growing in the dominant frontal, temporal, or parietal lobes were the subject of our study. Stimulation for finding areas of speech arrest were performed according to standard neurosurgical techniques for awake neurosurgical operations (10, 11). The intensity of the current used for cortical stimulation mapping varied from patient to patient, with the maximum intensity being that which did not produce after-discharges in the simultaneous ECoG traces. The position of the stimulating and recording electrodes was determined visually and recorded on a neuronavigator (Vectorvision Brain Lab). This allowed the unambiguous classification of all electrodes in the dominant frontal lobes either as electrodes corresponding to the anatomical Broca's area (24) or as electrodes that were outside it.

Neural and Audio Recordings. ECoG recordings were obtained using a multichannel electroencephalographer (System Plus; Micromed) with a sampling rate of 2,048 Hz and analog-to-digital converter (ADC) resolution of 16 bit. One or two HDM grids with an interelectrode distance of 5 mm measured from center to center were simultaneously used for ECoG recordings. Sound tracks and neural activities were simultaneously recorded during testing, and sound was acquired by an H1 X/Y stereo microphone (Zoom H1; Zoom Corp.) placed at the base of the neck on the same side of the operated hemisphere at sampling rates of 96 KHz.

Neuropsychological Testing.

Linguistic items. Linguistic expressions were based on standard Italian taken from the "Lessico di Frequenza dell'Italiano Parlato" (25) and included either simple words or sentences.

Reading aloud. The patients were asked to read standard Italian words and sentences aloud, pronouncing all the utterances as they are used to, while trying to keep the sound intensity as uniform as possible between different reading sessions. This resulted in a uniform mean absolute amplitude of the sound envelopes obtained from all patients (mean, 2.48 mV; SD, ±0.30 mV).

Reading mentally. After successfully completing the reading aloud testing, if the patient was still comfortable and attentive, we asked her or him to read mentally the same words and sentences she or he just read aloud, without changing his or her reading pace. The patient was also explicitly instructed to avoid lip movement or other voluntary movement mimicking sound emission.

Pushing the button. Three patients were also recorded while they were pushing a button with the hand contralateral to the operated dominant hemisphere. The button was pushed every time the drawing of a hand with the index finger pushing a red button over a black background was displayed on a computer screen. The image of the pushing finger was randomly alternated with black screens at the same pace rate used for the patient when she or he was tested for reading.

Signal Analysis. The frequency components of the high-resolution audio signal were separated into five contiguous bands (audio band 1: 750–1,250; audio band 2: 1,250–1,750; audio band 3: 1,750–2,250; audio band 4: 2,250–2,750, and audio band 5: 2,750–3,250 Hz). For each audio band, the sound envelope was extracted from the audio track by interpolation, followed by low-pass filtering. The ECoG and the envelope were separated into eight contiguous-frequency bands in octave ratio, ranging from 0.04 to 128 Hz (ECoG band 1: 0.04–1; ECoG band 2: 1–2; ECoG band 3: 2–4; ECoG band 4: 4–8; ECoG band 5: 8–16; ECoG band 6: 16–32; ECoG band 7: 32–64; and ECoG band 8: 64–128 Hz) by using standard numerical Gaussian filters. After normalization, the cross-correlation between the sound envelope and the corresponding ECoG traces of the different electrodes was separately computed for all frequency bands. We highlighted periodicities in the obtained cross-correlations by computing the periodogram of the signal by a standard fast Fourier transform algorithm (26, 27).

The ECoG signals averaged before evaluating the cross-correlation with the audio envelope started 500 ms before the audio track.

Shuffling and Statistical Analysis. Shuffling was implemented by slicing the electrode signal, corresponding to a given sentence or word, into time subintervals of increasing length, starting from 0.01 s with a 0.01 s step up to 0.4 s, and with a 0.05 step from 0.4 s up to the full time span. Time subintervals of equal duration were then randomly mixed before computing the cross-correlation of the shuffled signal with the unshuffled positive envelope of the audio track. Statistical analysis was performed on absolute values (modules) of cross-correlation. Statistical significance was assessed by parametric (unpaired Student t test) and nonparametric (Mann–Whitney test) methods when amplitude of correlations and delay of envelope of the audio over the ECoG over the envelope of the audio trace were compared. Distribution of electrodes simultaneously active was compared by χ^2 statistic. All statistical analyses were calculated by using MedCalc for Windows, version 14.8.1 (MedCalc Software).

