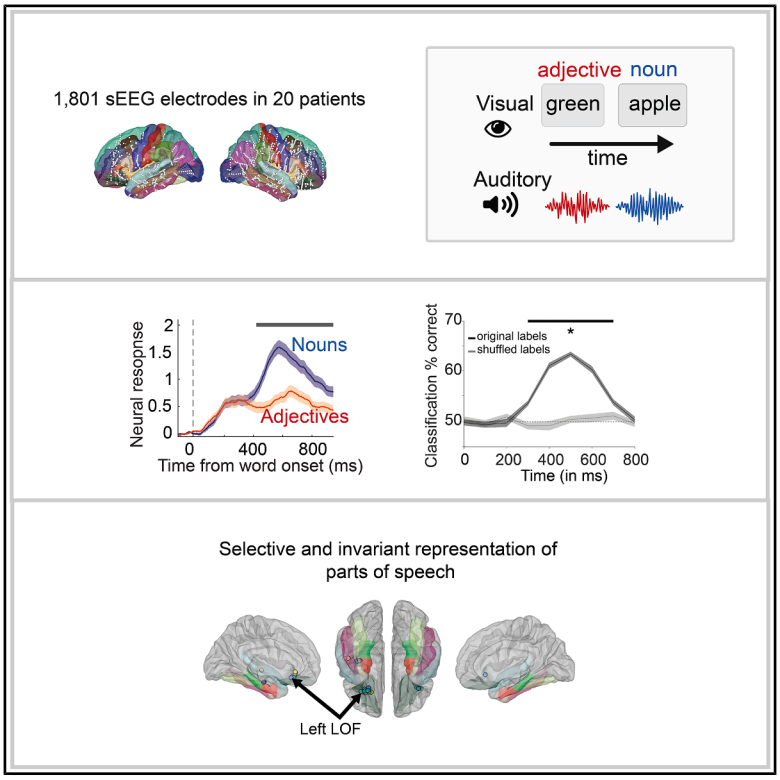
Article

Current Biology

# Invariant neural representation of parts of speech in the human brain

## Graphical abstract




## Authors

Pranav Misra, Yen-Cheng Shih, Hsiang-Yu Yu, Daniel Weisholtz, Joseph R. Madsen, Sceillig Stone, Gabriel Kreiman

## Correspondence

gkreiman@gmail.com

## In brief

Misra et al. use invasive neurophysiological recordings from the human brain to describe a highly circumscribed region within the left lateral orbitofrontal cortex that shows strong selectivity for parts of speech. The neural signals distinguish nouns from adjectives, generalizing across modalities, word properties, and languages.

## Highlights

- Neurophysiological recordings from 1,801 electrodes in 20 patients with epilepsy
- Left orbitofrontal cortex signals distinguish nouns from adjectives
- Part-of-speech signals generalize across visual and auditory modalities
- Part-of-speech representation is invariant to word properties and across languages

Misra et al., 2026, Current Biology *36*, 1–14
September 7, 2026 
https://doi.org/10.1016/j.cub.2026.07.029

CellPress

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029

Current Biology

CellPress
OPEN ACCESS

Article

# Invariant neural representation of parts of speech in the human brain

Pranav Misra,[1,5] Yen-Cheng Shih,[2,3] Hsiang-Yu Yu,[2,3] Daniel Weisholtz,[4] Joseph R. Madsen,[5] Sceillig Stone,[5] and Gabriel Kreiman[5,6,7,*]

[1]Harvard University, Cambridge, MA 02138, USA
[2]Department of Neurology, Taipei Veterans General Hospital, Taipei 112301, Taiwan
[3]School of Medicine, National Yang Ming Chiao Tung University College of Medicine, Taipei 112301, Taiwan
[4]Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
[5]Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA
[6]Center for Brains, Minds and Machines, Cambridge, MA 02137, USA
[7]Lead contact
*Correspondence: gkreiman@gmail.com
https://doi.org/10.1016/j.cub.2026.07.029

## SUMMARY

Elucidating the internal representation of language in the brain has major implications for cognitive science, brain disorders, and artificial intelligence. A pillar of linguistic studies is the notion that words have defined functions, often referred to as parts of speech. Here, we recorded invasive neurophysiological responses from 1,801 electrodes in 20 patients with pharmacologically resistant epilepsy while they were presented with two-word phrases consisting of an adjective and a noun. We observed neural signals that distinguished between these two parts of speech. The representation of parts of speech showed invariance across visual and auditory presentation modalities; robustness to word properties such as length, order, frequency, and semantics; and even generalization across different languages. The selective signals were circumscribed within a small region in the left lateral orbitofrontal cortex. This selective, invariant, and localized representation of parts of speech extends classic fronto-temporal language models, providing a target for mechanistic and causal tests of how the brain represents the basic building blocks of language, and introduces a systematic approach to elucidate the orchestration of more complex aspects of language.

## INTRODUCTION

Language plays a central role in most daily activities and in how we interact with others.[1,2] Early neurological and electrical stimulation studies highlighted specific brain regions that play essential roles in language understanding and production.[3–8] Despite the critical importance of language, progress toward elucidating the neural circuits underlying its representation has remained elusive due to the difficulties in investigating non-human animal models and the difficulty of examining neurophysiological responses in humans.

Neurophysiological experiments have begun to investigate neural signals associated with the presentation of individual words or short phrases,[9–17] the orthographic features of real versus pseudowords,[18–20] phonetic features of word comprehension[10,18,21,22] and production,[23–27] retrieval of semantic information for audiovisual naming,[10] and semantic encoding.[28] Beyond individual words, at the heart of linguistic structures is the notion that words serve specific functions within a sentence, including articles, nouns, adjectives, and verbs. These parts of speech (POS) are shared across languages, are combined according to defined grammatical rules, and play critical roles in AI algorithms.[1,11,29–37] Patients with brain lesions have shown deficits in the retrieval of nouns versus verbs.[17,19,38–42] However, previous studies lacked the spatial and temporal resolution necessary to resolve explicit neural circuits associated with POS processing and did not evaluate the degree of invariance in the representation of POS.

What would a representation for POS look like? Consider the adjective "green" and the noun "apple," combined to create the simple phrase "green apple." Fundamental constraints for such a representation should include the basic invariances underlying the cognitive understanding of this phrase. The basic desiderata for the representation of POS include invariance to (1) presentation modality (e.g., auditory versus visual), (2) specific noun or adjective (e.g., green or red), (3) position within a phrase (e.g., "green apple" versus "apple green"), (4) specific language in bilingual speakers (e.g., "green apple" in English versus "*manzana verde*" in Spanish), and (5) other word properties like their written length, number of syllables, and phoneme composition.

We investigated the representation of POS in humans by recording intracranial field potentials with high spatiotemporal resolution and high signal-to-noise ratio from 1,801 electrodes implanted in 20 participants with pharmacologically resistant epilepsy. We describe neural signals, especially in the left lateral orbitofrontal cortex (LOF), that selectively distinguish between nouns and adjectives. These selective signals for POS are robust when words are matched for orthography (e.g., word length), acoustic features (e.g., number of syllables), word sequence

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



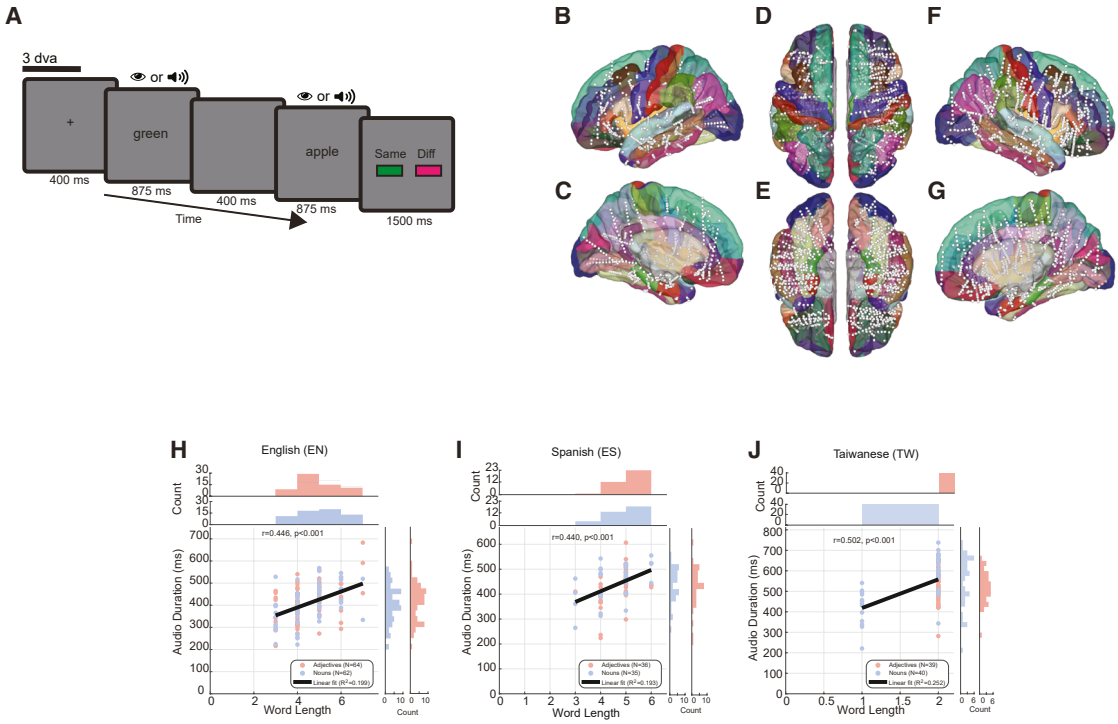



**Figure 1. Task schematic and electrode locations**

(A) Two words were sequentially presented either in visual modality or auditory modality. Participants indicated whether the two words were the same (e.g., "apple apple" or "green green," 8% of trials of each type) or different (e.g., "green apple" or "apple green": 42% of trials of each type, STAR Methods). In the 84% of trials where the two words were different, there was an adjective followed by a noun or a noun followed by an adjective.

(B–G) Location of all electrodes overlaid on the Desikan-Killiany atlas is shown with different views. Each white circle shows one electrode. (B) Left lateral view ($n$ = 693), (C) left medial view ($n$ = 693), (D) superior, whole brain view ($n$ = 1,801), (E) inferior, whole brain view ($n$ = 1,801), (F) right lateral view ($n$ = 1,108), (G) right medial view ($n$ = 1,108).

(H–J) Scatter of audio word duration ($y$ axis) in milliseconds and word length ($x$ axis) for nouns (blue) and adjectives (red) for English (EN) (H), Spanish (ES) (I), and Taiwanese (TW) words (J). For English, Spanish, and Taiwanese, mean audio durations were 421 ± 86, 438 ± 72, and 538 ± 103 ms, respectively. Nouns averaged 426 ± 89 ms (EN), 443 ± 70 ms (ES), and 549 ± 124 ms (TW), while adjectives averaged 414 ± 81 ms (EN), 434 ± 80 ms (ES), and 527 ± 81 ms (TW). Mean text lengths were 4.8 ± 1.2 characters (EN), 4.6 ± 0.8 characters (ES), and 1.8 ± 0.4 characters (TW). For nouns, text lengths averaged 4.9 ± 1.2 characters (EN), 4.5 ± 0.8 characters (ES), and 1.7 ± 0.5 characters (TW). For adjectives, text lengths averaged 4.5 ± 1.1 characters (EN), 4.7 ± 0.7 characters (ES), and 2.0 ± 0.0 characters (TW). Duration correlated with text length in all three languages ($R^2$ = 0.19–0.25, $p < 0.001$). Word-type comparisons between nouns and adjectives showed no differences in audio duration for any language ($p > 0.05$, $t$ test) (STAR Methods).

See also Figures S1–S3 and Tables S1–S3 and S6.

(e.g., noun or adjective at first or second position within a phrase), and frequency of occurrence. Remarkably, the representation of nouns versus adjectives generalizes across audio and visual modalities, across different semantic categories, and even across different languages.

## RESULTS

We recorded intracranial field potentials from 1,801 electrodes (gray matter: 840, white matter: 961) implanted in 20 participants with pharmacologically resistant epilepsy (Table S1). Participants heard (auditory modality) or read (visual modality) two words that were sequentially presented and indicated whether the words were the same or not (Figure 1A; STAR Methods). Participants' accuracy was 93.6% ±7.7% (here and throughout, mean ± SD, $p < 0.01$, two-tailed $t$ test). All electrode locations are shown in Figures 1B–1G (see also Tables S1 and S2). We used a bipolar reference, and we focus on the signals filtered in the gamma frequency band (65–150 Hz, STAR Methods), referred to as neural responses throughout and reported as normalized gamma power. Audio duration and word length for all stimuli are shown in Figures 1H–1J.

