Brain & Language 263 (2025) 105546

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

ERP representational similarity analysis reveals the prediction of semantic features in minimal phrasal contexts

Jingxiao Li [a], Mingdong Li [a], Wei Zhou [a,b,*], Qingqing Qu [c,d,*]

[a] Beijing Key Lab of Learning and Cognition School of Psychology Capital Normal University Beijing China
[b] Chinese Institute for Brain Research Beijing China
[c] Key Laboratory of Behavioral Science Institute of Psychology Chinese Academy of Sciences Beijing China
[d] Department of Psychology University of Chinese Academy of Sciences Beijing China

ARTICLE INFO

Keywords:
Semantic prediction
Pre-activation of semantic features
Adjective-noun phrase
EEG
Representational similarity analysis

ABSTRACT

Existing studies demonstrate that comprehenders can predict semantic information during language comprehension. Most evidence comes from a highly constraining context and it is less investigated that whether individuals predict following semantic information in a less constraining context. In the present study, we investigated semantic prediction when reading minimal adjective-noun phrases using electroencephalography (EEG) combined with representational similarity analysis (RSA). Native Chinese Mandarin comprehenders were presented with animate-constraining or inanimate-constraining adjectives, followed by animate-congruent or animate-incongruent nouns. EEG amplitude analysis revealed an N400 for incongruent conditions. Critically, we quantified the similarity between patterns of neural activity, and animate-constraining adjectives revealed greater similarity than inanimate-constraining adjectives before the presentation of the nouns. This pre-noun similarity effect suggests pre-activation of animacy-related semantic information of nouns, and provides evidence for the prediction of semantic features of upcoming words, even in minimal phrase contexts.

1. Introduction

Probabilistic prediction has been proposed as a crucial computational principle for language comprehension (Federmeier, 2007; Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pulvermüller & Grisoni, 2020). An increasing array of studies has illustrated that individuals are capable of predicting linguistic elements at multiple levels of representation, including semantic (Altmann & Kamide, 1999; Federmeier & Kutas, 1999; Lau et al., 2013; Wang et al., 2018, 2020), phonological (DeLong et al., 2005; Li et al., 2022; Vissers et al., 2006), orthographic (Kim & Lai, 2012; Laszlo & Federmeier, 2009), and morphosyntactic (Dikker et al., 2009, 2010; Van Berkum et al., 2005). Predominantly, such investigations, particularly those on downstream word-form representations, focus on contexts that relatively strongly constrain specific words. A less investigated issue is whether predictions can be effectively made in minimal contexts with weak constraints on word prediction. Revealing that predictions can be made under these conditions would underscore the robustness of predictive processing in language comprehension, indicating that individuals can still engage predictive mechanisms even in sparse contexts. Our study seeks to explore this by investigating the prediction of semantic properties in phrasal adjective-noun combinations through the use of electroencephalography (EEG) and representational similarity analysis (RSA). Moreover, while previous research has focused on semantic context effects that develop throughout a sentence, our study specifically examined prediction on the basis of cues that were relatively local or adjacent.

Early evidence for semantic prediction comes from anticipatory eye movements towards objects that are predictable from context, before hearing their names (Altmann & Kamide, 1999; Kamide et al., 2003). For instance, in a visual world task, Kamide et al. (2003) manipulated the agent of an action and reported that when the preceding context was "the man will ride …", listeners predictively looked at "motorbike" more than "carousel" before the presentation of the following nouns. The findings suggest that comprehenders use world knowledge about plausible actions by agents to predict semantic information about plausible actions. Similarly, predictable words are read faster (Ehrlich & Rayner, 1981) and processed more easily, as indicated by reduced neural signals, most commonly observed as a reduction in the N400 ERP component. In a seminal ERP (event-related potential) study by Federmeier and Kutas

* Corresponding authors.
E-mail addresses: zhouwei@cnu.edu.cn (W. Zhou), quqq@psych.ac.cn (Q. Qu).

https://doi.org/10.1016/j.bandl.2025.105546
Received 4 July 2024; Received in revised form 8 December 2024; Accepted 22 January 2025
Available online 5 February 2025
0093-934X/

(1999), participants were asked to read sentence pairs such as, "They wanted to make the hotel look more like a tropical resort. So along the driveway, they planted rows of …" The subsequent word could be a predictable word (palms), an unpredictable but semantically related word (pines), or an unpredictable word from a different semantic category (tulips). The study revealed that semantically related but unpredictable words (pines) triggered smaller N400 brain responses compared to completely unpredictable words from an unrelated category (tulips). This variation in the N400 response, which depended on the cloze probability rather than sentence plausibility, indicates pre-activation of subsequent words, not merely the integration of the word into the context. More recently, studies have identified prestimulus predictive brain activity (Grisoni et al., 2016, 2017, 2021; León-Cabrera et al., 2017; see Pulvermüller & Grisoni, 2020, for a review). These findings show that contexts with strong constraints can elicit brain activity before the expected words are presented, and that the magnitude of this activity correlates with word predictability (see Grisoni et al. 2021).

Further evidence supporting semantic prediction comes from representational similarity analysis (RSA, Kriegeskorte et al., 2008), which involves analysing patterns of neural activity for similarities. RSA operates on the assumption that similar items can elicit similar neural patterns. In a study utilizing RSA, Hubbard and Federmeier (2021) quantified the neural activity similarity between predictable target words and the preceding words within the same sentence. They hypothesized that if features of the final word are pre-activated, its neural representation is likely to emerge during the processing of the pre-final word, leading to a higher similarity between these pre-final and final words when the final word is highly expected. Their findings confirmed this hypothesis. In a more recent research, Wang et al. (2020) applied RSA to identify neural patterns associated with predicting the animacy features of forthcoming nouns, based on the constraints imposed by preceding verbs. The study posited that nouns with animate features would exhibit greater neural similarity than those with inanimate features because animate objects typically share more semantic characteristics. The key issue was whether this similarity effect would emerge before the nouns were actually encountered. The results showed a higher neural similarity for contexts that constrained animate nouns compared to those constraining inanimate nouns, prior to the appearance of the nouns. This pre-noun similarity effect was considered to reflect the pre-activation of upcoming nouns, although caution is needed to exclude the possibility that this effect merely reflects the similarity of the prior sentence context.