ACKNOWLEDGMENTS. We thank Franco Bottoni for help with sound recording tools and Mauro Tachimiri, together with all the EEG technicians, for help with intraoperative recordings; Dr. Mauro Benedetti for preliminary discussions; the Neurocognition and Theoretical Syntax Research Center at Institute for Advanced Studies - Pavia for endowing G.A.with a fellowship; the Department of Electrical, Computer, and Biomedical Engineering at the University of Pavia for endowing G.A. with a fellowship; and Professor Stefano Cappa, Professor Noam Chomsky, and Professor Ian Tattersall for useful comments on the manuscript.

1. Giraud A-L, Poeppel D (2012) Cortical oscillations and speech processing: Emerging computational principles and operations. *Nat Neurosci* 15(4):511–517.
2. Pulvermüller F, et al. (2006) Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci USA* 103(20):7865–7870.
3. Nourski K-V, et al. (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci* 29(49):15564–15574.
4. Bouchard K-E, Mesgarani N, Johnson K, Chang E-F (2013) Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495(7441):327–332.
5. Mesgarani N, Cheung C, Johnson K, Chang E-F (2014) Phonetic feature encoding in human superior temporal gyrus. *Science* 343(6174):1006–1010.
6. Pasley B-N, et al. (2012) Reconstructing speech from human auditory cortex. *PLoS Biol* 10(1):e1001251.
7. Monson B-B, Han S, Purves D (2013) Are auditory percepts determined by experience? *PLoS ONE* 8(5):e63728.
8. Kubanek J, Brunner P, Gunduz A, Poeppel D, Schalk G (2013) The tracking of speech envelope in the human cortex. *PLoS ONE* 8(1):e53398.
9. Egidi G, Caramazza A (2013) Cortical systems for local and global integration in discourse comprehension. *Neuroimage* 71:59–74.
10. Ojemann G, Ojemann J, Lettich E, Berger M (1989) Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg* 71(3):316–326.
11. Sanai N, Mirzadeh Z, Berger M-S (2008) Functional outcome after language mapping for glioma resection. *N Engl J Med* 358(1):18–27.
12. Sahin N-T, Pinker S, Cash S-S, Schomer D, Halgren E (2009) Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326(5951):445–449.
13. Hanna J, et al. (2014) Early activation of Broca's area in grammar processing as revealed by the syntactic mismatch negativity and distributed source analysis. *Cogn Neurosci* 5(2):66–76.
14. Saberi K, Perrott D-R (1999) Cognitive restoration of reversed speech. *Nature* 398(6730): 760.
15. Moro A, et al. (2001) Syntax and the brain: Disentangling grammar by selective anomalies. *Neuroimage* 13(1):110–118.
16. Friederici A-D (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6(2):78–84.
17. Musso M, et al. (2003) Broca's area and the language instinct. *Nat Neurosci* 6(7): 774–781.
18. Pallier C, Devauchelle A-D, Dehaene S (2011) Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci USA* 108(6):2522–2527.
19. Moro A (2008) *The Boundaries of Babel. The Brain and the Enigma of Impossible Languages* (The MIT Press, Cambridge, MA).
20. Smith J-D, Wilson M, Reisberg D (1995) The role of subvocalization in auditory imagery. *Neuropsychologia* 33(11):1433–1454.
21. MacSweeney M, Capek C-M, Campbell R, Woll B (2008) The signing brain: The neurobiology of sign language. *Trends Cogn Sci* 12(11):432–440.
22. du Feu M, McKenna PJ (1999) Prelingually profoundly deaf schizophrenic patients who hear voices: A phenomenological analysis. *Acta Psychiatr Scand* 99(6):453–459.
23. Atkinson J-R (2006) The perceptual characteristics of voice-hallucinations in deaf people: Insights into the nature of subvocal thought and sensory feedback loops. *Schizophr Bull* 32(4):701–708.
24. Keller S-S, Crow T, Foundas A, Amunts K, Roberts N (2009) Broca's area: Nomenclature, anatomy, typology and asymmetry. *Brain Lang* 109(1):29–48.
25. De Mauro T, Mancini F, Vedovelli M, Voghera M (1993) *Lessico di Frequenza dell'Italiano Parlato* (ETAS, Milano).
26. Oppenheim A-V, Schafer R-W (1989) *Discrete Time Signal Processing* (Prentice Hall, Englewood Cliffs, NJ).
27. Orfanidis S-J (1995) *Introduction to Signal Processing* (Prentice Hall, Englewood Cliffs, NJ).