### Neural signals reflect visual, auditory, and multimodal inputs

We observed 565 electrodes (31.4% of the total) that responded to auditory stimuli (Figures S1A–S1C and S1G–S1I; STAR Methods) and 532 electrodes (29.5% of the total) that responded to visual stimuli (Figures S1D–S1F and S1G–S1I). The overall proportions and dynamics of visual- and auditory-responsive signals are consistent with previous work.[24,43] Of these electrodes, 293 responded to *both* auditory and visual stimuli (Figures S1G–S1I, left hemisphere: 147 [50.2%], right hemisphere: 146 [49.8%]). These 293 electrodes represent 16.3% of the total, 51.9% of the auditory-responsive electrodes, and 55.0% of the visually responsive electrodes. This number of audiovisual electrodes is unlikely to arise by chance from the number of auditory and visual electrodes ($p < 10^{-4}$, permutation test,

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



CellPress
OPEN ACCESS

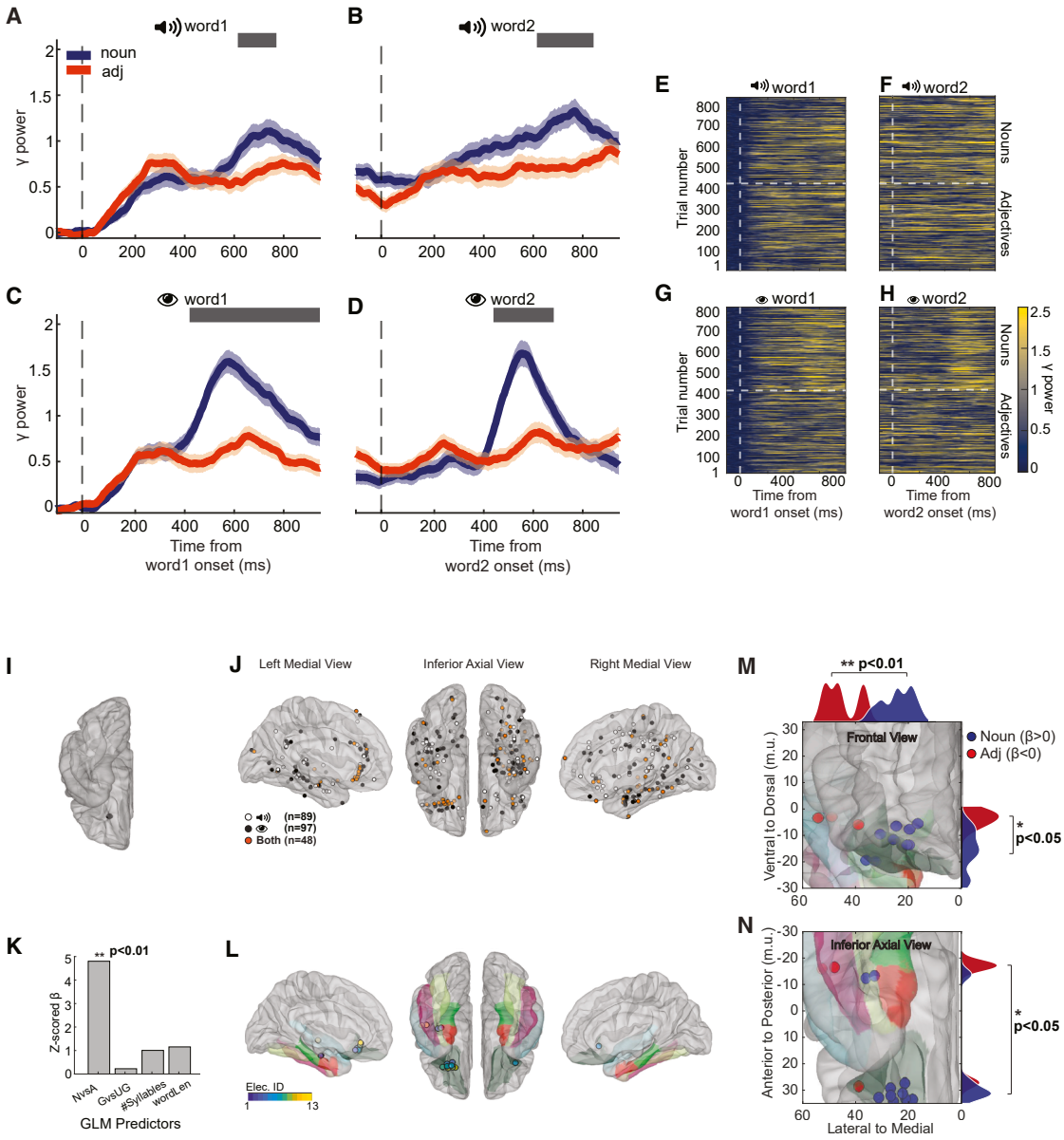



**Figure 2. Neural signals distinguish between different POS**

(A–D) Trial-averaged normalized gamma-band power of responses from an example electrode in the left LOF (see location in I) to nouns (blue) or adjectives (red) during presentation of auditory stimuli (A and B, $n$ = 435 grammatical and 432 ungrammatical trials) or visual stimuli (C and D, $n$ = 435 grammatical and 432 ungrammatical trials) aligned to the onset (vertical dashed line) of the first word (A and C) or second word (B and D). Shaded areas denote SEM. Horizontal gray lines denote windows of statistically significant differences between responses to nouns versus adjectives ($t$ test $p < 0.05$, Benjamini-Hochberg false detection rate, $q < 0.05$).

(E–H) Raster plots showing the responses in each individual trial (see color scale on bottom right). The blue and red curves in (A)–(D) correspond to the averages of noun and adjective trials, respectively, in (E)–(H).

(I) Location of the example electrode in the left LOF.

(J) Electrodes that were selective to differences between nouns and adjectives for audio-only (white; $n$ = 89; left = 30; right = 59), visual-only (black; $n$ = 97; left = 41; right = 56), and multimodal (orange; $n$ = 48; left = 21; right = 27) stimuli. Brain is shown in three views: left medial (left, $n$ = 92), inferior axial (center), and right medial views (right, $n$ = 142).

(K) $Z$ scored β coefficients for GLM used to predict AUC between 200 and 800 ms post-word onset, using four task predictors: noun versus adjectives, grammatically correct versus incorrect, number of syllables (auditory presentation), and word length (visual presentation). Asterisks denote statistically significant coefficients, corrected for multiple comparisons (STAR Methods).

(L) Electrodes that revealed statistically significant differences between nouns and adjectives for both audio and visual presentation ($n$ = 13 electrodes; colors indicate unique electrodes to match them across the different brain views). Electrodes whose responses were significantly explained only by the nouns versus adjective task predictor in the GLM are included in this plot. iELVis pullout factor = 0, opaqueness = 0.4. The fusiform (light green), parahippocampal (lime yellow),

*(legend continued on next page)*

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



$n = 10^6$ iterations). Figures S1J–S1M show the responses of an example audiovisual-responsive electrode in the left rostral middle-frontal gyrus, revealing strong evoked responses in trial-averaged signals (Figure S1J) and even in individual trials for both auditory (Figure S1K) and visual stimuli (Figure S1L).

To compare the response dynamics of auditory and visual responses, we calculated the time at which the neural signals reached half of the max amplitude (half-maximum time, arrows in Figure S1J). There was no significant difference between the half-maximum time for auditory-only electrodes (329 ± 187 ms, gray-left in Figure S1N) and visual-only electrodes (336 ± 174 ms, black-middle in Figure S1N) ($p > 0.05$, rank-sum test). Similarly, there was no significant difference between the half-maximum time for the audio and visual responses of audiovisual electrodes (379 ± 193 ms versus 341 ± 174 ms, half-black-right in Figure S1N, $p > 0.05$, rank-sum test). However, there was a small but significant difference between the half-maximum time for audio-only electrodes and the auditory responses of audiovisual electrodes ($p < 0.01$, rank-sum test). As expected, for the audio-only electrodes, the response area under the curve (AUC) for auditory stimuli (108 ± 100 $\mu V^2$/Hz-ms) was larger than for visual stimuli (44 ± 16 $\mu V^2$/Hz-ms) ($p < 10^{-4}$, rank-sum test, Figure S1O). Conversely, for the visual-only electrodes, the response AUC to auditory stimuli (40 ± 23 $\mu V^2$/Hz-ms) was smaller than to visual stimuli (53 ± 43 $\mu V^2$/Hz-ms) ($p < 10^{-4}$, rank-sum test, Figure S1P). For the audiovisual electrodes, the response AUC to auditory stimuli (71 ± 72 $\mu V^2$/Hz-ms) was slightly larger than to visual stimuli (54 ± 39 $\mu V^2$/Hz-ms) ($p < 0.01$, rank-sum test, Figure S1Q).

We compared the response effect sizes for visual versus auditory responses for all electrodes (Figure S2A; STAR Methods). As expected, for the audio-only electrodes (cyan hexagrams), the effect size for auditory stimuli (19.8% ± 5.4%) was larger than for visual stimuli (10.0% ± 4.4%, $p < 0.01$, rank-sum test). Similarly, for the visual-only electrodes (black squares), the effect size for visual stimuli (19.0% ± 4.8%) was larger than for auditory stimuli (10.2% ± 4.3%, $p < 0.01$, rank-sum test). For the audiovisual electrodes (orange triangles), the effect size for auditory stimuli (19.3% ± 5.5%) was not different from the one for visual stimuli (19.2% ± 4.8%, $p > 0.01$, rank-sum test). A similar scatterplot for the rostral middle-frontal region is shown in Figure S2B. Of the 41 regions in the Desikan-Killiany atlas where we had sampling (34 defined regions and 7 extra regions representing deep gray matter structures; STAR Methods; Figure 1; Tables S1 and S2), 13 regions had a significantly higher number of multimodal electrodes than expected by chance ($p < 0.01$, permutation test, $n = 10^6$ iterations, Figures S2A–S2C, indicated in bold in Table S2).

In sum, some electrodes responded exclusively to auditory stimuli, other electrodes exclusively on visual stimuli, and also other electrodes that revealed multimodal responses to both auditory and visual inputs. There were large amplitude differences for unimodal electrodes but only subtle differences in the latencies and amplitudes of responses to auditory versus visual stimuli for the multimodal electrodes.

## Multimodal neural signals distinguish different POS

We evaluated whether the neural signals differentiated between nouns and adjectives. Nouns and adjectives were matched for their number of syllables and word length to control for potential confounds not specific to POS (Table S3; STAR Methods). Figure 2 shows the responses of an example electrode in the orbital H-shaped sulcus within the left LOF (Figure 2I depicts the electrode location). The orbital H-shaped sulcus lies above the bone of the eye socket, where a butterfly-like gyrus can be seen, formed along H-shaped recessions of the sulcus. The neural responses are aligned to the word onset (vertical dashed line) for auditory presentation (Figures 2A and 2B) or visual presentation (Figures 2C and 2D) for the first (Figures 2A and 2C) or second (Figures 2B and 2D) word in each trial. This electrode showed multimodal responses triggered by both auditory and visual stimuli. The responses to nouns (blue) were stronger than adjectives (red) across all four conditions, including both word 1 and word 2, and both for visual and auditory stimuli. The differences between nouns and adjectives can be readily appreciated even in individual trials (Figures 2E–2H). These differences became significant at approximately 430 ms after word onset for visual presentation and 610 ms for auditory presentation.

In all, there were 89 electrodes (Figure 2J, white), 97 electrodes (Figure 2J, black), and 48 electrodes (Figure 2J, orange; total = 234) that showed a difference between nouns and adjectives for auditory stimuli only, visual stimuli only, or both modalities, respectively ($p < 0.05$, $t$ test, Benjamini-Hochberg false detection rate, $q < 0.05$, calculated per modality, STAR Methods). These electrodes represent 24% (89 + 48 = 137 out of 565), 27% (97 + 48 = 145 out of 532), and 16% (48 out of 293) of the auditory responsive, visually responsive, and multimodal responsive electrodes, respectively. It is unlikely that the 48 multimodal electrodes distinguishing nouns and adjectives could be ascribed to randomly sampling from the total of auditory and visual electrodes ($p < 10^{-4}$, permutation test, $n = 10^6$ iterations). Figure S2H shows the regions that had a larger number of multimodal electrodes distinguishing nouns versus adjectives compared with the null hypothesis obtained by randomly sampling from all auditory and visual electrodes, like the LOF shown in Figures S2E and S2G ($p < 10^{-4}$, permutation test, $n = 10^6$ iterations; see Figure S3A for a schematic of the electrode count at each step).

Next, we characterized the dynamics and effect size for selective responses. We calculated the selectivity response window

LOF (dark green-gray), entorhinal (red), inferior-temporal (magenta), and superior-temporal (cyan) regions are shown in the color scheme of the Desikan-Killiany atlas (Figures 1B–1G; STAR Methods).

