Heliyon 10 (2024) e39256

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

The effects of prediction representations on implicit learning: Evidence from sentence reading and perceptual identification

Xuliu Ren, Xin Xin *, Xiaorong Gao, Guiqin Ren **

College of Psychology, Liaoning Normal University, Dalian, 116029, China

A R T I C L E I N F O

Keywords:
Lexical prediction
Prediction error
Implicit learning
Implicit memory
Suppression
Inhibition

A B S T R A C T

Predicting errors can facilitate implicit learning, but the long-term consequences of prediction errors are not yet fully understood. Especially when predictions are disconfirmed, it remains unclear whether initially correct prediction representations persist or are suppressed. In this study, participants first engaged in a sentence reading task and then performed a perceptual identification task after completing an N-back task or after a 24-h delay. The perceptual identification task presented previously expected and unexpected words and previously predicted but not presented words to measure implicit memory for the critical items. This study aims to investigate the mechanisms underlying the persistence of prediction representations and the long-term effects of prediction errors on implicit learning. Our results indicate that prediction errors can promote implicit learning and can persist for more than 24 h. Furthermore, originally correct but not seen in reality prediction representations persist to facilitate performance on the implicit memory task after 24 h. This may reflect long-term changes in the internal representation probabilities of prediction representations.

1. Introduction

Predictions and the evaluation of those predictions are fundamental computations of the brain that support perception and action by continually attempting to match incoming sensory input with top-down expectations or predictions [1]. In language comprehension, prediction is particularly important as it helps comprehenders anticipate the likely occurrence of certain words, thereby facilitating lexical processing [2]. According to reinforcement learning and predictive coding theories, confirmed expectations may facilitate stimulus processing, with learners reinforcing their internal representations and thus processing predicted representations more efficiently. A vast quantity of literature demonstrates that either through behavioral response times [3,4], eye-tracking measurements [5,6], or event-related potentials techniques [7–9], the processing of predictable words in context is facilitated. Meanwhile, prediction errors play a crucial role in language processing. When a prediction is violated, the resulting prediction error may lead to increased attention, enhanced memory encoding, and deeper semantic processing, which ultimately contributes to a fuller understanding of the language input [10,11].

However, it is not entirely clear whether there are long-term effects of prediction errors and prediction representations when predictions are incorrect. Some researchers have suggested that prediction errors may have long-term effects on memory representations and that this mismatch between predicted and actual input allows learners to consider new input and possibly update internal

* Corresponding author.
** Corresponding author.
E-mail addresses: xinxin2695@126.com (X. Xin), renguiqin@126.com (G. Ren).

https://doi.org/10.1016/j.heliyon.2024.e39256
Received 14 December 2023; Received in revised form 14 September 2024; Accepted 10 October 2024
Available online 11 October 2024

representations, i.e., a process of continuous adaptation, and that the resulting error signals are thought to facilitate language processing and drive language learning [12–14]. Whereas this prediction error-driven learning is subsequently considered a form of implicit learning, the claim is justified by neural network modeling [15–17]. Most importantly, this implicit learning can be measured by implicit memory [18]. On the other hand, recent research has shown that when predictions are not confirmed, predictive representations (i.e., words that match the prediction but are not actually seen) may remain in memory with increased accessibility and elicit false memories even when words that previously matched the prediction are not present [19–21]. However, the mechanisms underlying the maintenance of predictive representations remain unclear. Therefore, the primary objective of this study is to examine whether prediction errors induced by unexpected sentence-final words during the encoding process enhance implicit learning in readers and have long-term effects. The secondary objective is to further investigate the mechanisms by which predictive representations of lure words, which align with predictions but are not actually encountered, are sustained in the absence of confirmation.