Much of the research reviewed above has focused on predictive processing in (highly constraining) sentence contexts, and the typical manipulation is to vary the predictability of the final word (or the very final) of a sentence. In these context, comprehenders gradually develop a prediction for the last word as more and more prior words accumulate. It remains uncertain about how individuals predict in minimal contexts where prediction is only generated based on immediate and sparse cues. Fruchter et al. (2015) used magnetoencephalography to study the prediction in adjective–noun phrases, and focused the preactivation of specific nouns triggered by preceding adjectives. They observed an increase in activity within the left middle temporal gyrus when participants were presented with highly predictive adjectives—those that strongly constrain the subsequent noun. This research underscores the capability of certain adjectives to generate the prediction of specific nouns. A recent study by Huang et al. (2023) explored such coarse-grained semantic prediction using classifier-noun phrases. In Mandarin Chinese, classifiers constrain animacy of the subsequent nouns. Huang et al. found that neural activity patterns following classifiers that constrained animate nouns were more similar than those classifiers constraining inanimate nouns, suggesting the preactivation of coarse-grained animacy features of the nouns by classifiers.

Rather than focusing on the preactivation of particular lexical items, in the present study, we employed adjective-noun phrases to explore the prediction of coarse-grained semantic features, and adopted the animacy constraints of adjectives to following nouns. Compared with Chinese classifiers as investigated by Huang et al. (2023), the adjective-noun phrases we used in the present study impose even fewer specific constraints than classifier-noun structures. For example, a noun following "one human classifier" ("一位") usually refers to a person (e.g., "teacher" or "doctor"), and one following "one animal classifier" ("一头") indicates a large animal (e.g., "cow" or "elephant"). In contrast, the adjective "strong" may modify nouns referring to both people (like "athlete") and animals (like "lion"), thus offering broader and less specific constraints. The structure of adjective-noun phrases, which has been widely used to study semantic integration processes (e.g., Barber & Carreiras, 2005; Hagoort, 2003; Kochari et al., 2021), provides a unique contribution for understanding how comprehenders process linguistic information. Typically, semantic inconsistencies in adjective-noun phrases elicit N400 effects, indicating neural responses to semantic incongruency (Barber & Carreiras, 2005; Hagoort, 2003), whereas grammatical gender mismatches trigger P600 effects (Loerts et al., 2013).

The present study aims to test whether semantic animacy features can be anticipated in contexts with minimal constraints. We recorded ERP data as participants were presented with adjective-noun phrases, focusing on whether the limited constraints provided by adjectives can generate predictions about the broader semantic features associated with the animacy of forthcoming nouns. Although the mapping between adjectives and nouns can be very arbitrary so that adjectives can modify a large number of different nouns, some adjectives are unambiguous in specifying animacy features of nouns, therefore constraining the animacy of following nouns. Some adjectives modify animate nouns only (e.g., brave, friendly) whereas some adjectives modify inanimate nouns (e.g., heavy, durable). In the study, we included animate-constraining and inanimate-constraining adjectives and both types of adjectives were matched well in terms of semantic and lexical similarity. We manipulated the similarity between subsequent nouns so that animate nouns are more semantically similar than inanimate nouns. We examined the similarity of brain activity patterns between items. If comprehenders predict the animacy of upcoming nouns, we should be able to observe the neural similarity effect, critically before the onset of following nouns.

2. Method

2.1. Participants

Thirty-seven native speakers of Mandarin Chinese (25 females, M_{age} = 21.4) from Beijing participated in the ERP experiment. The sample size was determined by recent EEG/MEG RSA studies of language prediction (Huang et al., 2023; Hubbard & Federmeier, 2021; Wang et al., 2018). All participants were right-handed, with normal or corrected-to-normal vision, and had no history of language disorders. Participants provided informed consent and were compensated 100 RMB for their participation. EEG data from five participants were excluded from the data analysis due to the high percentage of rejected trials (more than 30 %) after data preprocessing; thus, thirty-two participants were included in the ERP and RSA analyses. The study was approved by the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences.

2.2. Materials and design

Adjective-noun phrases were used as stimuli in the task. Adjectives varied in whether they constrained for an animate noun (animate-constraining, e.g., "brave") or an inanimate noun (inanimate-constraining, e.g., "durable"). To select stimuli, we first began with a large set of adjectives that did not constrain strongly for a specific upcoming noun. To establish the animacy constraints of adjectives, we retrieved five nouns that had the highest co-occurrence following each adjective from the corpus (Xun et al., 2016). An adjective was classified as animate-constraining if more than 50 % of its top five co-occurring nouns were animate; if not, it was classified as inanimate-constraining. Using this approach, a total of 140 adjectives were selected, half of which were animate-constraining adjectives, and the other half were inanimate-constraining adjectives. Each of the 140 adjectives was combined with a noun either a confirmed or violated the adjective's animacy constraint, rendering the congruent or incongruent conditions, to form 280 adjective-noun phrases in total. The crossing of adjective animacy constraint (animate- vs. inanimate- constraining adjective) and adjective-noun animacy congruency (congruent vs. incongruent) generated the four conditions. The objective was to explore whether participants could use the animacy information constrained by the adjectives to predict the nouns that followed, and assess the influence of semantic congruency on language comprehension.

To establish the constraint of the 140 adjectives, we have assessed the cloze probability of the adjectives using the cloze test. A group of 21 Chinese native speakers who did not take part in the EEG experiment were presented with adjectives and were asked to continue with the first word that came to mind (Taylor, 1953). Constraint was measured as the proportion of participants who gave the most frequent word. Overall, the cloze probability of the adjectives was low (27.2 %), indicating that adjectives used in our study do not produce specific word predictions. The cloze probability was slightly higher for inanimate-constraining classifiers (30.3 %) compared to animate-constraining classifiers (24.1 %), and this difference did not reach significance ($p = 0.057$). Besides, these adjectives can constrain animacy of subsequent nouns as expected (probability of animate nouns following animate-constraining

