

The Spatiotemporal Dynamics of Bottom-Up and Top-Down Processing during At-a-Glance Reading

Nigel Flower¹ and  Liina Pykkänen^{1,2}

Departments of ¹Linguistics and ²Psychology, New York University, New York, New York 10003

Like all domains of cognition, language processing is affected by top-down knowledge. Classic evidence for this is missing blatant errors in the signal. In sentence comprehension, one instance is failing to notice word order errors, such as transposed words in the middle of a sentence: “you that read wrong” (Mirault et al., 2018). Our brains seem to fix such errors, since they are incompatible with our grammatical knowledge, but how do our brains do this? Following behavioral work on inner transpositions, we flashed four-word sentences for 300 ms using rapid parallel visual presentation (Snell and Grainger, 2017). We compared magnetoencephalography responses to fully grammatical and reversed sentences (24 human participants: 21 females, 4 males). The left lateral language cortex robustly distinguished grammatical and reversed sentences starting at 213 ms. Thus, the influence of grammatical knowledge begun rapidly after visual word form recognition (Tarkiainen et al., 1999). At the earliest stage of this neural “sentence superiority effect,” inner transpositions patterned between grammatical and reversed sentences, showing evidence that the brain initially “noticed” the error. However, 100 ms later, inner transpositions became indistinguishable from grammatical sentences, suggesting at this point, the brain had “fixed” the error. These results show that after a single glance at a sentence, syntax impacts our neural activity almost as quickly as higher-level object recognition is assumed to take place (Cichy et al., 2014). The earliest stage involves detailed comparisons between the bottom-up input and grammatical knowledge, while shortly afterward, top-down knowledge can override an error in the stimulus.

Key words: bottom-up processing; error detection; magnetoencephalography; rapid parallel visual presentation; syntax; top-down processing

Significance Statement

Language processing, like all cognitive domains, is profoundly influenced by top-down knowledge, evident in the oversight of errors in the signal. For example, individuals often miss order errors, such as transposed words midsentence. Utilizing rapid parallel visual presentation, we investigated this phenomenon by exposing participants to four-word sentences for 300 ms. Magnetoencephalography revealed robust differentiation between grammatical and reversed sentences in the left lateral language cortex starting at 213 ms postpresentation. Intriguingly, initial neural responses to inner transpositions treated them as deviant, but 100 ms later, neural signals grouped them with grammatical sentences, indicating rapid error correction. These findings reveal the brain’s remarkable capacity to reconcile bottom-up input with linguistic knowledge almost instantaneously.

Introduction

Language comprehension involves both detailed, bottom-up analysis of a stimulus and top-down grammatical constraints (Gibson, 2006; Matar et al., 2021). The “sentence superiority effect,” or SSE, is an instance of top-down syntactic knowledge

guiding the interpretation of a linguistic stimulus. In studies involving the presentation of four-word stimuli at 200 ms using the rapid parallel visual presentation paradigm (RPVP), participants show facilitated processing of grammatical sentences relative to scrambled sentences (Snell and Grainger, 2017; Massol et al., 2021). When the stimuli obey the grammatical rules of the participants’ native language, this syntactic knowledge can be deployed to rapidly form a sentential representation of the stimulus, requiring no longer than 200 ms of presentation.

Mirault et al. (2018) present a second kind of behavioral effect seen during parallel presentation, known as the transposed-word effect (TWE). When presented with sentences containing minor word errors such as “you that read wrong”, readers easily

Received Feb. 26, 2024; revised June 21, 2024; accepted Aug. 26, 2024.

Author contributions: N.F. and L.P. designed research; N.F. performed research; N.F. contributed unpublished reagents/analytic tools; N.F. analyzed data; N.F. wrote the paper.

This work was supported by the National Science Foundation award #2335767 (LP) and award G1001 from NYUAD Institute, New York University Abu Dhabi (LP). We thank Alex White and Dustin Chacón for feedback and discussion.

Correspondence should be addressed to Nigel Flower at nff2102@nyu.edu.

<https://doi.org/10.1523/JNEUROSCI.0374-24.2024>

Copyright © 2024 the authors

interpret these as their grammatical counterparts. PEGADO and Grainger (2020) propose that this behavioral effect is evidence for an error of sensory, bottom-up processing of the stimulus that takes place in parallel, along the lines of reading models such as Snell et al. (2018). Both of these claims are, however, controversial, and other work has argued that instead, the TWE is compatible with serial processing models of language (Reichle et al., 2009) and is an instance of top-down grammatical knowledge revising the initial analysis of the sentence with the inner two words transposed (Huang and Staub, 2023). The claim that the recognition and processing of multiple words can take place in a parallel fashion (Snell et al., 2018) is also controversial given that most prominent theories of sentence processing assume incremental processing that takes place word-by-word (Frazier and Fodor, 1978; Hale, 2001; Lewis and Vasishth, 2005).

The literature employing the RPVP paradigm largely consists of behavioral and electroencephalography (EEG) studies (Wen et al., 2019, 2021; Dunagan et al., 2024). In this study, we make use of the superior spatial resolution of magnetoencephalography (MEG) to localize the neural source of the SSE. A key objective of this study is to gain insight into how the brain “fixes” erroneous stimuli containing an inner transposition. Is the bottom-up analysis of the input guided by top-down syntactic predictions in even the earliest stages as suggested by parallel models of reading (Snell et al., 2018; PEGADO and Grainger, 2020), or does top-down grammatical knowledge operate slightly later in processing (Huang and Staub, 2023) and revise an analysis that is unlikely given our knowledge of the structure of our language? Finally, we aim to offer insight into whether at-a-glance reading of multiple words involves serial (Reichle et al., 1998, 2009) or parallel (Snell et al., 2018) mechanisms.

To answer these questions, we conducted an MEG experiment with 24 native speakers of English using parallel presentation of linguistic stimuli. We presented subjects with grammatical sentences, sentences with inner transpositions, and reversed sentences (Fig. 1) to test which cortical regions show activity consistent with a neural SSE and at what point do these regions “correct” the inner transpositions and treat them as their grammatical counterparts. To query whether the brain uses serial or parallel mechanisms in at-a-glance reading of parallel input, we also used the regions discovered in the first analysis as functional regions of interest (fROIs) in a two-stage regression analysis to test whether bigram frequencies and word-to-word transition probabilities across the four-word stimulus impacted neural activity at the same time or in a more sequential fashion. We correlated bigram frequency and transition probability with the source estimates from the fROIs of each participant using simple linear models. Multiple bigram frequency effects in overlapping time windows were taken to suggest parallel processing and multiple transition probability effects in sequential time windows to reflect serial processing of the multiword stimulus (Fig. 5).

Materials and Methods

Participants

Thirty right-handed native speakers of English participated in the study. All participants had normal or corrected-to-normal vision and gave informed consent. Two recordings were excluded from analysis due to excessive noise, and three participants were excluded due to falling asleep. As a result, a total of 25 participants (21 women; 18–40 years old; mean age, 22.583; SD, 4.293) were included in the behavioral and MEG analyses.

Design

To investigate the neural bases of the SSE, we employed a contrast between clearly grammatical and clearly ungrammatical sentences,

both of which were then also compared with sentences containing an inner transposition. Specifically, to obtain a behavioral and a neural SSE, our sentence-type manipulation contrasted grammatical sentences such as “all cats are nice” to reversed versions of these stimuli such as “nice are cats all.” Given that the aim of our study was to test how top-down grammatical knowledge may serve to fix certain types of errors in the signal (such as inner transpositions), it was crucial that our choice of an ungrammatical control truly behaved as ungrammatical both behaviorally and in neural signals. To assure this, we piloted the grammatical versus reversed contrast before embarking on the full study, finding both a robust behavioral and a neural SSE, that is, faster and more accurate behavioral responses and increased neural signals for the grammatical sentences. Given the robust divergence, the reversed sentences were chosen for the clearly ungrammatical condition. The reversed sentences were in fact initially created as a double-transposition condition, reversing both the inner two words as in the inner-transposition condition and then adding a transposition of the first and last words, but since this in fact yields a reversed sentence, the label “reversed” will be used here for simplicity.

Each of our grammatical stimuli were formed by creating a set of 50 plural noun–adjective pairs (“cats–nice”), inserting a determiner to the left of the plural noun (“all cats–nice”), and finally inserting the verb “are” in between the noun and the adjective. To test whether inner transpositions are detected during the composition of a parallelly presented stimulus, an inner-transposition condition was created by taking the grammatical stimulus and simply swapping the second and third words of the sentence (“all are cats nice”). Finally, for the purposes of another project, we also varied the kind of determiner used for the sentences. A total of four determiners were chosen, “all,” “some,” “no,” and “the,” yielding a 3×4 experimental design (Fig. 1a). Because 50 noun–adjective pairs were used per condition, the total number of trials amounted to 600.