Supporting Information

Magrassi et al. 10.1073/pnas.1418162112

SI Materials and Methods

Participants and Cortical Mapping. The present study (named LIPS1) was approved by the Fondazione I.R.C.C.S. Policlinico S. Matteo institutional review board and ethics committee on human research. Patient testing, cortical mapping, and HDM placement were determined entirely by clinical criteria. All analyses, the results of which are presented in our study, were performed offline and did not interfere with the clinical management of the patients. Acute intraoperative recordings obtained from 16 Italian native speaker patients (12 males and 4 females) affected by primary or secondary malignant tumors growing in the dominant frontal, temporal, or parietal lobes were the subject of our study. All participants' clinical history was shorter than 3 mo, and none of the participants had a long-term history of epilepsy (only four of them had sporadic epileptic fits before surgery), making it unlikely that any long-term changes in connectivity and brain excitability resulting from chronic epilepsy could limit the relevance of our results for human physiology. Stimulation for finding areas of speech arrest were performed according to standard neurosurgical techniques for awake neurosurgical operations (1, 2). Biphasic square-wave pulses at 60 Hz delivered by handheld bipolar electrodes separated by a distance of 5 mm (Ojemann cortical stimulator; Radionics) were used. The intensity of the current used for cortical stimulation mapping varied from patient to patient, with the maximum intensity being that which did not produce after-discharges in the simultaneous ECoG traces. The position of the stimulating and recording electrodes was determined visually and recorded on a neuronavigator (Vectorvision Brain Lab). This allowed the unambiguous classification of all electrodes in the dominant frontal lobes either as electrodes corresponding to the anatomical Broca's area (3) or as electrodes that were outside it.

Neural and Audio Recordings. ECoG recordings were obtained using a multichannel electroencephalographer (System Plus; Micromed) with a sampling rate of 2,048 Hz and ADC resolution of 16 bit. One or two HDM grids with an interelectrode distance of 5 mm measured from center to center were simultaneously used for ECoG recordings. One HDM was made of 20 electrodes arranged in a rectangular array (1.5 × 2 cm), and a second was made of four electrodes arranged in a 2-cm-long row. The larger HDM was either centered over the main speech arrest area identified by direct cortical stimulation in the dominant frontal lobe of the hemisphere undergoing surgery (left, 13 patients; right, 2 patients) or the left superior parietal gyrus (two patients used as controls). The smaller HDM consisted of four electrodes arranged in a row and was positioned over the dominant temporal lobe in 10 patients who also had the large HDM positioned over the frontal lobe. The specific positions of all electrodes in the left hemisphere of all patients are shown in Fig. 1*B*. All ECoG signals were visually inspected by a standard viewer (EDF Browser), and traces and parts of traces, with obvious artifacts (including excessive electromagnetic noise from operating room equipment and poor contact with the cortical surface) were removed. The 50-Hz AC interference was suppressed, along with its harmonics (up to 150 Hz), by means of a digital comb filter.

