Title Page

Neural Sensitivity to Word Frequency Modulated by Morphological Structure: Univariate and Multivariate fMRI Evidence from Morphologically Complex Words

Joonwoo Kim[a,b], Solbin Lee[b], Kichun Nam[a,b*]

[a] School of Psychology, Korea University, Seoul, Republic of Korea
[b] Wisdom Science Center, Korea University, Seoul, Republic of Korea

[*] Correspondence
Kichun Nam
School of Psychology, Korea University
145, Anam-ro, Seongbuk-gu, Seoul, 02739, Republic of Korea
E-mail: kichun@korea.ac.kr

Abstract

A fundamental question in psycholinguistics concerns whether morphological decomposition is obligatory during visual word recognition or whether whole-word access can occur under conditions of high frequency and familiarity. The present study aimed to examine how morphological complexity and lexical frequency jointly influence neural representations during visual word recognition. We addressed this question using rapid event-related fMRI with both univariate and multivariate analyses. Twenty-five native Korean speakers performed lexical decisions on simple and inflected nouns that varied parametrically in surface frequency (token frequency of the full word form) and base frequency (cumulative stem frequency). Korean's transparent agglutinative morphology enabled a principled dissociation of these frequency measures. MVPA failed to decode morphological condition from activation patterns in any region, suggesting that simple and inflected forms are not represented as discrete neural categories. Crucially, RSA demonstrated robust encoding of surface frequency—but not base frequency—in inferior frontal gyrus (IFG) pars opercularis and supramarginal gyrus (SMG), with significantly stronger correlations for inflected than simple nouns. Univariate analyses confirmed this pattern: surface frequency interacted with morphological condition in inferior parietal lobule (IPL), such that higher surface frequency increased activation selectively for inflected forms, while base frequency showed no reliable effects. These results challenge obligatory decomposition and support a distributional framework in which lexical statistics guide the processing of morphologically complex words. Taken together, our findings suggest that morphological complexity modulates neural sensitivity to whole-word rather than stem-level statistics, underscoring how structural and statistical factors interact to support morphological processing.

1. Introduction

Human language comprehension requires efficient recognition of morphologically complex words. A central question in psycholinguistics is whether the cognitive system obligatorily decomposes these words into morphemes before meaning access or whether whole-word representations can be accessed directly for familiar forms.

Evidence for obligatory decomposition comes primarily from masked-priming studies, where briefly presented primes like *hunter* facilitate recognition of *hunt*, even when semantically opaque (e.g., *corner–corn*) relative to orthographic controls (Longtin et al., 2003; Rastle et al., 2004). Such priming, occurring within 200 ms, suggests rapid morpho-orthographic segmentation that precedes semantics. ERP and MEG work corroborates this view, showing early negativities before semantic divergence (Lavric et al., 2011; Lewis et al., 2011). In contrast, distributional perspectives argue that processing depends on lexical statistics, transparency, and predictability (Baayen et al., 2011; Bybee, 1995; Milin et al., 2017). High-frequency or predictable forms may be accessed holistically, whereas rare or novel forms are decomposed. In lexical decision and naming tasks, surface frequency, the token frequency of the whole form, often better predicts recognition than base frequency, the cumulative frequency of words sharing a stem (Bertram et al., 2000; Kuperman et al., 2009). Morphological priming effects are likewise modulated by transparency and familiarity (Feldman et al., 2015; Jared et al., 2017). Connectionist models further show how distributional learning can yield decomposition-like effects without an explicit parsing mechanism (Plaut & Gonnerman, 2000). Together, these findings suggest morphological processing may reflect both automatic segmentation and flexible, frequency-driven representations.

Neuroimaging evidence has shown mixed results. fMRI studies report that complex forms engage a left-lateralized network including the inferior frontal gyrus (IFG), inferior parietal lobule (IPL), and ventral occipitotemporal cortex, with selective IFG responses to inflected or transparent derivations interpreted as evidence for segmentation (Bozic et al., 2007; Lehtonen et al., 2006). Other work, however, suggests these effects overlap with general orthographic or semantic processes rather than reflecting a unique decomposition stage (Desai et al., 2006; Devlin et al., 2004). Findings on word frequency also remain inconsistent, with some studies reporting stronger activation for low-frequency words (Fiebach et al., 2002; Fiez et al., 1999), and others showing divergent patterns depending on task and subregion (Carreiras et al., 2009; Graves et al., 2010). Few studies have manipulated base frequency directly, and most report limited independent effects once surface morphology is controlled (Lehtonen et al., 2006; Pliatsikas et al., 2014).

Recent multivariate approaches offer powerful tools for examining how linguistic information is represented in the brain. Multivoxel pattern analysis (MVPA) assesses whether distributed activity patterns discriminate between linguistic conditions, while representational similarity analysis (RSA) relates neural dissimilarities to models derived from lexical statistics (Haxby et al., 2001; Kriegeskorte, 2008). Together, these methods reveal fine-grained representations of syntactic, semantic, and sublexical features (Bruffaerts et al., 2013; Fischer-Baum et al., 2017; Yue & Martin, 2021). These methods enable examinations of both categorical distinctions (e.g., morphological complexity) and graded encoding of continuous variables (e.g., frequency). To our knowledge, no study has directly examined whether morphological complexity modulates the neural encoding of surface versus base frequency.

The present study employed a rapid event-related fMRI design with a lexical decision task to test whether brain activation patterns track morphological structure or statistical frequency. Participants judged simple and inflected nouns varying in both measures. Korean provides an ideal test case. As an agglutinative language, Korean forms inflected nouns by attaching suffixes to invariant stems (e.g., 옥수수 /oksusu/ "corn" vs. 까닭–까닭은 /kkadak–kkadageun/ "reason–reason+TOPIC"). In such cases, surface frequency can diverge from base frequency, while simple nouns necessarily equate the two. We applied complementary approaches: univariate mixed-effects models, MVPA to test categorical morphology, and RSA to assess graded encoding of frequency. Our central question was whether the brain preferentially encodes statistical frequency or structure-based decomposition. A structure-driven view predicts robust base frequency effects and reliable classification of morphology. A statistics-driven view predicts a stronger influence of surface frequency, especially for inflected nouns, with weaker base frequency effects and limited morphological classification.

2. Methods

2.1. Participants

Twenty-five healthy adults (15 female; age: 26 ± 3.83 years) recruited from Korea University participated in the experiment. They were all native Korean speakers, right-handed (confirmed by the scores on the Edinburgh Handedness Test (Oldfield, 1971): $M = 7.76$, $SD = 0.28$) and had normal or corrected-to-normal vision. Participants were informed and provided written consent prior to the experiment and were monetarily compensated for their participation. Two participants withdrew before completing the experiment, and two others were excluded due to poor behavioral performance (accuracy < 60%). The sample size was determined based on prior

fMRI studies using similar multivariate and univariate ROI-based approaches in the language domain (Bozic et al., 2007; Carota et al., 2016; Kim et al., 2024), which typically report medium-to-large effects in comparable tasks with sample sizes ranging from 16 to 24. Given the within-subjects design and the use of trial-wise beta estimation, this sample size affords adequate power to detect region-level effects of morphology and lexical frequency using both traditional and multivariate analyses. Analyses were conducted on the remaining 21 participants (13 female; age: 26.13 ± 3.94 years) and are reported henceforth. The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Korea University (protocol code KUIRB-2021-0427-01, 21 December 2021).

2.2. Stimuli

A total of 120 Korean nouns were selected from the Korean Sejong corpus (Kang & Kim, 2009), which contains approximately 15 million words. Korean is an agglutinative language with a transparent morphological system, in which grammatical relations are marked by attaching case particles or topic markers to noun stems. Furthermore, the Korean writing system (Hangul) encodes syllables both orthographically and phonologically, enabling precise control of sublexical units. These features make Korean particularly well-suited for isolating effects of morphological complexity, stem-level access, and syllable-level frequency during visual word recognition.

Morphological complexity was orthogonally manipulated such that half of the stimuli were morphologically simple nouns, and the other half were inflected nouns. Each condition included an equal number of bi-syllabic and tri-syllabic words, ensuring that word length was fully balanced across conditions. Simple nouns consisted of bi- or tri-syllabic stems, while inflected

nouns were composed of mono- or bi-syllabic stems followed by a monosyllabic inflectional suffix (e.g., case or topic marker). For example, 숭어 (trout) and 뱀은 (snake + TOPIC) illustrate the contrast between simple and inflected forms, with 옥수수 (corn) and 까닭은 (reason + TOPIC) showing the contrast between a tri-syllabic simple noun and a tri-syllabic inflected noun consisting of a bi-syllabic stem (까닭) with the topic particle −은.