The nouns and adjectives in this study were selected based on character length to ensure that the sentences would be fully within the visual range of the fovea and the parafovea. Naturalness and plausibility were also considered when selecting the noun–adjective pairs for each stimulus. Each of the nouns in the study was a total of three characters long, so that they would only be four characters long in their plural form. The adjectives were either three or four characters long. The average length of the stimuli is 16.72 characters with a standard deviation of 0.84, with a minimum of 15 and a maximum of 18. The stimuli on average occupied 6.11° of the visual field, meaning that most of the stimuli were close within the central visual field.

The trials began with a fixation cross on for 200 ms and off screen for 200 ms. A sentence from the experimental design (Fig. 1) was then presented for 300 ms, followed by a blank screen for 500 ms. A second sentence was then presented. This sentence was either identical to the first or involved replacing one word of the first sentence with another word taken from the lexicon used to generate the stimuli. The second sentence remained on screen until the participant marked whether the target was a match or a mismatch with respect to the first stimulus. Both the first and second stimuli were enclosed in a dark, gray rectangular box with a width of 300 pixels and a height of 50 pixels. The box was placed around the stimuli to direct participants’ gaze to the center of the screen and to help discourage eye movements outside of the boundary. The box was not presented in the intervening 500 ms between Sentence 1 and 2. The structure of a complete trial is presented below in Figure 1.

Procedure

Before the MEG recording, each participant had their head shape scanned with a Polhemus FastSCAN three-dimensional laser digitizer to locate the positions of marker coils placed on the head during recording. The digitized head shape was used during the data preprocessing stage to constrain source localization data. Participants had the option of pausing the experiment and resting every 75 trials. The stimuli were presented to participants using the PsychoPy package in Python (Peirce et al., 2010) on a screen roughly 50 cm away from the participant’s face. The stimuli were presented in white Courier New font with a visual angle of $0^\circ 34'$ against a gray background. Trials were

a.

| Transposition | Quantifier | Example | Structure |
|-----------------|------------|--------------------|---|
| Grammatical | all | all cats are nice | <pre> graph TD all[all] --- cats[cats] all --- are[are] cats --- nice[nice] </pre> |
| | some | some cats are nice | |
| | no | no cats are nice | |
| | the | the cats are nice | |
| Inner Transpose | all | all are cats nice | <pre> graph TD all[all] --- are[are] are --- cats[cats] are --- nice[nice] style are stroke-dasharray: 5 5 </pre> |
| | some | some are cats nice | |
| | no | no are cats nice | |
| | the | the are cats nice | |
| Reversed | all | nice are cats all | <pre> graph TD nice[nice] --- are[are] are --- cats[cats] are --- all[all] style are stroke-dasharray: 5 5 </pre> |
| | some | nice are cats some | |
| | no | nice are cats no | |
| | the | nice are cats the | |

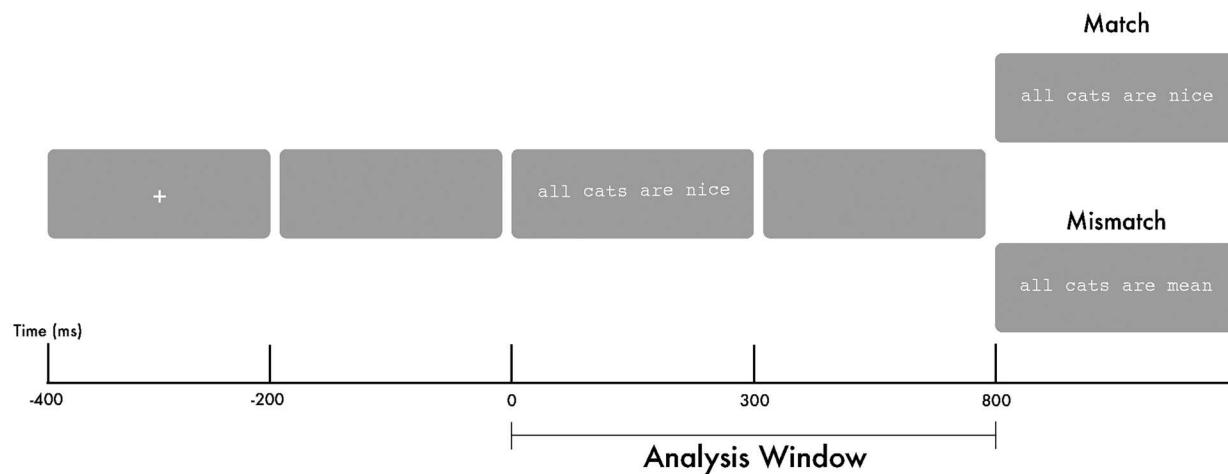
b.**Task: Match or Mismatch?**

Figure 1. Experimental design and trial structure. **a**, Full 3×4 factorial design with three sentence types (grammatical, inner-transposed, and reversed) using four different types of quantifiers/determiners. Apart from adding variability to the stimuli, the quantifier manipulation was not relevant for the aims of the current study. **b**, In our trial structure, the first, critical sentence was followed by either a matching stimulus (top) or a mismatching stimulus (bottom), prompting a match/mismatch task. MEG analysis targeted an epoch of 800 ms after the onset of the first sentence.

completely randomized for each participant. Participants had the option of taking a break every time they completed an eighth of the trials. Before the participants underwent the experiment, they each completed 10 practice trials to familiarize themselves with the procedure. The duration of the experiment for each the participant lasted roughly 25 min with breaks.

MEG data acquisition and preprocessing

The raw MEG data were collected using a whole-head 157-channel axial gradiometer system (Kanazawa Institute of Technology) with a sampling rate of 1,000 Hz. During data collection, the MEG data were filtered with a high-pass filter of 1 Hz, due to the high amount of NYC environmental noise, and a low-pass filter of 200 Hz. After collection, the raw MEG data

were then noise-reduced using the continuously adjusted least-square method algorithm (Adachi et al., 2001) in the MEG 160 software (Meg Laboratory 2.004A, Yokogawa Electric, Kanazawa Institute of Technology). All further preprocessing stages used MNE-Python (Gramfort et al., 2013, 2014). The noise-reduced data were further low-pass filtered at 40 Hz. Bad channels were removed by visual inspection, and independent component analysis was performed on the data to isolate and remove artifacts such as heartbeat and eyeblinks. The cleaned data were epoched from 100 ms before the onset of Sentence 1 to 800 ms after the presentation of the sentence, resulting in epochs of 900 ms. The first 100 ms of the epoch was used as a baseline in the calculation of the noise-covariance matrix. Individual epochs for which any of the sensor values exceeded 3,000 fT at any time point were rejected.

Estimates of source-level activity were computed from the evoked responses for each participant using dynamical statistical parameter mapping (dSPM; Dale et al., 2000). Each participant's head shape and fiducial landmarks that were collected prior to the experiment were used to morph and coregister the "fsaverage" brain using the FreeSurfer software (<http://surfer.nmr.mgh.harvard.edu/>). For each condition, MEG activity was averaged, and the forward solution was computed using the boundary element model (Bonnet, 1999; Mosher et al., 1999) as the source model. Covariance matrices were estimated using the 100 ms before the presentation of the stimulus. The inverse solution and activity at the source level were estimated by calculating minimum-norm estimates (Hämäläinen and Ilmoniemi, 1994).

Behavioral data analyses

The behavioral data served to assess whether the stimuli elicited a behavioral SSE, so that any difference in neural activation between the grammatical and reversed sentence conditions would reflect the activation difference between successfully and unsuccessfully building an abstract sentential representation of the stimulus. We thus chose to include reaction time and accuracy data from the participants whose MEG data were unusable while still excluding data from the participants who were excessively sleepy. Only one participant scored below 80%, as low as 69.1%. This participant was excluded from further statistical analyses. We first cleaned the data by eliminating trials with reaction times <200 ms or longer than 4,000 ms. We further removed trials from each participant that deviated 3 standard deviations from the participant's mean RT. To analyze the RTs, we fit a linear mixed-effect regression model to the log-transformed reaction times. We did this separately for both the match and the mismatch trials, as each type of trial likely reflects different cognitive demands of the participant. Each regression model had the categorical variable sentence type as the only fixed effect, with subjects and items being the random effects. To analyze the accuracy data, we used a generalized linear mixed-effect logistic regression model with all of the cleaned data using the same combination of fixed and random effects as the previous model. Pairwise comparisons between sentence-type conditions were performed by likelihood ratio tests and were corrected using Tukey's adjustment. The behavioral data were analyzed using the *lme4* (Bates et al., 2015) and *afex* (Singmann et al., 2023) packages in R (v4.3.1).