1.1. The mechanisms of prediction in language processing

Before exploring the facilitating effects of prediction on language processing and memory, the mechanisms of prediction inevitably need to be discussed. Most current researchers distinguish between two prediction mechanisms, predictions related to the association system and predictions related to the language production system [22,23]. The former is of the simple, automatic type, associated with mechanisms such as activation spread, is non-targeted and short-term, and relies on the prediction of associations between different elements in a given context. It involves predicting upcoming words or events based on the activation of related concepts or information stored in memory. The latter is the more complex, resource-intensive type that may be associated with language production, a prediction mechanism that relies on language knowledge and rules to predict upcoming words or events by generating possible language expressions. This prediction mechanism is closely related to an individual's ability to produce language and involves understanding and applying linguistic structures and rules. Pickering et al. categorize prediction into prediction-by-association and prediction-by-production, where prediction-by-association is driven by underlying associative mechanisms (e.g., activation spreading) [23]. Prediction is likely to be involved in the spreading of activation between related representations, such as semantic/associative priming or phonological priming (e.g. king spreads to queen), and activation spreads quickly to representations associated with the target word in long-term memory. For instance, in the sentence "The boy went out to the park to fly a kite," the word "kite" following the article "a" is highly predictable. When individuals predict the word "kite," they must activate certain components of their linguistic representation, such as the sound/k/or conceptual features related to being flyable. Successful prediction allows comprehenders to pre-activate the representations that will be used when they encounter the predicted input. This pre-activation enables them to engage in anticipatory processing, which explains how prediction can enhance comprehension [23]. In addition, since this type of activation also decays very rapidly [24], even if the speaker eventually produces the relevant word, it may appear too late for the prediction to be useful. Prediction-by-association can explain the facilitation of prediction for language comprehension, and the fact that the prediction (or pre-activation) perspective is usually associated with the lexical access perspective (i.e., facilitation of lexical access due to pre-activation of predicted words in memory) rather than with the integration perspective helps to explain the facilitation effect of the extraction of the predicted word in memory. However, it is important to recognize that this predictive mechanism has limitations in supporting language comprehension because it is non-targeted and transient, and, according to this mechanism, the prediction itself cannot be sustained for an extended period of time when the prediction is not confirmed.

Prediction-by-production is considered by some researchers to be the central mechanism of prediction [15,16,23], Prediction-by-production goes beyond simple activation spreads and posits a connection between language production and prediction. This theory argues that the same predictive mechanisms operate in both language comprehension and production, and involve long-term changes in internal representations. In the P-chain framework proposed by Dell and Chang, language acquisition and subsequent language knowledge adjustments are both driven by predictive behaviors [16]. These predictive behaviors occur during the gradual processing of language within a context, and predictions are considered to be generated by the language production system. Although prediction-by-production is often more accurate, they are not always available, especially in cognitively constrained populations. Multiple lines of research provide evidence for the shared mechanisms between prediction and language production [25,26]. In a study distinguishing prediction and integration, Federmeier suggests that prediction during language comprehension occurs predominantly in the left hemisphere while speech production is strongly left hemisphere lateralized [27]. If prediction is a form of language production, then this can explain the lateralization of prediction in the brain [7].

Based on the above two prediction mechanisms, first, we focus on whether inducing a prediction error when the prediction is incorrect strengthens or weakens the memory for the word. Currently, there is some inconsistency in the results of studies on prediction errors and the long-term effects of prediction on memory. Prediction error-driven plasticity is a fundamental feature of the brain. This plasticity mechanism supports both memory encoding (storage) and perception (memory retrieval) within the same region, arguing that the discrepancy between prediction and the forward transmission of sensory evidence - i.e., "prediction error" - -drives memory encoding and retrieval [11]. Some studies have also indicated that words that do not match the predictions are remembered better than words that do match the predictions in recollection memory tests, this coincides with novelty-driven memory and supports the claim that prediction errors enhance memory [28,29]. Highly predictable sentences may cause readers to encode sentences in a "top-down validation mode," which may come at the expense of thorough processing, leading to the formation of less obvious memory traces [29]. In contrast, however, some researchers have found that expected words show a memory advantage over unexpected words on recognition tests and that expected words are remembered more successfully on subsequent memory tasks [19,30], which contradicts the notion that prediction errors promote memory.

1.2. Error-driven implicit learning

However, the majority of the aforementioned studies have primarily examined the influence of prediction errors on explicit memory, while fewer have explored the implications for implicit memory. Implicit memory is defined as the influence of prior exposure without (or independent of) explicit memory [31–33]. The experimental behavioral evidence for learning based on errors initially came primarily from studies of structural priming (syntactic priming) and adaptation. Structural priming refers to the phenomenon where language users are more likely to use a similar structure in language production after encountering a specific syntactic structure compared to before encountering it [34]. This process was initially observed in children as young as 3 years old and maintained throughout the learning process [35,36]. The priming effect can lead to lasting adaptations over weeks or months, representing a long-term structural change [37,38]. The reason why this priming effect is considered a form of error-based implicit learning is because it goes beyond mere temporary activation of representations and actually alters the language system of the comprehender. This alteration is akin to the modulation of connection weights observed in connectionist models during the process of representational learning [39,40]. The effect of priming can persist for a sustained period of time and is not weakened by the passage of time or unrelated sentences [41]. This type of learning is considered implicit because it can occur in patients with amnesia who lack explicit memory [42].