Sound tracks and neural activities were simultaneously recorded during testing; sound was acquired by an H1 X/Y stereo microphone (Zoom H1; Zoom Corp.) placed at the base of the neck on the same side of the operated hemisphere at sampling rates of 96 KHz. In parallel, a low-resolution sound trace with a 2,048-Hz sampling rate was directly recorded in one channel of the ECoG and later exploited to synchronize the high-resolution sound trace and the electrode signals. To this end, a suitable triggering audio signal was used at the beginning of each recording session.

Neuropsychological Testing.

Linguistic items. Linguistic expressions were based on standard Italian taken from the "Lessico di Frequenza dell'Italiano Parlato" (4) and included either simple words or sentences. Words included six singular nouns depicting common objects that were either manipulable or not, two deverbal nouns with the same subfixation expressing actions, one complex name or number, one verb, and two words that were ambiguous and that could be interpreted either as objects or verbs. The number of syllables varied from two to a maximum of 10 (average, 3.5) and included all vocalic phonemes and all basic types of consonantic phonemes represented alphabetically. Sentences were all affirmative active present-tensed clauses and included six single and five complex clauses, including a declarative sentence. The number of syllables varied from seven to 13 (average, 10) and included all vocalic phonemes represented and all basic types of consonantic phonemes represented alphabetically: seven sentences included a nonexpressed subject as a well-known common property of Italian syntax. As for their semantics, all nouns and verbs were taken from high-frequency lexemes and were all describing simple conditions or states; only one sentence contained a semantic clash. All sentences were strictly unambiguous and were expressly designed not involve "garden path" effects to avoid rereading (5).

All linguistic items were randomly repeated at least five times (minimum, five; maximum, 12) during every testing section (each lasting about 5 min), and each testing section was also repeated at least two times, up to a maximum of 40 min.

All words and sentences were written in white characters on a black background shown on a computer screen, which was placed at a distance determined by the patient for effortless and comfortable reading.

Reading aloud. The patients were asked to read standard Italian words and sentences aloud, pronouncing all the utterances as they are used to while trying to keep the sound intensity as uniform as possible between different reading sessions. This resulted in a uniform mean absolute amplitude of the sound envelopes obtained from all patients (mean, 2.48 mV; SD, ±0.30 mV). The dimension of the fonts and the pace of presentation were determined by the patient for her or his maximal comfort in preliminary trials. These were administered before surgery to familiarize the patients with the procedure and the texts.

Reading mentally. After successfully completing the reading aloud testing, if the patient was still comfortable and attentive, we asked her or him to read mentally the same words and sentences she or he just read aloud without changing his or her reading pace. The patient was also explicitly instructed to avoid lip movement or other voluntary movement mimicking sound emission. During silent reading, an audio trace coming from the microphone on the neck was continuously recorded, and trials from patients that documented sound emission by the patient during mental reading were discarded.

Pushing the button. Three patients were also recorded while they were pushing a button with the hand contralateral to the operated dominant hemisphere. The button was pushed every time the drawing of a hand with the index finger pushing a red button over a black background was displayed on a computer screen. The image of the pushing finger was randomly alternated with black screens at the same pace rate used for the patient when she or he was tested for reading.

Signal Analysis. The frequency components of the high-resolution audio signal were separated into five contiguous bands (audio band 1: 750–1,250; audio band 2: 1,250–1,750; audio band 3: 1,750–2,250; audio band 4: 2,250–2,750; and audio band 5: 2,750–3,250 Hz). For each audio band, the sound envelope was extracted from the audio track by interpolation, followed by low-pass filtering. The ECoG and the envelope were separated into eight contiguous frequency bands in octave ratio, ranging from 0.04 to 128 Hz (ECoG band 1: 0.04–1; ECoG band 2: 1–2; ECoG band 3: 2–4; ECoG band 4: 4–8; ECoG band 5: 8–16; ECoG band 6: 16–32; ECoG band 7: 32–64; and ECoG band 8: 64–128 Hz) by using standard numerical Gaussian filters. The amplitudes of the ensuing signals were then normalized, imposing that the autocorrelation of each signal at zero delay was equal to one. After normalization, the cross-correlation between the sound envelope and the corresponding ECoG traces of the different electrodes was separately computed for all frequency bands.