Word frequency was operationalized using Zipf-measure, a log-transformed word frequency per billion words in the given corpus (Brysbaert et al., 2018). Stem frequency was similarly transformed using the same Zipf-measure to allow for direct comparison across lexical levels. Importantly, word frequency (Zipf) was matched across simple and inflected conditions ($t = -1.58$; $p = .117$), whereas stem frequency (Zipf-stem) was not explicitly controlled. This was due to the structural properties of Korean: in simple nouns, the stem and surface word are identical, yielding the same frequency value. In contrast, inflected nouns contain an additional morpheme, such that the word frequency reflects the full surface form, while the stem frequency corresponds to the base noun. Given this natural divergence, Zipf and Zipf-stem were modeled separately in all statistical analyses, allowing us to dissociate effects of surface-level lexical access and stem-level frequency. Additionally, first-syllable frequency was statistically matched across conditions to minimize sublexical confounds ($t = -1.03$; $p = .303$).

The same number of pseudowords ($N = 120$) were created to serve as fillers in the lexical decision task. These pseudowords matched the real word stimuli in both form and length (i.e., 2–3 syllable) and were evenly divided into morphologically simple and inflected forms. Inflected pseudowords consisted of a real suffix combined with a pseudo-stem while simple pseudowords contained only a pseudo-stem. All pseudowords were generated by randomly combining syllables from real word stems and suffixes, which were orthographically and phonologically

legal but none of the pseudowords carried meaning according to the Standard Korean Dictionary. Finally, 30 non-linguistic masks, consisting of a string of asterisks (*****), were used as a baseline condition for a comparison with experimental conditions.

2.3. Procedure

A rapid event-related fMRI design was employed using a visual lexical decision task (see **Figure 1**). Participants were instructed to respond, as quickly and accurately as possible, whether each presented stimulus is a real word or not. Each trial began with a fixation cross (+) displayed for 100 ms, followed by a target stimulus presented at the center of the screen for 1000 ms. After pressing response button, blank screen was shown during the inter-trial interval (ITI), which was jittered with an average duration of 4000 ms (range: 1000–7000 ms). Stimuli were presented in a pseudo-random order that included words, pseudowords, and a non-linguistic baseline mask. The trial order and ITI durations were optimized using optseq2 (https://www.nitrc.org/projects/optseq). All visual stimuli were presented in white Courier New font (size 34) on a black background, controlled using E-Prime 2.0 Professional (Psychology Software Tools, Pittsburgh, PA). Prior to the main experiment, participants completed a practice session consisting of 12 randomly ordered trials. Only those who achieved at least 70% of accuracy during practice proceeded to the main task.

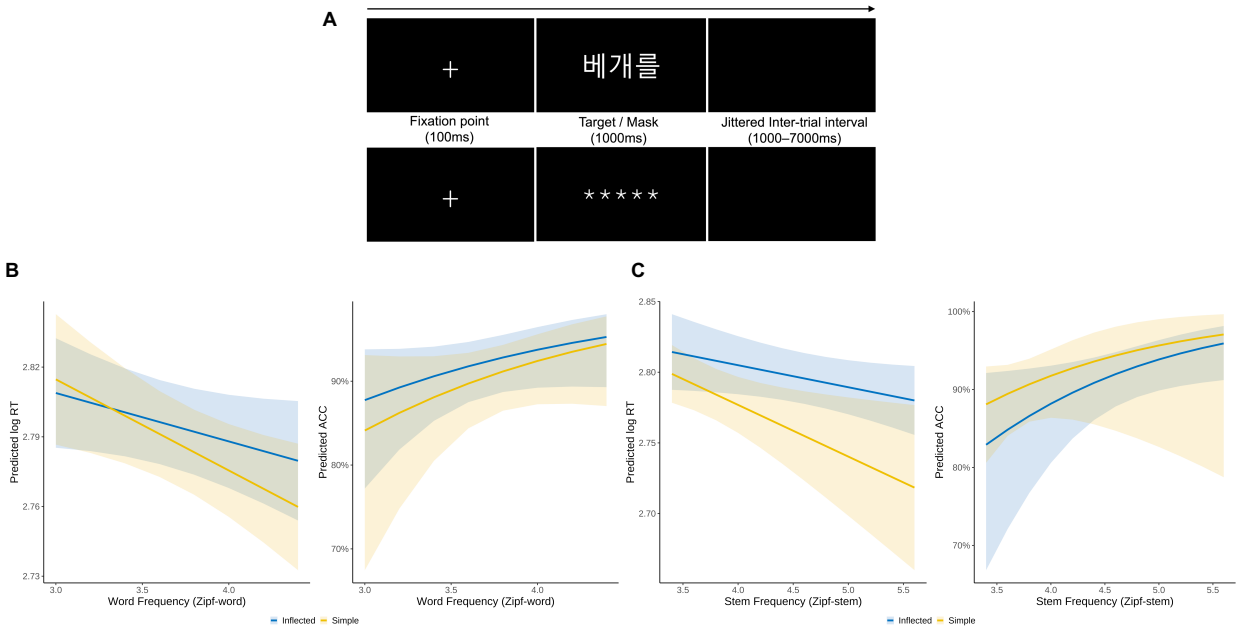

Figure 1. Experimental paradigm and behavioral results. (A) Timeline of a rapid event-related lexical-decision trial. (B) Mixed-effects predictions for log-transformed reaction time (left y-axis) and accuracy (right y-axis) across the observed range of surface word frequency (Zipf-word). (C) Corresponding predictions for stem frequency (Zipf-stem). Blue = Inflected; yellow = Simple; shaded ribbons indicate 95 % confidence intervals (CIs).

2.4. Image acquisition and preprocessing

MRI data were acquired at the Korea University Brain Imaging Center using a Siemens Magnetom Trio 3T MRI scanner (Erlangen, Germany). Functional images were acquired using T2*-weighted gradient EPI (echo-planar imaging) sequences (TR = 2000 ms; TE = 20 ms; flip angle = 90◦; field of view = 240 mm; slice thickness = 3 mm; no gap for 42 slices; matrix size = 80 × 80; and voxel size = 3 × 3 × 3 mm). A high-resolution structural image was acquired for each participant using a T1-weighted structural images were acquired with a 3D MP-RAGE (magnetization-prepared rapid gradient echo) sequence (TR = 1900 ms; TE = 2.52 ms; flip angle = 90◦; field of view = 256 ms; matrix size = 256 × 256; and voxel size = 1 × 1 × 1 mm) covering the whole head.

Preprocessing of the fMRI data was performed using SPM12 (Wellcome Centre for Human Neuroimaging, London, UK; https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) in conjunction with custom-built MATLAB scripts. The first three volumes of each session were discarded to allow for magnetic field stabilization. The remaining functional images were first realigned to the first volume to correct for head motion using a six-parameter rigid-body transformation. Slice timing correction was applied to adjust for differences in acquisition timing between slices, using the middle slice as reference. Each participant's structural image was then co-registered to the mean functional image and segmented. The deformation fields obtained during segmentation were subsequently applied to normalize the functional images to the standard MNI space, with resampling to 3 × 3 × 3 mm isotropic voxels. All normalized images were smoothed with a 6 mm full-width at half-maximum (FWHM) Gaussian kernel to improve signal-to-noise ratio. The six motion parameters from the realignment step were included as nuisance covariates in all first-level general linear models (GLMs).

2.5. Data analysis

2.5.1. Behavioral data

We analyzed reaction times (RT) and accuracy (ACC) from the lexical decision task using linear mixed-effects models (LMMs), respectively. For RT, only correct trials were included, and RTs were log-transformed to reduce skewness. Separate models were constructed for whole-word frequency (Zipf-word) and stem frequency (Zipf-stem), each including Morphological Condition (Simple vs. Inflected) and their interaction as fixed effects. Random intercepts for participants and items were included to account for repeated measures (Baayen et al., 2008). Accuracy data was analyzed using a binomial logistic regression in a general linear mixed effects model. These

analyses were performed in R (R Core Team, 2022) using the lme4 package (Bates et al., 2015), and the significance of fixed effects was assessed using Satterthwaite's approximation for degrees of freedom, as implemented in the lmerTest package (Kuznetsova et al., 2017).