It should be noted that several studies have highlighted the potential relationship between prediction errors and implicit memory processing. Event-related potential (ERP) studies have demonstrated that unexpected final words, counter to the predictions made, elicit pronounced N400 amplitudes during sentence comprehension. These findings suggest that prediction errors may also impact implicit memory processes, further emphasizing the need to investigate the role of prediction errors across different memory systems [43], whereas the decrease in N400 amplitude is primarily associated with implicit priming processes, as it occurs relatively independently of memory encoding depth [44] or recognition memory accuracy [21]. The interaction between word predictability and repetition priming has offered new perspectives on the possible relationship between N400 amplitude and implicit learning. In the context of repetition priming, facilitated processing of stimuli is a characteristic of the paradigm, and this advantage in processing due to prior experience is thought to reflect implicit memory. In these studies, incongruent sentences elicit a larger N400 amplitude upon first presentation compared to congruent sentences. However, when sentences are repeated after a delay, there is also a greater reduction in N400 amplitude [45]. Prediction errors and priming have also been used to explain longer-term implicit learning in language acquisition [46]. Hodapp and Rabovsky argue that the N400 reflects error-based implicit learning signals during language comprehension and that the processes that cause changes in the N400 also simultaneously drive implicit learning [18,47]. Therefore, compared to expected words, unexpected words that elicit larger N400 amplitudes result in greater levels of adaptation [18]. Adaptation refers here to a gradual and continuous process of implicit learning that updates the model's internal probability estimates based on specific (incompletely predicted) inputs to make better predictions in the future. The interaction of expectation and repetition is consistent with the notion that larger N400 amplitudes result in increased adaptability, and this has been successfully simulated in the sentence Gestalt model [17,18].

Therefore, an increase in N400 amplitude due to prediction error corresponds to greater adaptation, which can be measured through implicit memory [18,48]. This enhanced adaptability is associated with increased representation probability, which refers to the likelihood of a specific concept or information being robustly represented in the brain, resulting in heightened salience during cognitive processing. Consequently, in subsequent tasks, the corresponding word is more accessible (with shorter response times) [17,18,49].

Based on the aforementioned hypothesis, Hodapp and Rabovsky conducted a study in which participants performed a perceptual identification task measuring implicit memory after a sentence reading task [18]. In the sentence reading task, unexpected sentence-final words elicited larger N400 amplitudes compared to expected sentence-final words, indicating greater adaptation following prediction errors. Moreover, in the subsequent perceptual identification task, previously unexpected words were recognized more quickly than new words and previously expected words. Notably, previously expected words do not show an identification advantage over new words. These findings support the hypothesis that prediction errors alter the probability of internal word representations, leading to greater adaptation and facilitating implicit learning. However, the sentence presentation in the above study was presented word-by-word, and the interval between the two tasks was relatively short (around 10 min for an N-back task), which may have captured greater attention during encoding and resulted in the retention of non-matching words. In the study by Hubbard et al. participants first performed a sentence reading task, followed by a few minutes of a simple distractor task before a recognition test [20]. The results show that unexpected words exhibited greater accuracy. However, in a further study by Hoeltje and Mecklinger, the recognition test was administered 24 h later [19]. The results show that expected words performed better than unexpected ones, which contradicted the previous study. These studies indicate that the interval between learning and testing is an essential factor that may influence the results. Therefore, the first aim of our study is to explore the duration of implicit learning facilitated by prediction errors elicited by unexpected sentence-final words.