(M and N) All electrodes from I projected onto the left hemisphere are shown on the frontal plane (M) and the axial plane (N, same plane as L). All the electrodes that respond more strongly to nouns, i.e., nouns versus adjectives $\beta > 0$ ($n = 10$ electrodes), are shown in blue and electrodes that responded more strongly to adjectives ($\beta < 0$, $n = 3$ electrodes), are shown in red. All units are in MNI305 coordinates. Kernel density curves (bandwidth 2) outline the marginal distributions of noun-preferring (blue) and adjective-preferring (red) electrodes along the medial-lateral axis (M and N: $x$ axis, zero being more medial), ventral-dorsal axis (M: right $z$ axis), and posterior-anterior axis (N: right $y$ axis). $p$ values indicate significant differences between the coordinates for noun- and adjective-preferring electrodes (rank-sum test).

See also Figures S2–S6 and Tables S4 and S5.

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



(max window) as the length of the longest continuous period with a significant difference between nouns and adjectives across both words (Figure S2D, similar to gray horizontal bars in Figures 2A–2D) and the effect size as the percentage change in the response with respect to baseline (Figure S2F). As expected, the max window and effect sizes for unimodal electrodes were consistent with their modality preferences (Figures S2D and S2F; cyan: audio, black: vision) and were enhanced for both vision and audio in the multimodal electrodes (Figure S2D and S2F; orange). The electrodes in the LOF are separately shown in Figures S2E and S2G (the example electrode from Figures 2A–2H is shown with a green arrow).

In sum, we observed electrodes that selectively distinguished between nouns and adjectives, both for auditory stimuli and visual stimuli, and irrespective of the specific order in which the noun and adjective were presented.

## Neural selectivity for nouns versus adjectives was robust to word properties, phrase grammar, usage frequency, and word subcategory

Even though nouns and adjectives were matched in their average number of syllables and word length, we asked whether these variables could still contribute to the neural responses differentiating nouns and adjectives. Additionally, each trial could be grammatically correct (e.g., "green apple") or incorrect (e.g., "apple green"). Therefore, we asked whether grammar could contribute to the neural differences between nouns and adjectives. To address these questions, we built a generalized linear model (GLM) for each electrode to predict its response AUC between 200 and 800 ms after word onset using four predictors: nouns versus adjectives, grammatically correct or not, word length (vision), or number of syllables (audition) (STAR Methods). The predictor coefficients in the GLM model for the example electrode in Figures 2A–2D show that only the nouns versus adjectives label significantly explained the neural responses for both auditory and visual presentations (Figure 2K). A total of 14 electrodes showed nouns versus adjectives as the *only* statistically significant predictor in the GLM analysis. 13/14 (93%) of these electrodes distinguished nouns versus adjectives for both auditory and visual inputs, such as the example electrode in Figures 2A–2I (one additional POS visual-only, Figures S4K–S4T).

Table S4 (POS-selective electrodes across brain regions) shows the distribution of electrodes distinguishing POS as the only GLM predictor between the left and right hemispheres for all brain regions, and Table S5 (noun- and adjective-preferring electrodes across subjects) shows the distribution across different participants. The locations of electrodes that robustly distinguished nouns and adjectives as the only predictor in the GLM (Figure 2L) revealed a cluster enriched in the left LOF. Within the left LOF, all of the electrodes (8 out of 8) were in the posterior part of the orbital H-shaped sulcus. We recorded from 113 electrodes in the LOF (left hemisphere: 38; right hemisphere: 75; Figures 1B–1G; Table S2). Figure S3B shows the number of electrodes along each step of processing for the LOF (see Figure S3A for all regions). Of the 38 left-LOF electrodes, 21% distinguished nouns from adjectives during both audio and visual presentation. In stark contrast, only 1.3% of the 75 LOF electrodes in the right hemisphere distinguished nouns from adjectives in both audio and vision (these hemispheric differences were statistically significant: $p < 10^{-4}$, permutation test, $n = 10^6$ iterations).

We had hypothesized that distinguishing POS constitutes a core component of language and would therefore be reflected *exclusively in both* visual and auditory modalities. Indeed, 13/14 (93%) of electrodes differentiating nouns from adjectives in the GLM did so in both modalities. In addition to these 13 electrodes, there was a small number of electrodes (2 auditory only and 1 visual only) that showed differences between nouns and adjectives in one modality but not the other. Unlike the electrodes in Figure 2L, for the 2 auditory-only electrodes, the number of syllables also significantly contributed toward explaining the neural responses. Figures S4A–S4J show the responses of an example electrode located in the right insula that showed a difference between nouns and adjectives during auditory presentation but *not* during visual presentation. The number of syllables also contributed to the response prediction for this electrode (Figure S4J). Conversely, Figure S4K–S4T shows the responses of an example electrode located in the left LOF, but outside the posterior part of the orbital H-shaped sulcus, that showed a clear difference between nouns and adjectives during visual presentation but *not* during auditory presentation. This electrode showed a small difference between nouns and adjectives also in the auditory modality, but this difference was not statistically significant. Figures S4U and S4V show the locations of auditory-only (white circles) and visual-only electrodes (black circle) in the left (Figure S4U) and the right hemispheres (Figure S4V), distinguishing between nouns and adjectives.

In all, the responses of 52 out of 234 electrodes were predicted by at least one variable in the GLM (Figure S3A). For 16 out of 52 electrodes (31%), POS was a significant predictor, and the responses of the other 36 electrodes were correlated with other variables. Of those 16 electrodes, the responses of 14 electrodes (88%) were explained by POS only (Table S4, POS-selective electrodes across brain regions), and for the remaining 2, the number of syllables was also a significant predictor (Figures S4U and S4V; see also Figures S3A and S3B). Figure S3C shows which predictors in the GLM significantly explained the non-POS responses in Figure S3A.

We assessed whether separate GLMs for each modality may reveal electrodes that better capture the contributions from other predictors. There was a slight increase in the number of electrodes modulated by other predictors in addition to POS ($n$ = 19, total POS electrodes). However, the number of electrodes whose responses were accounted for by POS only remained similar to the one with the combined GLM analyses ($n$ = 13 out of 19 POS-only electrodes).

Nouns and adjectives differ in their usage frequency. We asked whether the differences in the neural responses to nouns versus adjectives depended on usage frequency. To address this question, we randomly subsampled the trials to match the distribution of Google Books Ngram frequency (STAR Methods). The Google Books Ngram documents the usage frequencies of words in printed sources published between 1500 and 2022. Figure S5A shows matched noun and adjective distributions for the example electrode shown in Figures 2A–2I and 2K. This electrode showed differential responses between POS

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



for auditory (Figures S5B, S5C, S5F, and S5G) and visual (Figures S5D, S5E, S5H, and S5I) stimuli during word 1 (Figures S5B, S5D, S5F, and S5H) and word 2 (Figures S5C, S5E, S5G, and S5I), even after nouns and adjectives were matched for their frequency of occurrence. Of the 13 audiovisual electrodes where nouns versus adjectives was the only significant predictor in the GLM analysis, 6 electrodes (46%, 4 in the left LOF, and 2 in the left superior-temporal gyrus) robustly distinguished nouns and adjectives matched for their frequency of occurrence, like the example electrode in Figure 2, whereas the other electrodes maintained their statistically significant selectivity in most but not all conditions (note that the statistical power is lower in the matched analyses due to subsampling).

We considered the possibility that the frequency of occurrence and word surprisal may also contribute to neural responses. There was a significant negative correlation between frequency of occurrence and word surprisal (calculated using GPT-2, STAR Methods) for both word 1 ($r = -0.62$, $p < 0.001$, two-tailed $t$ test) and word 2 ($r = -0.39$, $p < 0.001$, two-tailed $t$ test). Thus, we made separate GLMs for each of these predictors given that GLMs are less effective with correlated variables. POS was the only GLM predictor that significantly explained the neural responses of this electrode for both variables. A majority of electrodes from Figure 2L revealed POS as the only significant contributor for GLMs with frequency of occurrence ($n = 8$ with POS as the only significant predictor; $n = 5$ electrodes ceased to have any significant predictor instead of an alternative predictor) and for word surprisal ($n = 11$).

There were two subcategories of nouns and two subcategories of adjectives within our stimulus set: animals/food and concrete/abstract (Table S3). We asked whether the electrodes that showed differential responses generalized across different word subcategories. The example electrode in Figures 2A–2I did not show differences between the two noun or adjective subcategories for either auditory stimuli (Figures S6A, S6B, S6F, and S6G), visual stimuli (Figures S6C, S6D, S6H, and S6I), word 1 (Figures S6A, S6C, S6F, and S6H), or word 2 (Figures S6B, S6D, S6G, and S6I). Of the 13 audiovisual electrodes where noun versus adjective was the only significant predictor in the GLM analysis, 8 electrodes (62%) showed generalization across different noun or adjective subcategories. The remaining 5 electrodes (38%) showed a statistically significant difference between the two noun subcategories or between the two adjective subcategories (Table S5, noun- and adjective-preferring electrodes across subjects; note that the statistical power is also lower in this analysis, which includes approximately half the number of trials). Figures S6J–S6R show one of the exceptions—an electrode in the left LOF that showed a significant response only for food nouns. This selectivity was particularly pronounced for visual stimuli (Figures S6L, S6M, S6Q, and S6R) but was also apparent for auditory stimuli (Figures S6J, S6K, S6O, and S6P) and was evident for both word 1 and word 2.

In sum, differences in selective responses to nouns versus adjectives were particularly prominent and clustered in the left LOF; persisted across different word lengths, whether the word was used in a grammatically correct phrase or not, after equalizing word occurrence frequency; and generalized across different noun or adjective subcategories.

## Neural signals enhanced for nouns versus adjectives were anatomically segregated

Of those electrodes uniquely selective for POS, 79% showed responses that were significantly stronger for nouns compared with adjectives ($\beta_{NvsA} > 0$), as illustrated by the example in Figures 2A–2I. The remaining 21% showed responses that were stronger for adjectives compared with nouns ($\beta_{NvsA} < 0$), as illustrated by the example in Figures S5J–S5S (Table S5, noun- and adjective-preferring electrodes across subjects). There were also marked differences in the dynamics underlying the responses to nouns and adjectives for auditory stimuli but not for visual stimuli. For auditory stimuli, the difference in the onset time between nouns and adjectives was larger for noun-preferring electrodes (550 ± 107 ms) than for adjective-preferring electrodes (312 ± 94 ms, rank-sum test, $p < 0.05$). For visual stimuli, the difference in the onset time between nouns and adjectives was not different between noun-preferring electrodes (425 ± 107 ms) and adjective-preferring electrodes (437 ± 134 ms, rank-sum test, $p > 0.05$). There was a significant correlation between auditory and visual difference onset times for noun-preferring electrodes (Pearson $R = 0.80$, $p < 0.01$, $n = 10$) but not for adjective-preferring electrodes (Pearson $R = -0.70$, $p > 0.05$, $n = 3$).

We observed an anatomical separation between these two groups of responses (Figures 2M and 2N, $x$ axis: lateral to medial, $y$ axis: anterior to posterior, $z$ axis: ventral to dorsal). We compared noun- versus adjective-preferring electrodes along 3 axes of Montreal Neurological Institute 305 Coordinates (MNI305, units abbreviated as m.u.).[44] Along the medial to lateral axis ($x$ axis in Figures 2M and 2N, zero being more medial), noun-preferring electrodes had a mean of 25.3 ± 6.2 m.u., and adjective-preferring electrodes had a mean of 47.3 ± 7.7 m.u. ($p < 0.01$, rank-sum test). Along the ventral-dorsal axis ($z$ axis in Figure 2M), noun electrodes had a mean of −12.2 ± 5.3 m.u., and adjective electrodes had a mean of −3.7 ± 1.7 m.u. ($p < 0.05$, rank-sum test). Along the posterior-anterior axis ($y$ axis in Figure 2N), noun electrodes had a mean of 21.4 ± 18.9 m.u., and adjective electrodes had a mean of −2.7 ± 25.8 m.u. ($p < 0.05$, rank-sum test). Table S4 (noun- and adjective-preferring electrodes across brain regions) summarizes the locations of noun- versus adjective-preferring electrodes across brain regions. Electrodes in the LOF tended to show stronger responses to nouns (~90% $\beta_{NvsA} > 0$, $p < 10^{-4}$, permutation test, $n = 10^6$ iterations).

