COGNITIVE SCIENCE
A Multidisciplinary Journal

Check for updates

Cognitive Science 49 (2025) e70054
© 2025 Cognitive Science Society LLC.
ISSN: 1551-6709 online
DOI: 10.1111/cogs.70054

Semantic and Phonological Prediction in Language Comprehension: Pretarget Attraction Toward Semantic and Phonological Competitors in a Mouse Tracking Task

Wenting Ye,[a,b,c] Qingqing Qu[c,d]

[a]*Key Laboratory of Cognition and Personality (SWU), Ministry of Education*
[b]*Faculty of Psychology, Southwest University (SWU)*
[c]*Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences*
[d]*Department of Psychology, University of Chinese Academy of Sciences*

Received 10 November 2024; received in revised form 25 February 2025; accepted 1 March 2025

Abstract

Recent evidence increasingly suggests that comprehenders are capable of generating probabilistic predictions about forthcoming linguistic inputs during language comprehension. However, it remains debated whether language comprehenders predict low-level word forms and whether they always make predictions. In this study, we investigated semantic and phonological prediction in high- and low-constraining sentence contexts, utilizing the mouse-tracking paradigm to trace mouse movement trajectories. Mandarin Chinese speakers listened to high- and low-constraining sentences which resulted in high and low predictability for the critical target words. While listening, participants viewed a visual display featuring two objects: one corresponding to the critical target word (the target object) and the other being either semantically related, phonologically related, or unrelated to the target word. Participants were instructed to click on the target object. The analysis of mouse movement trajectories revealed two key findings: (1) In both high- and low-constraining contexts, there was a spatial attraction of the cursor toward semantic competitors, notably occurring before the target word was heard; (2) there are indications that phonological pretarget attraction effects were observed primarily in high-constraining contexts. These findings suggest that the constraints of sentences have the potential to modulate the representational contents of linguistic prediction during language comprehension. Methodologically, the mouse-tracking paradigm presents a promising tool for further exploration of linguistic prediction.

Keywords: Preactivation; Semantic prediction; Phonological prediction; Prediction; Mouse-tracking

Correspondence should be sent to Qingqing Qu, Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Chaoyang District, Beijing, China. E-mail: quqq@psych.ac.cn

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1. Introduction

Probabilistic prediction is widely regarded as an important computational principle in language comprehension (Federmeier, 2007; Huettig, 2015; Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pulvermüller & Grisoni, 2020). Over the past decade, a substantial body of empirical research has supported that comprehenders actively anticipate the meaning of forthcoming information (Federmeier & Kutas, 1999; Huang, Feng, & Qu, 2023; Wang, Kuperberg, & Jensen, 2018; Wicha, Moreno, & Kutas, 2004). A topic of ongoing debate is the depth of these predictions, particularly whether they extend to downstream word-form representations as compared to high-level semantic representations (see Nieuwland et al., 2020 for a comprehensive review; also see DeLong, Urbach, & Kutas, 2005; Ito, Pickering, & Corley, 2018; Kukona, 2020; Li, Li, & Qu, 2022; Wei, Huang, Feng, & Qu, 2023). Additionally, a relatively underexplored issue is highlighted by Huettig (2015) as the "when issue" (i.e., do individuals always predict upcoming information during language processing). Most existing studies in predictive processing focus on contexts that highly constrain toward specific target words. However, as Luke and Christianson (2016) noted, most words in natural language are not highly predictable, and it is unclear about predictive processing in low-constraining contexts. Our current study seeks to bridge this gap. We investigate both semantic and phonological prediction in high- and low-constraining sentence contexts, utilizing the mouse-tracking paradigm to provide new insights into dynamic aspects of language prediction.

Numerous studies support the notion that comprehenders can anticipate the semantic content of upcoming input. For instance, highly predictable words are read more quickly (Ehrlich & Rayner, 1981) and evoke a smaller N400 event-related potentials (ERP) response compared to unexpected words (Federmeier & Kutas, 1999; Kutas & Hillyard, 1984). However, one should be cautious in interpreting faster reading times and ERP data as indicators of prediction, considering that these effects might be confounded with integration with prior context (Federmeier & Kutas, 1999; Mantegna, Hintz, Ostarek, Alday, & Huettig, 2019; Nieuwland et al., 2020; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). More recent investigations have uncovered prestimulus predictive brain activity, also known as prediction potential (Grisoni, Dreyer, & Pulvermüller, 2016, 2017, 2021; León-Cabrera, Rodríguez-Fornells, & Morís, 2017; reviewed in Pulvermüller & Grisoni, 2020). These studies report that high-constraining contexts trigger anticipatory brain activity prior to the appearance of expected words, termed semantic prediction potential. Notably, the magnitude of this anticipatory activity correlates with the predictability of the expected word (see Grisoni, Tomasello, & Pulvermüller, 2021). Moreover, multivariate pattern analyses of neural data, which examine the time window before the presentation of predictable words, have revealed predictive effects (Huang et al., 2023; Hubbard & Federmeier, 2021; Wang et al., 2018, 2020). For example, representational similarity analysis (RSA) has shown that neural activity patterns following animate-constraining verbs were more alike compared to those following inanimate-constraining verbs, before the presentation of upcoming words, indicating preactivation of the animacy features of upcoming words (Wang et al., 2020). Similar findings were observed with Chinese classifiers, which constrain the animacy of following nouns (Huang et al., 2023). Additionally, it has been observed that the neural

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

pattern similarity for identical words in different contexts is higher than for distinct words, importantly, prior to the words being presented (Wang et al., 2018). These findings suggest that comprehenders anticipate semantic information before encountering upcoming words.

The debate concerning whether predictions extend to low-level word-form representations has been an issue of considerable interest, with mixed evidence emerging from various studies (DeLong et al., 2005; Ito et al., 2018; Nieuwland et al., 2018; for a comprehensive review, see Nieuwland et al., 2020; Li et al., 2022). ERP research has observed a reduction in the N400 component in response to unexpected words that share orthographic features with predicted words (DeLong, Chan, & Kutas, 2019; DeLong, Chan, & Kutas, 2021; Ito, Corley, Pickering, Martin, & Nieuwland, 2016). Beyond the N400, additional electroencephalography (EEG) or magnetoencephalography (MEG) studies have reported early components (M100, P130, N1, P2, N200/PMN, N250, P300) as potential indicators of word-form prediction (Dikker, Rabagliati, & Pylkkänen, 2009, 2010; Dikker & Pylkkänen, 2013; Gagnepain, Henson, & Davis, 2012; Kim & Lai, 2012; Sohoglu, Peelle, Carlyon, & Davis, 2012; Vespignani, Canal, Molinaro, Fonda, & Cacciari, 2010). However, the occurrence of these effects after the onset of critical linguistic input has led to questions about whether they truly demonstrate predictive processing (Nieuwland, 2019).

A landmark study by DeLong et al. (2005) investigated phonological prediction in language comprehension using the English phonological rule for articles "a" and "an." The study revealed an N400 amplitude associated with cloze probability for both nouns and their preceding articles, suggesting the preactivation of the upcoming target at the level of phonological representations. However, the findings were challenged by a large-scale study conducted by Nieuwland et al. (2018), which suggest that the limitations of the a/an paradigm, such as articles not always directly preceding nouns, may contribute to the inconsistent results, leaving the question of phonological prediction unresolved. Recently, Wei et al. (2023) explored phonological prediction in Chinese idioms using ERP RSA. They presented participants with four-character Chinese idioms, and varied the syllable of the third characters so that idiom pairs either shared a syllable (i.e., within-pairs) or not (between-pairs). RSA results revealed greater neural pattern similarity for idiom pairs that shared a syllable (within-pairs) compared to those that did not (between-pairs). Critically, this similarity effect was observed before the presentation of phonologically similar characters. This finding provides evidence for the preactivation of upcoming phonological information, at least under circumstances that encourage predictive processing, such as in structurally fixed Chinese idioms.