MEG data analyses

Spatiotemporal clustering tests. Spatiotemporal cluster-based permutation tests (Maris and Oostenveld, 2007) were performed over the entire left and right cortical surfaces to detect effects of our manipulation. The *p* value threshold for a cluster was set to be <0.05, and only clusters that spanned a minimum of 20 ms and spatial size of 10 sources were considered in the analysis. Corrected cluster *p* values were estimated using 10,000 permutations. The analysis was performed over the entire 800 ms period from the onset of Sentence 1 until the onset of Sentence 2, to capture both early and late emerging effects. Clustering analyses were run separately on the left and right hemispheres as opposed to running a full brain analysis to prioritize viewing results from the left hemisphere, as each individual analysis was computationally intensive and required multiple days to run. The cluster-based permutation tests were run using the Python package *eelbrain* (Brodbeck et al., 2023).

The sentence-type factor of our repeated-measure ANOVA for spatiotemporal clustering was used to identify cortical areas whose

activation shows a neural SSE. We define a neural SSE as a pattern of activity in which the source activation for grammatical sentences is significantly increased as compared with reversed sentences, consistent with neurons firing in response to detecting well-formed sentence structure, as has also been observed in studies on composition of serially presented words (Pylkkänen, 2019). Importantly, any interactions between sentence type and determiner were not used to assess whether a source location exhibits a neural SSE, as any pattern more complex than an increase for grammatical as compared with reversed sentences was deemed too complex for a straightforward reflex of a neural SSE. The second factor of determiner type addressed questions other than those under discussion in this paper and thus will not figure into any analyses reported here.

Clusters that were revealed by the spatiotemporal clustering test further underwent a series of planned pairwise temporal clustering tests to further determine their functional properties. We performed temporal clustering tests for each pair of conditions within the sentence-type factor (sentence and inner transposition, sentence and reversed sentence, inner transposition, and reversed sentence). The tests were carried out by using the spatial extent of the significant clusters as fROIs, which were then entered into temporal clustering tests within an interval comprising of the significant temporal cluster in the initial ANOVA plus an added 50 ms at the beginning and end of the interval. To reiterate, finding a significant difference between grammatical sentences and reversed sentences was the criterion for assessing whether the cluster displayed a neural SSE. Furthermore, if a cluster displayed a three-way pattern, in which the activation of the inner-transposed sentences is intermediary between the other two conditions, this was taken to indicate that the cluster shows a neural SSE that is sensitive to the minor error of the transposition. In contrast, if the cluster made only a two-way distinction between the grammatical sentences on the one hand and the inner-transposed and reversed sentences on the other, this was taken to index a neural SSE that does not detect such an error.

Generalized linear model (GLM) analysis. To probe the extent of serial versus parallel processing in any observed effects of sentence type, we carried out a two-stage multiple regression model analysis by first computing a linear model of single-trial source data for each subject using the features of log bigram frequency and log transition probability. If a neural SSE performs a purely serial, left-to-right kind of processing, we expect to see the pattern of results depicted in Figure 5a, in which we observe significant effects of either bigram frequency or transitional probability unfolding in a serial manner left to right: the first bigram, then the second bigram, and finally the third bigram. On the other hand, if a neural SSE performs parallel processing, then we should instead observe the behavior shown in Figure 5b, where the effects of bigram frequency or transitional probability unfold all within the same time window.

Bigram frequencies and transition probabilities were computed using the Corpus of Contemporary American English (Davies, 2008). Specifically, bigram frequencies were calculated by counting the number of times that the first and second word appeared adjacent in the corpus. For each subject, two models were estimated, one using a linear combination of the logged bigram frequencies of each bigram in the trial [$dSPM \sim \text{bigram}(1, 2) + \text{bigram}(2, 3) + \text{bigram}(3, 4)$] and one using a linear combination of the logged bigram transition probabilities of each bigram in the trial [$dSPM \sim \text{trans}(1, 2) + \text{trans}(2, 3) + \text{trans}(3, 4)$]. Each of these models estimated the single-trial source data in source regions that were significant in the 3×4 repeated-measure ANOVA of the spatiotemporal clustering analysis described above. We computed these two models for each condition separately (sentences, inner transpositions, reversed sentences), given the high degree of variability in the values for bigram frequency and transition probability across conditions. A grand total of six models were computed for each participant. The spatiotemporal clustering analysis had identified two clusters.

We thresholded the *F* values of the earlier cluster to only include source points containing *F* values >7.5. We did the same for the later cluster but with a threshold of 5.0. The value of 7.5 was used to threshold the earlier cluster, because 7.5 was the *F* value that most successfully constrained the spatial boundaries of the cluster upon visual inspection and

similarly for the later cluster with the F value threshold of 5.0. The resulting spatial extent of the two clusters had very similar distributions to the original clusters but was much more focalized to certain regions. This was done to constrain the spatial extent of the clusters, as they both covered a wide stretch of the left-lateralized language cortex. The dependent measure for each of the models was the averaged dSPM values across all of the source points in the thresholded clusters.

The first stage of the regression analysis consisted of constructing a linear model of single-trial source activity for a single subject, following the methodology of Gwilliams et al. (2016). From the thresholded clusters, we took the single-trial source activations for a single participant and averaged them across the spatial extent of the cluster, yielding a single time course for each trial. We then fit the linear models described above to the dSPM values at each time point. This stage of regression resulted in a β coefficient for every time point of the epoch. After computing a series of β coefficients for each participant, the second stage of the regression consisted of a one-sample temporal clustering t test on the coefficients across subjects at each time point to see whether the

coefficients were significantly different from zero. The p value threshold for a cluster was set to be <0.05 , and only temporal clusters that spanned a minimum of 20 ms were considered in the analysis. Cluster p values were estimated using 10,000 permutations. We used the significant time windows from the repeated-measure ANOVA analysis to constrain the search for significant effects in the temporal clustering analysis with an additional 50 ms of padding in the beginning and end of the window. For the early thresholded cluster, the time window was 163–519 ms, and for the late thresholded cluster, the window was 398–775 ms. The GLM analyses were carried out using eelbrain (Brodtbeck et al., 2023).

Results

Behavioral results

The behavioral results in Figure 2 show a clear behavioral SSE in RT data both for match and mismatch trials. Accuracy in our simple task was across the board very high and showed no reliable condition differences. Numerically, the grammatical