## A population of electrodes in the lateral orbitofrontal cortex can distinguish nouns from adjectives in individual trials and generalizes across word positions and modalities

The raster plots in Figures 2E–2H and S4–S6 qualitatively show that the responses are strong enough to be discerned in individual trials. To assess whether information about POS was available in individual trials, we used a machine learning pseudopopulation approach by combining electrodes within anatomically defined brain regions in the Desikan-Killiany atlas.[45] We binned the responses in 100 ms windows and used the top-N principal components that explained more than 70% of the variance in the training data for all the electrodes. We trained a support vector machine (SVM) classifier with a linear kernel to distinguish

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



CellPress
OPEN ACCESS

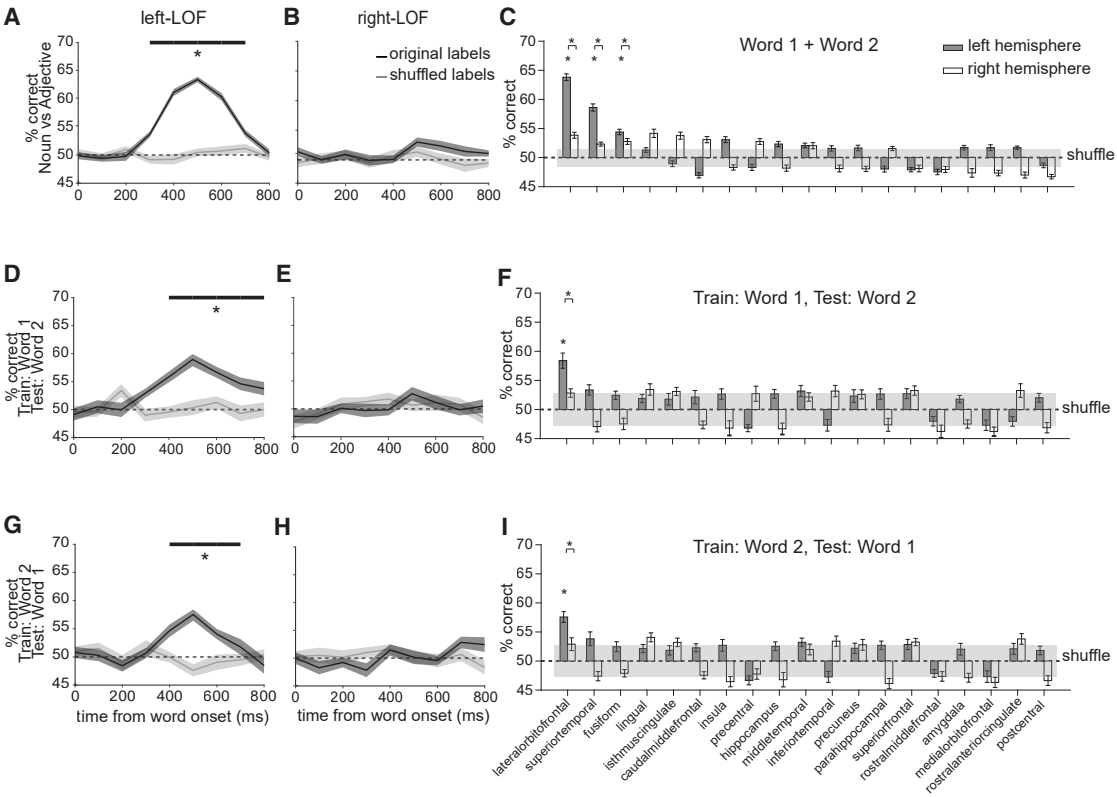



**Figure 3. Neural signals distinguishing nouns and adjectives in single trials generalize across word 1 and word 2**

(A, B, D, E, G, and H) Average cross-validated performance of a support vector machine (SVM) classifier (80% training/20% test) decoding nouns versus adjectives for all electrodes in the left lateral orbitofrontal cortex (LOF) (A, D, and G) or the right LOF (B, E, and H). The dotted horizontal black line shows the chance level. Shaded areas denote SEM. Solid horizontal black bar shows time points where performance significantly differed from chance (100 random shuffles, rank-sum test, $p < 0.01$). The inputs to the SVM included the top-N principal components of the electrode response that explained >70% variance for the training data at each time bin (STAR Methods). (A and B) Features from auditory and visual responses were combined and used for training and testing on a dataset of both word 1 and word 2 trials. (D and E) Generalization across word order was evaluated on a dataset where word 1 trials were used for training and word 2 trials were used for testing. (G and H) Training on word 2 and testing on word 1. Black: original labels; gray: shuffled labels (see Figure 4 for decoding performance when the number of electrodes was the same across all regions and both hemispheres).

(C, F, and I) Summary of the average of max-decoding performance for distinguishing nouns versus adjectives in each hemisphere (dark: left; white: right) for different brain regions. Bottom asterisks denote regions with significant decoding performance with respect to chance and performance from the real and null distribution do not overlap within 3 standard deviations of each other ($p < 0.01$, rank-sum test, corrected for multiple comparisons, STAR Methods). Shaded box: maximum of the mean ± SD. For the null distribution across all regions. Top asterisks with a U-bracket denote significant differences between decoding accuracy of the left versus the right hemisphere ($p < 0.01$, rank-sum test, corrected for multiple comparisons). Regions are sorted in descending order of performance in (C). (C) Classifiers were trained and tested with features from both word 1 and word 2 trials. (F) Classifiers were trained on word 1 trials and tested on word 2 trials. (I) Classifiers were trained on word 2 trials and tested on word 1 trials (see Figure 4 for controls on word 1-only, word 2-only, audio-only, visual-only, audio-to-vision, and vision-to-audio performance).

between nouns and adjectives and tested the classifier on held-out data (STAR Methods). In Figure 3, the same word could randomly be part of the training and test set while still using independent train and test data across trials or word positions. Figure 3 shows decoding accuracy for the left (Figures 3A, 3D, and 3G) and the right (Figures 3B, 3E, and 3H) LOF as a function of time from word onset. When trained using data from both word 1 and word 2 with combined auditory and visual features, there was a statistically significant decoding performance starting approximately at 300 ms after word onset and reaching a peak of 63.6% ± 1.1% at ~500 ms after word onset in the left LOF (Figure 3A). Statistical significance was assessed by comparing with a control where noun and adjective labels were randomly shuffled (STAR Methods). Even though there were almost twice as many electrodes in the right LOF compared with the left LOF (Figures 1B–1G; Table S2), decoding performance was higher for the left LOF compared with the right LOF (compare Figure 3A versus Figure 3B). The differences between the left and right LOF persisted after randomly subsampling to equalize the number of electrodes across hemispheres for all regions (Figures 4A and 4B).

In Figures 3A–3C, word 1 and word 2 were combined. Decoding performance in the left LOF was also significant when separately considering word 1 (Figures 4D–4F) or word 2 (Figures 4G–4I). Furthermore, the machine learning classifier could generalize across word positions, as evidenced by training on the responses to word 1 and testing on the responses to word 2 (Figures 3D–3F), and vice versa (Figures 3G–3I). Similarly,

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



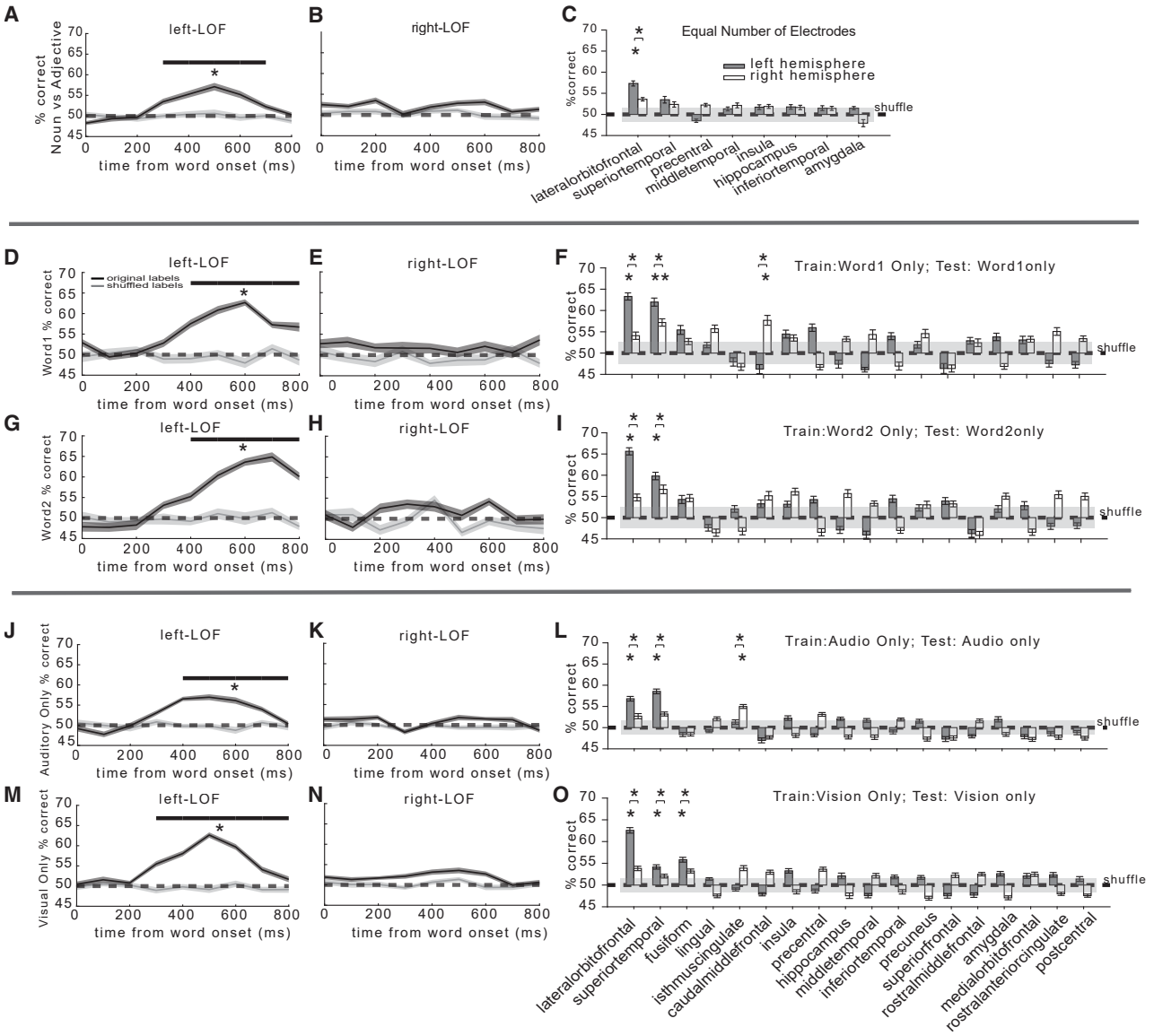



**Figure 4. Neural signals from left-LOF distinguish nouns and adjectives for (1) regions normalized for electrode numbers, (2) given word position in phrase, and (3) given modality**

(A and B) Average cross-validated performance of a support vector machine (SVM) classifier (80% training/20% test) decoding nouns versus adjectives for 8 randomly subsampled electrodes in the left lateral orbitofrontal cortex (LOF) (A) and in the right LOF (B).

Black: original labels; gray: shuffled labels. The dotted horizontal black line shows the chance level. Solid horizontal gray bar shows time points where decoding from correct labels significantly differed from that of shuffled labels (100 random shuffles of the data, rank-sum test, $p < 0.01$). The inputs to the SVM were 100 ms time bins from word onset containing the top-N principal components of the electrode response at each bin that explained >70% variance for the training data (STAR Methods).

(C) Summary of the average of max-decoding performance for distinguishing nouns versus adjectives across both hemispheres (left hemisphere: dark gray bars; right hemisphere: white bars) for different brain regions when a total of 8 electrodes were taken from each hemisphere in each region for the decoding. Regions with fewer than 8 electrodes in either hemisphere were omitted.

Asterisk: significant hemisphere within a Desikan-Killiany-defined brain region ($p < 0.01$, rank-sum test, corrected for multiple comparisons, and performance from the real and null distributions do not overlap within 3 standard deviations of each other) (STAR Methods). Gray box: maximum mean ± SD for the null distribution across all regions. Asterisk with a U-bracket: significant difference between decoding accuracy of the left versus the right hemisphere ($p < 0.01$, rank-sum test, corrected for multiple comparisons). Regions are sorted in descending order of performance in (C).

(D, E, G, H, J, K, M, and N). Average cross-validated performance of a SVM classifier (80% training/20% test) decoding nouns versus adjectives for all electrodes in the left LOF (D, G, J, and M) and in the right LOF (E, H, K, and N). The dotted horizontal black line shows the chance level. Shaded areas denote SEM. Solid horizontal black bar shows time points where performance significantly differed from chance (100 random shuffles, rank-sum test, $p < 0.01$). The inputs to the SVM included the top-N principal components of the electrode response that explained >70% variance for the training data at each time bin (STAR Methods).