Another line of evidence for linguistic prediction comes from studies using the visual world paradigm, involving anticipatory eye movements toward objects before the corresponding expressions are heard. In the task, participants listen to sentences while viewing a display of objects related to the spoken stimuli. These studies revealed anticipatory fixations toward critical objects before hearing it. For example, Altmann and Kamide (1999) and Kamide, Altmann, and Haywood (2003) showed that participants exhibited more fixations on objects like a cake after hearing "the boy will eat the…" compared to "the boy will move the…," indicating prediction of semantic information. Word-form prediction has also been demonstrated in this paradigm. In Ito et al. (2018), participants listened to highly predictable sentences while viewing a visual display that contained a target object corresponding to the predictable word and a phonological competitor object that was phonologically related to the predictable word. Ito et al. found that participants fixated more on phonological competitor objects before hearing the target words, indicating the preactivation of phonological information. While some replications supported these findings (Kukona, 2020; Li et al., 2022; Li & Qu, 2024; Xu, Ji, Li, & Pickering, 2024; Zhao et al., 2024), others did not (Ito & Husband, 2017; Ito & Sakai, 2021). Very recently, meta-analytic evidence on visual-world eye-tracking studies from Ito (2024) found a small but significant phonological prediction effect and highlighted that the variables driving these predictions are not yet well understood. In particular, Ito (2024) identifies cloze probability as a significant mediator of prediction, strengthening the motivation to systematically manipulate this factor in experimental settings. Phonological prediction, as reviewed above, remains a topic of controversy.

In addition to the "what issue" of linguistic prediction (what is predicted, as discussed by Huettig, 2015), the "when issue" remains less explored. This concerns whether individuals always predict upcoming information during language processing, especially whether this occurs exclusively in high-constraining contexts or also in those that are less constraining. Addressing this is crucial, as most words in natural language are not highly predictable, and insights into this area could deepen our understanding of predictive processes (Luke & Christianson, 2016). Despite its theoretical importance, the "when issue" has not been extensively explored, and the prevailing body of research on predictive processing has focused predominantly on high-constraining contexts. It is underexplored whether and how people predict in less constraining contexts. In a seminal study by Federmeier and Kutas (1999), participants were shown sentences word-by-word, ending with a predictable word (e.g., "palms"), an unexpected word from the same semantic category (e.g., "pines"), or an unexpected word from a different category (e.g., "tulips"). Federmeier and Kutas found that N400 amplitude was reduced for unexpected words from the same semantic category compared to unrelated ones, in both high- and low-constraining contexts, suggesting facilitation in word processing is not limited to high-constraining contexts. Subsequent studies have further investigated the impact of context constraints on behavioral or neural responses of words (e.g., DeLong et al., 2005; Ding et al., 2023; see Brothers & Kuperberg, 2021 for details). Using the visual world task, Shen et al. (2021) found anticipatory fixations on words phonologically related to expected words in the predictive context. In contrast, there was no such anticipatory phonological effect in the nonpredictive context (cloze probability = 0). In another study, Ding et al. (2023) examined predictive processing across sentences with varying levels of constraints—strong, moderate, and weak—using ERPs. Their findings indicated that strong constraints led to an ERP positive deflection preceding the nouns, specifically in participants with a high working memory capacity, while no similar pattern was observed in moderate constraints. These studies indicate that linguistic predictions are influenced by contextual constraints.

However, it is unclear whether contextual constraints differentially affect the prediction of meaning and word form. It is plausible to hypothesize that semantic and phonological predictions might not be governed by a uniform cognitive mechanism, each sensitive to varying degrees of contextual information. Based on the observed prevalence of semantic prediction effects in the literature, semantic prediction is hypothesized to be a fundamental component of language processing, and potentially engaged across different contexts. In contrast,

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

phonological prediction may be more context-sensitive, being more likely to be engaged in high-constraining contexts where the linguistic input suggests a limited set of forthcoming phonological forms.

The present study aims to explore these hypotheses of semantic and phonological prediction in language comprehension, by investigating semantic and phonological prediction in high- versus low-constraining sentence contexts. We adopted an experimental approach, the mouse-tracking paradigm (MT; Freeman & Ambady, 2010; Schoemann, O'Hora, Dale, & Scherbaum, 2021), which records the trajectories of a computer mouse while participants process spoken language. The dynamic trajectory of mouse movements provides a data-rich and real-time window into how people process spoken language dynamically. MT offers a promising approach to studying the activation of spoken words. In a foundational study, Spivey, Grosjean, and Knoblich (2005) used the mouse-tracking paradigm to explore the real-time processing in spoken word recognition by examining mouse movement trajectories. Participants were instructed to move the cursor from the bottom center of the screen toward a target picture upon hearing a spoken target word. The trajectory was influenced by unselected alternative competitors, resulting in deflections toward the competitors. In Spivey et al., larger deflections were observed toward phonologically related competitors. For example, the target object "candle" and a competitor object whose name shares initial phonemes with the target word, such as "candy," appeared in the top left and top right corners of the screen, respectively. Then, participants heard the instruction "Click the candle." There was a larger deflection toward "candy" (phonological competitor) while clicking on "candle," compared to an unrelated competitor. This finding was interpreted as evidence of parallel activation and competition among multiple related items during spoken word recognition.

The mouse-tracking paradigm has been employed in exploring various aspects of language processing, including language learning and the perception of foreign accents (Bartolotti & Marian, 2012; Darley, Kent, & Kazanina, 2020; Farmer, Anderson, & Spivey, 2007a; Farmer, Cargill, Hindy, Dale, & Spivey, 2007b; Freeman, Dale, & Farmer, 2011; Incera, Shah, McLennan, & Wetzel, 2017; Loy, Rohde, & Corley, 2017; Spivey & Dale, 2006). Expanding its application beyond linguistic tasks, Bruhn, Huette, and Spivey (2014) used the mouse-tracking paradigm in a spatial cueing task. Their findings highlighted that mouse-tracking could sensitively detect the influence of cue certainty on anticipatory hand movements. More recently, mouse-tracking has been applied to the study of linguistic prediction, with a particular focus on how different speech rates influence predictive language processing (Kukona, 2023). These studies underscore the sensitivity of mouse-tracking in capturing predictive processing. Moreover, in a recent study by Kukona (2025), the mouse-tracking paradigm has been employed to investigate phonological prediction during sentence processing. Participants were instructed to click on the object (i.e., target) referred to in sentences (Experiment 1), or to perform this task interleaved with a cloze procedure designed to enhance lexical prediction. Despite these efforts to elicit predictive behaviors, no evidence of attraction toward phonological competitors was observed. Based on these findings, Kukona concluded that phonological prediction is not a necessary component of language comprehension and may be optional, depending on various factors. One such factor is lexical predictability (Ito, 2024; Kukona, 2025), that is, the degree to which a word is predictable in context may influence the likelihood of phonological prediction occurring.