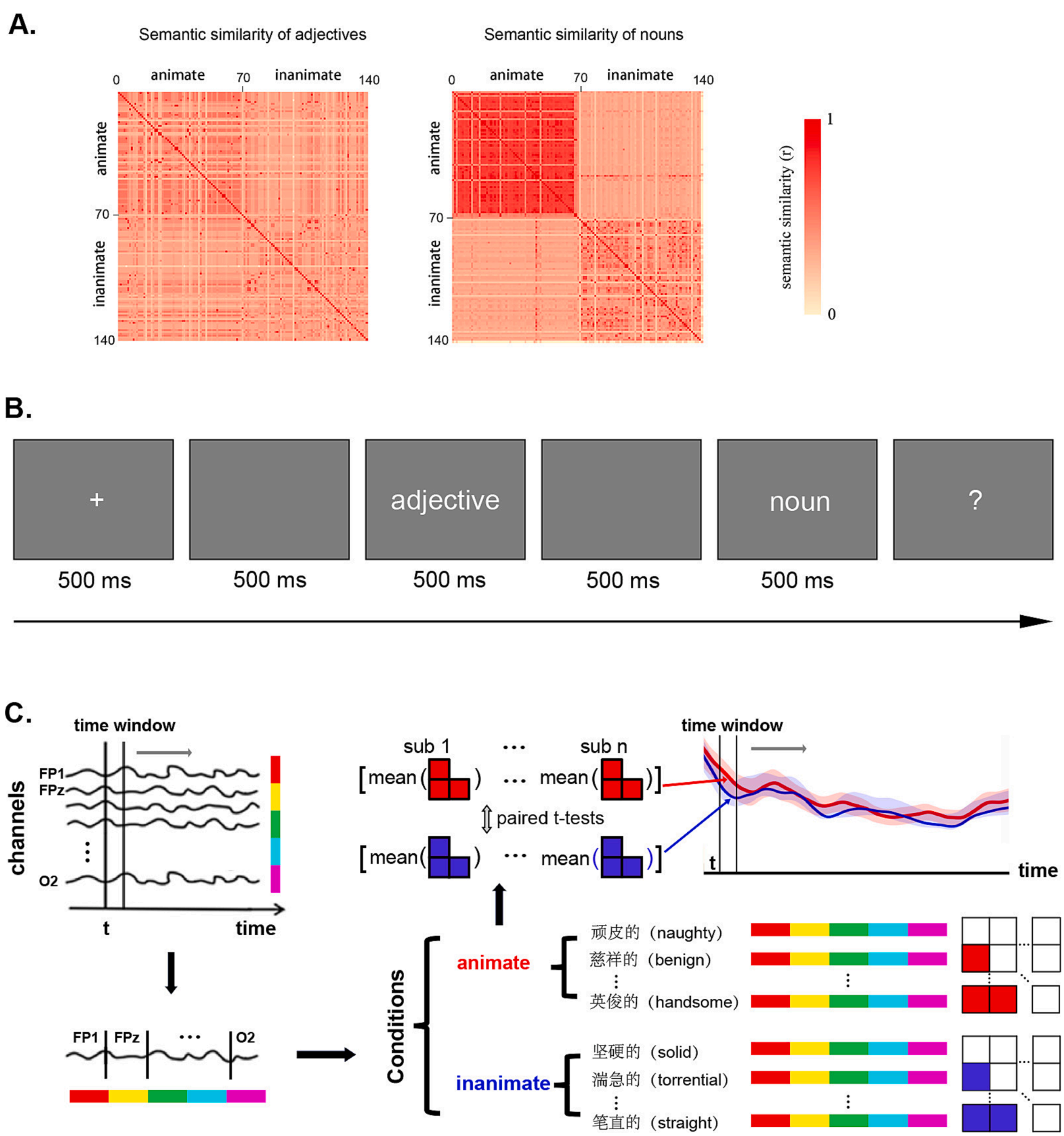

Fig. 1. A. Semantic similarity R values for adjectives and nouns. B. Demonstration of the procedure in a trial. C. Schematic illustration of RSA.

adjectives: 95.0 %, probability of inanimate nouns following inanimate-constraining adjectives: 97.1 %), $p = 0.334$.

In addition, we conducted a semantic plausibility rating for all phrases used in the study. A total of 16 participants rated the plausibility of each phrase on a 7-point scale (1 = semantically implausible, 7 = semantically plausible). As expected, congruent phrases were rated highly plausible ($M = 6.39$), while incongruent phrases were rated as highly implausible ($M = 1.86$). The semantic plausibility ratings for animate-congruent and inanimate-congruent phrases were closely matched (animate-congruent vs. inanimate-congruent: 6.37 vs. 6.40; $t = -0.48$, $p = 0.640$). However, animate-incongruent phrases were rated as less plausible than inanimate-incongruent phrases (animate-incongruent vs. inanimate-incongruent: 1.73 vs. 1.99; $t = -2.45$, $p = 0.027$).

2.3. Matching semantic similarity and lexical properties between animate- vs. inanimate- constraining adjectives

To ensure that the differences observed in the similarity of neural activity patterns between animate and inanimate conditions in adjectives come from the prediction of subsequent nouns, we verified that the two types of adjectives matched in terms of semantic and other lexical properties. To quantify the semantic similarity between word pairs among animate vs. inanimate- constraining adjectives, we used HowNet, an online database that calculates the interconceptual and interattribute relationships of Chinese lexicons (Dong et al., 2010). Semantic similarity values for all possible adjective phrases were measured via a path-based approach by Wu and Palmer (1994). These pairwise Wu–Palmer semantic similarity values in a 140 × 140 matrix are presented in Fig. 1A. The mean semantic similarity values for the animate- vs. inanimate-constraining adjectives were matched ($t = 1.03$, $p = 0.303$). We also performed statistical analysis on character complexity (measured by stroke numbers of characters) and word frequency measured by the Modern Chinese Word Frequency Dictionary (Beijing Language Institute, 1986) associated with the animate- vs. inanimate- constraining adjectives, and character complexity ($t = 0.143$, $p = 0.886$) and word frequency ($t = -0.788$, $p = 0.432$) were matched between the two types of adjectives.

2.4. Quantifying semantic similarity and lexical properties between animate vs. inanimate nouns

The experimental hypothesis rested on the assumption that animate nouns would be more semantically similar to each other than inanimate nouns. We used the same approach as described above to examine the difference in semantic similarity between the two groups of nouns. As shown in Fig. 1A, the mean semantic similarity within the animate group was greater than that within the inanimate group ($t = 3.986$, $p < 0.001$). Moreover, it is important to confirm that any differences in neural similarity produced by predicted animate and inanimate nouns were not generated by differences in similarity of lexical properties. Statistical analysis showed that character complexity and word frequency were matched between animate vs. inanimate nouns (stroke count: $t = 0.415$, $p = 0.428$; word frequency: $t = -0.796$, $p = 0.135$).

2.5. Procedure

The experiment was conducted using E-Prime software. The experimental procedure is illustrated in Fig. 1B. Participants were instructed that they would be presented with an adjective and a noun on the computer screen. Participants were instructed to judge the semantic plausibility of adjective-noun combinations by pressing keyboard buttons. On each trial, participants sequentially viewed a fixation cross (500 ms), a blank screen (500 ms), a display of an adjective (500 ms), a blank screen (500 ms), a display of a noun (500 ms), and a question mark "?" which would disappear once participants responded. Eight practice trials before 280 experimental trials were presented in seven experimental blocks and separated by a short break. The experimental task lasted 40 min.