1.3. Disconfirmed prediction

Besides investigating the impact of prediction errors on implicit learning, the second objective of this study was to explore the influence of lure words on implicit memory when predictions are disconfirmed. Earlier research has suggested that, in situations where a prediction error occurs, the predicted word tends to be actively suppressed or repressed in memory [27,50,51]. Related studies in ERP have shown that word repetition reduces the amplitude of the N400 [52,53]. However, an ERP study conducted by Rommers et al. on the phenomenon of repetition priming revealed that, similar to repeated words, words that were previously predicted but not actually presented also elicit a reduction in the N400 amplitude [21]. This "pseudo-repetition" effect demonstrates that despite being disconfirmed, the originally correctly predicted representations persist at certain stages of information processing. Hubbard et al. found that lure words also produced greater false alarm rates in an explicit recognition task following a sentence reading task, suggesting that unobserved words are not completely suppressed but remain in memory and can have some effect on episodic memory [20]. Hoeltje and Mecklinger further demonstrated that lure words still show higher false alarm rates after a 24-h interval, demonstrating that activation of predictive representations may be maintained for longer periods even if predictions are not confirmed [19]. However, Gibson's influential theory on Syntactic Prediction Locality in the domain of linguistic complexity posits that, "the longer a predicted category must be kept in memory before the prediction is satisfied, the greater is the cost for maintaining that prediction" [54]. It appears improbable that the activation of such predictive representations is actively maintained for hours. Certain researchers have suggested that this may be related to the pre-updating of sentence context with predicted words in working memory, meaning that the brain activates predicted words in advance and stores them in working memory before they appear [51,55]. The pre-updating process is believed to enhance language processing by reducing the cognitive load required for processing new information. By pre-activating predicted words in working memory, the brain can more efficiently process actual words when encountered, resulting in faster and more accurate sentence comprehension. However, if pre-updated sentence representations are not thoroughly revised after predicted words are disconfirmed, predicted but not actual presented words may persist in memory for a longer duration [19]. However, the above explanation lacks consideration for the possibility that accurate predictive representations may not be solely or simply enhanced through activation. As mentioned earlier, prediction-by-association decays rapidly, and the process that leads to N400 changes may also facilitate implicit learning [17,56]. Thus, the increased accessibility of lure words in memory may be due not only to the sustained activation of predictive representations but also to an increase in their internal probabilistic representations.

1.4. The present study

In this study, two experiments were conducted to investigate the impact of prediction errors and prediction representations on implicit learning, as well as the duration of these effects. Specifically, experiment 1 aimed to determine if prediction errors and correct prediction representations could facilitate implicit learning when prediction is confirmed. Experiment 2 explored whether this facilitation effect would persist after a longer time interval of 24 h. The difference between the two experiments is that Experiment 1 consisted of a 10-min n-back task between the learning and testing phases, whereas Experiment 2 consisted of a 24-h interval. The experiment was divided into two phases: learning and testing. In the learning phase of the experiment, subjects were presented with strongly constrained sentences with either expected or unexpected but reasonable endings. After a while, subjects were asked to perform a perceptual identification task [57], which is considered a reliable measure of implicit memory [58]. In this paradigm, subjects are asked to identify as quickly as possible an increasingly clear word presented on a screen. It could be an ending word that appeared in a previous learning phase, a word that is expected but not actually present, or a completely unrelated new word. This experimental design allowed us to investigate whether words subjected to different sentence processing displayed different speeds on implicit memory identification.

We hypothesize that, on the one hand, if prediction error supports implicit learning, then unexpected ending words under strong contextual constraints should be recognized more quickly than expected ending words or new words, with lower recognition error rates, and maintain a facilitation effect after a 24-h interval. On the other hand, if lure words also affect implicit memory, they should exhibit faster recognition speed compared to new words, with lower recognition error rates, and continue to exhibit a facilitating effect after a 24-h interval.

2. Experiment 1

2.1. Methods

This study was conducted in accordance with the Declaration of Helsinki and approved by the Human Research Ethics Committee of the College of Psychology, Liaoning Normal University (protocol number: LL2023076, date of approval: 30 October 2023).

2.1.1. Participants

In this study, we recruited 38 college students who were native Chinese speakers. Data analysis led to the removal of 2 subjects' data due to high error rates and false trigger rates (the " false trigger " refers to the participants' habitual act of pressing the space bar at the end of a trial, attempting to prematurely advance to the next trial, resulting in this trial being judged as clearly seen, even though in reality, they were still in a fully masked state), leading to a sample size of 36 participants in the final analyses, consisting of 17 males and 19 females. To ensure the statistical power of the experimental design, a priori sample size estimation was conducted. Based on the anticipated medium effect size (Cohen's $f = 0.5$), a significance level ($\alpha = 0.05$), and the desired statistical power ($1 - \beta = 0.95$), calculations were performed using the pwr package in R (R Core Team, 2023) [59,60]. The results indicated that a minimum of 19 participants per group is required. All participants have either normal or corrected-to-normal vision and are exclusively right-handed. None of the participants have a reported history of neurological or mental disorders. Monetary compensation was provided to all participants as remuneration for their participation in the experiment.

2.1.2. Stimulates and procedure

The experimental design was derived from a previous investigation conducted by Hodapp and Rabovsky [18]. It entailed a two-part assessment comprising a sentence-reading task and an implicit memory task. To reduce the possibility of confounding factors arising from explicit memory, an interposed n-back task was integrated between the primary tasks (see Fig. 1).