In the present study, we conducted two experiments to examine semantic and phonological prediction in language comprehension. In Experiment 1, we examined semantic and phonological prediction in high-constraining sentences that included a predictable target word. While listening to the sentences, the target object referring to the predictable target word (e.g., 蜡烛, /la4zhu2/, English translation: *candle*) was presented alongside one of three types of competitors, either a semantically related competitor (e.g., 火柴, /huo3chai2/, English translation: *match*), a phonologically related competitor (e.g., 辣椒, /la4jiao1/, English translation: *pepper*), or an unrelated competitor. If listeners predict the semantic and/or phonological information of upcoming words, there would be greater movement deflection toward the corresponding competitors before the predictable target word is heard. In Experiment 2, we assessed semantic and phonological prediction in low-constraining sentences, using the same set of spoken target words and competitor words to make the effects more comparable across contextual constraints. As reviewed above, semantic and phonological predictions might rely on different mechanisms, with semantic prediction being a fundamental component of language processing across various contexts, and phonological prediction being more dependent on high-constraining contexts. Therefore, we hypothesize that semantic prediction occurs consistently across different contexts, whereas phonological prediction is more likely to emerge in high-constraining contexts.

2. Experiment 1. Semantic and phonological effects in high-constraining sentences

In Experiment 1, we investigated the preactivation of semantic and phonological representations in high-constraining sentences using the mouse-tracking paradigm.

2.1. Method

2.1.1. Participants

A power analysis showed that a sample size of 52 would be required to achieve 80% power for detecting an average two-level within-participants psychological effect size ($d_z = 0.40$; $\alpha = 0.05$, Brysbaert, 2019). For the current study, 59 native Mandarin Chinese speakers (48 women, age $M = 20.54$, $SD = 1.42$, range = 18–24) were recruited from Southwest University. All participants confirmed that they had normal or corrected-to-normal vision, and they used the mouse with their right hand in daily activities. Ethical approval was granted by Southwest University.

2.1.2. Materials, design, and procedure

Seventy-two sentences were selected from Li et al. (2022). Thirty-six of them were high-constraining sentences that each predicted a specific upcoming word, and the remaining 36 were neutral sentences as fillers. We used a cloze probability test to assess the predictability of the target words, in which target sentences were truncated before the critical words (Taylor,

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1953). Thirteen native Chinese speakers who did not participate in the mouse-tracking experiment were asked to complete each sentence fragment with the first word that came to mind. The mean cloze probability of the target words was 92.5% (SD = 7.23%). The mean length of target sentences was 26.3 characters (range: 21–34 characters, SD = 2.8), and the mean position of target words was at the 17.9th character (range: 8–25th character, SD = 3.4).

In each trial, participants were presented with spoken sentences and a visual display of black-and-white line drawings of two objects (one target and one competitor). For each target object, three types of competitors (i.e., semantic, phonological, and unrelated) were constructed. In the semantic competitor condition, the competitor object (e.g., 火柴, */huo3chai2/*, English translation: *match*) was semantically related to the target object (e.g., 蜡烛, */la4zhu2/*, English translation: *candle*). In the phonological competitor condition, the competitor (e.g., 辣椒, */la4jiao1/*, English translation: *chilli*) shared the first syllable (regardless of tone) with the name of the target object. In the unrelated condition, the competitor was semantically/phonologically unrelated to the target object. Name agreement on the line drawings, word frequency, and visual complexity of the objects were matched across conditions. All visual objects were adjusted to a size of 263 × 188 pixels. We also assessed the plausibility of target words, semantic, phonological, and unrelated competitors within the sentences, in which all sentences were truncated after the target or competitor words. A group of 20 participants rated the plausibility of each sentence on a scale ranging from 1 (very implausible) to 7 (very plausible). The plausibility of the target words was higher than that of competitor words (target: 6.84; semantic competitor: 3.18; phonological competitor: 1.22; unrelated distractor: 1.34). The experiment consisted of a total of 80 trials, which included eight practice trials and 72 experimental trials (36 critical trials and 36 filler trials). Three lists were generated with 12 trials in each type of three competitors, so that no item was repeated within any of the lists. Participants were randomly assigned to one of the three lists. The positions of the two objects were fully randomized.

The experiment was operated via MouseTracker software (Freeman & Ambady, 2010). Participants were seated approximately 60 cm away from a computer screen (DELL P19175S; 19 in. Flat screen monitor with a screen resolution of 1280 × 1024) and completed the task using a computer mouse (DELL, MS116c). The trial structure is displayed in Fig. 1. Participants were instructed to click on a "Start" button at the bottom center of the screen to initiate a trial. Once participants initiated a trial, a sentence was presented via headphones, and two visual objects appeared in the upper left and upper right corners of the screen, which were presented 2000 ms before the onset of the spoken target word (mean duration of spoken target words: 740 ms). That is, the preview time of visual objects preceding spoken target words was fixed. Participants were instructed to mouse-click the visual objects mentioned in the sentence as quickly and accurately as possible. Feedback was provided for 1000 ms if participants made an incorrect response. To prevent any premature mouse movements and ensure that participants' reactions were directly in response to the visual stimuli, the mouse cursor immediately disappeared after the "Start" button was clicked. It then relocated to the bottom center and remained locked until the visual objects were presented. That is, the first time point in a mouse trajectory represented the onset of visual objects. MouseTracker recorded the raw

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

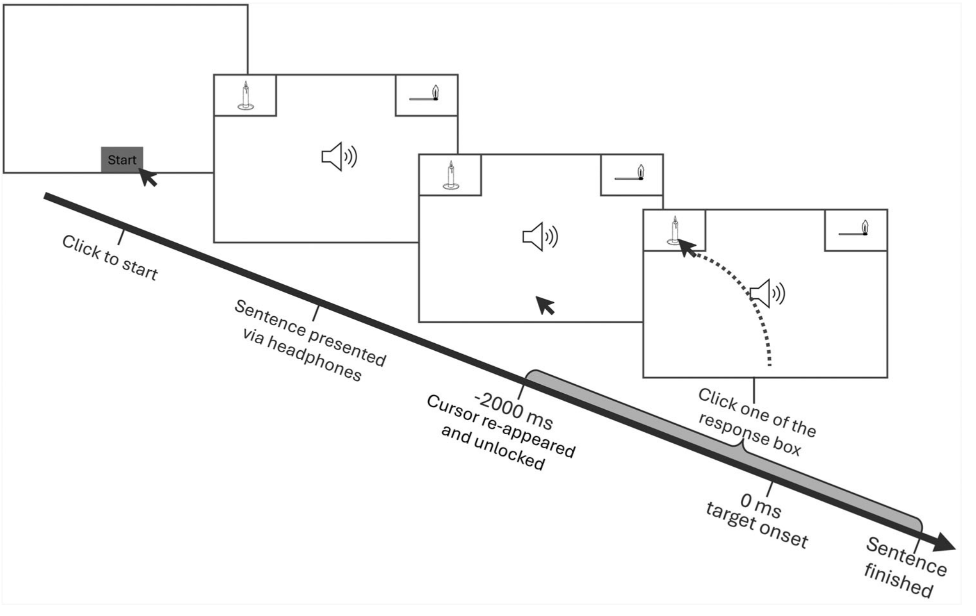

Fig. 1. Trial structure.

data of each mouse trajectory, capturing the *x* and *y* coordinates at a sampling rate of 50 Hz. Cursor speed was set to a value of 12 within MouseTracker, with acceleration turned off.