2.6. EEG recordings and preprocessing

EEG signals were collected from 64 electrodes secured in an elastic cap and recorded with Neuroscan software. A vertical electrooculogram (VEOG) was taken via two electrodes above and below the left eye. A horizontal electrooculogram (HEOG) was recorded via two electrodes on the left and right external cantus. The left mastoid electrode was used as a reference. The EEG data were rereferenced to the average of both mastoids. All electrode impedances were kept below 8 kΩ. EEG signals were recorded with a bandpass filter between 0.05 and 70 Hz and sampled at 1,000 Hz. The EEG data were segmented into 1,700 ms epochs, which included a 200 ms prestimulus baseline, a 500 ms adjective display, a 500 ms interval and a 500 ms noun display. The MATLAB EEGLAB kit was used to preprocess the EEG data. The raw data were filtered with a high-pass cutoff point of 0.1 Hz and a low-pass cutoff point of 30 Hz. Independent component analysis (ICA) using the infomax ICA algorithm (Bell & Sejnowski, 1994) was used to remove artefacts from the segmented data. The ICLabel plug-in (Pion-Tonachini et al., 2019) of EEGLAB was utilized to assess the likelihood that an independent component was noisy. Components with a probability greater than 90 % were not identified as EEG data and were removed by the ICLabel plug-in, resulting in an average removal of one component per participant. Epochs with amplitudes greater than ± 90 μV (approximately 9.6 % of all epochs) and trials with incorrect responses (3.1 % of all trials: animate-congruent: 2.7 %, animate-incongruent: 1.7 %, inanimate-congruent: 4.5 %, inanimate-incongruent: 3.6 %) were excluded from the analyses.

2.7. EEG amplitude analyses

Mean amplitude analyses for EEG signals after noun presentation were conducted to investigate the semantically incongruent effect evoked by the incongruence between adjectives and nouns. We expected an N400 effect evoked by the incongruency of the adjective-noun phrase, which has been frequently reported in previous studies (Federmeier & Kutas, 1999; Kutas & Federmeier, 2000; Zhang et al., 2012; Zhou et al., 2010). Based on previous comparable studies, nine electrodes from frontal-central areas (Fz, F1, F2, FCz, FC1, FC2, Cz, C1, and C2) were selected as spatial regions of interest (ROIs). and the time window of 300–500 ms was selected based on visual inspection for N400 effects. The time window of 300–500 ms, which corresponds to the typical latency range of the N400, was selected based on visual inspection to capture the N400 effects. Mean amplitudes on this time window were entered into a 2 × 2 repeated measures ANOVA with the factors the animacy of adjectives (animate- vs. inanimate-constraining), congruency (congruent vs. incongruent).

2.8. Representational similarity analysis

For RSA analysis, following previous studies (He et al., 2022; Wang et al., 2023), we performed temporal-spatial RSA. The analysis procedure is illustrated in Fig. 1C. We defined a 30 ms time window (15 sampling points) with a 2 ms step size (1 sampling point) and applied a sliding window approach to segment EEG and extract EEG data for each trial, participant, and condition. For each participant, the voltage waveforms within each time window from each channel for each trial were concatenated, producing a single vector of voltage values capturing both spatial and temporal variations in neural activity. The similarity between pairs of trials was quantified using Pearson correlation (r) between the voltage vectors of two trials, resulting in a trial-by-trial correlation matrix for each time window. To analyze these matrices, we excluded the diagonal (representing self-similarity) and the upper triangle (as the matrix is symmetric), averaging the values in the lower triangle to compute the mean similarity for each time window, participant, and condition. To visualize differences between conditions, we calculated the grand-average similarity values by averaging across participants for each time window and condition. To statistically test the differences in similarity between animate- and inanimate-constraining adjectives, we conducted a cluster-based permutation test (Maris and Oostenveld, 2007). During this test, condition labels were randomly shuffled within each time window for all participants. Paired-sample *t*-tests were performed for each time window, and adjacent time windows with significant *t*-values ($p < 0.05$) were considered as clusters. Every permutation test may generate multiple clusters, and we summed the maximum *t*-values of multiple clusters as cluster-level statistic. This procedure was repeated 10,000 times to generate a null distribution of cluster statistics. A significant effect was identified if the observed cluster *t*-value from the actual data exceeded the 95th percentile of this null distribution.

To assess whether the semantic similarity of adjectives is explanatory for the observed pre-noun RSA effect, we conducted an additional analysis. Specifically, we computed a Spearman's rho correlation between the semantic similarity matrix of adjectives and the EEG representational similarity matrix. The analysis procedure is illustrated in Fig. 2. If the semantic similarity of adjectives contribute to any pre-noun RSA effect, we should observe the correlation between adjective similarity and pre-noun EEG similarity. First, we quantified the pairwise semantic similarity of adjectives, creating a 140 × 140 matrix representing all adjective pairs. This semantic similarity matrix was the same across participants and time windows. Then, we constructed single-participant neural similarity matrix: For each participant and time window (a given time window has 30 ms), we concatenated the waveforms across all channels for each trial, producing a single vector of voltage values that captured both spatial and temporal variations. We then computed Pearson correlations between these vectors for all pairs of trials, generating a neural similarity matrix for each participant and for a given time window. Note since the remaining trials of each participant after data preprocessing were different, the ERP similarity matrix was subject-individualized and established based on the remaining trials of each participant. Subsequently, we compared the semantic similarity matrix to the single-participant ERP similarity matrix at a given time window. we conducted independent-sample *t*-tests on the correlations across all time windows, to determine whether the mean Spearman rho was significantly greater than zero (false discovery rate correction to *p*-values was applied).

3. Results[1]

3.1. N400 effect of adjective-noun congruency

The ERPs from nine representative electrodes were averaged for each of the four conditions. Fig. 3A shows the grand average ERPs for the four conditions, and there was a prominent negative wave within the 300–500 ms time frame, identified as the N400 component elicited by noun stimuli based on its polarity, latency, and scalp distribution. ANOVAs were conducted on the mean amplitude of nouns within the 300–500 ms time window, with the factors animacy of adjectives and congruency. The main effect of congruency was significant ($F(1,31) = 16.55$, $p < 0.001$), reflecting that the incongruent condition (−3.26 μV) elicited a larger negativity than the congruent condition (−2.76 μV). The main effect of animacy was not significant ($F(1,31) = 1.74$, $p = 0.196$). The two-factor interaction was significant ($F(1,31) = 4.65$, $p = 0.039$). Following the interaction, separate analyses were conducted for each animate condition. For the animate-constraining condition, the incongruent phrases elicited larger negativity than the congruent ones (animate-incongruent vs. animate-congruent: −3.51 μV vs. −2.52 μV, $t = -4.33$, $p < 0.001$), whereas for the inanimate-constraining condition, there was no significant difference between the congruent and incongruent phrases, although the incongruent phrases showed the trend of larger negativity (−3.01 μV vs. −2.69 μV, $t = -1.50$, $p = 0.144$).