*(legend continued on next page)*

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



CellPress
OPEN ACCESS

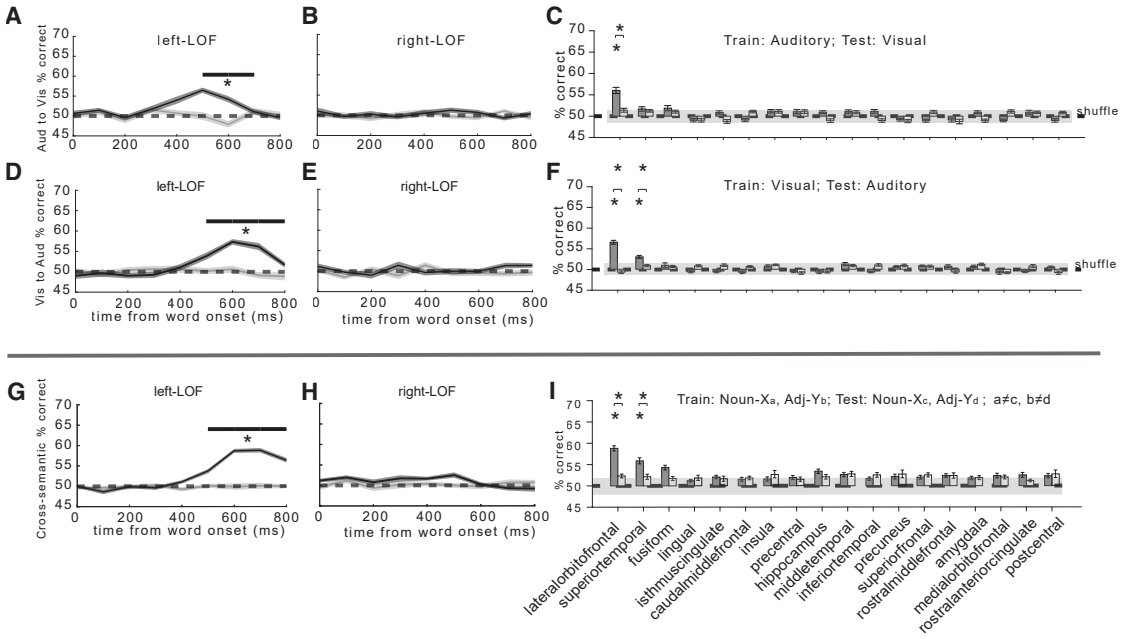



**Figure 5. Neural signals from left-LOF distinguish nouns and adjectives (1) across modalities, and (2) word subclasses**

(A, B, D, and E) Average cross-validated performance of a support vector machine (SVM) classifier (80% training/20% test) decoding nouns versus adjectives for all electrodes in the left LOF (A and D) and the right LOF (B and E). Using a combined dataset of word 1 and word 2 trials, the classifier was trained on auditory features and tested on visual features (A and B) and trained on visual features and tested on auditory features (D and E). SVM inputs were the top-N principal components of the electrode response that explained >70% variance for the training data at each time bin (STAR Methods). The dotted horizontal black line shows the chance level. Shaded areas denote SEM. Solid horizontal black bar shows time points where performance significantly differed from chance (100 random shuffles, rank-sum test, $p < 0.01$).

(C and F) Max-decoding accuracy across regions (left: dark; right: white) for the audio (train) → vision (test) (C) and vision → audio (F) generalization conditions. Bottom asterisks: significant above chance; top U-bracket asterisks: left-right differences ($p < 0.01$, corrected).

(G and H) Average cross-validated performance of an SVM in the left (G) and right LOF (H), trained on noun-Xa versus adjective-Yb and tested on noun-Xc versus adjective-Yd, where {Xa, Xc} and {Yb, Yd} are different subclasses of nouns and adjectives, respectively (e.g., Xa = animal nouns, Xc = food nouns). This yields 4 combinations for testing noun-versus-adjective decoding across semantic subcategories. The average across these 4 tests for 100 random shuffles is plotted. Solid horizontal black bar shows time points where performance significantly differed from chance (100 random shuffles, rank-sum test, $p < 0.01$).

(I) Summary of average max-decoding performance for distinguishing nouns versus adjectives across subclasses in each hemisphere (dark: left; white: right) for different brain regions. Bottom asterisks: significant above chance; top U-bracket asterisks: left-right differences ($p < 0.01$, corrected).

See also Figure S6.

auditory and visual trials were combined in Figures 3A–3C. Decoding performance in the left LOF was also significant when separately considering auditory trials (Figures 4J–4L) and visual trials (Figures 4M–4O). The machine learning classifier could generalize across modalities, as evidenced by training on auditory trials and testing on vision trials (Figures 5A–5C) and vice versa (Figures 5D–5F). Furthermore, the classifier also generalized when we trained on given subclasses of nouns and adjectives (e.g., food nouns versus abstract adjectives) and tested with different subclasses (e.g., animal nouns versus concrete adjectives) (Figures 5G–5I). To investigate the contribution of multimodal sites in decoding, we removed the multimodal responsive electrodes (Table S2) from the left LOF before performing the decoding analyses. After the removal of the multimodal responsive electrodes, the performance of the decoder using left LOF was not different from chance (51.9% ± 3.4%, rank-sum test, $p > 0.05$). However, when we exclusively used multimodal electrodes, the performance was significantly greater than chance (63.8% ± 2.6%, rank-sum test, $p < 0.01$). In sum, the encoding of POS generalized across word position and sensory modality.

We extended the analyses in Figures 3A, 3B, 3D, 3E, 3G, and 3H to all other regions. In addition to the left LOF, the left superior-temporal cortex and the left fusiform cortex also showed statistically significant decoding performance (Figure 3C). However, in contrast to the results for the left LOF, the decoding results for other regions were less robust (Figure 4C) and did not consistently generalize across word position (Figures 3F and 3I) or across modalities (Figures 5C and 5F).

Features from auditory and visual responses were combined and used for training and testing on datasets of word 1 (D and E) and word 2 trials (G and H). Using a combined dataset of word 1 and word 2 trials, the decoding performance was evaluated for audio-only (J and K) and vision-only features (M and N).

(F, I, L, and O) Max-decoding accuracy across regions (left: dark; right: white) for the same conditions. Bottom asterisks: significant above chance; top U-bracket asterisks: left-right differences ($p < 0.01$, corrected).

See also Figure S4.

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



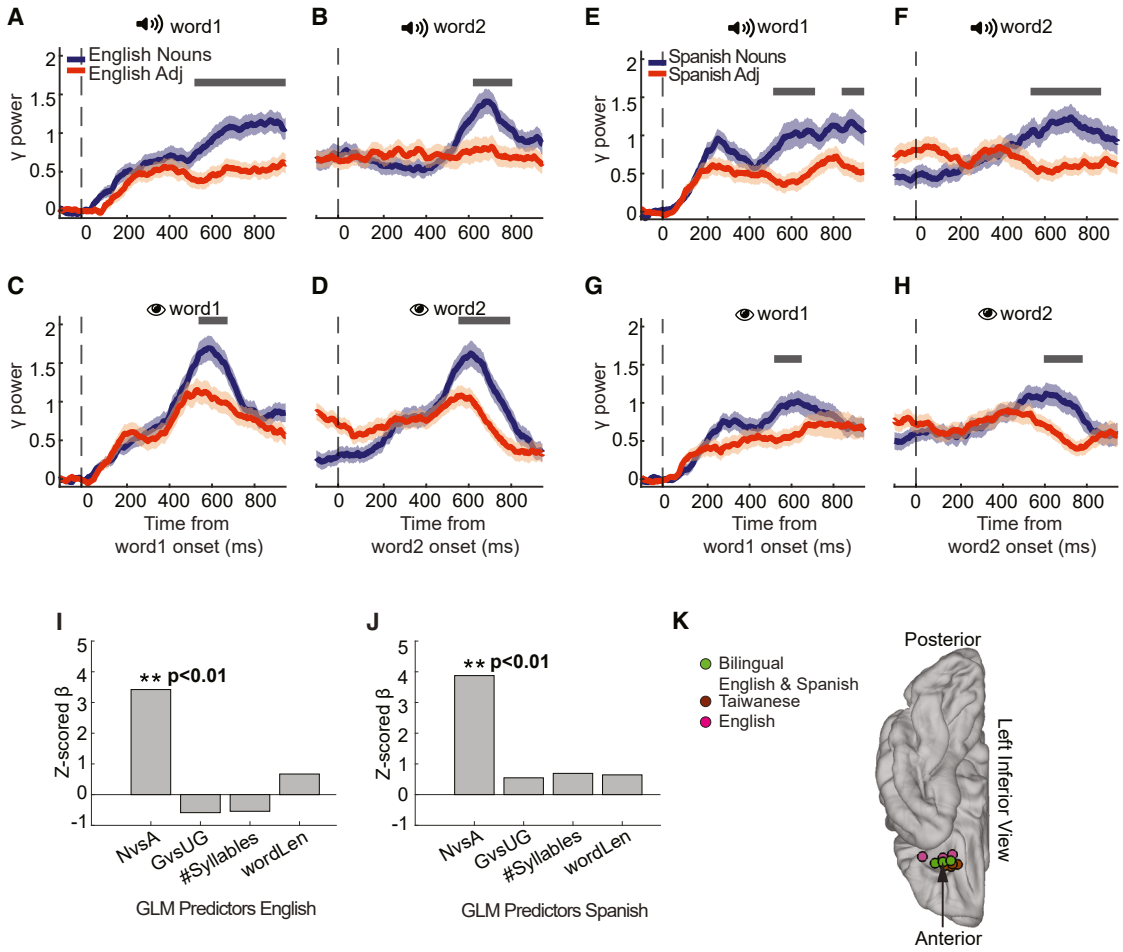



**Figure 6. Neural signals in left LOF generalize across languages in a bilingual participant and in monolingual participants**

(A–H) Trial-averaged responses of an electrode in the left LOF from a bilingual patient. The format follows Figures 2A–2D. (A–D) English words (audio: $n = 190$ grammatical and 185 ungrammatical trials; vision: $n = 189$ grammatical and 191 ungrammatical trials). (E–H) Spanish words (audio: $n = 184$ grammatical and 184 ungrammatical trials; vision: $n = 184$ grammatical and 186 ungrammatical trials). Auditory responses (A, B, E, and F). Visual responses (C, D, G, and H). Word 1 (A, C, E, and G) and word 2 (B, D, F, and H).

(I and J) $Z$ scored β coefficients for GLM to predict area under the curve (AUC) for the English experiment (I) and for the Spanish experiment (J). The AUC was computed between 200 and 800 ms post-word onset using four task predictors: noun versus adjectives, grammatical versus ungrammatical, number of syllables (auditory presentation), and word length (visual presentation). Asterisks denote statistically significant coefficients corrected for multiple comparisons (STAR Methods). The word order for grammatically correct trials in English is an adjective followed by a noun, such as "green apple." This word order gets flipped in grammatically correct Spanish trials.

(K) Inferior view of all the 9 out of 38 electrodes (8 audiovisual: Figure 2L, 1 visual-only: Figures S4U and S4V; see Tables S4 and S5) in the left LOF that showed noun-versus-adjective differences across different languages in which the experiment was conducted (significant nouns versus adjectives β, $p < 0.01$ corrected for multiple comparisons). These electrodes come from 4 different subjects. Electrodes from the bilingual patient are in green with a black arrow indicating the example electrode. Electrodes from one monolingual English patient are in pink, and those from 2 monolingual Taiwanese patients are in brown.

## Multimodal neural signals distinguishing different POS are conserved across languages

One of the participants was fluent in two languages, English and Spanish, providing an opportunity to ask whether the neural signals discriminating between different POS were language-specific or showed invariance across languages. All the words were translated into Spanish by a native Spanish speaker, and the task was repeated in both languages. Figures 6A–6H show the responses of an example electrode located in the left LOF (Figure 6K). This electrode showed a stronger response to nouns compared with adjectives for auditory stimuli (Figures 6A, 6B, 6E, and 6F), for visual stimuli (Figures 6C, 6D, 6G, and 6H), for word 1 (Figures 6A, 6C, 6E, and 6G), and for word 2 (Figures 6B, 6D, 6F, and 6H). Interestingly, the separation between nouns and adjectives was evident both when the words were presented in English (Figures 6A–6D) and Spanish (Figures 6E–6H). The GLM analysis showed that nouns versus adjectives were the only significant predictors in English (Figure 6I) and Spanish (Figure 6J). In all, there were three electrodes in this participant showing a multimodal response selective for POS. All three of these electrodes were in the left orbital H-shaped sulcus within the LOF (Figure 6K, green).

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029





In addition to this bilingual participant, the task was run with monolingual participants who spoke English ($n$ = 16 participants) and monolingual participants who spoke the Taiwanese dialect ($n$ = 3 participants, Table S1). In Figure 6K, we show all electrodes from the left LOF that showed POS encoding (Table S5, distribution of POS-selective electrodes in the left LOF across subjects). LOF electrodes in 4 out of 8 participants (50%) with coverage in the region exhibited differences between nouns and adjectives. We also indicate the language in which this difference was observed (English: pink; Taiwanese: brown; or bilingual English/Spanish: green). Electrodes separating POS from monolingual participants were also clustered in the same region.