2.1.3. Data recording and analysis

Data were processed in R 4.3.0 (R Core team, 2023) using the *mousetrap* package (Kieslich et al., 2019; Wulff et al., 2023). For each trial, response accuracy, initiation time (IT), and response time (RT) were computed and recalibrated relative to the onset of spoken target words. As miniscule movements after clicking on the “Start” region were often observed and then erroneously recorded as extremely short ITs by MouseTracker, we redefined IT as the first time sample when the cursor left the starting region relative to the onset of spoken target words. This could be negative if participants moved the cursor before the onset of spoken target words. Following the recommendation of Wirth, Foerster, Kunde, and Pfister (2020), the starting region was recomputed as a virtual circular starting region with a diameter equal to the height of the “Start” box (34 pixels), as a rectangular starting region might be confounded with the starting angle of a movement. *Response Times* were measured as the time interval between the onset of spoken target words and the click on the response field. To examine the effects of semantic and phonological competitors on IT and RT, we computed generalized linear mixed-effects models using a gamma distribution with a log link. These models were run with IT and RT as dependent variables using the *mixed()* function from the *afex* package in R (Singmann et al., 2023). We first constructed the maximal model with the competitor condition (semantic competitor vs. phonological competitor vs. unrelated competitor) as a fixed factor, and the full random effect structure including random

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

intercepts and slopes for the competitor condition by participants and by items (Barr, Levy, Scheepers, & Tily, 2013). However, the model was simplified, due to the failure of convergence. The final model included the competitor condition as a fixed factor, and by-participant and by-item random intercepts. Sum-to-zero contrasts were used for coding the competitor condition. Post-hoc comparisons were computed using the *emmeans* package (Lenth, 2023).

More critically, to investigate the temporal dynamics of semantic and phonological effects, *time course analyses* were performed on the deflection of mouse movements on the horizontal axis (i.e., x-coordinates; e.g., Incera & McLennan, 2018; Spivey et al., 2005; Ye & Damian, 2022). In the time course analysis, x-coordinates at each sampled time step (for every 20-ms time bin at 50 Hz sampling frequency) were aggregated, dependent on the conditions. At each time step, paired t-tests were performed on x-coordinates for phonological competitor versus unrelated conditions, and semantic competitor versus unrelated conditions, to determine when the two competitor effects began to emerge and the duration of the competitor effects on the mouse movement trajectory. We performed this analysis on raw trajectories so the onset of the spoken target words could be identified from the raw time steps (Ye & Damian, 2022). As time steps in mouse movement trajectories are not independent, correction for multiple comparison was necessary to address whether consecutive significant trajectory segments in our experiments exceeded what would occur by chance. We followed the Monte Carlo simulations approach of Dale, Kehoe, and Spivey (2007); cited in Scherbaum, Gottschalk, Dshemuchadse, & Fischer, 2015) to establish a criterion for the multiple t-tests on the results of trajectories. We generated 10,000 simulations for the smallest effect (i.e., the phonological competitor effect) found in both experiments. To generate synthetic x-coordinates of phonological competitor and unrelated conditions, we sampled each time step from a normal distribution with the mean and standard deviation of this time step from the actual phonological and unrelated conditions. This procedure aimed to preserve the interdependence between time steps in the synthetic data. Then, as in the actual data analysis, a paired t-test was computed at each time step and the longest significant segments were recorded. For each experiment, we constructed the same number of model participants and aggregated time steps as were used in the actual analyses. This established a criterion of five consecutive significant time steps for Experiment 1, and five consecutive time steps for Experiment 2, ps $< .001$, for the analysis of the actual data (i.e., 10,000 simulations for each experiment).

RT records the onset of the response and is a measure of processing speed. It does not capture the complexity of the cognitive process during language comprehension. In contrast, the time course analysis of mouse movement trajectory focuses on the entire course of the mouse movement, providing a richer and more detailed picture of the process. Previous mouse-tracking studies have shown that mouse movement trajectory analysis could be more sensitive than RT in detecting effects, with some effects observed only in the mouse trajectory analysis (e.g., Damian, Ye, Oh, & Yang, 2019; Incera & McLennan, 2016). In our study, the time course analysis captures how effects develop dynamically over time, which might allow us to better examine whether there are spatial attraction effects and when these effects emerge, and thus we will mainly focus on the results of the time course of mouse trajectory.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

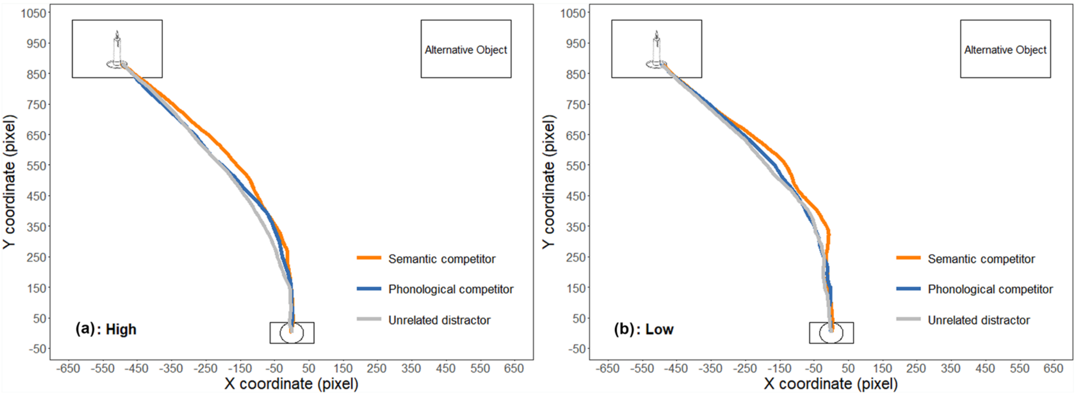

Fig. 2. Visualization of time-normalized average trajectories (a) in the high-constraining sentences and (b) in low-constraining sentences. Rightward trajectories were flipped to appear as pointing toward the left response. The rectangles surrounding the objects (target vs. competitor) indicate response regions where a mouse should click to end the trial.

2.2. Results

Data from trials with errors were excluded from further analysis (1.5%), as well as data from trials following errors (1.5%) to avoid potential post-error adjustments (e.g., Danielmeier & Ullsperger, 2011). Errors were not further analyzed due to the low error rate. Trials with RTs exceeding 2.5 standard deviations from the mean (2.4%) were also excluded from further analysis. Fig. 2a displays time-normalized average trajectories for each of the three types of competitors in the high-constraining sentences. The time course of trajectories was normalized into 101 time steps by linear interpolation, after which trajectories were aggregated at each time step within each condition. Rightward trajectories were flipped to appear as pointing toward the left response. The average trajectories show the effects of the competitors, with more curved trajectories in the semantic and phonological competitor conditions, compared to the unrelated condition. This reflects an attraction toward semantic and phonological competitors, with a greater attraction toward semantic competitors than phonological competitors.