3.2. Representation similarity analysis

Grand average similarity waveforms over time are displayed in Fig. 3B for animate- vs. inanimate- constraining adjectives. Statistical analyses revealed that during the presentation of adjectives (−1000 ms– −500 ms), there was no greater neural similarity for animate-constraining adjectives than for inanimate-constraining adjectives. The neural similarity within the animate-constraining condition was greater than that within the inanimate-constraining condition, from −480– −414 ms, −302– −246 ms, and −136– −92 ms before the onset of the noun. Statistical analyses confirmed that the greater neural similarity for the animate-constraining condition was observed before noun onset, relative to neural similarity within the inanimate-constraining condition in the time windows above (cluster-based permutation test: $ps < 0.001$), reflecting the preactivation of animacy features associated with the upcoming nouns. Moreover, we conducted an additional analysis to assess whether the semantic similarity of adjectives is explanatory for the observed pre-noun RSA effect. The results show that there was no correlation during the entire −500 to 0 ms time window ($ps > 0.05$). This finding confirms that adjective similarity does not explain the observed RSA effect.

4. Discussion

The present study investigated how individuals predict in minimal contexts where the prediction is generated based only on the immediate preceding word. In adjective-noun phrases, adjectives were varied in animacy constraints (animate-constraining or inanimate-constraining) and were followed by animacy-congruent, or animacy-incongruent nouns. A larger N400 was observed for the incongruent condition compared to the congruent one. The N400 effect was significant for animate-constraining adjectives but did not reach conventional significance for inanimate-constraining adjectives, although a clear trend in the expected direction was observed. Critically, the RSA results revealed that the similarity among patterns of neural activity following the animate-constraining adjectives initiated to be greater than that following the inanimate-constraining adjectives before the presentation of nouns, which suggests that comprehenders can use the constraints of adjectives to set expectations about the animacy feature of nouns that follows the adjectives.

As discussed in the introduction, previous research has primarily focused on the preactivation of specific lexical items. For instance, Wang et al. (2018) demonstrated that participants are capable of making specific lexical predictions when comprehending sentences. For instance, in paired sentences like "In the crib there is a sleeping…" and "In the hospital there is a newborn…", the predictable target word

[1] In the present study, we collected response latencies, but they are not a primary focus of the analysis. In the study, participants were instructed to delay their responses until a response signal (a question mark) appeared immediately after the offset of the nouns. Upon seeing the response signal, participants pressed a button to judge the semantic plausibility of the phrase as quickly as possible. Response latencies were measured from the onset of the response signal rather than the onset of the nouns. This delayed response procedure was implemented to minimize EEG motion artifacts caused by button presses. However, by the time the response signal was presented, the cognitive processes of interest—such as lexical access, semantic integration, and prediction—were likely completed. Consequently, the response latencies primarily reflect more peripheral factors, such as decision-making and response execution, rather than the core linguistic processes under investigation. Given these considerations, we are cautious in interpreting the response latency data (Inanimate congruent: 375 ms; Inanimate incongruent: 332 ms; Animate congruent: 341 ms; Animate incongruent: 338 ms) and we emphasize the primary focus on the neural results in the present study.

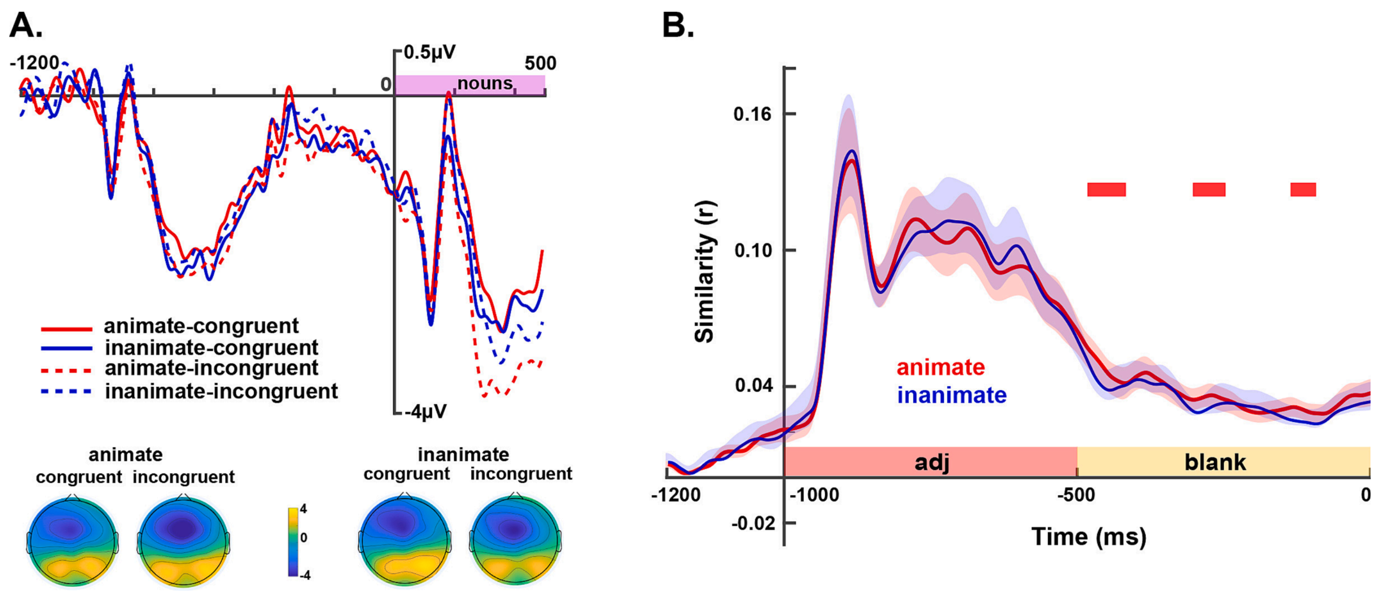

Fig. 2. Representational similarity between the ERP matrix and the semantic similarity matrix. For each participant, we used Pearson correlation to separately compute their adjective semantic similarity matrix and EEG representational similarity. Subsequently, we performed Spearman's rho correlation between the adjective semantic similarity matrix and the EEG similarity matrix for each time window (e.g., the intervals −500 ms to −470 ms and −498 ms to −468 ms represent the time spans used in the analysis), generating a time series of rho values for the −500 to 0 ms interval. Because we used a 30 ms sliding time window, there was insufficient data after the −30 to 0 ms window. To resolve this, we padded the data with 28 ms (14 time points) of zeros after the segment's end, enabling the 30 ms window to slide continuously within the −500 to 0 ms time range for analysis. A one-sample *t*-test was conducted on the rho values across participants for each time window. Finally, FDR correction was applied for multiple comparisons, and the results showed no significant correlated time windows after correction (*ps* > 0.05).

"baby" elicited greater neural pattern similarity compared to non-paired sentences such as "On Valentine's day, he sent his girlfriend a bouquet of red…". This increased RSA similarity in neural activity was observed when participants could predict words in sentences that were closely related. The study highlighted the use of highly constraining sentences that facilitated predictions of specific target words. However, in everyday language use, most contexts do not tightly constrain a specific lexical item but are more likely to specify broader semantic features. It is therefore crucial, from a theoretical perspective, to address whether comprehenders are capable of predicting broader semantic information about upcoming words that extends beyond individual lexical items.