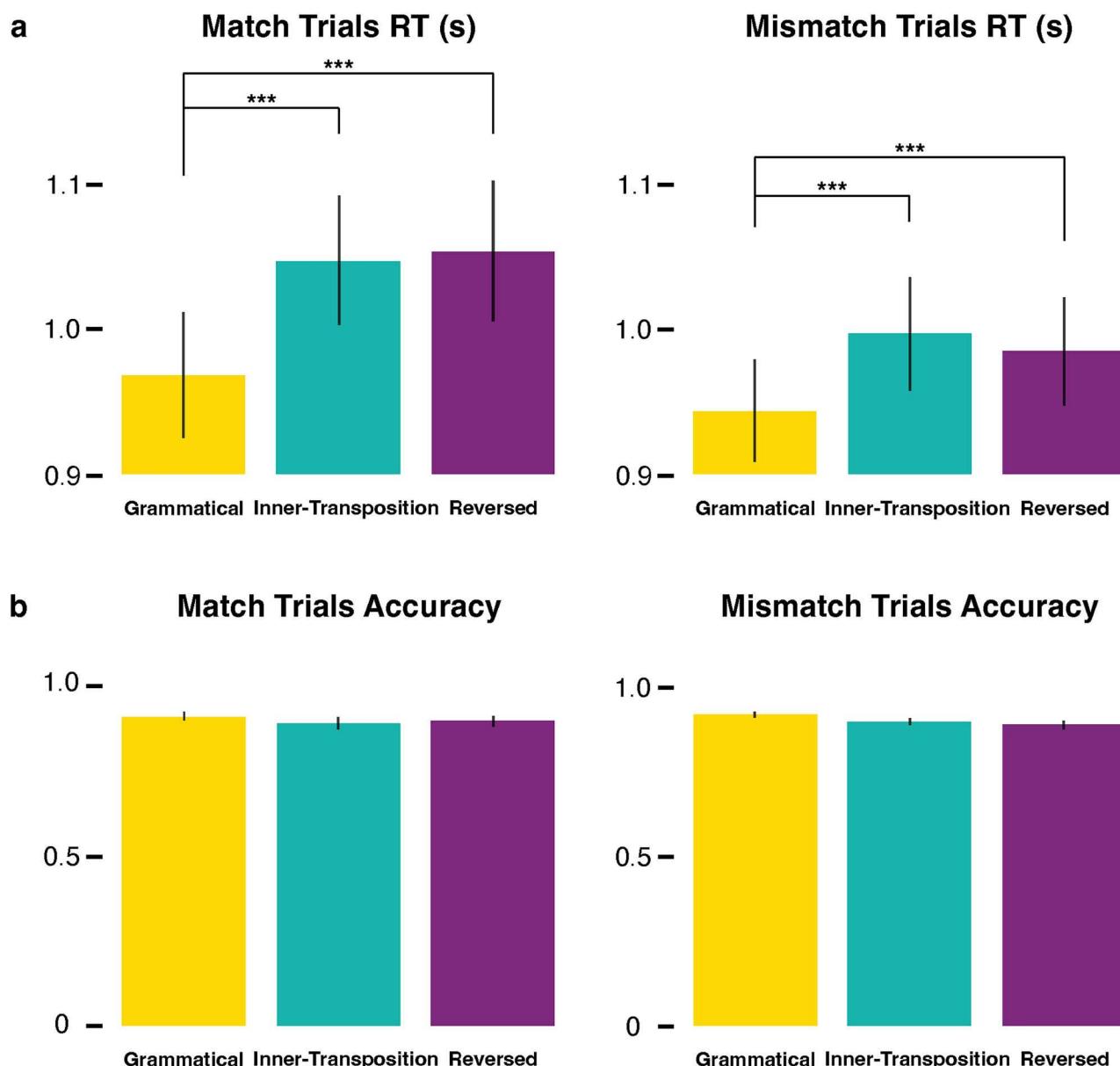


Figure 2. Behavioral results demonstrating SSEs. **a**, In both match and mismatch trials, we observe a behavioral SSE in reaction time for grammatical sentences compared with both inner-transposed and reversed stimuli. **b**, Regarding accuracy, the same behavioral SSE pattern is observed in mismatch trials; however, in match trials, a significant difference is only seen between grammatical and inner-transposed stimuli.

sentences had the highest accuracy (mean \pm SD, $91.27 \pm 28.23\%$) and the lowest RT (951.1 ± 433.7 ms), followed by the inner-transposed stimuli (RT, $1,015.4 \pm 478.2$ ms; accuracy, $89.42 \pm 30.76\%$) and finally the reversed sentences (RT, $1,011.8 \pm 479.1$ ms; accuracy $89.35 \pm 30.85\%$). Match trials elicited longer RTs ($1,016.3 \pm 492.1$ ms) but higher accuracies ($90.08 \pm 29.89\%$) relative to mismatch trials (968.5 ± 434.8 ms; accuracy, $89.95 \pm 30.07\%$).

Within the match trials (Fig. 2a, left), pairwise comparisons reveal that the RTs of the grammatical sentence stimuli were significantly faster than those of the inner-transposed stimuli ($p < 0.0001$) as well as the reversed stimuli ($p < 0.0001$). The mismatch trials (Fig. 2a, right) show the same pattern, where the sentences elicited shorter RTs compared with inner-transposed ($p = 0.0002$) and reversed stimuli ($p = 0.0044$). The significant differences in RT between the sentences and the

reversed sentences for both match and mismatch trials is a clear example of the SSE. The fact that inner transpositions patterned with the reversed sentences as opposed to the grammatical sentences reveals that in this matching task, the behavioral SSE extends to inner transpositions.

MEG results

Spatiotemporal clustering results

Two spatiotemporal cluster-based 3×4 ANOVAs were performed on source-estimated data. One was performed on the entire left cortical surface, and the other was performed on the entire right cortical surface. No significant clusters were found in the right hemisphere. In the left hemisphere, however, significant effects of the sentence type were found in two separate clusters. The first cluster (Fig. 3) broadly spans over the left-lateralized language cortex from 213 to 469 ms ($p < 0.001$)

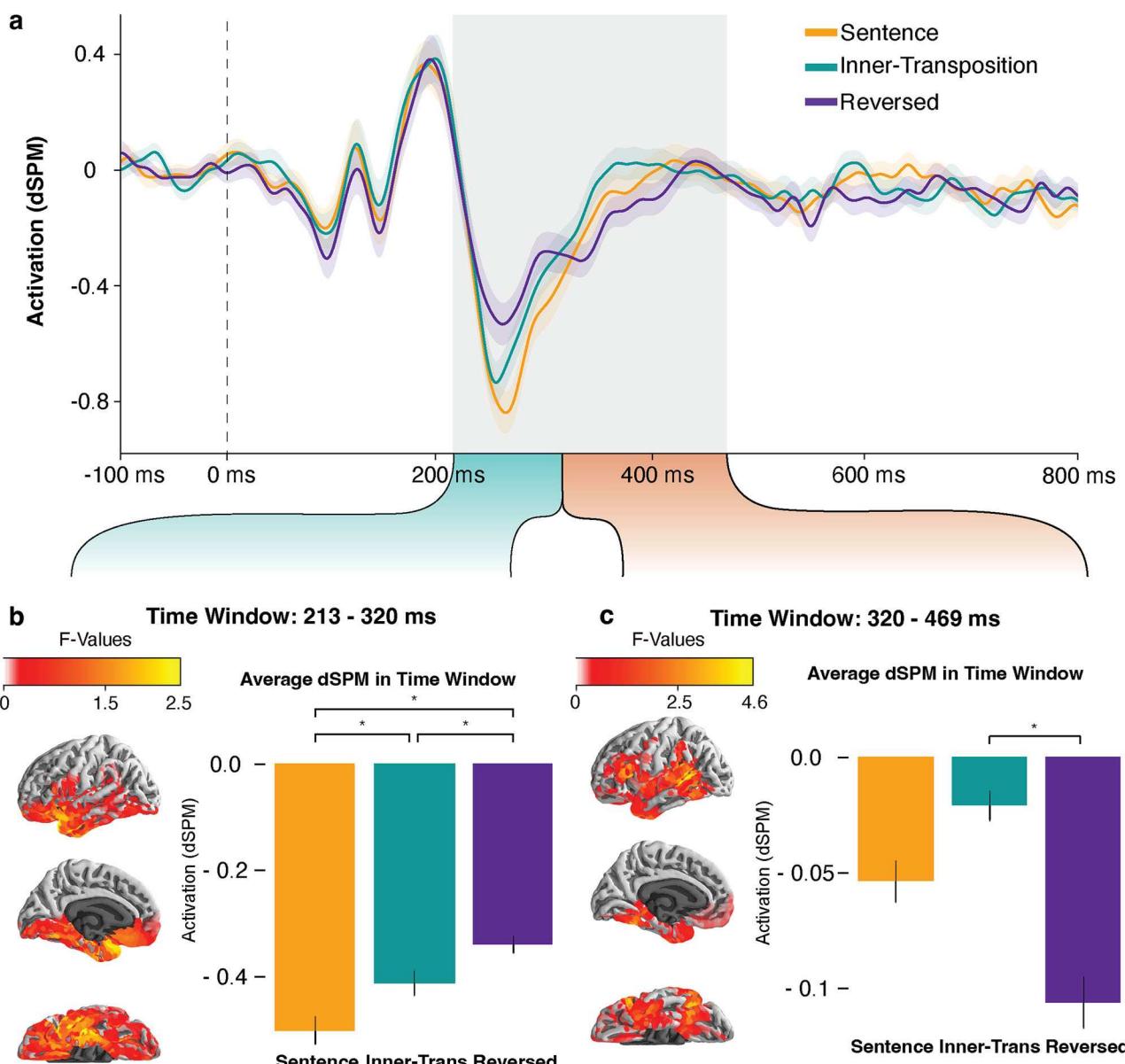


Figure 3. The first cluster displaying a significant effect of sentence-type manipulation, occurring between 213 and 469 ms. **a**, Average waveform across all sources showing significant effects. The initial half of the cluster exhibits a distinct trough-shaped pattern. **b**, On the left are the subset of sources first half of the cluster. On the right are the results of pairwise *t* tests, showing a three-way pattern suggestive of early detection of an inner transposition. **c**, On the left are the subset of sources and average activations in the later cluster, lacking the error detection profile observed in the earlier half.