Table 1 provides a descriptive summary of the two dependent measures and the effects. The generalized linear mixed-effects model showed no significant effect of competitor conditions on ITs, $\chi^2(2) = 1.16$, $p = .560$, as participants typically initiated the movements pre-emptively. There was a main effect of competitor conditions on RTs, $\chi^2(2) = 126.38$, $p < .001$. The post-hoc tests revealed that RTs in the semantic competitor condition were significantly longer than those in the unrelated condition ($p < .001$), but the phonological competitor condition did not differ from the unrelated condition ($p = .369$). While contrasting between the two competitor conditions, the semantic competitor condition showed longer RTs compared to the phonological competitor condition, $\Delta = 373$ ms, $p < .001$.

In the time course analysis, we aggregated the x-coordinates at each sampled time step for each competitor condition, using a 20-ms time bin based on a 50 Hz sampling frequency.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Initiation times (IT) and response times (RT), separately for each condition (semantic competitors, phonological competitors, and unrelated objects), in the high-constraining sentences (Experiment 1) and low-constraining sentences (Experiment 2)

	IT (ms)		RT (ms)	
	High	Low	High	Low
Condition [95% CI]				
Semantic	−1129 [−1257, −978]	−724 [−940, −464]	1002 [674, 1370]	1735 [1360, 2152]
Phonological	−1166 [−1289, −1022]	−748 [−960, −493]	629 [342, 951]	1446 [1100, 1830]
Unrelated	−1159 [−1283, −1014]	−858 [−1051, −626]	579 [297, 895]	1392 [1052, 1770]
Effect (log diff. [95% CI])				
Semantic	30 (0.04 [−0.07, 0.14])	134 (0.11 [0.00, 0.22])	423***(0.15 [0.12, 0.19])	343***(0.10 [0.07, 0.12])
Phonological	7 (−0.01 [−0.11, 0.09])	110 (0.09 [−0.02, 0.20])	50 (0.02 [−0.01, 0.05])	54 (0.02 [−0.01, 0.04])

Note. IT and RT are in milliseconds. Results for conditional means are back-transformed to the original scale for interpretation with 95% confidence intervals [95% CI]. Pairwise comparisons were performed on the log scale, with Tukey-adjusted p-values reported. Effects are presented as raw differences (in milliseconds) and their corresponding log-scale differences (log diff.) with 95% confidence intervals [95% CI] were reported in brackets.

*** $p < .001$.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Fig. 3. Results of time course analysis. Time course analysis of the effects of phonological competitors and semantic competitors in (a) high-constraining sentence, and (b) low-constraining sentence contexts. 0 ms at the x-axis represents the onset of spoken target word. The left panels depict the x-coordinates over time (dashed lines) for the semantic competitor (in orange), phonological competitor (in blue), and unrelated (in gray) conditions in the two contexts. In the right panels (solid lines), the orange lines show the semantic competitor effects, and the blue lines show the phonological competitor effects. The average time points of movement IT, RT, and target word onset are represented with vertical lines. The horizontal dashed line represents at $t = \pm 2$, with values exceeding this threshold indicating statistical significance ($p = .05$).

Since raw trajectories have variable end points and an increased degree of data loss occurs at later time steps (i.e., toward the right side of Fig. 3), we stopped conducting tests at the time step when less than 70% of data points remained (i.e., 1060 ms since the onset of spoken word). X-coordinates were compared for semantic versus unrelated conditions, and phonological versus unrelated conditions by paired t-tests. Fig. 3a presents the results, with the y-axis representing the t values and the x-axis representing the time locked to the onset of spoken target words. Compared to the unrelated objects, trajectories were likely to be attracted by both the semantic competitors and phonological competitors. The semantic effect and phonological effect both emerged *before* the onset of spoken target words. The semantic competitor effect emerged approximately 1380 ms *before* the onset of spoken target words and persisted until 220 ms after the onset, spanning 81 consecutive significant time steps. The phonological competitor effect also emerged *before* the onset of spoken target words, from −600 to −440 ms relative to the onset, spanning nine consecutive significant time steps. With a criterion of five consecutive significant time steps, the results of 81 consecutive time steps for semantic

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

effect and nine consecutive time steps for phonological effect were accepted as significant segments with $p < .001$.

In Experiment 1, while comprehending high-constraining sentences, listeners demonstrated attraction toward both semantic and phonological competitors. The dynamic time course analysis revealed that listeners' attraction toward semantic and phonological competitors occurred *before* encountering the spoken target word. Moreover, semantic competitors generated an earlier attraction than phonological competitors.

3. Experiment 2. Semantic and phonological effects in low-constraining sentences

In Experiment 2, we investigated the preactivation of semantic and phonological representations in low-constraining sentences using the mouse-tracking paradigm.

3.1. Method

3.1.1. Participants

A power analysis showed that a sample size of 52 is required to achieve 80% power for detecting an average two-level within-participants psychological effect size ($d_z = 0.40$; $\alpha = 0.05$, Brysbaert, 2019). A separate group of 56 native Mandarin Chinese speakers (45 women, age $M = 22.25$, $SD = 2.44$, range = 18–30) were enrolled at Southwest University.

3.1.2. Materials, design, and procedure

All aspects of Experiment 2 were the same as those of Experiment 1, with the exception that spoken target words were embedded within low-constraining sentences. The same set of critical words were used to construct low-constraining sentences. As in Experiment 1, cloze probabilities for these sentences were obtained from a group of 20 Chinese native speakers who did not take part in the mouse-tracking experiment. The mean cloze probability was 23.9% ($SD = 11.1\%$). The mean length of each target sentence was 26.7 characters (range = 19–34 characters, $SD = 3.5$), and the mean position of target words was at the 16.7th character (range = 10–28th character, $SD = 5.0$). In addition, the plausibility rating on a scale ranging from 1 (very implausible) to 7 (very plausible) showed that the plausibility of the target words was significantly higher than competitor words (target words: 6.05; semantic competitor: 3.88, phonological competitor: 1.66, unrelated competitor: 1.66). The average sentence length, the position of target words within the sentences, and the plausibility pattern sentences were closely matched across experiments.

3.1.3. Data analysis

The data analysis was identical to Experiment 1.

3.2. Results

Data from trials with errors (0.9%) and trials following errors (0.9%) were excluded from the analysis. Trials with RTs exceeding 2.5 standard deviations from the mean (2.6%) were also excluded. Fig. 2b displays time-normalized average trajectories for each of the three types of competitors in the low-constraining sentences. Rightward trajectories were flipped to appear as pointing toward the left response. Average trajectories show the effects of competitors, with more curved trajectories in the semantic competitor conditions, compared to the other two conditions, reflecting an attraction toward semantic competitors.

Table 1 provides a descriptive summary for ITs and RTs. The generalized linear mixed-effects model showed that there was a significant main effect of competitor conditions on ITs, $\chi^2(2) = 6.09$, $p = .048$, but the post-hoc tests did not show any significant phonological effect or semantic effect. A main effect of competitor conditions was also observed on RTs, $\chi^2(2) = 89.39$, $p < .001$, with longer RTs in the semantic competitor ($p < .001$) but not in the phonological competitor conditions ($p = .370$) compared to the unrelated condition. While contrasting between the two competitor conditions, the semantic competitor condition showed longer RTs compared to the phonological competitor condition, $\Delta = 289$ ms, $p < .001$.