In sum, electrodes in the left LOF robustly distinguished between POS for both auditory and visual presentations of stimuli across participants speaking English, Taiwanese, or Spanish. Additionally, in one bilingual subject, the same electrodes showed robust POS selectivity that generalized across English and Spanish.

## DISCUSSION

We described neurophysiological signals that selectively discriminate between two POS, nouns and adjectives (Figure 2). This selectivity was robust to word length, number of syllables, and word occurrence statistics (Figure 2). This selectivity for POS generalized across sensory modalities (Figures 2, 4, and 5), word position, grammar and motor outputs (Figures 2, 3, 4, 5, and 6), and semantic subgroups of nouns and adjectives (Figure S6). These neurophysiological signals enabled discrimination between POS in single trials (Figures 2, 3, 4, and 5). Electrodes that uniquely distinguished nouns from adjectives were clustered within a small region of the left LOF (posterior H-shaped sulcus, Figures 2 and 6). POS selectivity and invariance were evident in English-speaking and Taiwanese-speaking participants (Figure 6). Interestingly, in a bilingual participant, the same left-LOF electrodes distinguished nouns and adjectives in both English and Spanish (Figure 6).

Clinically, a focal LOF locus carrying invariant POS signals may inform language mapping and risk assessment for resections. Given the strong lateralization for language, it is extremely important to precisely understand the relevant neural structures in these patients. These results could help guide surgical approaches for epilepsy.

Some words can be used both as a noun *and* as an adjective (e.g., a *light* color versus turn on the *light*). In most instances, one usage is more frequent than the other. We selected nouns and adjectives that are highly overrepresented in their labeled POS (Table S6). Similarly, some words can be used both as a noun and as a verb (e.g., "long *race*" versus "*race* you to the top"). All the nouns in this study are highly overrepresented in their usage as nouns versus verbs (Table S6).

In English, adjectives precede nouns. Other languages reverse this order. In Spanish, adjectives typically follow nouns (though the English order can also be used). Many electrodes demonstrated strong selectivity for POS, irrespective of their position within the two-word phrases. Furthermore, in one bilingual participant, the neural responses separated nouns and adjectives in both languages. This bilingual experiment emphasizes that POS selectivity and invariance cannot be explained solely by word order, first-word surprisal, or word frequency.

Non-invasive scalp E/MEG has revealed correlates of language processing with latencies from 100 to 600 ms.[46] We show an invariant distinction between nouns and adjectives in the LOF commencing at 400 ms after stimulus onset, well after the onset of modality-specific purely visual or auditory signals.

We can interpret the word *cat* when uttering the word, writing it, listening to it, reading it, and even when examining a photograph of a cat. It is tempting to speculate that there may be an invariant representation of language concepts in the brain. Several studies have examined putative correlates of language processing using only unimodal signals.[11–15,18,24,29,34,35] While some electrodes distinguished between POS only in the auditory or visual stimuli, those responses could be partly explained by other variables, including number of syllables, word frequency, or grammar. Using strict criteria and after controlling for confounding variables, most electrodes that distinguished POS showed selectivity during *both* auditory and visual presentation. Future work should evaluate whether the same electrodes also distinguish POS when participants utter words, write them, or when examining images. An intriguing study described neurons in the medial temporal lobe responding selectively to images and their corresponding text or sound.[47,48] However, those neurons do not show POS selectivity and instead are connected with memory formation.[49,50]

The LOF spans Brodmann areas (BAs) 10, 11, 12, 13, and 47.[51–53] Neuroanatomical tracings from rats, mice, and macaques have identified LOF as a nexus of many inputs,[53] conveying multisensory information. The LOF has been associated with many cognitive functions, including multisensory integration, working memory, long-term memory consolidation, reward processing, social interactions, decision-making, and emotion processing.[50,52,54–58] This heterogeneity might be partly ascribed to investigations probing different cognitive tasks. Given the prominent role of language in cognition, it is conceivable that previous studies that describe other roles of the LOF did not probe its possible associations with language. However, it is even more likely that descriptors like LOF that refer to such large brain areas would inevitably fail to uncover specific functionality. These results point to a highly circumscribed location within the left LOF, the posterior part of the H-shaped sulcus. This location overlaps with BA13-lateral and BA47-medial and has strong convergence of auditory and visual inputs.[52,53,59] Work on primary progressive aphasia and frontotemporal lesions implicates the orbitofrontal cortex in word and sentence comprehension deficits.[7,28,59–61] The LOF, dorsal premotor cortex, temporoparietal junction (canonical Wernicke's area), and pars opercularis were associated with sentence comprehension and grammatical production aphasias. Word comprehension and naming deficits were assessed using binary perceptual choice tasks, implicating the LOF and the anterior temporal lobe (ATL). Consistent with extensive work documenting the lateralization of language functions, the results presented here also show a strong predominance of the left hemisphere in the representation of POS, despite the fact that there were more electrodes in the right hemisphere.

Several limitations are worth noting. First, the results are derived from patients with epilepsy, the predominant way to

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



access neurophysiological signals from the human brain.[62,63] While all patients used language fluently, epilepsy could potentially impact the representation of language. Second, electrode locations are strictly dictated by clinical criteria. Our sampling of brain activity is extensive but not exhaustive (Figure 1; Tables S1 and S2). Other areas not examined here may also reveal neural correlates of POS. Additionally, the limitations in the number of participants and electrodes, especially a single bilingual participant, should be considered when interpreting the results. Third, this work focuses on two POS; future work should examine the representation of pronouns,[64] verbs, adverbs, prepositions, and conjunctions. Finally, our work provides a *correlate* of the representation of POS, but future work should evaluate whether such signals are *causally* required for online language interpretation.

These results provide initial glimpses into highly localized structures that represent a fundamental component of language. The representation of nouns versus adjectives is invariant to the presentation modality, word properties, grammar, and the noun and adjective subcategories tested here and even generalizes across different languages. These findings highlight interesting questions about how next-token prediction in language models relates to the brain's implicit biases that give rise to stable, invariant representations of POS. These observations open the doors to begin to elucidate the neural representation of more complex language concepts and to bridge work in linguistics to their underlying neural representations.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Gabriel Kreiman (gkreiman@gmail.com).

### Materials availability

All the stimuli are publicly available at the repository listed below.

### Data and code availability

All data and code are publicly available at the following site: https://kreimanlab.com/code/neural-coding-for-language-representations-in-the-human-brain/.

The pseudocode is included in the accompanying *Readme.docx* file in the same repository.

## ACKNOWLEDGMENTS

We thank Marcelo Armendariz and Yen-Ling Kuo for helping translate the task to Spanish and Taiwanese, respectively. We thank Professors Alfonso Caramazza and Boris Katz for suggestions and insightful comments throughout this work. This work was supported by NIH grant R01EY026025 and NSF grant CCF-1231216.

## AUTHOR CONTRIBUTIONS

The experiment was designed by P.M. and G.K. All the surgeries were performed by Y.-C.S., H.-Y.Y., J.R.M., and S.S. D.W. helped coordinate all the experiments. All the data were collected and analyzed by P.M. The manuscript was written by P.M. and G.K. with feedback from all the authors.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Preregistration
  - Participants
  - Recordings and electrode locations
  - Experiment design
- METHOD DETAILS
  - Preprocessing
  - Responsive electrodes
  - Part-of-speech selectivity
  - Generalized linear model (GLM)
  - Surprisal calculation
  - Anatomical comparisons
  - Decoding analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical criteria for electrode responsiveness
  - *t*-tests for noun vs adjective selectivity and false discovery rate correction
  - Ranksum tests
  - GLM coefficient significance testing
  - Permutation tests for anatomical specificity
  - Decoding significance tests
  - Correlation *p*-values

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2026.07.029.

Received: August 11, 2025
Revised: April 28, 2026
Accepted: July 13, 2026

## REFERENCES

1. Chomsky, N. (1995). The Minimalist Program (MIT Press).
2. Scott, S.K. (2019). From speech and talkers to the social world: the neural processing of human spoken language. Science *366*, 58–62. https://doi.org/10.1126/science.aax0288.
3. K.M. Heilman, and E. Valenstein, eds. (1993). Clinical Neuropsychology, Third Edition (Oxford University Press).
4. Ojemann, G.A., Ojemann, J.G., Lettich, E., and Berger, M.S. (1989). Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in s117 patients. J. Neurosurg. *71*, 316–326. https://doi.org/10.3171/jns.1989.71.3.0316.
5. Petrides, M. (2014). Neuroanatomy of Language Regions of the Human Brain (Academic Press).
6. Mahon, B.Z., and Caramazza, A. (2009). Concepts and categories: a cognitive neuropsychological perspective. Annu. Rev. Psychol. *60*, 27–51. https://doi.org/10.1146/annurev.psych.60.110707.163532.
7. Mesulam, M.M., Thompson, C.K., Weintraub, S., and Rogalski, E.J. (2015). The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. Brain *138*, 2423–2437. https://doi.org/10.1093/brain/awv154.
8. Warrington, E.K., and Shallice, T. (1984). Category specific semantic impairments. Brain *107*, 829–854. https://doi.org/10.1093/brain/107.3.829.
9. Nourski, K.V., Steinschneider, M., Rhone, A.E., Kovach, C.K., Kawasaki, H., and Howard, M.A., 3rd. (2022). Gamma activation and alpha suppression within human auditory cortex during a speech classification task.

J. Neurosci. *42*, 5034–5046. https://doi.org/10.1523/JNEUROSCI.2187-21.2022.

10. Forseth, K.J., Kadipasaoglu, C.M., Conner, C.R., Hickok, G., Knight, R.T., and Tandon, N. (2018). A lexical semantic hub for heteromodal naming in middle fusiform gyrus. Brain *141*, 2112–2126. https://doi.org/10.1093/brain/awy120.

11. Murphy, E., Woolnough, O., Rollo, P.S., Roccaforte, Z.J., Segaert, K., Hagoort, P., and Tandon, N. (2022). Minimal phrase composition revealed by intracranial recordings. J. Neurosci. *42*, 3216–3227. https://doi.org/10.1523/JNEUROSCI.1575-21.2022.

12. Keshishian, M., Akkol, S., Herrero, J., Bickel, S., Mehta, A.D., and Mesgarani, N. (2023). Joint, distributed and hierarchically organized encoding of linguistic features in the human auditory cortex. Nat. Hum. Behav. *7*, 740–753. https://doi.org/10.1038/s41562-023-01520-0.

13. Sinai, A., Bowers, C.W., Crainiceanu, C.M., Boatman, D., Gordon, B., Lesser, R.P., Lenz, F.A., and Crone, N.E. (2005). Electrocorticographic high gamma activity versus electrical cortical stimulation mapping of naming. Brain *128*, 1556–1570. https://doi.org/10.1093/brain/awh491.

14. Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. *19*, 158–164. https://doi.org/10.1038/nn.4186.

15. Cometa, A., d'Orio, P., Revay, M., Bottoni, F., Repetto, C., Lo Russo, G.L., Cappa, S.F., Moro, A., Micera, S., and Artoni, F. (2023). Event-related causality in stereo-EEG discriminates syntactic processing of noun phrases and verb phrases. J. Neural Eng. *20*, 026042. https://doi.org/10.1088/1741-2552/accaa8.

16. Artoni, F., d'Orio, P., Catricalà, E., Conca, F., Bottoni, F., Pelliccia, V., Sartori, I., Lo Russo, G.L., Cappa, S.F., Micera, S., et al. (2020). High gamma response tracks different syntactic structures in homophonous phrases. Sci. Rep. *10*, 7537. https://doi.org/10.1038/s41598-020-64375-9.

17. Crepaldi, D., Berlingeri, M., Paulesu, E., and Luzzatti, C. (2011). A place for nouns and a place for verbs? A critical review of neurocognitive data on grammatical-class effects. Brain Lang. *116*, 33–49. https://doi.org/10.1016/j.bandl.2010.09.005.

18. Woolnough, O., Donos, C., Rollo, P.S., Forseth, K.J., Lakretz, Y., Crone, N.E., Fischer-Baum, S., Dehaene, S., and Tandon, N. (2021). Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. Nat. Hum. Behav. *5*, 389–398. https://doi.org/10.1038/s41562-020-00982-w.

19. Vigliocco, G., Vinson, D.P., Druks, J., Barber, H., and Cappa, S.F. (2011). Nouns and verbs in the brain: a review of behavioural, electrophysiological, neuropsychological and imaging studies. Neurosci. Biobehav. Rev. *35*, 407–426. https://doi.org/10.1016/j.neubiorev.2010.04.007.

20. Castellucci, G.A., Kovach, C.K., Howard, M.A., 3rd, Greenlee, J.D.W., and Long, M.A. (2022). A speech planning network for interactive language use. Nature *602*, 117–122. https://doi.org/10.1038/s41586-021-04270-z.