2.5.2. Whole-brain analysis

First-level analyses modeled four contrasts for each participant: Inflected > Simple, Inflected > Baseline, Simple > Baseline, and Word > Baseline. Events were modeled with durations of 2 s and convolved with the canonical hemodynamic response function (HRF), and six motion parameters were included as nuisance regressors. At the second level, one-sample *t*-tests were performed on first-level contrast images. Whole-brain results were thresholded at a voxel-wise $p < .001$ (uncorrected) with a minimum cluster size of 30 voxels, consistent with exploratory standards in prior morphological and lexical processing fMRI studies. Activation patterns were further evaluated using anatomical masks from the Harvard-Oxford Atlas.

2.5.3. ROI selection

Our main analyses, including univariate contrasts, multivariate pattern analysis (MVPA), and representational similarity analysis (RSA), focused on six anatomically defined, left-hemisphere regions of interest (ROIs, see **Fig. 3A**). These included two subregions of the inferior frontal gyrus (pars triangularis, pars opercularis), two temporal-occipital regions including the middle temporal gyrus (MTG), fusiform gyrus (FuG), and two regions within the inferior parietal lobule (IPL): the angular gyrus (AG) and supramarginal gyrus (SMG). ROI selection was motivated by prior literature implicating these regions in morphological and lexical-semantic processing, allowing us to examine convergent and dissociable effects across multiple analytical

frameworks. All ROIs were defined using the probabilistic Harvard-Oxford cortical atlas (1 mm resolution). Each region was thresholded at 25% probability and binarized and resampled to match functional image resolution (3 mm) using SPM12. All masks were visually inspected to ensure accurate anatomical localization.

2.5.4. Univariate analysis

To examine morphological effects and lexical frequency modulations within key language-related regions, we first conducted ROI-based univariate analyses. Trial-wise beta values were extracted using MarsBaR (Brett et al., 2002) and submitted to linear mixed-effects models in R. Separate models were estimated for each frequency measure (Zipf or Zipf-stem) and ROI, with fixed effects for morphological condition (simple vs. inflected), frequency, and their interaction. Random intercepts were included for both participants and items to account for repeated measures and individual variability. These models assessed whether frequency effects varied by morphological complexity and whether either measure explained additional variance in regional activation.

2.5.5. MVPA (Multivariate pattern analysis)

We performed multivariate pattern classification to determine whether local voxel-wise activation patterns within predefined regions of interest (ROIs) could reliably distinguish between morphologically simple and inflected nouns. Trial-wise beta images were estimated from single-trial GLMs (see Section 1.5.3) and served as input for all MVPA analyses. All multivariate classification analyses were implemented using scikit-learn (Pedregosa et al., 2011) and nilearn library (http://nilearn.github.io) in Python (version 3.13.3) (Abraham et al., 2014).

Voxel values were extracted from anatomically defined ROIs (see Section 2.5.4) using NiftiMasker, as implemented in the nilearn package. Prior to classification, all beta images were z-scored across voxels to standardize the input while preserving relative pattern structure within each trial. Within each ROI, we trained a linear support vector classifier (SVC) with a regularization parameter $C = 1$ to discriminate trials by morphological condition (Simple vs. Inflected). Classification was performed using stratified 5-fold cross-validation (CV), ensuring class balance across folds. To avoid data leakage, voxel standardization and classifier training were performed separately within each fold.

Only participants and ROIs with at least five valid trials per condition (i.e., correct word trials with available beta estimates) were included to ensure sufficient training data. Classification performance was assessed using both mean decoding accuracy and area under the ROC curve (AUC), averaged across the five folds. AUC was used as the primary outcome metric due to its robustness to class imbalance and interpretability across regions.

Group-level statistical analyses were conducted using one-sample *t*-tests against chance-level decoding (0.5) for each ROI. We tested whether classification performance (AUC) across participants was significantly greater than chance using two-tailed t-tests for each ROI and corrected for multiple comparisons using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995). For visualization, we also report mean decoding accuracies across ROIs along with 95% confidence intervals.

2.5.6. Representational Similarity Analysis (RSA)

Representational similarity analysis (RSA) was used to test whether local activation patterns reflected graded similarity based on word or stem frequency. All representational similarity

analyses (RSA) were implemented using custom Python scripts in combination with scikit-learn (Pedregosa et al., 2011), SciPy (Virtanen et al., 2020), and Nilearn (Abraham et al., 2014). Analyses were performed within the same six anatomically defined ROIs used in the univariate and MVPA analyses.

Six theoretical representational dissimilarity matrices (RDMs) were constructed based on continuous lexical frequency measures: surface-level word frequency (Zipf) and base-form stem frequency (Zipf-Stem). Each RDM encoded the absolute pairwise frequency difference between trials (i.e., Euclidean distance in a 1D lexical space), producing symmetrical matrices where larger values reflect greater dissimilarity in frequency. For each frequency measure, we created separate model RDMs for three trial subsets: (1) all word trials (**Figures 2A, 2D**), (2) morphologically inflected nouns (**Figures 2B, 2D**), and (3) only simple nouns (**Figures 2C, 2F**). This yielded six theoretical RDMs in total (2 frequency models × 3 morphological conditions). As shown in **Figure 1**, the Zipf-based model showed greater frequency spread among inflected forms, while the Zipf-Stem model exhibited greater dispersion overall, particularly for morphologically complex items.

To construct neural RDMs, we extracted trial-wise beta images per participant and ROI using NiftiMasker (standardized across voxels) and computed pairwise Pearson correlations between activation vectors for each pair of trials. Neural dissimilarity was defined as 1 minus Pearson r, resulting in a neural RDM for each subject, ROI, and condition (All, Simple, Inflected). Spearman rank correlation was used to compute the correspondence between neural and model RDMs. Correlation values were computed separately for each subject, ROI, condition, and model.

At the group level, we assessed whether neural RDMs significantly aligned with theoretical frequency structure using one-sample t-tests against zero (within-subject Fisher-z transformed Spearman values). To test whether representational structure differed across morphological conditions, we performed paired-sample *t*-tests comparing Simple vs. Inflected correlations within each ROI and frequency model. All resulting *p*-values were corrected for multiple comparisons using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995).

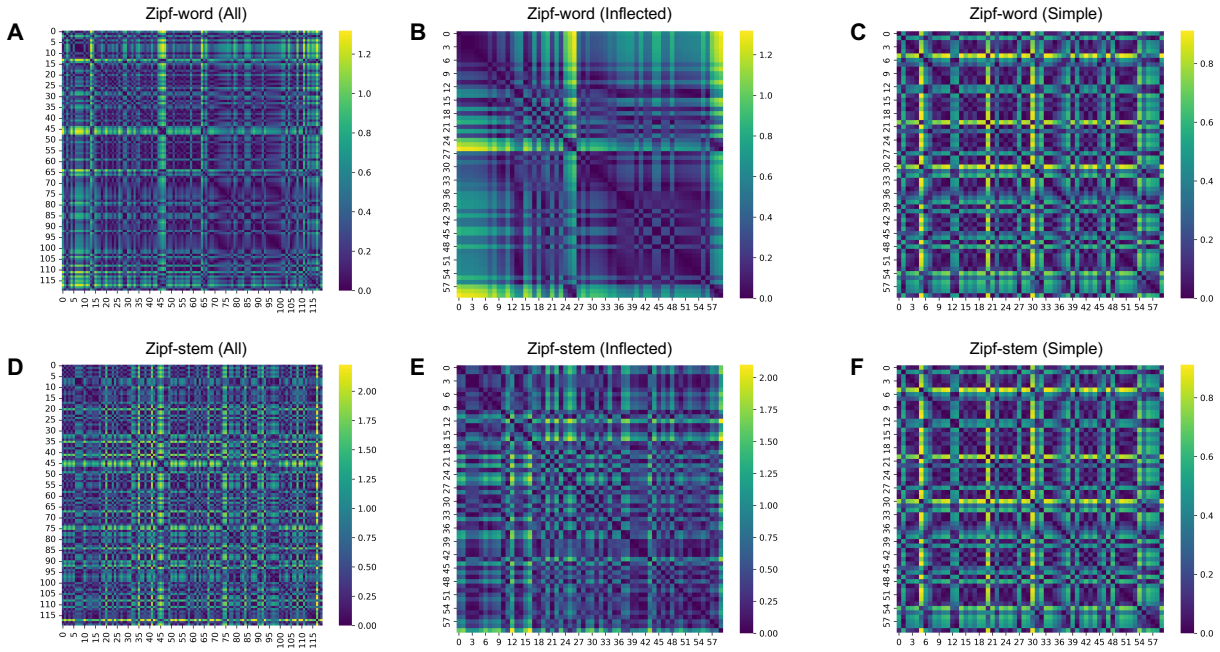

Figure 2. Theoretical frequency-based representational dissimilarity matrices (RDMs). Each matrix encodes the pairwise lexical distance that served as a model predictor in the RSA. Word-frequency RDMs (Zipf-word; panels A–C) and stem-frequency RDMs (Zipf-Stem; panels D–F) were computed as the absolute difference (i.e., Euclidean distance in a 1-D frequency space) in Zipf-scale frequency between two items, yielding symmetrical matrices whose values increase with greater frequency disparity. For both frequency measures we constructed three matrices: all nouns, only morphologically inflected nouns, and only simple nouns. Warmer colors denote larger frequency gaps; cooler colors denote smaller gaps.