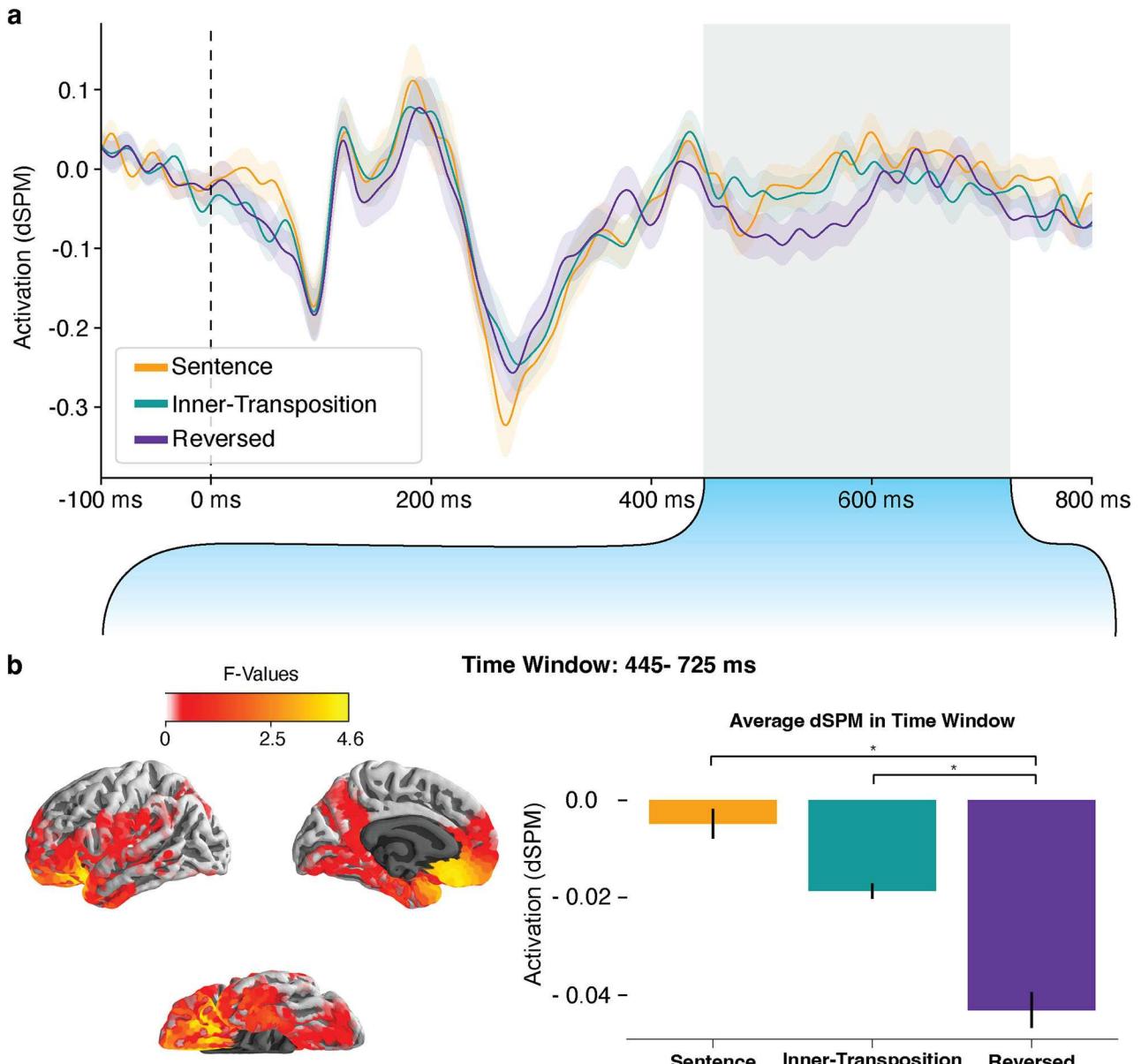


Figure 4. The second cluster displaying a significant effect of sentence-type manipulation, occurring between 445 and 725 ms. *a*, Waveform averaged across all sources reflecting the effect of sentence type. *b*, More focal activation observed, primarily in the left inferior frontal gyrus, ventromedial prefrontal cortex, and left anterior temporal pole. Temporal clustering tests reveal a two-way pattern: sentences and inner-transposed sentences show significantly different activation compared with reversed sentences, indicating top-down grammatical correction by this stage of processing.

after the onset of the stimulus. The second cluster (Fig. 4) spans over much of the same regions as the first cluster, but with a much more focal distribution centered in the ventromedial prefrontal cortex starting from 448 ms to 725 ms ($p < 0.001$) after the onset of the stimulus.

Follow-up pairwise temporal clustering tests revealed that the initial stage of the early cluster showed a three-way distinction between each of the conditions in the sentence-type factor, with sentences eliciting the highest activity, followed by inner transpositions and, finally, reversed sentences. Specifically, there was a significant difference between the sentence and the inner-transpose conditions from 258 to 288 ms ($p = 0.033$; Fig. 3*b*), between the sentence and the reversed conditions from 239 to 313 ms ($p < 0.001$), and, finally, between the inner-transposition and reversed conditions from 241 to 271 ms ($p = 0.0475$) and from 333 to 412 ms ($p = 0.0013$). The significant increase for

grammatical sentences as compared with the reversed sentences meets our definition of a neural SSE. The three-way distinction revealed by the pairwise tests (Fig. 3*b*) shows early detection of the inner transposition during this processing stage. In sum, the early cluster (210–320 ms) shows evidence of a neural SSE that is sensitive to inner transpositions in the stimulus, that is, this activity detects the error as opposed to missing it.

However, as is apparent from the waveform of the cluster revealed by the initial ANOVA test (Fig. 3*a*), the behavior of the cluster was not unitary but rather included at least two different response profiles. The three-way distinction reported above is apparent only during the earlier half of the time window of the cluster as reflected by the initial trough of the waveform (Fig. 3*b*). However, starting from ~ 320 ms after the onset of the stimulus, the activity pattern of the waveform changes and no longer reflects the directionality of the three-way pattern

mentioned above. Instead, the pairwise tests within this later period show only a significant difference between the inner transpositions and the reversed sentences (Fig. 3c) following the 320 ms mark ($p=0.0013$) specifically between 333 and 412 ms. Given that there was no significant difference between the grammatical sentences and the reversed sentences during this time window, the activity observed in this part of the cluster does not correspond to our definition of a neural SSE, and thus we do not speculate on what kind of processing may be taking place here.

In the later cluster (Fig. 4), the follow-up pairwise tests show a different functional pattern involving a clear two-way distinction. The temporal clustering tests find a significant difference between the sentence and reversed conditions (Fig. 4b) from 508 to 533 ms ($p=0.0491$) and 536 to 605 ms ($p=0.0016$) as well as in the inner-transpose and reversed conditions from 563 to 583 ms ($p=0.0386$), but no significant effect between the sentence and inner-transpose conditions. The significant increase of activity for the grammatical sentences relative to the reversed sentences again fits our definition of a neural SSE. However, the lack of significant difference in activity between the grammatical sentences and the inner transpositions indicates that the neural activity at this stage of processing has “missed” the presence of the inner transposition.

GLM results

In order to shed light on whether the ANOVA results reported above reflect serial or parallel processing, we used the observed spatiotemporal clusters as fROI/TOIs in subsequent regression analyses testing whether bigram frequencies and word-to-word transition probabilities across the sentence affect the neural signals simultaneously or in a left-to-right sequential matter, as visualized in Figure 5. The results from the two-stage regression analyses indicate separate functional profiles for the two clusters arising from the main ANOVA analysis. All correlations that we report are signal increases as a function of logged bigram frequencies and transition probabilities. We do not make explicit any hypotheses as to the directionality of the effects, and so we do not speculate about their directionality.