21. Yi, H.G., Chandrasekaran, B., Nourski, K.V., Rhone, A.E., Schuerman, W.L., Howard, M.A., 3rd, Chang, E.F., and Leonard, M.K. (2021). Learning nonnative speech sounds changes local encoding in the adult human cortex. Proc. Natl. Acad. Sci. USA *118*, e2101777118. https://doi.org/10.1073/pnas.2101777118.

22. Gwilliams, L., King, J.-R., Marantz, A., and Poeppel, D. (2022). Neural dynamics of phoneme sequences reveal position-invariant code for content and order. Nat. Commun. *13*, 6606. https://doi.org/10.1038/s41467-022-34326-1.

23. Forseth, K.J., Pitkow, X., Fischer-Baum, S., and Tandon, N. (2021). What the brain does as we speak. Preprint at bioRxiv. https://doi.org/10.1101/2021.02.05.429841.

24. Forseth, K.J., Hickok, G., Rollo, P.S., and Tandon, N. (2020). Language prediction mechanisms in human auditory cortex. Nat. Commun. *11*, 5240. https://doi.org/10.1038/s41467-020-19010-6.

25. Bhaya-Grossman, I., and Chang, E.F. (2022). Speech computations of the human superior temporal gyrus. Annu. Rev. Psychol. *73*, 79–102. https://doi.org/10.1146/annurev-psych-022321-035256.

26. Hamilton, L.S., Oganian, Y., Hall, J., and Chang, E.F. (2021). Parallel and distributed encoding of speech across human auditory cortex. Cell *184*, 4626–4639.e13. https://doi.org/10.1016/j.cell.2021.07.019.

27. Khanna, A.R., Muñoz, W., Kim, Y.J., Kfir, Y., Paulk, A.C., Jamali, M., Cai, J., Mustroph, M.L., Caprara, I., Hardstone, R., et al. (2024). Single-neuronal elements of speech production in humans. Nature *626*, 603–610. https://doi.org/10.1038/s41586-023-06982-w.

28. Jamali, M., Grannan, B., Cai, J., Khanna, A.R., Muñoz, W., Caprara, I., Paulk, A.C., Cash, S.S., Fedorenko, E., and Williams, Z.M. (2024). Semantic encoding during language comprehension at single-cell resolution. Nature *631*, 610–616. https://doi.org/10.1038/s41586-024-07643-2.

29. Chomsky, N., Gallego, A.J., and Ott, D. (2019). Generative grammar and the faculty of language: insights, questions, and challenges. Catalan J. Linguist. Special Issue, 229–261.

30. OpenAI (2023). GPT-4 technical report. Preprint at arXiv. https://doi.org/10.48550/arXiv.2303.08774.

31. Hagoort, P. (2019). The neurobiology of language beyond single-word processing. Science *366*, 55–58. https://doi.org/10.1126/science.aax0289.

32. Hagoort, P., and Indefrey, P. (2014). The neurobiology of language beyond single words. Annu. Rev. Neurosci. *37*, 347–362. https://doi.org/10.1146/annurev-neuro-071013-013847.

33. Călinescu, L., Ramchand, G., and Baggio, G. (2023). How (not) to look for meaning composition in the brain: A reassessment of current experimental paradigms. Front. Lang. Sci. *2*, 1096110. https://doi.org/10.3389/flang.2023.1096110.

34. Goldstein, A., Zada, Z., Buchnik, E., Schain, M., Price, A., Aubrey, B., Nastase, S.A., Feder, A., Emanuel, D., Cohen, A., et al. (2022). Shared computational principles for language processing in humans and deep language models. Nat. Neurosci. *25*, 369–380. https://doi.org/10.1038/s41593-022-01026-4.

35. Cai, J., Hadjinicolaou, A.E., Paulk, A.C., Soper, D.J., Xia, T., Wang, A.F., Rolston, J.D., Richardson, R.M., Williams, Z.M., and Cash, S.S. (2025). Natural language processing models reveal neural dynamics of human conversation. Nat. Commun. *16*, 3376. https://doi.org/10.1038/s41467-025-58620-w.

36. Tenney, I., Das, D., and Pavlick, E. (2019). BERT rediscovers the classical NLP pipeline. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Association for Computational Linguistics), pp. 4593–4601. https://doi.org/10.18653/v1/P19-1452.

37. Elazar, Y., Ravfogel, S., Jacovi, A., and Goldberg, Y. (2021). Amnesic probing: behavioral explanation with amnesic counterfactuals. Trans. Assoc. Comput. Linguist. *9*, 160–175. https://doi.org/10.1162/tacl_a_00359.

38. Rapp, B., and Caramazza, A. (2002). Selective difficulties with spoken nouns and written verbs: A single case study. J. Neurolinguist. *15*, 373–402. https://doi.org/10.1016/S0911-6044(01)00040-9.

39. Caramazza, A., and Hillis, A.E. (1991). Lexical organization of nouns and verbs in the brain. Nature *349*, 788–790. https://doi.org/10.1038/349788a0.

40. Woolnough, O., Forseth, K.J., Rollo, P.S., Roccaforte, Z.J., and Tandon, N. (2022). Event-related phase synchronization propagates rapidly across human ventral visual cortex. Neuroimage *256*, 119262. https://doi.org/10.1016/j.neuroimage.2022.119262.

41. Aflalo, T., Zhang, C.Y., Rosario, E.R., Pouratian, N., Orban, G.A., and Andersen, R.A. (2020). A shared neural substrate for action verbs and observed actions in human posterior parietal cortex. Sci. Adv. *6*, eabb3984. https://doi.org/10.1126/sciadv.abb3984.

42. Damasio, A.R., and Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. Proc. Natl. Acad. Sci. USA *90*, 4957–4960. https://doi.org/10.1073/pnas.90.11.4957.

43. Bansal, A.K., Madhavan, R., Agam, Y., Golby, A., Madsen, J.R., and Kreiman, G. (2014). Neural dynamics underlying target detection in the human brain. J. Neurosci. *34*, 3042–3055. https://doi.org/10.1523/JNEUROSCI.3781-13.2014.

44. Groppe, D.M., Bickel, S., Dykstra, A.R., Wang, X., Mégevand, P., Mercier, M.R., Lado, F.A., Mehta, A.D., and Honey, C.J. (2017). iELVis: an open source MATLAB toolbox for localizing and visualizing human intracranial electrode data. J. Neurosci. Methods *281*, 40–48. https://doi.org/10.1016/j.jneumeth.2017.01.022.

45. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage *31*, 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021.

46. Tager-Flusberg, H., and Seery, A.M. (2013). Early development of speech and language: cognitive, behavioral, and neural systems. In Neural Circuit Development and Function in the Healthy and Diseased Brain, J.L.R. Rubenstein, and P. Rakic, eds. (Academic Press), pp. 315–330.

47. Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. Nature *435*, 1102–1107. https://doi.org/10.1038/nature03687.

48. Kriegeskorte, N., and Kreiman, G. (2011). Visual Population Codes (MIT Press).

49. Quiroga, R.Q., Kreiman, G., Koch, C., and Fried, I. (2008). Sparse but not 'grandmother-cell' coding in the medial temporal lobe. Trends Cogn. Sci. *12*, 87–91. https://doi.org/10.1016/j.tics.2007.12.003.

50. Tang, H., Yu, H.Y., Chou, C.C., Crone, N.E., Madsen, J.R., Anderson, W.S., and Kreiman, G. (2016). Cascade of neural processing orchestrates cognitive control in human frontal cortex. eLife *5*, e12352. https://doi.org/10.7554/eLife.12352.

51. Wojtasik, M., Bludau, S., Eickhoff, S.B., Mohlberg, H., Gerboga, F., Caspers, S., and Amunts, K. (2020). Cytoarchitectonic characterization and functional decoding of four new areas in the human lateral orbitofrontal cortex. Front. Neuroanat. *14*, 2. https://doi.org/10.3389/fnana.2020.00002.

52. Kringelbach, M.L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. Nat. Rev. Neurosci. *6*, 691–702. https://doi.org/10.1038/nrn1747.

53. Öngür, D., and Price, J.L. (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb. Cortex *10*, 206–219. https://doi.org/10.1093/cercor/10.3.206.

54. Xiao, Y., López, P.S., Wu, R., Srinivasan, R., Wei, P.-H., Shan, Y.-Z., Weisholtz, D., Cosgrove, G.R., Madsen, J.R., Stone, S., et al. (2023). Neurophysiological and computational mechanisms of non-associative and associative memories during complex human behavior. Preprint at bioRxiv. https://doi.org/10.1101/2023.03.27.534384.

55. Hunt, L.T., Malalasekera, W.M.N., de Berker, A.O., Miranda, B., Farmer, S.F., Behrens, T.E.J., and Kennerley, S.W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. Nat. Neurosci. *21*, 1471–1481. https://doi.org/10.1038/s41593-018-0239-5.

56. Nogueira, R., Abolafia, J.M., Drugowitsch, J., Balaguer-Ballester, E., Sanchez-Vives, M.V., and Moreno-Bote, R. (2017). Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. Nat. Commun. *8*, 14823. https://doi.org/10.1038/ncomms14823.

57. Noonan, M.P., Walton, M.E., Behrens, T.E.J., Sallet, J., Buckley, M.J., and Rushworth, M.F.S. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc. Natl. Acad. Sci. USA *107*, 20547–20552. https://doi.org/10.1073/pnas.1012246107.

58. de Araujo, I.E.T., Rolls, E.T., Kringelbach, M.L., McGlone, F., and Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. Eur. J. Neurosci. *18*, 2059–2068. https://doi.org/10.1046/j.1460-9568.2003.02915.x.

59. Dronkers, N.F., Wilkins, D.P., Van Valin, R.D., Jr., Redfern, B.B., and Jaeger, J.J. (2004). Lesion analysis of the brain areas involved in language comprehension. Cognition *92*, 145–177. https://doi.org/10.1016/j.cognition.2003.11.002.

60. Mesulam, M.M., Coventry, C.A., Bigio, E.H., Sridhar, J., Gill, N., Fought, A.J., Zhang, H., Thompson, C.K., Geula, C., Gefen, T., et al. (2022). Neuropathological fingerprints of survival, atrophy and language in primary progressive aphasia. Brain *145*, 2133–2148. https://doi.org/10.1093/brain/awab410.

61. Mesulam, M.-M., Rogalski, E.J., Wieneke, C., Hurley, R.S., Geula, C., Bigio, E.H., Thompson, C.K., and Weintraub, S. (2014). Primary progressive aphasia and the evolving neurology of the language network. Nat. Rev. Neurol. *10*, 554–569. https://doi.org/10.1038/nrneurol.2014.159.

62. I. Fried, U. Rutishauser, M. Cerf, and G. Kreiman, eds. (2014). Single Neuron Studies of the Human Brain: Probing Cognition (MIT Press).

63. Mukamel, R., and Fried, I. (2012). Human intracranial recordings and cognitive neuroscience. Annu. Rev. Psychol. *63*, 511–537. https://doi.org/10.1146/annurev-psych-120709-145401.

64. Dijksterhuis, D.E., Self, M.W., Possel, J.K., Peters, J.C., van Straaten, E.C.W., Idema, S., Baaijen, J.C., van der Salm, S.M.A., Aarnoutse, E.J., van Klink, N.C.E., et al. (2024). Pronouns reactivate conceptual representations in human hippocampal neurons. Science *385*, 1478–1484. https://doi.org/10.1126/science.adr2813.

65. Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. *10*, 433–436. https://doi.org/10.1163/156856897X00357.

66. Fischl, B. (2012). FreeSurfer. Neuroimage *62*, 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021.

67. Radford, A., Wu, J., Child, R., Luan, D., and Amodei, D. (2019). Sutskever I. Language Models Are Unsupervised Multitask Learners. OpenAI Blog.

68. Wang, J., Tao, A., Anderson, W.S., Madsen, J.R., and Kreiman, G. (2021). Mesoscopic physiological interactions in the human brain reveal small-world properties. Cell Rep. *36*, 109585. https://doi.org/10.1016/j.celrep.2021.109585.

69. Dale, A.M., Fischl, B., and Sereno, M.I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage *9*, 179–194. https://doi.org/10.1006/nimg.1998.0395.

70. Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L.H., and Papademetris, X. (2011). Unified framework for development, deployment and robust testing of neuroimaging algorithms. Neuroinformatics *9*, 69–84. https://doi.org/10.1007/s12021-010-9092-8.