3. Results

3.1. Behavioral performance

In both the Zipf-word and Zipf-stem models, morphological condition did not show a significant main effect (Zipf-word: $\beta = -0.027$, $p = .457$; Zipf-stem: $\beta = -0.028$, $p = .405$), nor was the interaction between morphology and frequency significant (Zipf-word: $p = .403$; Zipf-stem: $p = .216$). However, both frequency predictors showed significant main effects, with faster responses to higher-frequency words. Specifically, word frequency (Zipf-word) was associated with reduced RTs ($\beta = -0.029$, $p = .003$), and the same pattern was observed for stem frequency ($\beta = -0.026$, $p = .003$).

Accuracy was modeled using binomial generalized linear mixed-effects models with a logit link. No significant effects were found for morphological condition or its interaction with frequency (all *p*s > .83). Word frequency showed a marginal positive effect ($\beta = 0.707$, $p = .063$), while stem frequency had a significant effect ($\beta = 0.697$, $p = .036$), suggesting greater accuracy for higher-frequency items, particularly at the stem level.

3.2. Whole-brain analysis results

Table 1 and **Fig. 3** present the results of the univariate whole-brain analysis. All contrasts were computed at a voxel-level threshold of $p < .001$ (uncorrected) with a cluster extent threshold of $k > 30$. Anatomical labels were assigned using the Harvard-Oxford cortical atlas (HOA), and all reported clusters exceeded the statistical threshold.

When collapsed across all word conditions (*Word > Baseline*), robust activation was observed in a broad left-lateralized network associated with visual and lexical-semantic processing. Significant clusters were found in the left inferior frontal gyrus (pars triangularis and pars opercularis), frontal orbital cortex, temporal occipital fusiform cortex, and lateral occipital cortex (inferior division) bilaterally, with peak activity extending into the occipital pole.

The contrast Inflected > Rest revealed a largely overlapping activation pattern, including the left inferior frontal gyrus, middle and posterior temporal regions, temporal occipital fusiform cortex, frontal orbital cortex, and paracingulate gyrus. In addition, bilateral lateral occipital cortex (inferior division) showed significant engagement, reflecting visual and morphological processing demands in the inflected noun condition.

The Simple > Rest contrast similarly activated the inferior frontal gyrus, frontal orbital cortex, and temporal occipital fusiform cortex, along with strong bilateral activation in the lateral occipital cortex and occipital pole. However, the extent of frontal and temporal involvement was more restricted compared to the inflected condition.

Notably, the direct contrast Inflected > Rest did not yield any suprathreshold clusters, suggesting that the two morphological conditions engaged similar neural systems when contrasted directly, despite observable differences in activation extent relative to baseline.

Together, these results indicate that both simple and inflected nouns recruited a left-lateralized fronto-temporo-occipital network, with no statistically significant differences in overall activation magnitude between conditions at the whole-brain level.

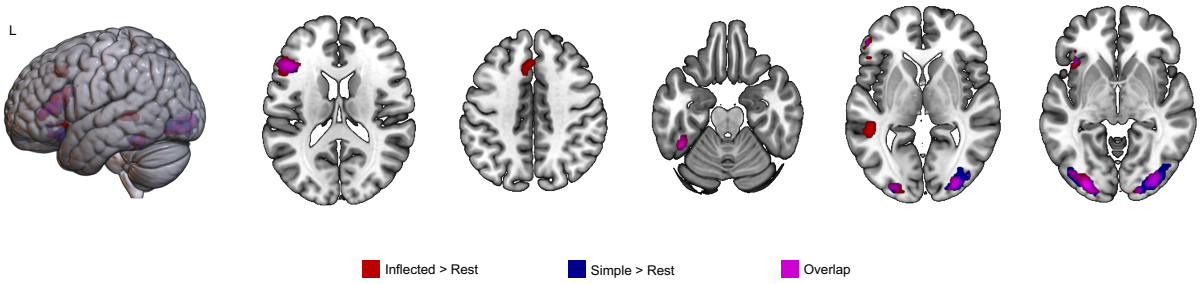

Figure 3. Whole-brain task activation for lexical decision. Brain activation maps depict voxels significantly engaged during word recognition in the whole-brain level (voxel-level threshold uncorrected $p < .001$, cluster extent threshold $k \geq 30$). Red = Inflected > Rest, Blue = Simple > Rest, and Magenta = overlap of the two contrasts.

Table 1. MNI coordinates of significant activation identified in the whole-brain analysis.

Brain Regions		Peak MNI coordinates	Z	Cluster Size
Word > Rest				
Lateral Occipital Cortex (inferior), Occipital Pole, Occipital Fusiform Gyrus	LH	[–30, –91, –4]	8.2	2321
Lateral Occipital Cortex (inferior), Occipital Pole	RH	[33, –88, –1]	7.4	2002
Frontal Orbital Cortex	LH	[–42, 26, –16]	6.8	145
Inferior Frontal Gyrus (pars triangularis, opercularis)	LH	[–48, 26, 17]	6.5	1395
Temporal Occipital Fusiform Cortex	LH	[–43, –49, –22]	5.8	293
Inflected > Rest				
Lateral Occipital Cortex (inferior), Occipital Pole, Occipital Fusiform Gyrus	LH	[–30, –91, –4]	7.6	2293
Temporal Occipital Fusiform Cortex	LH	[–42, –46, –22]	6.5	391
Lateral Occipital Cortex (inferior)	RH	[33, –88, –1]	6.4	1421
Inferior Frontal Gyrus (pars triangularis, opercularis)	LH	[–48, 26, 17]	6.3	1759
Middle Temporal Gyrus (post)	LH	[–49, –43, 2]	5.9	430
Frontal Orbital Cortex	LH	[–40, 20, –7]	5.4	240
Paracingulate Gyrus	LH	[–3, 27, 42]	5.2	347
Simple > Rest				
Lateral Occipital Cortex (inferior), Occipital Pole, Occipital Fusiform Gyrus	LH	[–30, –91, –4]	7.7	2388
Lateral Occipital Cortex (inferior), Occipital Pole	RH	[33, –88, –1]	7.3	2732
Frontal Orbital Cortex	LH	[–42, 26, –16]	6.6	228
Inferior Frontal Gyrus (pars triangularis, opercularis)	LH	[–48, 26, 15]	6.3	1488
Temporal Occipital Fusiform Cortex	LH	[–43, –49, –22]	5.2	234

Note. Anatomical labels are from the Harvard–Oxford Cortical Atlas. All clusters survive whole-brain correction at voxel-level threshold of uncorrected $p < .001$ and cluster extent threshold of $k \geq 30$. LH = left hemisphere; RH = right hemisphere.

3.3. ROI analysis results

The results of the ROI analysis and illustration are presented in **Figure 4**. We examined whether morphologically simple versus inflected nouns elicited differential activation across eight predefined ROIs and whether this activation was modulated by surface-level (Zipf-word) (**Fig. 4B**) or stem-level (Zipf-stem) frequency (**Fig. 4C**). Trial-level beta values extracted from each

ROI were analyzed using linear mixed-effects models with fixed effects for Morphology (Simple vs. Inflected), Frequency (Zipf-word or Zipf-stem), and their interaction, and random intercepts for Subject and Trial. Separate models were estimated for each ROI and frequency measure.