Moreover, most previous studies have focused on the prediction in large-scale contexts. For example, in Wang et al. (2020), stimuli consisted of three-sentence scenarios, where the first two sentences established a discourse context, and the final sentence introduced a pronominal subject referring back to a previously mentioned protagonist, followed by a critical verb that constrains either an animate or inanimate noun. The level of discourse constraint varied, with some sentences having strong lexical constraints and others being more vague (low constraint). In this context, animacy features and lexical constraints of the nouns could be shaped by the full discourse context, which is a broader context than the immediate preceding adjective in our study. In contrast, our study focused on minimal adjective-noun contexts, with the goal of investigating whether comprehenders make immediate and time-sensitive predictions based solely on the preceding adjective. We tested the pre-activation of animacy-related semantic features that are triggered by adjectives in this minimal context. Our study demonstrates that neural activity patterns following animate-constraining adjectives are more similar than those following inanimate-constraining adjectives, and critically this neural similarity effect was detected before the presentation of the pre-noun. This finding supports the notion of preactivation of coarse-grained semantic features, distinguishing between upcoming animate and inanimate entities. This finding extended previous effects of prediction which develops with the incremental processing of the preceding context, rather than being solely reliant on immediate verbs. By contrast, our study minimized contextual influence by focusing solely on adjective-noun combinations. Our findings indicate that prediction can indeed occur based on the immediate lexical context alone.

Could be the semantic similarity of adjectives explanatory for the observed pre-noun RSA effect here? We think it is very unlikely. If semantic similarity differences between animate- and inanimate-constraining adjectives were driving the effect, we would expect differences in neural similarity during the presentation of adjectives (−1000 ms to −500 ms). However, as reported above, our statistical analyses revealed that there was no greater neural similarity for animate-constraining adjectives than for inanimate-constraining adjectives during this time window. The neural similarity effect was not observed during the presentation of the adjectives themselves but emerged only after the offset of adjectives, ruling out the possibility that adjective similarity explains the effects. Additionally, we conducted an additional analysis to directly assess whether the semantic similarity of adjectives contribute to the observed pre-noun RSA effect. Specifically, we computed a Spearman's rho correlation between the semantic similarity matrix of adjectives and the EEG representational similarity matrix. Across the entire predictive time window (−500 ms to 0 ms), no significant correlations were observed. This finding confirms that differences in adjective semantic similarity do not explain the observed RSA effect. These results suggest that the effect is not a byproduct of processing the adjectives themselves but instead reflects genuine anticipatory cognitive processes related to animacy-based predictions of the upcoming noun.

Our findings provide insights into *when* comprehenders begin to predict the meaning of upcoming words in a limited adjective-noun context. In our study, we observed the pre-activation of coarse-grained animacy-related semantic features approximately 500 ms after the onset of the preceding adjectives. This time course is about 200 ms earlier than the typical onset of semantic pre-activation observed in sentence or discourse contexts in previous RSA studies (Hubbard & Federmeier, 2021; Wang, Kuperberg, et al., 2018; Wang et al., 2020; Wang et al., 2023). These studies suggest that comprehenders begin generating semantic predictions roughly 300 ms after the onset of the

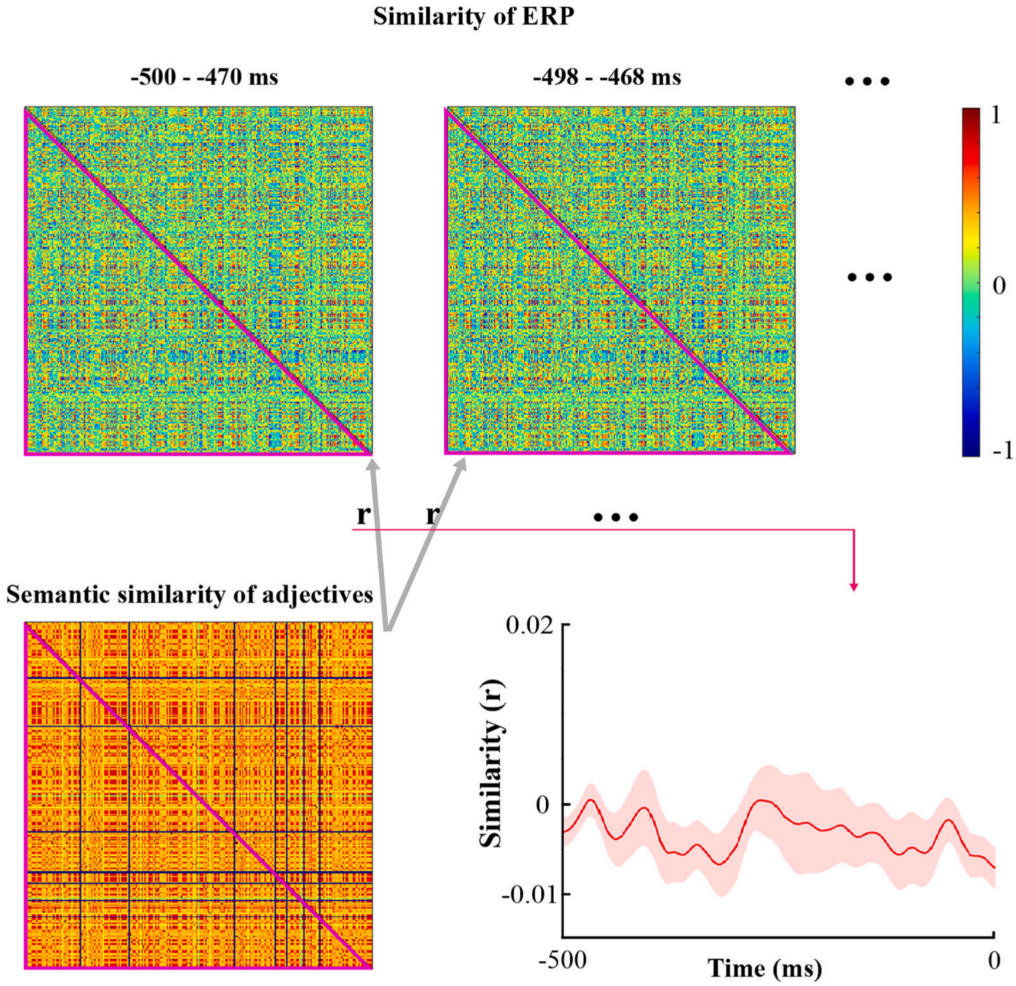

Fig. 3. ERP and RSA results. A. The average ERP waveforms from fronto-central electrodes (Fz, F1, F2, FCz, FC1, FC2, Cz, C1, C2) are shown for four conditions: animate-congruent (solid red line), animate-incongruent (dashed red line), inanimate-congruent (solid blue line), and inanimate-incongruent (dashed blue line). Compared to the congruent conditions, the incongruent conditions elicit larger N400 in the time window of 300 ms to 500 ms, as revealed by larger negativity for the incongruent condition (dashed line) compared to the congruent condition (solid line). B. Pre-noun RSA results. It shows similarity values of animate-constraining condition (red line) and inanimate condition (blue line) before the onset of nouns. The similarity between patterns of neural activity following animate-constraining condition was greater than following inanimate-constraining condition around −480– −414 ms, −302– −246 ms, and −136– −92 ms before the onset of nouns. Standard errors are indicated with shading. The significant time window is indicated by the red horizontal bar.