The early cluster (Fig. 6) showed behavior consistent with a parallel processor, showing sensitivity to the bigram frequency of the first bigram (238–274 ms; $p=0.0216$) and the second bigram (234–275 ms; $p=0.0181$) in overlapping intervals, with the third bigram showing a trending effect shortly afterward (297–318 ms; $p=0.0659$). These effects were only found in the grammatical sentences; in the inner-transposed and reversed sentences, there were no significant effects. This cluster also showed trending effects for transition probability in the grammatical sentences, but only for the second (242–267 ms;

Parser Predictions for Bigram Frequency and Transition Probability Effects

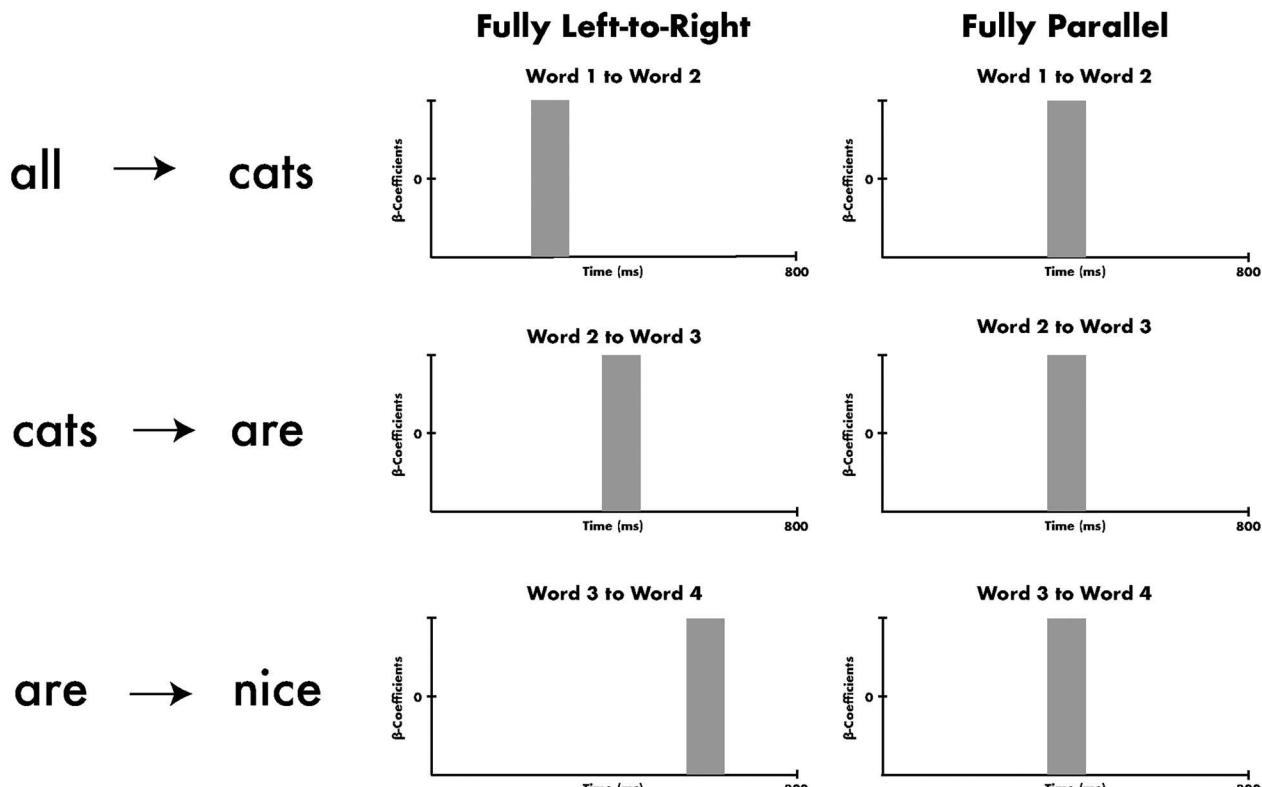


Figure 5. Visualization of serial and parallel hypotheses for the regression analyses of bigram frequency and transition probability reported in Figures 6 and 7. The fully left-to-right parser hypothesis predicts to see incremental effects of a predictor variable for each bigram in a strictly linear fashion, whereas the fully parallel processor hypothesis predicts to see overlapping effects of bigram frequency within a common time window.

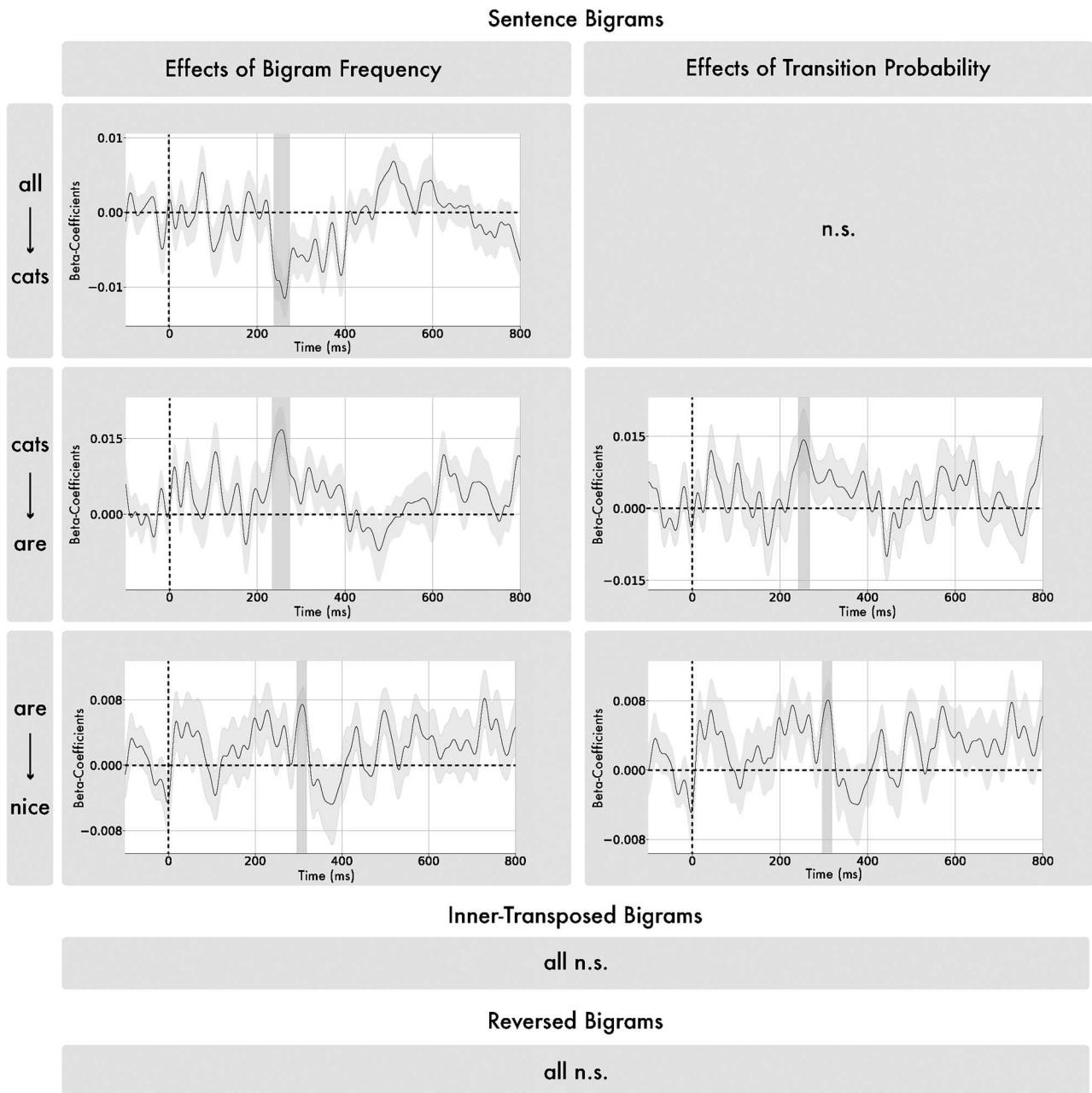


Figure 6. Bigram frequencies and transition probabilities affected activation in the early cluster but only for the grammatical sentences. The temporal dynamics of the significant effects fit the predictions of a parallel processor: the frequencies of the first and second bigrams were simultaneous with the third bigram immediately following. Transition probability effects were observed for the second and third word-to-word transitions and were also temporally very close.

$p = 0.0892$) and the third (297–319 ms; $p = 0.0597$) bigram. As with the effects of bigram frequency, no effects were found for transition probability in the inner-transposed and reversed sentences.

The later cluster (Fig. 7), on the other hand, showed signs of serial left-to-right processing but only for inner-transposed stimuli. The cluster showed effects of bigram frequency for the second bigram (568–623 ms; $p = 0.0118$) and the third bigram (719–772 ms; $p = 0.0041$). Similarly for transition probability, the cluster also showed effects for the second (567–622 ms; $p = 0.005$) and third (724–768 ms; $p = 0.0134$) bigram in roughly the same time windows. The cluster additionally exhibited an effect of transition probability for the first bigram (719–747 ms; $p = 0.0362$).