Our primary analysis focused on Zipf frequency. We observed significant main effects of word frequency in both regions in the inferior parietal lobule (IPL), including the angular gyrus ($\beta = 0.90$, $p = .023$) and supramarginal gyrus ($\beta = 0.92$, $p = .023$), indicating increased activation for higher-frequency words. No other regions showed significant frequency effects after FDR correction.

Importantly, a significant interaction of Condition and Zipf-word emerged in the SMG ($\beta = -1.30$, $p = .042$), reflecting stronger frequency effects for inflected compared to simple nouns. The AG also showed a marginal interaction ($\beta = -1.14$, $p = .070$), although this did not survive correction. No significant main effects or interactions were found in IFG (pars opercularis/triangularis), or FuG.

We then tested the influence of stem frequency (Zipf-stem). A significant main effect was observed in the angular gyrus ($\beta = 0.69$, $p = .0125$), indicating greater activation for higher-frequency stems across conditions, though this effect did not remain significant after correction. No other regions showed reliable stem frequency effects or interactions with morphological condition. Notably, SMG, which showed a robust interaction with surface frequency, did not exhibit significant modulation by stem frequency ($\beta = -0.79$, $p = .172$).

Together, these findings suggest that surface-level frequency exerts a stronger and more morphology-dependent influence on activation in parietal regions than stem-level frequency, pointing to a dissociation between surface-form and base-form frequency effects in morphological processing.

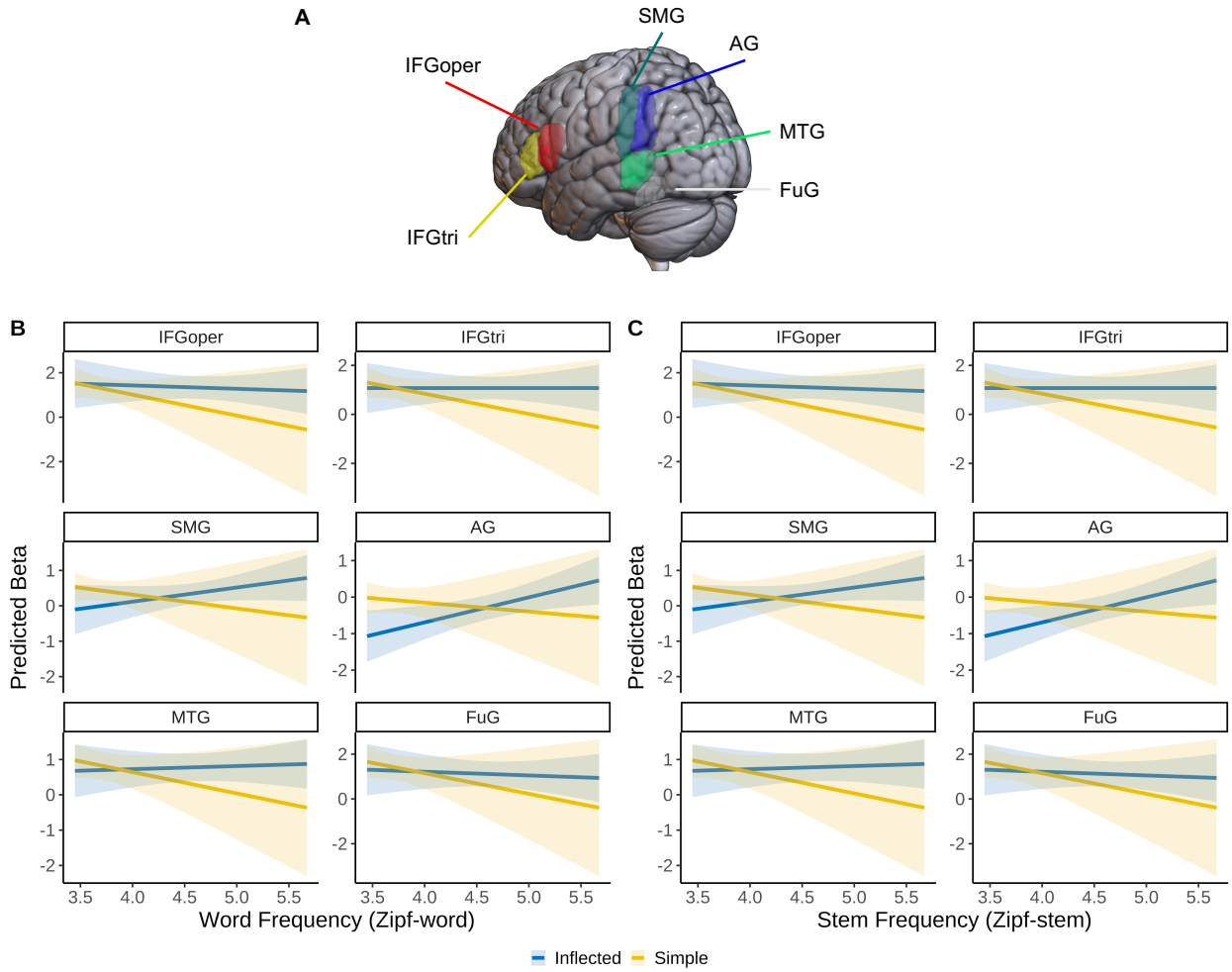

Figure 4. Regions of interest univariate analysis results. (A) Locations of anatomically-defined ROIs: IFG pars opercularis (IFGoper), IFG pars triangularis (IFGtri), supramarginal gyrus (SMG), angular gyrus (AG), middle temporal gyrus (MTG), and fusiform gyrus (FuG). (B) Predicted BOLD β as a function of Zipf word frequency for Inflected (blue) and Simple (yellow) nouns, extracted from mixed-effects models. (C) The same analysis for Zipf stem frequency. Shaded ribbons represent 95 % confidence intervals.

3.4. MVPA results

To test whether local activation patterns discriminated between morphologically simple and inflected nouns, we conducted multivoxel pattern classification using six anatomically defined ROIs. As shown in **Fig. 5**, classification accuracy measured by area under the curve (AUC) did not exceed chance level (0.5) in any ROI. One-sample *t*-tests against chance yielded no

significant effects after correction (all *p*s > .82), indicating that morphological condition could not be reliably decoded from multivoxel patterns in these regions.

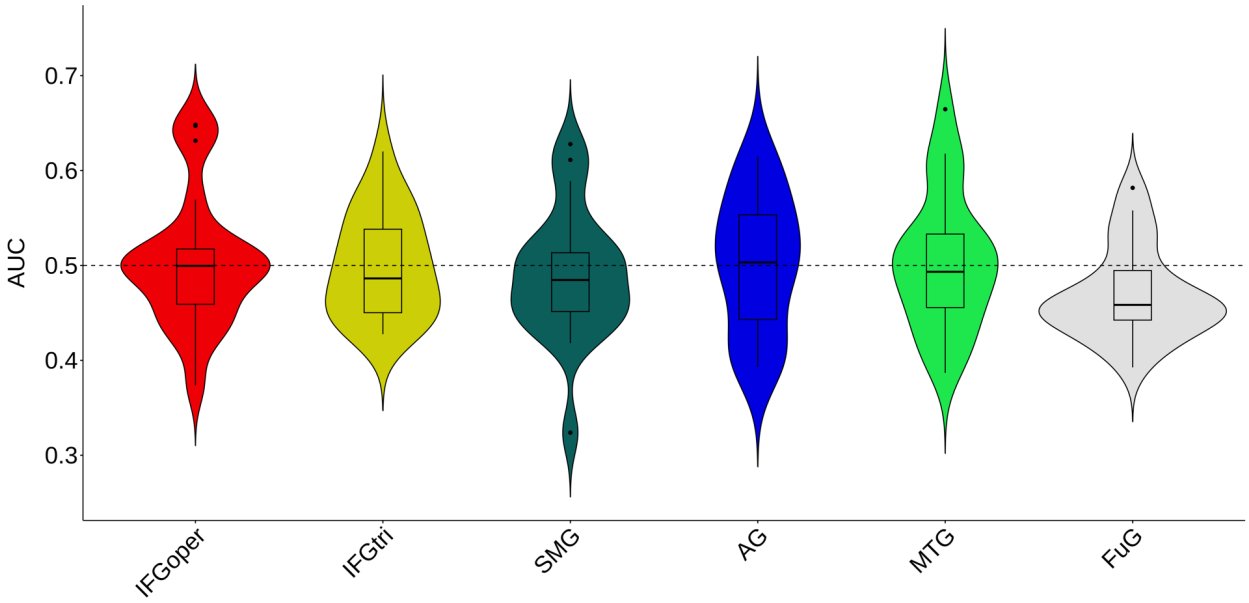

Figure 5. ROI-based multivariate voxel pattern analysis (MVPA) results. Violin/box plots show the distribution of subject-level area-under-the-curve (AUC) scores for classifying Inflected vs. Simple nouns from single-trial beta patterns within each ROI. The dotted line marks chance performance (AUC = 0.5).

