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.