In the time course analysis, *x*-coordinates were aggregated at each sampled time step for each competitor condition. We stopped conducting tests at the time step when less than 70% of data points remained (i.e., 1720 ms after the onset of the spoken word). Critically, the semantic effect emerged *before* the onset of spoken target words, whereas the phonological effect emerged *after* the onset of spoken target words. As shown in Fig. 3b, the semantic competitor effect emerged approximately 900 ms before the onset of spoken target words and persisted until 1140 ms after the onset of spoken target words, spanning 103 consecutive significant time steps. The phonological competitor effect emerged after the onset of spoken target words: from 560 to 740 ms after the onset of spoken target words, with 10 consecutive significant time steps. Correction for multiple comparisons was achieved by accepting segments of more than five consecutive significant *t*-tests ($p <.001$). According to this criterion, both the semantic and phonological effects were robustly significant.

In Experiment 2, while comprehending low-constraining sentences, listeners demonstrated attraction toward both semantic and phonological competitors. The dynamic time course analysis revealed that listeners' attraction toward semantic competitors occurred *before* encountering the spoken target word, whereas attraction toward phonological competitors emerged *after* encountering the spoken target word.

To compare semantic and phonological effects between Experiment 1 and Experiment 2, we performed ANOVAs for each time point with factors semantic/phonological effects × Experiment to assess whether these effects significantly differed across the high- and low-constraining contexts. The analyses revealed significant main effects for both semantic and phonological predictions prior to the onset of the target (ps $< .05$). Interaction effects showed marginal significance for both semantic and phonological effects and across specific time windows before the target onset (ps $< .09$). Specifically, semantic effects interacted with Experiment during the time window of −1080 to −980 ms, −700 to −640 ms, −560 to −440 ms, and −320 to −280 ms. These results are consistent with our observation that semantic effects in high-constraining contexts (Experiment 1) initiated approximately 1400 ms before the target onset, whereas in low-constraining contexts (Experiment 2), semantic effects began 900 ms before the target onset. Additionally, phonological effects interacted with Experiment during the time windows of −480 to −460 ms, corresponding to the later onset of phonological

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

effects in Experiment 1 (600 ms before target onset) and their absence in Experiment 2 before the target. After the target onset, the semantic effect significantly interacted with Experiment (ps $<$ .05), from 640 to 1140 ms. Phonological effects marginally interacted with Experiment from 620 to 640 ms, which aligns with the emergence of phonological effects in low-constraining sentences 560 ms after the target onset. While the ANOVAs revealed marginally significant interactions between the semantic/phonological effects and Experiment before target onset, we acknowledge that the marginal significance limits definitive conclusions about the differential phonological effects across experiments. Nonetheless, there are indications that phonological prediction effects are context-dependent, as t-tests showed.

4. Discussion

In the present study, we investigated semantic and phonological prediction in high- versus low-constraining sentence contexts, by recording mouse movement trajectories in a mouse-tracking task. Mandarin Chinese speakers listened to high- or low-constraining sentences while viewing a visual display of two objects. The results of mouse movement trajectories revealed that the cursor was attracted toward semantic competitors before the target word was heard, in both high- and low-constraining contexts. Phonological attraction effects emerged before the onset of the spoken target word in high-constraining sentence comprehension, but emerged after the onset of the target word in low-constraining sentence comprehension. These findings suggest that semantic prediction occurs in both high- and low-constraining sentences. Having said that, while semantic prediction occurs in both contexts, it begins earlier in high-constraining sentences than in low-constraining sentences. On the other hand, although the interaction between the effects of phonological prediction and the degree of context constraints was merely marginally significant, there are indications that phonological prediction occurs primarily in high-constraining sentences, indicating that phonological prediction is context-dependent.

As reviewed in the Introduction, research on whether comprehenders can predict phonological information of upcoming words in language comprehension has yielded mixed findings. A likely origin of these inconsistent results lies in the varying strength of contextual constraints across studies. The present study indicates that phonological prediction may depend on high-constraining context. This insight helps interpret the mixed literature: one plausible reason for the absence of phonological prediction effects in some studies could be insufficiently constraining contexts. In high-constraining sentences, robust contextual cues effectively narrow down potential targets, enabling comprehenders to anticipate the phonological form of the target word before it is heard. Conversely, low-constraining sentences offer limited contextual support, which fails to specify potential targets enough for anticipatory phonological processing. Consequently, in low-constraining contexts, listeners rely more on bottom-up processing, attending to the target word's phonological features only after its onset, rather than engaging in top-down predictive processing. This difference underscores the role of context strength in enabling or limiting phonological prediction in language comprehension. Interestingly, the phonological prediction effects observed in high-constraining sentences contrast with the findings of a mouse-tracking study by Kukona (2025), in which, despite efforts to enhance lexical prediction, no evidence was found for pretarget attraction toward phonological competitors. In a meta-analysis by Ito (2024), which examined studies with mean cloze probabilities ranging from 85% to 98%, a small but significant phonological effect was found. In the present study, the mean cloze probability of the target words was 92.5%, falling within this range. However, in Kukona (2025), the mean cloze probability was 78%, which is lower. This difference in cloze probability may help explain the divergent findings. However, a limitation of the present study is the lack of response time evidence for phonological effects. This may be because the duration of the phonological effect is relatively short, and, therefore, does not lead to a significant influence in overall response time. Moreover, the interactions between semantic effects and Experiment, as well as between phonological effects and Experiment, before the target onset, were marginally significant. Additional studies are needed to further validate these results.

Our findings may offer preliminary insights into the debate on whether linguistic prediction is an "essential" or "optional" function of the language system. Some perspectives argue that prediction is an essential component of language processing (Kuperberg & Jaeger, 2016; Levy, 2008; Federmeier, 2007), while others consider it optional (Pickering & Gambi, 2018). Both views agree that prediction is likely to occur in high-constraining contexts, but they diverge in their predictions for low-constraining contexts. The "essential function" view suggests that the language system actively computes probabilities for high-entropy upcoming words, even in less supportive contexts. In contrast, the "optional function" view proposes that prediction only occurs when strongly cued, with no active computation in low-constraining contexts. Our study aligns with this distinction, showing that while listeners engage in semantic prediction across both high- and low-constraining contexts, phonological prediction appears to rely on stronger contextual support to be activated. This distinction suggests that semantic and phonological prediction may be governed by different mechanisms in sentence comprehension: semantic prediction serves as a broad, generalized anticipatory mechanism, whereas phonological prediction is more context-dependent, requiring robust cues to become active.

Within high-constraining contexts (in Experiment 1), we observed anticipatory mouse movements toward semantic competitors initiating approximately 1380 ms (roughly two words) before the target word onset. This finding is consistent with previous eye-tracking visual world studies that used the same set of stimuli and the same preview time, as evidenced by Li et al. (2022) and Li and Qu (2024), who reported a comparable onset time of 1400 ms. As suggested by Li et al., this early effect likely stems from the initial contextual constraints imposed on the target words. This interpretation is further supported by a cloze probability test for the sentence fragment truncated 1400 ms before the target word in their study, which revealed a moderate cloze probability (mean: 33%, range: 20–45%) for the anticipated target words, and hence eliciting an early preactivation effect. In low-constraining contexts (Experiment 2), anticipatory movements toward semantic competitors began 900 ms before the target word onset, delayed compared to the high-constraining contexts. This indicates that the timing of semantic prediction is influenced by the predictability of the context. Specifically, the onset of anticipatory movements is modulated by the degree of constraint

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

imposed by the linguistic contexts, highlighting the adaptability of the language comprehension system to varying levels of predictability.