3.5. RSA results

The results of representational similarity analysis (RSA) are presented in **Figure 6**. We assessed whether local multivoxel patterns reflected lexical frequency structure by correlating neural RDMs with theoretical RDMs based on either whole-word frequency (Zipf) or stem frequency (Zipf-Stem). RSA was conducted within six anatomically defined ROIs (IFGoper, IFGtri, AG, SMG, MTG, FuG). Group-level RSA effects were tested using one-sample t-tests (against zero) for each model (i.e., Zipf and Zipf-Stem), condition (All, Inflected, Simple), and ROI. To directly compare encoding strength across conditions, we additionally conducted paired two-

sample t-tests between Inflected and Simple nouns. All p-values were corrected using false discovery rate (FDR) correction.

In RSA using the word frequency (Zipf-word) model (**Fig. 6A**), several ROIs showed significant encoding for word frequency. One-sample *t*-tests revealed significant encoding of Zipf frequency for inflected nouns in IFGoper ($t(20) = 3.30$, $pFDR = .025$). Notably, AG ($t(20) = 2.50$, uncorrected $p = .021$) and SMG ($t(20) = 2.22$, uncorrected $p = .038$) also showed trend-level effects for inflected nouns, although these did not survive FDR correction. In contrast, RSA correlations for simple nouns significantly differed from zero in IFGoper ($t(20) = -3.80$, $pFDR = .020$) and SMG ($t(20) = -3.33$, $pFDR = .025$), indicating negative similarity effects. Paired t-tests revealed significantly stronger encoding of word frequency for inflected compared to simple nouns in IFGoper ($t(20) = 6.44$, $pFDR = .0001$) and SMG ($t(20) = 3.95$, $pFDR = .0096$), suggesting that morphological complexity enhances the neural encoding of frequency information in these regions.

In contrast, RSA using the stem frequency model (Zipf-Stem; **Fig. 6B**) showed generally weaker effects. IFGoper ($t(20) = -3.80$, $pFDR = .020$) and SMG ($t(20) = -3.33$, $pFDR = .025$) showed significant negative RSA correlations for simple nouns. Trend-level negative RSA effects were observed in SMG for the All condition ($t(20) = -2.41$, uncorrected $p = .026$). However, no ROI showed significant above-zero encoding of Zipf-Stem values in the All or Inflected conditions, nor were significant differences found between inflected and simple nouns.

In sum, RSA revealed that whole-word frequency (Zipf) was significantly encoded in frontoparietal regions, especially IFGoper and SMG, with stronger encoding for inflected compared to simple nouns. Stem frequency (Zipf-Stem), in contrast, was only negatively

encoded in IFGoper and SMG for simple nouns, suggesting that morphological complexity enhances sensitivity primarily to whole-word rather than stem-level frequency structure.

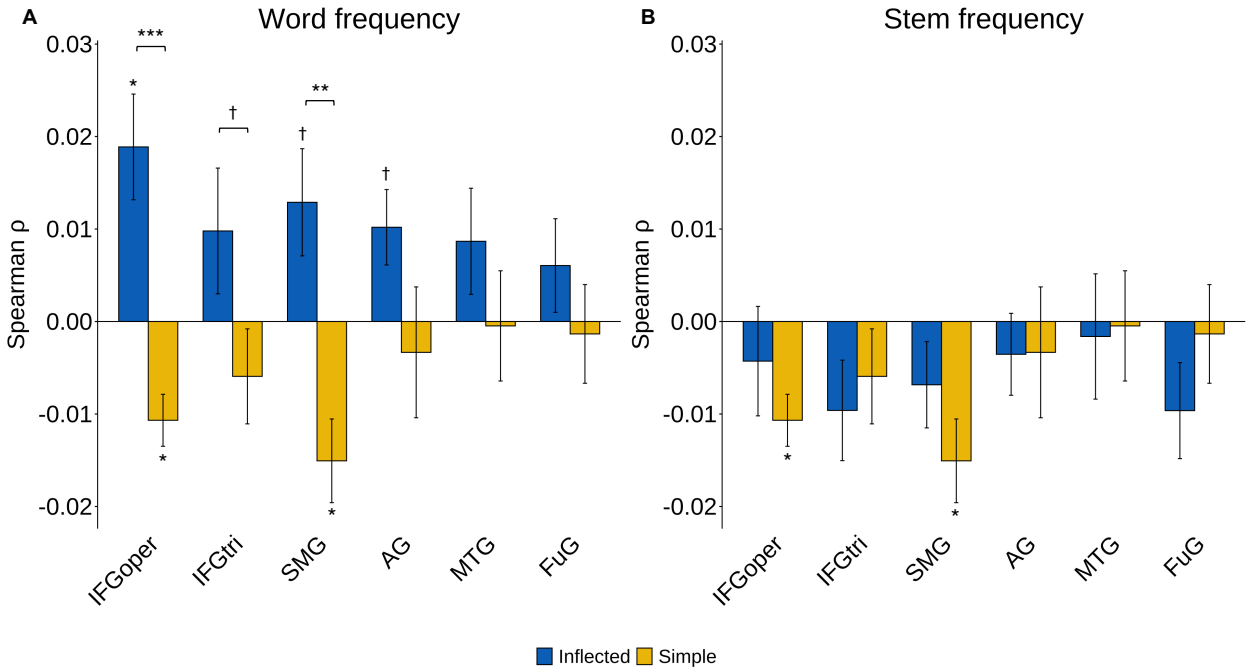

Figure 6. Representational similarity analysis (RSA) results. Mean Spearman ρ values between neural RDMs and (A) word-frequency (Zipf) or (B) stem-frequency (Zipf-Stem) model RDMs are plotted for each ROI (blue = Inflected; yellow = Simple); vertical bars represent ± SEM. Significance of each individual bar was tested with a one-sample t-test against zero and is indicated directly above the bar (* $p < .05$; ** $p < .01$; *** $p < .001$; † uncorrected $p < .05$, FDR-corrected within panel). Horizontal brackets linking the blue-yellow bar pairs denote paired-sample t-tests that compared Inflected and Simple correlations within the same ROI, using the same symbol conventions for significance.

4. Discussion

The present study aimed to examine how morphological complexity and lexical frequency interact in the human brain, testing competing predictions of obligatory decomposition versus flexible-access models of visual word recognition. Using a rapid event-related fMRI paradigm with Korean nouns, we combined univariate mixed-effects modeling, multivoxel pattern analysis (MVPA), and representational similarity analysis (RSA) within six left-hemisphere language regions (IFG pars opercularis and triangularis, MTG, fusiform gyrus, supramarginal gyrus, and angular gyrus). Our key findings were as follows: (1) MVPA failed to decode simple versus inflected nouns from local activation patterns in any ROI; (2) RSA revealed robust encoding of surface (whole-word) frequency but not base (stem) frequency, with significantly stronger correlations for inflected than simple forms in IFG pars opercularis and SMG; and (3) univariate ROI analyses showed a significant interaction of morphology and surface frequency—higher surface frequency elicited greater activation for inflected nouns in SMG and AG, whereas stem frequency did not modulate these regions. Collectively, these results favor a model in which morphological access is surface-driven and context-sensitive, challenging theories that posit an obligatory, automatic decomposition into stems and affixes.

First, the null MVPA results suggest that morphologically simple and inflected forms are not represented as discrete categories in the examined language ROIs. If obligatory decomposition yielded qualitatively distinct activation patterns for simple versus inflected nouns, one would expect above-chance classification of morphological condition. Instead, AUCs clustered around chance level across all ROIs. This absence of categorical pattern separability complements RSA and univariate findings by suggesting that morphological complexity does not engender a binary neural code in fronto-temporo-parietal regions. Rather, morphological

differences may be reflected in graded modulation by surface properties, specifically frequency, rather than in distinct pattern templates.