In the sentence reading task, the sentences consist of 96 strongly constraining sentences and 8 practice sentences. The selection of expected and unexpected words was based on their cloze probabilities, which were determined through a separate cloze norming procedure. These probabilities reflect the proportion of participants who successfully completed sentences with those words in an offline cloze test. The length of the sentences and the frequency of the target words were strictly matched under each condition (detailed information about the experimental materials can be found in Table 2). In the expected and unexpected conditions, the critical words in the test phase match the sentence-ending words from the reading phase. In contrast, for the lure and new word conditions, the test phase's critical words differ from the reading phase's endings. The lure words reflect the predicted words that were expected but not presented, while the new words are unrelated to the reading task, matched only by lexical properties (see Table 1 for examples of experimental materials).[1] In each sentence, each word only appears once, and words used as sentence endings in one condition will not be used as possible sentence endings in other conditions. To further avoid the influence of critical words repeating in other parts of the sentence stem, among all 96 sentences, only one critical word in the lure condition and one critical word in the new word condition had appeared in the stem of other trials before. The other 94 critical words had not appeared during the sentence reading phase. In the formal experimental sentences, half of the sentence endings are expected, and the other half are unexpected but plausible. Specifically, out of the 96 sentences, 24 predictable sentences correspond to the "previously expected" condition in the implicit memory task, 24 predictable sentences correspond to the "new words" condition, and 24 unpredictable sentences correspond to the "previously unexpected" condition in the implicit memory task, while the remaining 24 unpredictable sentences correspond to "lure words." To avoid the influence of associative priming, the final words in the sentences associated with the lure words were selected to be semantically unrelated to the corresponding lure words. The 96 sentences were divided equally into two blocks, and each block had the same number of sentences in each of the four conditions and was pseudo-randomized. Sentences in the same condition will not appear consecutively for 3 times. To avoid primacy effects and recency effects, the first and the last sentences of the experiments were those corresponding to the new words.

In our experiment, we adopted a non – RSVP (Rapid Serial Visual Presentation) narrative paradigm for sentence presentation, using the format of "sentence frame" + "target word." In addition to providing a more natural reading experience, the greatest advantage of this method of presenting sentences is that it allows participants to have a stronger expectation for the upcoming final word, and to enhance the memory encoding of critical words through this enhanced prediction. In our study, we are more concerned with the fate of predictive representations that are not actually seen, so we have adopted this sentence presentation method that can enhance prediction [19]. Moreover, this method of presenting sentences allows participants to receive more complete semantic information at one time, promoting deep semantic processing. Deep semantic processing of the overall sentence semantics can better activate and utilize existing semantic patterns, thereby enhancing memory encoding. In the RSVP paradigm, due to the rapid presentation of words one by one, readers may not have enough time to form a complete prediction of the sentence. The non-RSVP paradigm allows readers to consider the context of the entire sentence while reading each word, better facilitating the integration of contextual information, which is crucial for understanding sentence meaning and forming stable memory representations. Sentences were presented in the form of "sentence frame" + "target word", and each experiment started with a 500 ms fixation point, followed by a 3000 ms sentence frame and a 500 ms blank screen, followed by a 1500 ms target word, and finally a 500 ms blank screen. Participants are instructed to carefully read the sentences during this phase and make no other responses. To prevent the influence of explicit memory factors such as mnemonic strategies, we clearly instructed the participants not to deliberately memorize any words, but simply to understand the meaning of the sentences.

Following the sentence task, participants completed a four-block n-back task (increasing from $n = 1$ to $n = 4$) and used numbers to reduce any potential explicit memory effects of sentence reading. The numbers are displayed in the center of the screen for 1000 ms with a stimulus interval of 2000 ms. The task lasts approximately 10 min. In the n-back task, participants are required to respond when the stimulus matches the one that appeared n items ago. We do not analyze the accuracy of the n-back task; it is only necessary to ensure that participants make a genuine effort to solve the task.

The N-back task is followed by a perceptual identification task, which is used to measure implicit memory. Implicit memory refers to the influence of prior experiences on current task performance without (or independent of) explicit memory [31–33]. This type of memory typically does not come with conscious recollection but is reflected in the improvement of task performance, such as reduced reaction time in perceptual recognition tasks. Perceptual identification task were used to assess implicit memory because these tasks rely on the concept of perceptual fluency, which is the automated influence of prior experiences on the perception and processing of stimuli, a typical manifestation of implicit memory [18]. Following the work of Stark and McClelland [57], the repetition priming effect (i.e., items previously perceived becoming easier to recognize in subsequent perception) is a reliable indicator of implicit memory. Secondly, the gradual clarification of stimuli in the task simulates the perceptual process in the real world, allowing us to measure participants' reaction times at different levels of perceptual clarity, thereby assessing the impact of implicit memory. Unlike explicit memory tasks, this is an unconscious process that does not require participants to consciously recall previously encountered words but rather relies on the automated advantages of the perceptual process. The target word may be the ending word (expected or unexpected words) from the previous sentence reading task or new words that have not been seen before (lure and new words). The