In our study, we observed that significant cluster effects for phonological competitors commenced 600 ms before the target word in high-constraining sentences (Experiment 1), and approximately 560 ms after the target word in low-constraining sentences (Experiment 2). This phonological effect persisted for around 200 ms, much shorter than that of semantic effects. This brief duration of phonological predictive effects is consistent with findings from several recent studies using different paradigms and techniques. For instance, visual world studies like those conducted by Ito et al. (2018), Li et al. (2022), and Li and Qu (2024) have reported similar durations of phonological prediction effects, around 200 ms. In addition, a study integrating EEG with RSA by Wei et al. (2023) also documented predictive effects lasting approximately 200 ms. This brief duration of phonological predictive effects could be attributed to the relative ease of disconfirming phonologically related items in a sentence context. In essence, once a phonologically related but incorrect item is presented, it can be rapidly excluded from consideration, thus shortening the duration of the phonological effect. In contrast, disconfirming semantically related items may pose a greater challenge due to their higher semantic plausibility within the sentence. This complexity likely contributes to the more extended duration of semantic effects observed. However, it is worth noting that the predictive effects before the onset of target words in our study should theoretically be uninfluenced by the plausibility of the yet-to-be-encountered words.

The present study represents the application of the mouse-tracking paradigm in the exploration of linguistic prediction. This approach investigates linguistic prediction in sentence comprehension through the detailed analysis of mouse movement trajectories (see also Kukona, 2023 for the application of the mouse-tracking paradigm in linguistic prediction). The mouse-tracking task, alongside the eye-tracking visual world paradigm, introduces a set of possible upcoming referents in the visual display, which can influence linguistic processing. For example, this visual element may foster a predictive context within the task, encouraging participants to actively engage in prediction. This type of experimental paradigm involves restrictions imposed by the context with the available referents in the display, and possibly could not reflect the whole picture of the probability distribution of upcoming words. For instance, in low-constraining contexts, multiple words have a low probability, but they might not be detected when only a small number of options are presented in this type of paradigm. However, an important consideration is that scenarios involving both visual and spoken processing closely mimic real-world situations where auditory information and visual referents are commonly encountered in tandem. This aspect of the mouse-tracking and visual world paradigms offers a significant advantage, providing a more ecologically valid approach to studying linguistic prediction in environments where multiple sensory modalities converge (Huettig, 2015). Another consideration in the mouse-tracking paradigm is the potential issue of similarity between stimuli when presenting multiple options. To address this, an alternative could involve a Go/No-Go paradigm, where only a single stimulus is presented, and participants confirm or reject its appropriateness. This approach effectively minimizes potential interference from similar stimuli but does so at the expense of ecological validity.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The time course analysis was the main analysis approach used in the present study, as it is particularly useful for capturing the dynamics of variables. However, future research could benefit from the time continuous multiple regression (TCMR) approach (Scherbaum, Dshemuchadse, Fischer, & Goschke, 2010), especially if the study involves multiple interacting variables. TCMR allows for simultaneous consideration of multiple predictors and their continuous influence on responses over time, so the effects of multiple predictors could be dissected to determine whether they are independent or overlapping. Furthermore, the mouse-tracking paradigm is particularly advantageous for online testing, owing to its cost-effectiveness and ease of implementation. This makes it an ideal tool for large-scale studies that explore individual differences in predictive processing and associated cognitive skills. In addition, its simplicity and ease of operation also makes it accessible to a wide range of demographic groups, as participants need only the ability to use a mouse.

The findings of our study shed light on the dynamics of semantic and phonological prediction in language comprehension. While semantic predictions occurred in both high- and low-constraining contexts, their timing was influenced by the level of constraint, with earlier predictions in high-constraining sentences. Phonological predictions tend to occur predominantly in high-constraining contexts. These findings suggest that the constraints within sentences have the potential to modulate the representational contents of linguistic prediction during language comprehension. Methodologically, the mouse-tracking paradigm presents a promising tool for further exploration of linguistic prediction.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 32171058), Youth Elite Scientist Sponsorship Program (No. YESS20200138, China Association for Science and Technology), and Beijing Nova Program (20220484081) to Qingqing Qu.

Data availability statement

Data and analysis script are available at the Open Science Framework website: https://osf.io/thjzd/

Ethics approval statement

Ethical approval was granted by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences in Beijing (no. H21105).

References

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bartolotti, J., & Marian, V. (2012). Language learning and control in monolinguals and bilinguals. *Cognitive Science*, *36*(6), 1129–1147.

Brothers, T., & Kuperberg, G. R. (2021). Word predictability effects are linear, not logarithmic: Implications for probabilistic models of sentence comprehension. *Journal of Memory and Language*, *116*, 104174.

Bruhn, P., Huette, S., & Spivey, M. (2014). Degree of certainty modulates anticipatory processes in real time. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(2), 525.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*(1), 16.

Dale, R., Kehoe, C., & Spivey, M. J. (2007). Graded motor responses in the time course of categorizing atypical exemplars. *Memory & Cognition*, *35*, 15–28.

Damian, M. F., Ye, W., Oh, M., & Yang, S. (2019). Bilinguals as "experts"? Comparing performance of mono- to bilingual individuals via a mousetracking paradigm. *Bilingualism: Language and Cognition*, *22*(5), 1176–1193.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233.

Darley, E. J., Kent, C., & Kazanina, N. (2020). A 'no' with a trace of 'yes': A mouse-tracking study of negative sentence processing. *Cognition*, *198*, 104084.

DeLong, K. A., Chan, W. H., & Kutas, M. (2019). Similar time courses for word form and meaning preactivation during sentence comprehension. *Psychophysiology*, *56*(4), e13312.

DeLong, K. A., Chan, W. H., & Kutas, M. (2021). Testing limits: ERP evidence for word form preactivation during speeded sentence reading. *Psychophysiology*, *58*(2), e13720.

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.

Dikker, S., & Pylkkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, *127*(1), 55–64.

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.

Dikker, S., Rabagliati, H., & Pylkkänen, L. (2009). Sensitivity to syntax in visual cortex. *Cognition*, *110*(3), 293–321.

Ding, J., Zhang, Y., Liang, J., & Li, X. (2023). Modulation of working memory capacity on predictive processing during language comprehension. *Language, Cognition and Neuroscience*, *38*(8), 1133–1152.

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.

Farmer, T. A., Anderson, S. E., & Spivey, M. J. (2007a). Gradiency and visual context in syntactic garden-paths. *Journal of Memory and Language*, *57*(4), 570–595.

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007b). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, *31*(5), 889–909.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44*(4), 491–505.

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.

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods*, *42*(1), 226–241.

Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in motion reveals mind in motion. *Frontiers in Psychology*, *2*, 59.

Gagnepain, P., Henson, R. N., & Davis, M. H. (2012). Temporal predictive codes for spoken words in auditory cortex. *Current Biology*, *22*(7), 615–621.

Grisoni, L., Dreyer, F. R., & Pulvermüller, F. (2016). Somatotopic semantic priming and prediction in the motor system. *Cerebral Cortex*, *26*(5), 2353–2366.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

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.

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.

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.

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.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, *1626*, 118–135.

Incera, S., & McLennan, C. T. (2016). Mouse tracking reveals that bilinguals behave like experts. *Bilingualism: Language and Cognition*, *19*(3), 610–620.

Incera, S., & McLennan, C. T. (2018). The time course of within and between-language interference in bilinguals. *International Journal of Bilingualism*, *22*(1), 88–99.