The second key finding is that RSA provided strong evidence that local activation patterns systematically encode surface frequency, particularly for inflected nouns, while base frequency remains unrepresented or even negatively correlated in simple forms. IFG pars opercularis and SMG exhibited the strongest RSA correlations with the surface-frequency RDM for inflected nouns, and paired-sample tests confirmed significantly stronger encoding for inflected than simple forms. These RSA results align with univariate interactions and underscore that graded frequency structure is embedded in multi-voxel patterns. In distributional frameworks, high-frequency complex forms can be directly accessed via full-form representations (Baayen et al., 2011; Plaut & Gonnerman, 2000). Our data extend this view to the neural level, demonstrating that surface frequency—not base frequency—drives representational similarity in key language areas.

Third, the univariate ROI analysis yielded convergent evidence: SMG and AG in the inferior parietal lobule showed significant interactions of surface frequency and morphology, such that activation increased with surface frequency for inflected but not for simple nouns. The lack of reliable stem-frequency effects across ROIs further emphasizes that base frequency does not exert the predicted influence under obligatory-decomposition models. In classical dual-route accounts, stem frequency should modulate activation in morpho-syntactic regions whenever decomposition occurs (Marslen-Wilson et al., 2008). Instead, our findings suggest that surface frequency selectively enhances retrieval processes in parietal regions during the recognition of morphologically complex forms.

Taken together, these results challenge purely decompositional theories (Rastle et al., 2004; Taft & Forster, 1975) that posit mandatory parsing of all complex words. Although masked-priming and early ERP studies have demonstrated rapid form-based segmentation (Lavric et al., 2011; Solomyak & Marantz, 2010), such paradigms may tap initial pre-lexical processing stages that do not translate into sustained neural representations detectable by fMRI. Instead, our data support distributional or hybrid models (Baayen et al., 2011; Milin et al., 2017; Plaut & Gonnerman, 2000) in which both decomposition and whole-form access routes exist, with the engagement of each pathway determined by lexical statistics such as surface frequency.

Region-specific interpretations further clarify functional roles within the language network. The RSA effects in the IFG pars opercularis suggest that this region encodes graded lexical-frequency structure rather than serving solely as a competition-resolution hub. Posterior IFG has been associated with selection among competing candidates (Nagel et al., 2008; Thompson-Schill et al., 1997), but our RSA results, showing stronger frequency encoding for inflected words, imply a broader role in mapping statistics-driven access pathways. The SMG and AG univariate and RSA interactions highlight the IPL involvement in integrating lexical-semantic information weighted by whole-form frequency (Binder et al., 2009; Graves et al., 2010). These parietal areas may support retrieval of established lexical entries for frequent complex forms rather than decomposition of novel or low-frequency items.

However, the MTG, though implicated in both semantic control and representation (Lau et al., 2008; Whitney et al., 2011), did not show consistent frequency effects across analyses, suggesting a more nuanced or context-dependent role that may emerge under different task demands (e.g., semantic judgment versus lexical decision) or on sentence-level context not present here. Findings in fusiform gyrus and ventral occipitotemporal cortex have been

inconsistent across studies. Some report stronger responses to low-frequency words (Bruno et al., 2008; Fiebach et al., 2002; Kronbichler et al., 2004), while others find null effects (Fiebach et al., 2002). In our study, whole-brain analysis revealed reliable activation of the fusiform gyrus for all words, confirming its role in visual word processing. However, the anatomically broad FuG ROI showed no reliable modulation by surface or stem frequency. This apparent discrepancy may reflect both the task and the ROI definition. Classic word frequency effects in the mid-fusiform visual word-form area are most robust during passive reading or naming and when longer letter strings extend across several degrees of visual angle. Our two- to three-syllable nouns presented in a decision-driven context likely reduced reliance on ventral occipitotemporal coding, shifting frequency sensitivity to frontoparietal circuits. In addition, the Harvard–Oxford mask spans anterior and posterior fusiform regions, which may have diluted the signal from the mid-fusiform strip most sensitive to lexical frequency. Future studies employing functionally defined vOT ROIs, higher-resolution acquisition, and tasks that minimize decisional load may better reveal the canonical frequency gradient that remained sub-threshold in the present design.

Finally, the absence of categorical discriminability in our MVPA results aligns with previous neuroimaging findings on Korean inflectional morphology (Kim et al., 2022, 2024; Yim et al., 2006). Specifically, two fMRI works on Korean inflected verbs showed an equivalent IFG activation for regular and irregular verbs (Kim et al., 2024; Yim et al., 2006). Notably, Kim et al. (2024) observed a graded effect of verb regularity in the fusiform gyrus modulated by semantic homonymy, in accordance with our finding that IPL activation was associated with surface frequency, a lexical-semantic factor. Likewise, Kim et al. (2022) demonstrated that early ERP components (N100/N250) elicited by inflected verbs displayed systematic variations in relation to surface frequency and stem length, consistent with our RSA findings of frequency-

based representational geometry. Together, these studies suggest that both noun and verb inflections in Korean are encoded along continuous, experience-driven dimensions rather than as discrete morphological categories, thereby providing evidence for a distributional framework of morphological processing.

5. Limitations and Conclusion

Several limitations warrant consideration. First, the temporal resolution of fMRI precludes direct observation of the rapid decomposition processes indexed by ERP. Obligatory segmentation may occur only in the first 200 ms and then be overwritten by whole-form access; our BOLD measures cannot adjudicate that transient stage. Second, the sample size ($N = 21$) and ROI approach yield adequate power for medium-to-large effects but may miss subtle interactions—particularly in large regions such as MTG and the FuG where signal-to-noise is lower. Third, MTG and fusiform nulls should be interpreted cautiously: our anatomically broad masks likely diluted signal from mid-fusiform patches most sensitive to frequency, and the lexical-decision context may have shifted processing demands away from vOT. High-resolution, functionally defined vOT ROIs or passive reading designs may yet reveal the classic frequency gradient. Fourth, the exclusive focus on Korean, a transparent, agglutinative writing system facilitates the dissociation of surface and stem frequency but limits extension to fusional or infixing languages in which orthographic change often accompanies inflection. Finally, the lexical-decision task emphasizes decisional and attentional control; converging evidence from speeded naming, semantic judgment, or eye-tracking will be needed to determine the generality of the surface-frequency advantage.

Future research should therefore (i) combine EEG/MEG with fMRI to capture both early segmentation and sustained representational geometry, (ii) employ larger and cross-linguistic samples with language-specific ROI definitions, (iii) use tasks that parametrically vary decomposition demands, and (iv) incorporate computational modelling to trace how distributed codes evolve as a joint function of morphology and frequency (Plaut & Gonnerman, 2000).

In conclusion, our findings provide the first fMRI evidence comparing base and surface frequency effects for morphologically complex Korean words, combining univariate and multivariate analyses. The results suggest that morphological complexity modulates neural encoding of surface—rather than not base—frequency, and that morphological condition does not yield discrete category representations in classic language areas. These findings suggest that rather than a single, obligatory decomposition process, morphological processing may arise from the dynamic interplay between a word's structure and its statistical properties, shaped by both form-based and meaning-based constraints (Weissbart & Martin, 2024).

Reference

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*(FEB). https://doi.org/10.3389/fninf.2014.00014

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Milin, P., rević, D. F., Hendrix, P., & Marelli, M. (2011). An Amorphous Model for Morphological Processing in Visual Comprehension Based on Naive Discriminative Learning. *Psychological Review*, *118*(3), 438–481. https://doi.org/10.1037/a0023851

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. In *J. R. Statist. Soc. B* (Vol. 57, Issue I). https://academic.oup.com/jrsssb/article/57/1/289/7035855

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The Balance of Storage and Computation in Morphological Processing: The Role of Word Formation Type, Affixal Homonymy, and Productivity. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(2), 489–511. https://doi.org/10.1037/0278-7393.26.2.489

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Bozic, M., Marslen-Wilson, W. D., Stamatakis, E. A., Davis, M. H., & Tyler, L. K. (2007). Differentiating morphology, form, and meaning: Neural correlates of morphological complexity. *Journal of Cognitive Neuroscience*, *19*(9), 1464–1475. https://doi.org/10.1162/jocn.2007.19.9.1464

Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., & others. (2002). Region of interest analysis using the MarsBar toolbox for SPM 99. *Neuroimage*, *16*(2), S497.