Discussion

Our perception of the world draws from both sensory-driven analysis of a stimulus and prior, domain-specific knowledge of the stimulus and of the context more broadly. If a stimulus is impoverished visually, as in the case of fully gray-scaled images (Ramachandran, 1994), or auditorily, as in the case of sine-wave speech (Remez et al., 1981), then top-down knowledge is known to guide the perception of the degraded stimulus (Möttönen et al., 2006). Conversely, top-down knowledge can also “correct” our perception, causing us to misperceive or ignore what would otherwise be highly salient properties of the stimulus if attention had been cued to them beforehand (Simons and Chabris, 1999; Mack and Rock, 2000). In the domain of language, a similar mechanism of top-down knowledge overwriting the literal

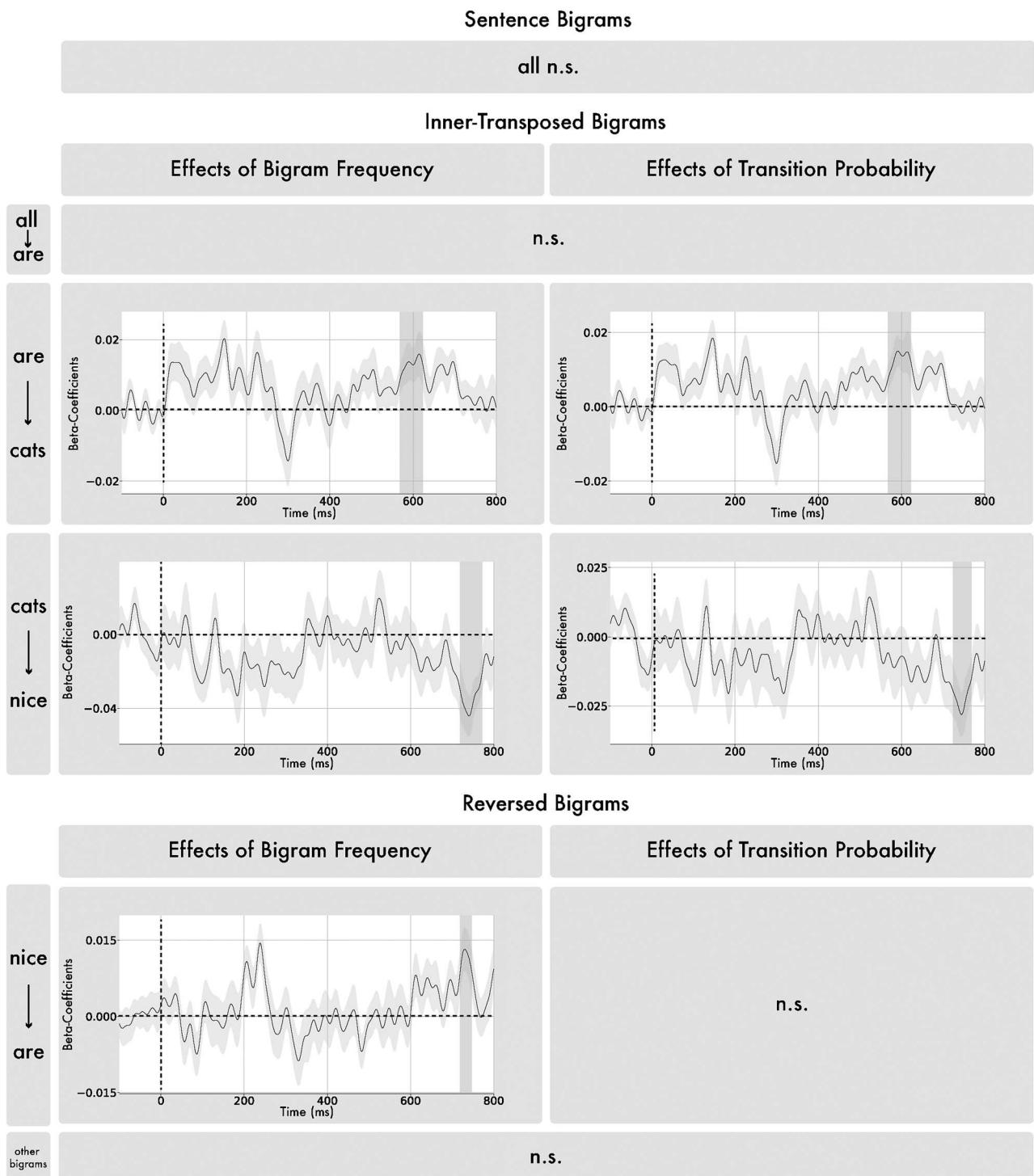


Figure 7. In the later cluster, bigram frequencies and transition probabilities mainly just affected the inner transpositions, where the second and third bigrams showed reliable and sequential effects of frequency of transition probability, with the third bigram following the second one by ~100 ms. Reversed sentences also showed one effect with the frequency of the first bigram having a late effect at ~750 ms.

interpretation of a stimulus is exemplified by not recognizing a transposition of an inner bigram in a sentence (Mirault et al., 2018). The underlying causes of this effect are heavily debated within the literature, with some authors arguing that the effect comes from noise during bottom-up encoding of word position (Mirault et al., 2018; Snell and Grainger, 2019) and others arguing for a postperceptual inference mechanism fixing the transposition (Huang and Staub, 2021, 2023; Hossain and White, 2023).

This study provides a neural time course of bottom-up and top-down mechanisms involved in the composition of multi-word expressions by comparing grammatical sentences to sentences with the inner two words transposed and fully reversed sentences. The spatiotemporal clustering analysis revealed two separate clusters with distinct functional profiles. The earlier of the two clusters was sensitive to both inner transpositions and reversals in the stimuli, suggesting that during early bottom-up

processing, the brain indeed detects phrase structure errors that may be corrected at a later stage of processing. Conversely, the later cluster showed only a two-way distinction, with the sentences and inner-transposed stimuli eliciting more activity than the reversed sentences. This suggests that during this stage of processing, the transposition has been “fixed,” because such stimuli elicited similar patterns of activation as the grammatical stimuli.

The psycholinguistic literature on the TWE posits two possible causes for the effect. In essence, the two hypotheses differ with respect to “when” the misperception of the word order arises: one positing that the bottom-up encoding of the stimulus is noisy (Mirault et al., 2018; Snell and Grainger 2019; Pegado and Grainger 2020) and the other positing that the top-down interpretation of the stimulus fixes it such that it conforms to the reader’s knowledge of phrase structure (Huang and Staub, 2021, 2023; Hossain and White, 2023). The first of these hypotheses contends that word recognition takes place in parallel and that when a reader encounters a sentence such as “the white was cat big,” the noisy encoding of word position makes it possible for the reader to recognize the word “cat” before the word “was,” resulting in a TWE. Under this account (Mirault et al., 2018; Snell and Grainger 2019; Pegado and Grainger 2020), the misordering arises at the level of perception, and thus the sentence would be processed bottom-up as the grammatically correct sentence “the white cat was big.”

The other account of the TWE proposes that bottom-up processing of the stimulus does encode the error but that postperceptual inference of the meaning of the stimulus converts it into its grammatical counterpart (Huang and Staub, 2021, 2023; Hossain and White, 2023). Following Gibson et al. (2013), this approach assumes that readers use syntactic and semantic knowledge to retrieve the most likely intended message given the noisy perceived message. The findings from this study are consistent with this postperceptual account. The behavior of the early cluster suggests that bottom-up analysis of the stimulus is sensitive to the presence of word order errors, whether they are transpositions or full reversals of the stimulus. Furthermore, the late cluster treats the sentences and the inner-transposed stimuli identically, which suggests that at this stage or earlier, the transposition in the stimulus has been fixed such that the stimulus can be interpreted as a grammatical sentence.

Our regression analysis showed that the two clusters had complementary profiles as regards their sensitivity to bigram frequencies and transition probabilities, which we regressed against the neural data to learn about the degree of parallel versus serial processing in the early bottom-up and later top-down clusters. Although there is prior evidence that two words cannot be simultaneously recognized (White et al., 2018, 2019, ***2020), the stimulus words of these studies did not linguistically compose with each other, but the sentence superiority phenomenon shows the critical impact of linguistic composition for our ability of rapid processing. Here our stimuli were clearly combinatorial (grammatical sentences) and clearly noncombinatorial (reversed sentences) and of ambiguous nature (inner transpositions). We took overlapping effects of bigram frequencies and transition probabilities across the stimulus to be suggestive of parallel processing and sequential effects of serial processing. The early bottom-up cluster had a more parallel profile and the late top-down cluster a more serial profile.

In the early cluster, we observed sensitivity to the bigram frequencies of the first and second bigrams in overlapping time windows (238–274 ms for the first bigram and 234–275 ms for the

second bigram) with the third bigram showing a trending effect shortly after (297–318 ms). These early effects were only present for grammatical stimuli and were not found for either the inner-transposed or reversed stimuli, which suggests that this early, bottom-up processing stage prefers inputs that are high in “form typicality” (Matar et al., 2021). The late cluster showed a very different profile, in which no bigram frequency or transition probability effects were elicited at all for the grammatical stimuli. Instead, this later cluster only showed effects for the inner-transposed stimuli as well as the reversed stimuli, and the temporal windows of these effects unfolded in a clearly serial fashion, unlike that of the early cluster. In sum, the regression analysis revealed a largely parallel profile limited to grammatical sentences in the early cluster and a more serial profile limited to the ungrammatical stimuli in the later cluster. These results suggest that when presented with multiple words at once, bottom-up combinatorial processing unfolds in a parallel fashion, not serially; the brain is able to make use of the parallel availability of the visual, linguistic stimulus, which is not the case in the auditory modality. This is consistent with recent EEG work showing nonleft-to-right effects of contextual surprisal during parallel presentation (Dunagan et al., 2024).