Incera, S., Shah, A. P., McLennan, C. T., & Wetzel, M. T. (2017). Sentence context influences the subjective perception of foreign accents. *Acta Psychologica*, *172*, 71–76.

Ito, A. (2024). Phonological prediction during comprehension: A review and meta-analysis of visual-world eye-tracking studies. *Journal of Memory and Language*, *139*, 104553.

Ito, A., Corley, M., Pickering, M. J., Martin, A. E., & Nieuwland, M. S. (2016). Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, *86*, 157–171.

Ito, A., & Husband, E. M. (2017). How robust are effects of semantic and phonological prediction during language comprehension? A visual world eye-tracking study. *Mental Architecture for Processing and Learning of Language*, *13*(7), 1–6.

Ito, A., Pickering, M. J., & Corley, M. (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, *98*, 1–11.

Ito, A., & Sakai, H. (2021). Everyday language exposure shapes prediction of specific words in listening comprehension: A visual world eye-tracking study. *Frontiers in Psychology*, *12*, 607474.

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

Kim, A., & Lai, V. (2012). Rapid interactions between lexical semantic and word form analysis during word recognition in context: Evidence from ERPs. *Journal of Cognitive Neuroscience*, *24*(5), 1104–1112.

Kukona, A. (2020). Lexical constraints on the prediction of form: Insights from the visual world paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(11), 2153.

Kukona, A. (2023). Predictive sentence processing at speed: Evidence from online mouse cursor tracking. *Cognitive Science*, *47*(4), e13285.

Kukona, A. (2025). Lexical influences on predictive mouse cursor movements. *Language and Cognition*, *17*, e2.

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

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*(5947), 161–163.

Lenth, R. V. (2023). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (R package version 1.8.8). Retrieved from https://CRAN.R-project.org/package=emmeans

León-Cabrera, P., Rodríguez-Fornells, A., & Morís, J. (2017). Electrophysiological correlates of semantic anticipation during speech comprehension. *Neuropsychologia*, *99*, 326–334.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177.

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.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Li, X., & Qu, Q. (2024). Verbal working memory capacity modulates semantic and phonological prediction in spoken comprehension. *Psychonomic Bulletin & Review*, *31*(1), 249–258.

Loy, J. E., Rohde, H., & Corley, M. (2017). Effects of disfluency in online interpretation of deception. *Cognitive Science*, *41*, 1434–1456.

Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, *88*, 22–60.

Mantegna, F., Hintz, F., Ostarek, M., Alday, P. M., & Huettig, F. (2019). Distinguishing integration and prediction accounts of ERP N400 modulations in language processing through experimental design. *Neuropsychologia*, *134*, 107199.

Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews*, *96*, 367–400.

Nieuwland, M. S., Barr, D. J., Bartolozzi, F., Busch-Moreno, S., Darley, E., Donaldson, D. I., Ferguson, H. J., Fu, X., Heyselaar, E., Huettig, F., Husband, E. M., Ito, A., Kazanina, N., Kogan, V., Kohút, Z., Kulakova, E., Mézière, D., Politzer-Ahles, S., Rousselet, G., … Von Grebmer Zu Wolfsthurn, S. (2020). Dissociable effects of prediction and integration during language comprehension: Evidence from a large-scale study using brain potentials. *Philosophical Transactions of the Royal Society B*, *375*(1791), 20180522.[CrossRef]

Nieuwland, M. S., Politzer-Ahles, S., Heyselaar, E., Segaert, K., Darley, E., Kazanina, N., Von Grebmer Zu Wolfsthurn, S., Bartolozzi, F., Kogan, V., Ito, A., Mézière, D., Barr, D. J., Rousselet, G. A., Ferguson, H. J., Busch-Moreno, S., Fu, X., Tuomainen, J., Kulakova, E., Husband, E. M., … Huettig, F. (2018). Large-scale replication study reveals a limit on probabilistic prediction in language comprehension. *eLife*, *7*, e33468.

Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, *144*(10), 1002.

Pulvermüller, F., & Grisoni, L. (2020). Semantic prediction in brain and mind. *Trends in Cognitive Sciences*, *24*(10), 781–784.

R Core Team. (2023). *R: A language and environment for statistical computing* (Version 4.3.0). R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Scherbaum, S., Dshemuchadse, M., Fischer, R., & Goschke, T. (2010). How decisions evolve: The temporal dynamics of action selection. *Cognition*, *115*(3), 407–416.

Scherbaum, S., Gottschalk, C., Dshemuchadse, M., & Fischer, R. (2015). Action dynamics in multitasking: The impact of additional task factors on the execution of the prioritized motor movement. *Frontiers in Psychology*, *6*, 934.

Schoemann, M., O'Hora, D., Dale, R., & Scherbaum, S. (2021). Using mouse cursor tracking to investigate online cognition: Preserving methodological ingenuity while moving toward reproducible science. *Psychonomic Bulletin & Review*, *28*(3), 766–787

Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2023). *afex: Analysis of Factorial Experiments* (R package version 1.3-0). Retrieved from https://CRAN.R-project.org/package=afex

Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, *32*(25), 8443–8453.

Spivey, M. J., & Dale, R. (2006). Continuous dynamics in real-time cognition. *Current Directions in Psychological Science*, *15*(5), 207–211.

Spivey, M. J., Grosjean, M., & Knoblich, G. (2005). Continuous attraction toward phonological competitors. *Proceedings of the National Academy of Sciences*, *102*(29), 10393–10398.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, *30*, 415–433.

Van Berkum, J. J., 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.

Vespignani, F., Canal, P., Molinaro, N., Fonda, S., & Cacciari, C. (2010). Predictive mechanisms in idiom comprehension. *Journal of Cognitive Neuroscience*, *22*(8), 1682–1700.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Wang, L., Kuperberg, G., & Jensen, O. (2018). Specific lexico-semantic predictions are associated with unique spatial and temporal patterns of neural activity. *eLife*, *7*, e39061.

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.

Wei, W., Huang, Z., Feng, C., & Qu, Q. (2023). Predicting phonological information in language comprehension: Evidence from ERP representational similarity analysis and Chinese idioms. *Cerebral Cortex*, *33*(15), 9367–9375.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272–1288.

Wirth, R., Foerster, A., Kunde, W., & Pfister, R. (2020). Design choices: Empirical recommendations for designing two-dimensional finger-tracking experiments. *Behavior Research Methods*, *52*, 2394–2416.

Wulff, D. U., Kieslich, P. J., Henninger, F., Haslbeck, J. M. B., & Schulte-Mecklenbeck, M. (2023). Movement tracking of psychological processes: A tutorial using mousetrap. *PsyArXiv*. https://doi.org/10.31234/osf.io/v685r

Xu, X., Ji, C., Li, T., & Pickering, M. J. (2024). The prediction of segmental and tonal information in Mandarin Chinese: An eye-tracking investigation. *Language, Cognition and Neuroscience*, *40*(1), 56–70.

Ye, W., & Damian, M. F. (2022). Exploring task switch costs in a color-shape decision task via a mouse tracking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *48*(1), 8–20.

Zhao, Z., Ding, J., Wang, J., Chen, Y., & Li, X. (2024). The flexibility and representational nature of phonological prediction in listening comprehension: Evidence from the visual world paradigm. *Language and Cognition*, *16*(2), 481–504.

15516709, 2025, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.70054 by Institute Of Psychology, Wiley Online Library on [22/05/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License