pre-target word in more complex sentence contexts. The later onset of pre-activation in the present study may be attributed to the more limited context, with only the preceding adjective (rather than a full sentence) guiding prediction. This suggests that when the context is more limited, the brain perhaps requires more time to initiate semantic predictions. Regarding the duration of the effect, previous RSA studies have typically reported transient effects with one time window—lasting ∼ 100 ms (e.g., Wang, Kuperberg, et al., 2018; Wang et al., 2020; Wang et al., 2023) or ∼ 200 ms (Huang et al., 2023). In contrast, the present study identified three time windows, each lasting around 50 ms. However, we do not interpret these time windows as reflecting multiple cognitive processes. Instead, they likely represent the temporal dynamics of a single cognitive process—the pre-activation of animacy-related semantic features—captured across different moments in time. The variability in these time windows may be attributable to multiple factors, such as individual differences in processing speed, fluctuations in neural responses across trials and participants, or the sensitivity of our RSA analysis in detecting subtle shifts in neural similarity over time.

The present study manipulated semantic congruency between adjectives and the nouns that followed, categorizing the nouns as either congruent or incongruent with the preceding adjectives. Adjective-noun incongruency generated larger N400 effects, which is consistent with prior research, reflecting the difficulty in semantic integration when the expected features suggested by the adjective do not match with the actual features of the noun. These findings align with existing literature that establishes the N400 component as a marker of lexical-semantic processing difficulty, particularly when encountering semantic anomalies or incongruities within a linguistic context (e.g., Federmeier & Kutas, 1999; Kutas & Federmeier, 2000; Zhang et al., 2012; Zhou et al., 2010). Interestingly, this incongruency effect only emerged for animate-constraining adjectives, but did not reach conventional significance for inanimate-constraining adjectives, although a clear trend in the expected direction was observed. This finding is in line with similar findings from Huang et al. (2023) that semantic incongruence elicited N400 effects, mainly with animate-constraining classifiers. One possible reason for the discrepancy of the incongruence effect between animate and inanimate-constraining conditions is that animate adjectives impose stronger constraints on animacy expectations compared to inanimate adjectives. However, this possibility is less likely given the animacy constraint was almost the same between animate- and inanimate-constraining conditions as measured by the cloze probability test. Another potential explanation is that the semantic plausibility of phrases was not well matched across conditions. The semantic plausibility ratings show that while animate-congruent and inanimate-congruent phrases were well matched in terms of plausibility, animate-incongruent phrases were rated as less plausible than inanimate-incongruent phrases (1.73 vs. 1.99). The less pronounced N400 effect observed for inanimate-incongruent phrases may be due to the fact that these phrases were not sufficiently incongruent. In other words, the discrepancy in N400 effects between animate and inanimate constraining conditions might be attributable to the mismatch in semantic plausibility between the animate and inanimate incongruent conditions. Future studies should aim to better match this factor to ensure more robust comparisons.

This study's focus on minimal contexts, specifically adjective-noun combinations, highlights how adjectives alone can trigger anticipatory cognitive processes. The finding that animacy-constraining adjectives facilitate the prediction of noun animacy without the support of extensive contextual cues underline the important role of adjectives as predictive cues in language prediction. This research extends beyond the findings regarding the predictions based on classifiers (Huang et al., 2023). In Chinese, the roles of adjectives and classifiers in conveying semantic information, including animacy, are distinct. Adjectives, typically describing qualities or states, contrast with classifiers, which categorize and quantify nouns. While classifiers provide a direct and explicit indication of categories, including animacy, adjectives constraint animacy less. This distinction is critical for understanding predictive processing in language, as it relates to how comprehenders utilize various linguistic cues to form expectations about forthcoming information.

In summary, we provide neural evidence for the prediction of coarse-grained animacy-related semantic features driven by isolated Chinese adjectives, revealing that the predictive processing occurs in minimal and immediate contexts.

CRediT authorship contribution statement

Jingxiao Li: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Mingdong Li:** Writing – original draft, Methodology, Formal analysis, Conceptualization. **Wei Zhou:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Data curation, Conceptualization. **Qingqing Qu:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Formal analysis, Data curation, Conceptualization.

Funding

This work was supported by the National Natural Science Foundation of China (No. 32171058 and No. 62061136001), Youth Innovation Promotion Association (Chinese Academy of Sciences), Youth Elite Scientist Sponsorship Program (No. YESS20200138, China Association for Science and Technology), and the Scientific Foundation of Institute of Psychology (No. E2CX3625CX, Chinese Academy of Sciences) to Qingqing Qu and the Cooperation Project with Chinese Institute for Brain Research, Beijing (No. 2021-NKX-XM-05) to Wei Zhou. This work was also supported by the Ministry of Education Humanities and Social Sciences Research Project (24YJCZH468).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Barber, H., & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience, 17*(1), 137–153. https://doi.org/10.1162/0898929052880101

Beijing Language Institute. (1986). *Modern Chinese Frequency Dictionary (in Chinese)*. Beijing: Beijing Language Institute press.

Bell, A. J., & Sejnowski, T. J. (1994). A Non-linear Information Maximisation Algorithm that Performs Blind Separation. *Neural Information Processing Systems.* https://api.semanticscholar.org/CorpusID:14295333.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early Occipital Sensitivity to Syntactic Category Is Based on Form Typicality. *Psychological Science, 21*(5), 629–634. https://doi.org/10.1177/0956797610367751

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition, 110*(3), 293–321. https://doi.org/10.1016/j.cognition.2008.09.008

Dong, Z., Dong, Q., & Hao, C. (2010). HowNet and its computation of meaning. COLING 2010, 23rd International Conference on Computational Linguistics, Demonstrations Volume, 23-27 August 2010, Beijing, China.