[1] We recruited 15 psychology experts to score the plausibility of presentation conditions during the sentence reading task on a five-point scale. The scores for words that were expected ranged from 4.58 to 5.00 with a mean of 4.92 and a standard deviation of 0.12. The scores for unexpected words ranged from 2.48 to 4.00 with a mean of 3.08 and a standard deviation of 0.44.

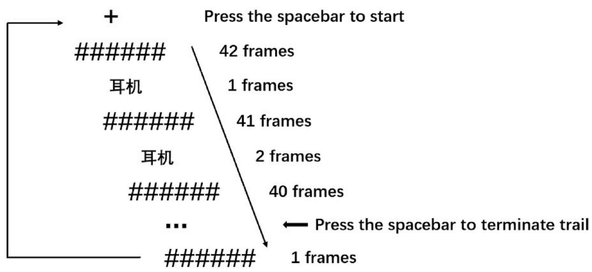

Fig. 1. The design of the perceptual identification task was adapted from Hodapp and Rabovsky's research [18]. In this task, the word gradually becomes clearer over time until the participant presses the spacebar. They are then prompted to enter the word they have recognized. This design enables the capture of continuous reaction time data, which has been demonstrated to be influenced by implicit memory.

Table 1
Example sentences and their approximate translations into English.

Condition	Sentence frame	Ending word	Critical word
Expected	在图书馆听音乐需要插上() Listening to music in the library requires you to plug in your ()	耳机 headphones	耳机 headphones
Unexpected	玛丽今晚睡觉枕着新买的() Mary sleeps tonight on the newly purchased ()	网球 tennis	网球 tennis
Lure	健身房教练有着发达的() The gym instructor has well developed ()	头脑 mind	肌肉 muscle
New	精神分裂患者丽莎拥有两个独立的() The schizophrenic Lisa has two separate ()	人格 personalities	法庭 court

Note. To minimize the effect of word frequency, the distributions of word frequencies were strictly matched. In the above example, "headphones," "tennis," "mind," and "muscle". The word frequencies of "personality" and "court" in Chinese are 17.81, 17.80, 20.49, 20.13, 17.11, and 20.02 respectively. The ending word is the end of the sentence in the sentence reading task, and the critical words are the test words in the perceptual identification task.

Table 2
Lexical properties of test words.

Condition	Frequency	Word length	contextual constraint	cloze probability
Expected	25.03(19.72)	12.33(1.99)	0.99(0.018)	0.99(0.018)
Unexpected	25.09(19.89)	12.29(2.14)	0.97(0.037)	0.06(0.022)
Lure	25.12(20.31)	12.50(2.20)	0.99(0.004)	0.99(0.004)
New	24.98(19.58)	12.44(2.05)	0.98(0.029)	–

Note. Values represent means across items (Values in parentheses are standard errors). There was no significant difference between each condition. The word frequency data comes from Sun's research conducted in 2018 [61]. The contextual constraint here is derived from the best completion probability of the sentence in a cloze test.

stimulus presentation order is pseudorandomized but maintains the overall sequence (compared to the sentence reading task) to control for temporal effects. Participants started each trial by pressing the space bar. The initial presentation consists of 42 frames of the masking stimulus, featuring six hash marks (######), sufficient to cover the presented word, immediately followed by one frame of the first target word presentation (1 frame = 6.06 ms). This process continued until participants pressed a button indicating that they had recognized the word. With each turn, the masking stimulus duration decreased by one frame, while the target stimulus presentation increased by one frame (see the inset in Fig. 2). From the subjects' point of view, this setup induced the perception of words that flickered on the screen and became clearer with each repetition, allowing the recording of reaction times corresponding to the perceived fluency of the word. They are asked to identify the word as early and as accurately as possible. After pressing the button, a prompt appears on the screen, requesting participants to enter their perceived answers. If the answer was incorrect or the response was too slow, feedback was given on the screen (after 25 repetitions, the word became fully visible, and the trial was discontinued). Subjects were asked to keep their head and hand in the same position as in the first trial, to record their responses, and to analyze only the correct trials.