The cross-correlations were computed over time intervals of several minutes, inclusive of many sentences and words, by standard algorithms (6, 7) on ECoG and audio envelope signals that were not averaged. This allowed a robust analysis of the spectral contributions of the cross-correlation. We then highlighted periodicities in the obtained cross-correlations by computing the periodogram of the signal by a standard fast Fourier transform algorithm (6, 7).

To increase the temporal resolution in the study of the correlation between sound envelope and ECoG during production of the linguistic items, the periods of the ECoG and audio traces containing the same linguistic item in a single trial were put in register and averaged before calculating the cross-correlation. Utterances were recognized by direct listening aided by suitable software (Audacity). The ECoG signals averaged before evaluating the cross-correlation with the audio envelope started 500 ms before the audio track. We extended by 500 ms the ECoG trace because the starting of Broca's area activity anticipates speech production by about 400 ms (8), and 100 ms were added as compensation for the patients' differences in speech rates.

Shuffling and Statistical Analysis. Shuffling was implemented by slicing the electrode signal, corresponding to a given sentence or word, into time subintervals of increasing length, starting from 0.01 s with a 0.01 s step up to 0.4 s, and with a 0.05 step from 0.4 s up to the full time span. Time subintervals of equal duration were then randomly mixed before computing the cross-correlation of the shuffled signal with the unshuffled positive envelope of the audio track. Depending on the number of the available subintervals, we calculated the average of the correlation values of 50 independent shuffled versions of the original utterance or the maximum number of the possible combinations of subintervals. We determined the maximum number of independent shuffled versions after we verified on selected words and sentences that the obtained correlation mean did not significantly change by averaging it over 20, 30, 50, and 100 shuffled cases. The minimum time interval of the slices (0.01 s) was selected because it is typically considered an appropriate sampling frame for human speech in artificial speech recognition technology and is below the 50 ms frame, the reversion of which does not affect speech comprehension (9).

Statistical analysis was performed on absolute values (modules) of cross-correlation. Only values of cross-correlation within a window of a 500-ms delay of the spoken utterances with respect to the ECoG trace were considered in the analyses. Statistical significance was assessed by parametric (unpaired Student *t* test) and nonparametric (Mann–Whitney test) methods when the amplitude of correlations and delay of envelope of the audio over the ECoG over the envelope of the audio trace were compared. Distribution of electrodes simultaneously active was compared by χ^2 statistic. All statistical analyses were calculated by using MedCalc for Windows, version 14.8.1 (MedCalc Software).

1. Ojemann G, Ojemann J, Lettich E, Berger M (1989) Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg* 71(3):316–326.
2. Sanai N, Mirzadeh Z, Berger M-S (2008) Functional outcome after language mapping for glioma resection. *N Engl J Med* 358(1):18–27.
3. Keller S-S, Crow T, Foundas A, Amunts K, Roberts N (2009) Broca's area: nomenclature, anatomy, typology and asymmetry. *Brain Lang* 109(1):29–48.
4. De Mauro T, Mancini F, Vedovelli M, Voghera M (1993) *Lessico di Frequenza dell'Italiano Parlato* (ETAS, Milano).
5. Hickok G (1993) Parallel parsing: Evidence from reactivation in garden-path sentences. *J Psycholinguist Res* 22:239–250.
6. Oppenheim A-V, Schafer R-W (1989) *Discrete Time Signal Processing* (Prentice Hall, Englewood Cliffs, NJ).
7. Orfanidis S-J (1995) *Introduction to Signal Processing* (Prentice Hall, Englewood Cliffs, NJ).
8. Sahin N-T, Pinker S, Cash S-S, Schomer D, Halgren E (2009) Sequential processing of lexical, grammatical, and phonological information within Broca's area. *Science* 326(5951):445–449.
9. Saberi K, Perrott D-R (1999) Cognitive restoration of reversed speech. *Nature* 398(6730):760.