71. Mitra, P., and Bokil, H. (2008). Observed Brain Dynamics (Oxford University Press).

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029





## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited data | | |
| Neurophysiological recordings, stimuli, and analysis code | This paper | https://kreimanlab.com/code/neural-coding-for-language-representations-in-the-human-brain/ |
| Experimental models: Human participants | | |
| 20 neurosurgical patients undergoing epilepsy monitoring | Collaborating Hospitals | IRB approved; see Table S1 |
| Software and algorithms | | |
| MATLAB | MathWorks, Natick, MA | https://www.mathworks.com; RRID: SCR_001622 |
| Psychtoolbox-3 | Brainard[65] | http://psychtoolbox.org; RRID: SCR_002881 |
| FreeSurfer | Fischl[66] | https://surfer.nmr.mgh.harvard.edu; RRID: SCR_001847 |
| iELVis | Groppe et al.[44] | https://github.com/iELVis/iELVis |
| GPT-2 | Radford et al.[67] | https://github.com/openai/gpt-2 |
| Other | | |
| XLTEK neurophysiological recording system | Natus Medical, Oakville, ON, Canada | N/A |
| Ad-Tech sEEG depth electrodes | Ad-Tech, Racine, WI, USA | N/A |

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

#### Preregistration

This study was preregistered on the Open Science Framework (OSF) website. The preregistration DOI is: https://doi.org/10.17605/OSF.IO/8TU2G.

#### Participants

We recorded data from 20 participants (9 male, 9-53 years old, 2 left-handed, 2 ambidextrous, Table S1) with pharmacologically resistant epilepsy (Figure 1A). All experiments were conducted while participants stayed at Children's Hospital Boston (CHB), Brigham and Women's Hospital (BWH), or Taipei Veterans General Hospital (TVGH). All studies were approved by each hospital's institutional review boards and were carried out with the participants' informed consent.

#### Recordings and electrode locations

Participants were implanted with intracranial electrodes via stereo electroencephalography (sEEG) (Ad-Tech, Racine, WI, USA). Neurophysiological data were recorded using XLTEK (Oakville, ON, Canada), Bio-Logic (Knoxville, TN, USA), Nihon Kohden (Tokyo, Japan), and Natus (Pleasanton, CA). The sampling rate was 2048 Hz at BCH and TVGH, and 1024 Hz or 512 Hz at BWH. All data were referenced in a bipolar montage. There were no seizure events in any of the sessions. Electrode locations were decided based on clinical criteria for each participant. Electrodes in the epileptogenic foci, as well as pathological areas, were removed from analyses. The total number of electrodes after bipolar referencing and removing electrodes with no signal, line noise or recording artifacts was 1,801.[68]

Following implantation, electrodes were localized by co-registration of pre-operative T1 MRI and post-operative CT scans using the iELVis software.[44] We used FreeSurfer to segment MRI images, upon which post implant CT was rigidly registered.[69] Electrodes were marked in the CT aligned to pre-operative MRI using the Bioimage Suite.[70] The Desikan-Killiany (DK) atlas was used to assign the electrodes locations. Figures 1B–1G and Table S2 show the locations of all the electrodes.

#### Experiment design

All visual stimuli were displayed on a 15.4 inch 2,880 × 1,800 pixel LCD screen using the Psychtoolbox in MATLAB (Natick, MA) and a MacBook Pro laptop (Cupertino, CA). The stimuli were positioned at eye level at about 80 cm from the participant and each word

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



subtended approximately 3 degrees of visual angle. Sounds were played from the speakers of a MacBook Pro 15.4 at 80% loudness using the Psychtoolbox in MATLAB.[65] We used the USB-1208FS-Plus device from Measurement Computing Corporation (Norton, Massachusetts) to send trigger pulses that enabled us to align stimuli onsets and behavioral responses to neural recordings.

A schematic of the task is shown in Figure 1. Participants were presented two words, 875 ms presentation time, with a 400 ms blank screen between them. At the end of each trial, participants were asked to indicate via a button press whether the two words were same or different. Word presentation was either visual or auditory. On average, we presented 1500 ± 710 trials (Table S1 shows the number of trials per participant).

There were four types of trials: Noun followed by Adjective (42% of trials, e.g., ''apple green''), Adjective followed by Noun (42% of trials, e.g., ''green apple''), Repeated Noun (8% of trials, e.g., ''apple apple''), and Repeated Adjective (8% of trials, e.g., ''green green''). The order of trials (stimulus presentation modality and noun/adjective structure) was randomly interleaved. Each word combination was presented in a randomized manner 5 times in the audio modality and 5 times in the visual modality. The nouns belonged to two categories, animals (e.g., ''cat'') and food (e.g., ''apple''). The adjectives belonged to two categories, concrete adjectives (e.g., ''big'') and abstract adjectives (e.g., ''good''). A list of all the nouns and adjectives is included in Table S3. We selected only high frequency English words that were more frequent than $10^{-6}$ in Google Ngram and were 3 to 7 letters long and had no more than 1 or 2 syllables. We used the max frequency of a word between 2006 and 2019. Finally, we created a balanced selection of nouns and adjectives such that noun and adjectives were indistinguishable from each other using word length or number of syllables ($p>0.05$ rank-sum test). We conducted the experiment in 3 languages, English (16 monolingual and 1 bilingual participant), Spanish (1 bilingual participant) and Taiwanese (3 monolingual participants). Two bilingual international scholars whose native language was Spanish (MAG), and Taiwanese (YLK) translated the words in the task. For non-English languages, we also kept nouns and adjectives indistinguishable based on word-length and number of syllables. Within each language in our data, audio duration was significantly correlated with text length ($R^2 = 0.19–0.25$, $p < 0.001$, Figures 1H–1J). Word-type comparisons showed no differences in audio duration for any language (t-tests, $p>0.05$). Only Taiwanese nouns and adjectives differed in text length (t-test, $p<0.01$), but this did not impact any of the results (see control analyses using General Linear Model (GLM)).

Participants had to indicate whether the two words in a trial were the same or not. The motor responses were the same for nouns or adjectives. The motor responses were also the same for noun followed by adjective or adjective followed by noun trials. Thus, the motor responses were orthogonal to parts of speech and grammar and differences between nouns and adjectives cannot be attributed to motor signals.

## METHOD DETAILS

### Preprocessing

A total of 2,428 electrode contacts were implanted, 627 of which were excluded from analysis due to bipolar referencing, presence of line noise or recording artifacts.[68] We removed 60 Hz line noise and its harmonics using a fifth-order Butterworth filter. We focus on the high-gamma band of the intracranial field potential signals obtained by bandpass filtering raw data of each electrode in the 65–150 Hz range (fifth-order Butterworth filter). The high gamma band (65-150 Hz) power was computed using the Chronux toolbox.[71] We used a time-bandwidth product of 3 and 4 leading tapers, a moving window size of 200 ms, and a step size of 5 ms. For every trial, we computed the normalized high gamma activity by subtracting the mean activity from -150 to 50 ms from the onset of the first fixation and then dividing by the standard deviation. This normalized response is reported as ''gamma power'' on the y-axis when showing electrode responses and referred to throughout the manuscript as *neural responses*.

### Responsive electrodes

We evaluated whether an electrode was responsive to visual or auditory stimuli by comparing the response magnitude during the 100 to 400 ms post stimulus onset time to the -400 to -100 ms before stimulus onset (e.g., Figures S1J–S1M). The responsiveness threshold was set using Cohen's d prime coefficient and based on the number of trials for a statistical power of 80% and $p<0.01$ (one-tailed z-test). We also computed the time at which the neural signals reach half of the maximum amplitude.

### Part-of-speech selectivity

We compared the neural responses to nouns versus adjectives. We normalized the responses in each trial by subtracting the mean and dividing by the standard deviation of the high gamma power in the baseline period 100 to 400 ms pre-word1 onset. Periods of significant selective activation were tested using a one-tailed t-test with $p<0.05$ at each time point to differentiate between nouns and adjectives and were corrected for multiple comparisons with a Benjamini-Hochberg false detection rate (FDR) corrected threshold of $q<0.05$, separately for auditory and visual trials. After fixing the FDR with $q<0.05$, an electrode was considered to be selective for part of speech if there was a significant difference between nouns and adjectives for a minimum contiguous window of 65 ms.

### Generalized linear model (GLM)

We created a GLM to tease out the experiment variables that significantly contribute to explaining the responses of a given electrode.

$$AUC = \beta_0 + \beta_{NvsA} NvsA + \beta_{GvsUG} GvsUG + \beta_{NSyllables} NSyllable + \beta_{WordLen} WordLen \quad \text{(Equation 1)}$$

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029





where AUC is the area under the response curve (e.g., Figure 2A) from 200 ms to 800 ms after word onset, $\beta_0$is a constant additive term, NvsA is 1 for Nouns and -1 for Adjectives, GvsUG is 1 for Grammatical trials and -1 for Ungrammatical trials, NumberOfSyllables is 1 or 2 (and 0 for visual trials), or WordLength goes from 3 to 7 (and 0 for auditory trials) as the task predictors. We fit this GLM model for each electrode separately using the MATLAB function fitglm and report the corresponding β coefficients (e.g., Figure 2K). We assessed whether each coefficient was significantly different from zero when compared to β coefficients generated from shuffled labels ($p<0.01$, corrected for multiple comparisons).

### Surprisal calculation

Surprisal was computed using pre-trained GPT-2 autoregressive language models on Google Colab. The sequence began with a beginning-of-sequence token (BOS), so each pair was encoded as [BOS, w1, w2] using the model's sub-word tokenizer. Token surprisal was defined as −log p(token | preceding tokens). Word surprisal was summed across tokens for a given word. In a separate analysis, we also added surprisal for each word as a predictor in the GLM.

### Anatomical comparisons

To assess the degree of anatomical specificity in the neural responses, we compared the percentage of significant electrodes in each brain region to the null distribution expected given the number of electrodes in each area using a permutation test ($p<0.01$, $10^6$ iterations). A similar approach was followed to compare the same region between the left and right hemispheres.

### Decoding analysis

We performed a machine learning decoding analysis to decode parts of speech in individual words combining all the electrodes in each brain region as defined by the Desikan-Killiany atlas[45] (Figures 1B–1G). The top-N principal components of all electrodes that explained more than 70% of the variance in the training data for the area under curve of non-overlapping 100 ms time-windows of the signal following word onset were used for decoding. The signal for decoding comprised of features from different frequency bands (beta:12-30 Hz, low gamma:30-65 Hz, and high gamma power: 65-150Hz). The analysis was repeated for 30 random splits of the data with 80% of the data used for training a Support Vector Machine with a linear kernel. Significant decoding performance was found by comparing performance from the original data at each time-window with a null distribution obtained by shuffling labels ($p<0.01$, ranksum test). Regions with statistically significant decoding performance were found by comparing the average of the maximum decoding performance across time for 30 random iterations of the original data with that of the null distribution, separately for both hemispheres ($p<0.01$, ranksum test corrected for multiple comparisons) (Figures 3C, 3F, 3I, 4C, 4F, 4I, 4L, 4O, 5C, 5F, and 5I). We also applied a threshold such that for a given region R:

$$[\mu_R - 3 * \sigma_R]_{original\ data} > [\mu_R + 3 * \sigma_R]_{null\ data} \quad \text{(Equation 2)}$$

where μ and σ represent the average and standard deviation in region R. For the significant regions, the average max-performance between the left and right hemispheres was compared to find if decoding performance was lateralized ($p<0.01$, ranksum test, corrected for multiple comparisons).

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Statistical criteria for electrode responsiveness

Electrodes were labeled responsive if high-gamma activity from 100–400 ms post-stimulus exceeded baseline (−400 to −100 ms) using Cohen's d–based thresholds corresponding to 80% power and $p < 0.01$ (one-tailed z-test).

### *t*-tests for noun vs adjective selectivity and false discovery rate correction

Responses to nouns versus adjectives were compared at each time point using one-tailed t-tests ($p < 0.05$) separately for auditory and visual trials. Significant time points were corrected using Benjamini–Hochberg FDR ($q < 0.05$). Selectivity required ≥65 ms of contiguous significance at both words.

### Ranksum tests

Non-parametric comparisons, including decoding-based regional comparisons and laterality assessments, used Wilcoxon ranksum tests with a significance threshold of $p < 0.01$ (multiple-comparison corrected).

### GLM coefficient significance testing

For each electrode, GLM coefficients were compared to coefficients derived from shuffled-label models. Coefficients exceeding the shuffled distribution at $p < 0.01$ (corrected for multiple comparisons, ranksum test) were considered significant.

### Permutation tests for anatomical specificity

The fraction of selective electrodes in each anatomical region was compared to a null distribution obtained by randomly permuting region assignments ($10^6$ iterations, $p < 0.01$, ranksum test, corrected for multiple comparisons).

Please cite this article in press as: Misra et al., Invariant neural representation of parts of speech in the human brain, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.07.029



### Decoding significance tests

Decoding accuracy in each 100-ms window was compared against a null distribution generated by label shuffling. Time windows and regions were significant when decoding exceeded the null at $p < 0.01$ (corrected for multiple comparisons, ranksum test)

### Correlation *p*-values

Significance of correlations was computed using the standard analytic test in which the correlation coefficient is converted to a t-statistic, and a two-tailed p-value is obtained from the corresponding t-distribution given the sample size.