Due to our experimental design, which places the implicit memory test tasks after all reading tasks have been completed, participants are unlikely to deduce the purpose of the experiment based on the practice effect. Even if participants might gradually become aware during the implicit memory test phase that the words we are testing could be related to the sentences they read earlier, the design of our experiment, where target words are presented as increasingly clear words with rapid flashing, requires them to identify as quickly as possible. This design makes it improbable for them to have the time to engage in conscious recall. After the experiment, we conducted a post-experimental inquiry with all participants. All reported that they did not have the opportunity to utilize conscious

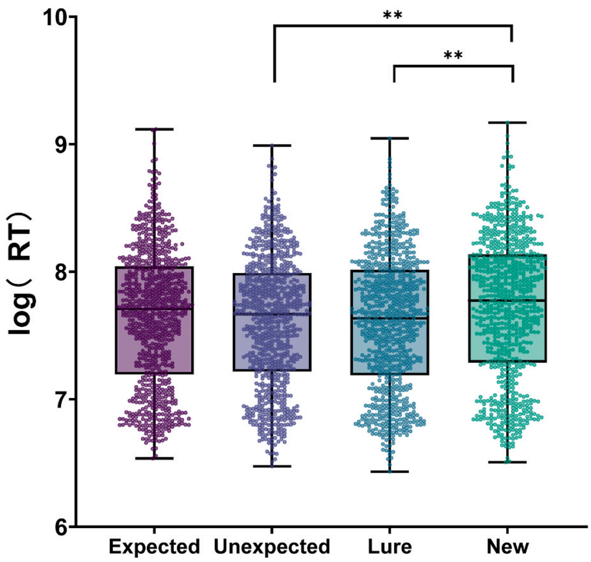

Fig. 2. Reaction times (log-transformed) in the perceptual identification task across participants and items by condition of experiment 1.

recall and retrieval processes. This further supports the effectiveness of our experimental design in minimizing the impact of explicit memory strategies on the implicit memory test outcomes. Therefore, the experiment was divided into three tasks, whereby all stimuli were run through the E-prime software, and the duration of the experiment was approximately 40 min.

2.1.3. Statistical analyses

We conducted analyses using lmertest in R (R Core Team, 2023) to examine the impact of experimental conditions on implicit memory through linear mixed-effects model (LMM) and generalized linear mixed-effects model (GLMM) [59,62]. To account for the non-normal distribution of reaction time data, we applied a logarithmic transformation.

During the model fitting process, we initially attempted to fit the maximal model that included all random effects. If the model could not be successfully fitted, we gradually reduced the complexity of the model. We first removed the random slope effect of the items, followed by the random slope effect of the participants. If the model still could not be successfully fitted, we sequentially removed the random intercept effect of the items and then the random intercept effect of the participants, until the model was successfully fitted. The fully fitted model that was successfully fitted was $logRT \sim Condition + Freq\_z + Length + (1|Participant) + (1|Item)$, and it was ultimately simplified to $logRT \sim Condition + (1|Participant) + (1|Item)$. For the random effects of the model, the $(1|Participant)$ term is used to account for the variation in logRT across different participants, while the $(1|Item)$ term is used to account for the variation in logRT across different items. In the first model we constructed, we used a three-level fixed effects model. We modeled the conditions (expected, unexpected, new words) as fixed factors to differentiate between these three distinct categories, and we used contrast coding to handle categorical variables in order to compare the response patterns of participants under different conditions. The second model used a two-level fixed effects model, modeling the conditions (lure, new words) as fixed factors, and also used contrast coding to handle categorical variables to compare the response patterns of participants under different conditions. The analysis of the model was conducted at the item level, using a likelihood ratio test to determine the significance of the fixed effects by comparing the fit of models with the same random effect structure but without the corresponding fixed effects. And the analysis of simple effects is conducted using the emmeans function [63].

Furthermore, we utilized a generalized linear mixed-effects model to analyze the error rates of reactions across different conditions. The model fitting process mirrored that of the linear mixed-effects model used for reaction time.

2.2. Experiment 1 results

We conducted data processing and filtering to ensure the accuracy and reliability of the data used. Regarding the analysis of reaction time (RT), we initially grouped the data by participants and retained only the RT data that were smaller than the mean RT plus three times the standard deviation. We also excluded any erroneous responses to obtain more accurate and reliable data for subsequent analysis. Simple typographical errors, such as an extra space or period after the correct answer, are not considered mistakes. When reviewing the results, we will correct answers that the program has judged as incorrect but participants have actually identified correctly.

Approximately 5.46 % of the data were excluded, and the remaining 94.44 % of the data were included in the analysis. For the analysis of error rates, since no data is judged to be wrong due to false trigger, we retained all erroneous responses.