While our neural data revealed an effect of top-down linguistic knowledge on the processing of inner transpositions starting at ~350 ms after the stimulus onset, at which point inner transpositions began to pattern with grammatical sentences, this was not the case in our behavioral data, where inner transpositions patterned together with the ungrammatical, reversed sentences. This is evidence of yet another stage of processing, at which the error introduced by the transposition is detected. The combination of these results and prior behavioral studies clearly suggests that the effect of inner transposition is highly task dependent, as the previous studies on TWEs with the matching task elicited the effect by first presenting a grammatical sentence followed by the same sentence containing an inner transposition (Mirault et al., 2018; Pegado and Grainger, 2020). However, our version of the matching task did not involve introducing a transposition on the second presentation. Our version included the possibility of participants first seeing an inner-transposed sentence to match with a second inner-transposed sentence. We made this decision to ensure a clean neural recording of a transposed sentence.

In summary, our results reveal two regions of the left-lateralized language cortex exhibiting robust SSEs. Initially, we observe increased activity for grammatical sentences in the anterior and posterior temporal cortex and inferior frontal regions, followed by more concentrated activity in the medial and inferior frontal cortex. At the earliest stages of multiword comprehension, combinatorial regions of the language cortex appear to operate on the stimulus in parallel. Later on, we see evidence of serial processing of ungrammatical stimuli. The fastest sentence-sensitive signals starting at ~200 ms show a bottom-up processing profile, immediately detecting any deviation from grammaticality, but by 450 ms, top-down processing guided by knowledge of linguistic structure rescues stimuli that only deviate minimally from a grammatical form.

References

- Adachi Y, Shimogawara M, Higuchi M, Haruta Y, Ochiai M (2001) Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Trans Appl Supercond* 11:669–672.
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* 67:1–48.

Bonnet M (1999) Boundary integral equation methods for solids and fluids. *Meccanica* 34:301–302.

Brodbeck C, Das P, Gillis M, Kulasingham JP, Bhattachari S, Gaston P, Resnik P, Simon JZ (2023) Eelbrain, a Python toolkit for time-continuous analysis with temporal response functions. *Elife* 12:e85012.

Cichy RM, Pantazis D, Oliva A (2014) Resolving human object recognition in space and time. *Nat Neurosci* 17:455–462.

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26:55–67.

Davies M (2008) The corpus of contemporary american english (COCA). Available at <https://www.english-corpora.org/coca/>

Dunagan D, Jordan T, Hale JT, Pylkkänen L, Chacón DA (2024) Rapid visual form-based processing of (some) grammatical features in parallel reading: an EEG study in English. *bioRxiv*, 2024-04.

Frazier L, Fodor JD (1978) The sausage machine: a new two-stage parsing model. *Cognition* 6:291–325.

Gibson E (2006) The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity. *J Mem Lang* 54:363–388.

Gibson E, Bergen L, Piantadosi ST (2013) Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proc Natl Acad Sci U S A* 110:8051–8056.

Gramfort A, et al. (2013) MEG and EEG data analysis with MNE-Python. *Front Neurosci* 267:1–13.

Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen MS (2014) MNE software for processing MEG and EEG data. *Neuroimage* 86:446–460.

Gwilliams L, Lewis GA, Marantz A (2016) Functional characterisation of letter-specific responses in time, space and current polarity using magnetoencephalography. *Neuroimage* 132:320–333.

Hale J (2001) A probabilistic Earley parser as a psycholinguistic model. In: *Second meeting of the North American Chapter of the Association for Computational Linguistics*.

Hämäläinen MS, Ilmoniemi RJ (1994) Interpreting magnetic fields of the brain: minimum norm estimates. *Med Biol Eng Comput* 32:35–42.

Hossain J, White AL (2023) The transposed word effect is consistent with serial word recognition and varies with reading speed. *Cognition* 238:105512.

Huang KJ, Staub A (2021) Using eye tracking to investigate failure to notice word transpositions in reading. *Cognition* 216:104846.

Huang KJ, Staub A (2023) The transposed-word effect does not require parallel word processing: failure to notice transpositions with serial presentation of words. *Psychon Bull Rev* 30:393–400.

Lewis R, Vasishth S (2005) An activation-based model of sentence processing as skilled memory retrieval. *Cogn Sci* 29:375–419.

Mack A, Rock I (2000) *Inattentional blindness*. Cambridge, MA: MIT Press.

Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG-and MEG-data. *J Neurosci Methods* 164:177–190.

Massol S, Mirault J, Grainger J (2021) The contribution of semantics to the sentence superiority effect. *Sci Rep* 11:20148.

Matar S, Dirani J, Marantz A, Pylkkänen L (2021) Left posterior temporal cortex is sensitive to syntax within conceptually matched Arabic expressions. *Sci Rep* 11:1–14.

Mirault J, Snell J, Grainger J (2018) You that read wrong again! A transposed-word effect in grammaticality judgments. *Psychol Sci* 29:1922–1929.

Mosher JC, Leahy RM, Lewis PS (1999) EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng* 46:245–259.

Möttönen R, Calvert GA, Jääskeläinen IP, Matthews PM, Thesen T, Tuomainen J, Sams M (2006) Perceiving identical sounds as speech or non-speech modulates activity in the left posterior superior temporal sulcus. *Neuroimage* 30:563–569.

Pegado F, Grainger J (2020) A transposed-word effect in same-different judgments to sequences of words. *J Exp Psychol Learn Mem Cogn* 46:1364.

Peirce J, Gray J, Halchenko Y, Britton D, Rokem A, Strangman G (2010) PsychoPy—A psychology software in Python.

Pylkkänen L (2019) The neural basis of combinatory syntax and semantics. *Science* 366:62–66.

Ramachandran VS (1994) 2-D or not 2-D—that is the question. In: *The artful eye* (Gregory RL, Harris J, eds, pp 249–267. Oxford, England: Oxford University Press.

Reichle ED, Pollatsek A, Fisher DL, Rayner K (1998) Toward a model of eye movement control in reading. *Psychol Rev* 105:125.

Reichle ED, Liversedge SP, Pollatsek A, Rayner K (2009) Encoding multiple words simultaneously in reading is implausible. *Trends Cogn Sci* 13: 115–119.

Remez RE, Rubin PE, Pisoni DB, Carrell TD (1981) Speech perception without traditional speech cues. *Science* 212:947–950.

Simons DJ, Chabris CF (1999) Gorillas in our midst: sustained in attentional blindness for dynamic events. *Perception* 28:1059–1074.

Singmann H, Bolker B, Westfall J, Aust F, Ben-Schachar M (2023) afex: Analysis of factorial experiments (R package version 1.3-0) [Computer software]. Available at <https://CRAN.R-project.org/package=afex>

Snell J, Grainger J (2017) The sentence superiority effect revisited. *Cognition* 168:217–221.

Snell J, Grainger J (2019) Word position coding in reading is noisy. *Psychon Bull Rev* 26:609–615.

Snell J, van Leipsig S, Grainger J, Meeter M (2018) OB1-reader: a model of word recognition and eye movements in text reading. *Psychol Rev* 125:969.

Tarkiainen A, Helenius P, Hansen PC, Cornelissen PL, Salmelin R (1999) Dynamics of letter string perception in the human occipitotemporal cortex. *Brain* 122:2119–2132.

Wen Y, Snell J, Grainger J (2019) Parallel, cascaded, interactive processing of words during sentence reading. *Cognition* 189:221–226.

Wen Y, Mirault J, Grainger J (2021) Fast syntax in the brain: electrophysiological evidence from the rapid parallel visual presentation paradigm (RPVP). *J Exp Psychol Learn Mem Cogn* 47:99.

White AL, Palmer J, Boynton GM (2018) Evidence of serial processing in visual word recognition. *Psychol Sci* 29:1062–1071.

White AL, Palmer J, Boynton GM, Yeatman JD (2019) Parallel spatial channels converge at a bottleneck in anterior word-selective cortex. *Proc Natl Acad Sci U S A* 116:10087–10096.

White AL, Palmer J, Boynton GM (2020) Visual word recognition: evidence for a serial bottleneck in lexical access. *Atten Percept Psychophys* 82: 2000–2017.