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. Journal of Verbal Learning and Verbal Behavior, 20(6), 641–655. doi: 10.1016/S0022-5371(81)90220-6.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44*(4), 491–505. https://doi.org/10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other Name: Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language, 41*(4), 469–495. https://doi.org/10.1006/jmla.1999.2660

Fruchter, J., Linzen, T., Westerlund, M., & Marantz, A. (2015). Lexical Preactivation in Basic Linguistic Phrases. *Journal of Cognitive Neuroscience, 27*(10), 1912–1935. https://doi.org/10.1162/jocn_a_00822

Grisoni, L., Dreyer, F. R., & Pulvermüller, F. (2016). Somatotopic Semantic Priming and Prediction in the Motor System. *Cerebral Cortex (New York, NY), 26*, 2353–2366. https://api.semanticscholar.org/CorpusID:18635819.

Grisoni, L., Miller, T. M., & Pulvermüller, F. (2017). Neural Correlates of Semantic Prediction and Resolution in Sentence Processing. *Journal of Neuroscience, 37*(18), 4848–4858. https://doi.org/10.1523/JNEUROSCI.2800-16.2017

Grisoni, L., Tomasello, R., & Pulvermüller, F. (2021). Correlated Brain Indexes of Semantic Prediction and Prediction Error: Brain Localization and Category Specificity. *Cerebral Cortex, 31*(3), 1553–1568. https://doi.org/10.1093/cercor/bhaa308

Hagoort, P. (2003). How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *NeuroImage, 20*(SUPPL. 1), S18–S29. https://doi.org/10.1016/J.NEUROIMAGE.2003.09.013

He, T., Boudewyn, M. A., Kiat, J. E., Sagae, K., & Luck, S. J. (2022). Neural correlates of word representation vectors in natural language processing models: Evidence from representational similarity analysis of event-related brain potentials. *Psychophysiology, 59*(3). https://doi.org/10.1111/psyp.13976

Huang, Z., Feng, C., & Qu, Q. (2023). Predicting coarse-grained semantic features in language comprehension: Evidence from ERP representational similarity analysis and Chinese classifier. *Cerebral Cortex, 33*(13), 8312–8320. https://doi.org/10.1093/cercor/bhad116

Hubbard, R. J., & Federmeier, K. D. (2021). Representational Pattern Similarity of Electrical Brain Activity Reveals Rapid and Specific Prediction during Language Comprehension. *Cerebral Cortex, 31*(9), 4300–4313. https://doi.org/10.1093/cercor/bhab087

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626*, 118–135. https://doi.org/10.1016/J.BRAINRES.2015.02.014

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory & Language, 49*(1), 133–156.

Kim, A. E., & Lai, V. T. (2012). Rapid Interactions between Lexical Semantic and Word Form Analysis during Word Recognition in Context: Evidence from ERPs. *Journal of Cognitive Neuroscience, 24*, 1104–1112. https://api.semanticscholar.org/CorpusID:608094.

Kochari, A. R., Lewis, A. G., Schoffelen, J. M., & Schriefers, H. (2021). Semantic and syntactic composition of minimal adjective-noun phrases in Dutch: An MEG study. *Neuropsychologia, 155*, Article 107754. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2021.107754

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32–59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Laszlo, S., & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language, 61*(3), 326–338. https://doi.org/10.1016/j.jml.2009.06.004

Lau, E. F., Holcomb, P. J., & Kuperberg, G. R. (2013). Dissociating N400 Effects of Prediction from Association in Single-word Contexts. *Journal of Cognitive Neuroscience, 25*(3), 484–502. https://doi.org/10.1162/jocn_a_00328

León-Cabrera, P., Rodríguez-Fornells, A., & Morís, J. (2017). Electrophysiological correlates of semantic anticipation during speech comprehension. *Neuropsychologia, 99*, 326–334. https://doi.org/10.1016/j.neuropsychologia.2017.02.026

Li, X., Li, X., & Qu, Q. (2022). Predicting phonology in language comprehension: Evidence from the visual world eye-tracking task in Mandarin Chinese. *Journal of Experimental Psychology. Human Perception and Performance, 48*(5), 531–547. https://doi.org/10.1037/xhp0000999

Loerts, H., Stowe, L. A., & Schmid, M. S. (2013). Predictability speeds up the re-analysis process: An ERP investigation of gender agreement and cloze probability. *Journal of Neurolinguistics, 26*(5), 561–580. https://doi.org/10.1016/J.JNEUROLING.2013.03.003

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin, 144*(10), 1002–1044. https://doi.org/10.1037/BUL0000158

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage, 198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Pulvermüller, F., & Grisoni, L. (2020). Semantic Prediction in Brain and Mind. *Trends in Cognitive Sciences, 24*(10), 781–784. https://doi.org/10.1016/j.tics.2020.07.002

Taylor, W. L. (1953). "Cloze Procedure": A New Tool for Measuring Readability. Journalism Quarterly, 30(4), 415–433. doi: 10.1177/107769905303000401.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(3), 443–467. https://doi.org/10.1037/0278-7393.31.3.443

Vissers, C. Th. W. M., Chwilla, D. J., & Kolk, H. H. J. (2006). Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. Brain Research, 1106(1), 150–163. doi: 10.1016/j.brainres.2006.05.012.

Xun, E., G. Rao, X. Xiao and J. Zang. (2016). "The construction of the BCC Corpus in the age of Big Data". Corpus Linguistics 3(1). 93–109+118.

Wang, L., Brothers, T., Jensen, O., & Kuperberg, G. R. (2023). Dissociating the pre-activation of word meaning and form during sentence comprehension: Evidence from EEG representational similarity analysis. Psychonomic Bulletin & Review. doi: 10.3758/s13423-023-02385-0.

Wang, L., Kuperberg, G. R., & Jensen, O. (2018). *Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity., ELife*, 7. https://api.semanticscholar.org/CorpusID:57761716.

Wang, L., Wlotko, E., Alexander, E., Schoot, L., Kim, M., Warnke, L., & Kuperberg, G. R. (2020). Neural Evidence for the Prediction of Animacy Features during Language Comprehension: Evidence from MEG and EEG Representational Similarity Analysis. *Journal of Neuroscience, 40*(16), 3278–3291. https://doi.org/10.1523/JNEUROSCI.1733-19.2020

Wu, Z., & Palmer, M. (1994). Verb Semantics and Lexical Selection. In *32nd Annual Meeting of the Association for Computational Linguistics* (pp. 133–138). https://doi.org/10.3115/981732.981751

Zhang, Y., Zhang, J., & Min, B. (2012). Neural dynamics of animacy processing in language comprehension: ERP evidence from the interpretation of classifier-noun combinations. *Brain and Language, 120*(3), 321–331. https://doi.org/10.1016/j.bandl.2011.10.007

Zhou, X., Jiang, X., Ye, Z., Zhang, Y., Lou, K., & Zhan, W. (2010). Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: An ERP study. *Neuropsychologia, 48*(6), 1551–1562. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2010.02.001