Bruffaerts, R., Dupont, P., Peeters, R., De Deyne, S., Storms, G., & Vandenberghe, R. (2013). Similarity of fMRI activity patterns in left perirhinal cortex reflects semantic similarity between words. *Journal of Neuroscience*, *33*(47), 18597–18607. https://doi.org/10.1523/JNEUROSCI.1548-13.2013

Bruno, J. L., Zumberge, A., Manis, F. R., Lu, Z. L., & Goldman, J. G. (2008). Sensitivity to orthographic familiarity in the occipito-temporal region. *NeuroImage*, *39*(4), 1988–2001. https://doi.org/10.1016/j.neuroimage.2007.10.044

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111

Carota, F., Bozic, M., & Marslen-Wilson, W. (2016). Decompositional Representation of Morphological Complexity: Multivariate fMRI Evidence from Italian. *Journal of Cognitive Neuroscience*, *28*(12), 1878–1896. https://doi.org/10.1162/jocn_a_01009

Carreiras, M., Riba, J., Vergara, M., Heldmann, M., & Münte, T. F. (2009). Syllable congruency and word frequency effects on brain activation. *Human Brain Mapping*, *30*(9), 3079–3088. https://doi.org/10.1002/hbm.20730

Desai, R., Conant, L. L., Waldron, E., & Binder, J. R. (2006). fMRI of Past Tense Processing: The Effects of Phonological Complexity and Task Difficulty. *Journal of Cognitive Neuroscience*, *18*(2), 278–297. https://doi.org/10.1162/089892906775783633

Devlin, J. T., Jamison, H. L., Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences*, *101*(41), 14984–14988. https://doi.org/10.1073/pnas.0403766101

Feldman, L. B., Milin, P., Cho, K. W., Moscoso del Prado Martín, F., & O'Connor, P. A. (2015). Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00111

Fiebach, C. J., Friederici, A. D., Müller, K., & Cramon, D. Y. von. (2002). fMRI Evidence for Dual Routes to the Mental Lexicon in Visual Word Recognition. *Journal of Cognitive Neuroscience*, *14*(1), 11–23. https://doi.org/10.1162/089892902317205285

Fiez, J. A., Balota, D. A., Raichle, M. E., & Petersen, S. E. (1999). Effects of Lexicality, Frequency, and Spelling-to-Sound Consistency on the Functional Anatomy of Reading. *Neuron*, *24*(1), 205–218. https://doi.org/10.1016/S0896-6273(00)80833-8

Fischer-Baum, S., Bruggemann, D., Gallego, I. F., Li, D. S. P., & Tamez, E. R. (2017). Decoding levels of representation in reading: A representational similarity approach. *Cortex*, *90*, 88–102. https://doi.org/10.1016/j.cortex.2017.02.017

Graves, W. W., Desai, R., Humphries, C., Seidenberg, M. S., & Binder, J. R. (2010). Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex*, *20*(8), 1799–1815. https://doi.org/10.1093/cercor/bhp245

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430. https://doi.org/10.1126/science.1063736

Jared, D., Jouravlev, O., & Joanisse, M. F. (2017). The effect of semantic transparency on the processing of morphologically derived words: Evidence from decision latencies and event-related potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(3), 422–450. https://doi.org/10.1037/xlm0000316

Kang, B., & Kim, H. (2009). *Frequency of Korean language: based on 15 million size Eojeol Sejong syntactic and semantic corpus*. Handkook munhwasa.

Kim, J., Kang, J., Kim, J., & Nam, K. (2022). Temporal dynamics of form and meaning in morphologically complex word processing: An ERP study on Korean inflected verbs. *Journal of Neurolinguistics*, *64*, 101098. https://doi.org/10.1016/j.jneuroling.2022.101098

Kim, J., Kim, S., & Nam, K. (2024). Neural Dynamics of Processing Inflectional Morphology: An fMRI Study on Korean Inflected Verbs. *Brain Sciences*, *14*(8), 752. https://doi.org/10.3390/brainsci14080752

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*(NOV). https://doi.org/10.3389/neuro.06.004.2008

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage*, *21*(3), 946–953. https://doi.org/10.1016/j.neuroimage.2003.10.021

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876–895. https://doi.org/10.1037/a0013484

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/JSS.V082.I13

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. In *Nature Reviews Neuroscience* (Vol. 9, Issue 12, pp. 920–933). https://doi.org/10.1038/nrn2532

Lavric, A., Rastle, K., & Clapp, A. (2011). What do fully visible primes and brain potentials reveal about morphological decomposition? *Psychophysiology*, *48*(5), 676–686. https://doi.org/10.1111/j.1469-8986.2010.01125.x

Lehtonen, M., Vorobyev, V. A., Hugdahl, K., Tuokkola, T., & Laine, M. (2006). Neural correlates of morphological decomposition in a morphologically rich language: An fMRI study. *Brain and Language*, *98*(2), 182–193. https://doi.org/10.1016/j.bandl.2006.04.011

Lewis, G., Solomyak, O., & Marantz, A. (2011). The neural basis of obligatory decomposition of suffixed words. *Brain and Language*, *118*(3), 118–127. https://doi.org/10.1016/j.bandl.2011.04.004

Longtin, C.-M., Segui, J., & Hallé, P. A. (2003). Morphological priming without morphological relationship. *Language and Cognitive Processes*, *18*(3), 313–334. https://doi.org/10.1080/01690960244000036

Marslen-Wilson, W., Bozic, M., & Randall, B. (2008). Early decomposition in visual word recognition: Dissociating morphology, form, and meaning. *Language and Cognitive Processes*, *23*(3), 394–421. https://doi.org/10.1080/01690960701588004

Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLoS ONE*, *12*(2). https://doi.org/10.1371/journal.pone.0171935

Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *NeuroImage*, *43*(4), 801–807. https://doi.org/10.1016/j.neuroimage.2008.07.017

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Pedregosa, F., Varoquaux, G., Gramfort, A., Micheal, V., Thirion, B., Grisel, O., Bloundel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., & Duchesnay, M. P. É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830. http://scikit-learn.sourceforge.net.

Plaut, D. C., & Gonnerman, L. M. (2000). Are non-semantic morphological effects icompatible with a distributed connectionist approach to lexical processing? *Language and Cognitive Processes*, *15*(4–5), 445–485. https://doi.org/10.1080/01690960050119661

Pliatsikas, C., Wheeldon, L., Lahiri, A., & Hansen, P. C. (2014). Processing of zero-derived words in English: An fMRI investigation. *Neuropsychologia*, *53*(1), 47–53. https://doi.org/10.1016/j.neuropsychologia.2013.11.003

R Core Team. (2022). *R: A language and environment for statistical computing* (4.3.1). R Foundation for Statistical Computing. https://www.r-project.org/

Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, *11*(6), 1090–1098. https://doi.org/10.3758/BF03196742

Solomyak, O., & Marantz, A. (2010). Evidence for Early Morphological Decomposition in Visual Word Recognition. *Journal of Cognitive Neuroscience*, *22*(9), 2042–2057. https://doi.org/10.1162/jocn.2009.21296

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638–647. https://doi.org/10.1016/S0022-5371(75)80051-X

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*(26), 14792–14797. https://doi.org/10.1073/pnas.94.26.14792

Weissbart, H., & Martin, A. E. (2024). The structure and statistics of language jointly shape cross-frequency neural dynamics during spoken language comprehension. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-53128-1

Whitney, C., Jefferies, E., & Kircher, T. (2011). Heterogeneity of the left temporal lobe in semantic representation and control: Priming multiple versus single meanings of ambiguous words. *Cerebral Cortex*, *21*(4), 831–844. https://doi.org/10.1093/cercor/bhq148

Yim, H., Park, C., Lim, H., & Nam, K. (2006). Mental Representation and Processing Involved in Comprehending Korean Regular and Irregular Verb Eojeols: An fMRI and Reaction Time Study. *Neural Information Processing: 13th International Conference, ICONIP 2006, Hong Kong, China, October 3-6, 2006. Proceedings, Part I 13*, *4232*, 247–254.

Yue, Q., & Martin, R. C. (2021). Maintaining verbal short-term memory representations in non-perceptual parietal regions. *Cortex*, *138*, 72–89. https://doi.org/10.1016/j.cortex.2021.01.020