2.2.1. Results of RT

The arithmetic mean of the reaction time data (Table 3)indicates that both expected words and unexpected words show a tendency towards faster identification speeds compared to new words. Specifically, compared to new words, the identification speed of unexpected words increased on average by 316 ms, while the identification speed of expected words increased on average by 209 ms. Additionally, lure words show a facilitative effect of 310 ms compared to new words.

Fig. 2 depicts the data after logarithmic transformation. In the LMM analysis of the logarithmically transformed data, compared to new words, the identification speed of unexpected words was significantly faster ($\beta = -0.12$, $SE = 0.04$, $t = -2.71$, $p = 0.0085$),[2,3]. It is worth noting that the difference in identification speed between expected words and new words was only marginally significant ($\beta = -0.08$, $SE = 0.04$, $t = -1.89$, $p = 0.0623$). Additionally, there was no significant difference in identification speed between previously unexpected words and previously expected words ($\beta = 0.04$, $SE = 0.04$, $t = 0.813$, $p = 0.4189$). Crucially, lure words show a significant facilitative effect on identification speed compared to new words ($\beta = -0.11$, $SE = 0.03$, $t = -3.33$, $p = 0.0017$).[4]

2.2.2. Results of error rates

Based on the analysis results presented in Fig. 3, there are no significant differences in error rates across the different conditions for either model, as indicated by the chi-square values and corresponding p-values: Model 1 ($\chi^2 = 1.39$, p = 0.4994) and Model 2 ($\chi^2 = 1.06$, p = 0.3026).

2.3. Experiment 1 discussion

Experiment 1 investigated whether prediction error and prediction representation influence implicit learning when predictions are incorrect. Participants are presented with a set of strongly constrained sentences that conclude with either an expected or unexpected word. They subsequently completed a 10-min n-back task, followed by an implicit perceptual identification task to assess their memory for target words.

According to the results of the reaction time analysis, the identification speed of unexpected words is faster compared to new words, while expected words do not show any advantage in terms of identification speed compared to new words. This finding is consistent with the results reported by Hodapp and Rabovsky [18]. However, we do not observe a significant difference in identification speed between unexpected words and expected words. In other words, the difference in the repetition benefit of words used in sentence reading tasks in the subsequent implicit memory task is not large enough. The marginally significant results for previously expected words imply that we cannot completely rule out whether this benefit is due to the influence of prediction errors or the enhancement of short-term memory caused by lexical repetition. Therefore, in Experiment 2, we consider adjusting the distractor task to minimize the influence of simple lexical repetition and further increase the difference in identification speed between expected words and unexpected words in the subsequent perceptual identification task.

On the other hand, we are also interested in investigating whether the representation of predictions would influence the performance of lure words in the implicit memory task. Previous literature has shown that words predicted during the learning phase but not actually presented lead to higher rates of false alarms in a delayed lexical recognition task, indicating that lure words are more likely to be incorrectly identified as previously seen [20]. Our results show that lure words exhibit faster identification speed in the subsequent implicit memory task, even when both lure words and new words were not actually seen during the sentence reading task. This suggests that when a prediction is disconfirmed, the initially correct prediction continues to persist, even in the absence of actual presentation, and influences subsequent implicit memory tasks. On one hand, this may be due to the sustained activation of prediction representations in the brain. On the other hand, it could also be a result of long-term changes in implicit memory representations, rather than simple short-term activation. The relatively short time interval cannot fully distinguish between these potential reasons. Finally, the analysis of error rates shows no significant differences across conditions, which aligns with the findings of Hodapp and Rabovsky [18].

In summary, the results of Experiment 1 show that both unexpected words and lure words can increase the speed of implicit identification. However, we cannot completely exclude the influence of repetition effects and spreading activation from the results. Furthermore, the duration of the effect of prediction errors on implicit learning remains unclear. Therefore, we conduct a second experiment with the same critical conditions but replacing the 10-min N-back task with an implicit identification task 24 h later. In Experiment 2, we aim to observe differences in identification speed between expected words and unexpected words. We also aim to investigate whether the faster identification of lure words is due to short-term activation or long-term changes in implicit memory representations.

[2] To control for the influence of word plausibility, we employed linear mixed models to fit the data from the expected and unexpected conditions in Experiment 1 and Experiment 2, including plausibility as one of the fixed effects in the models. The results showed that in Experiment 1, the fixed effect of plausibility on reaction time was not statistically significant (β = -0.01, SE = 0.03,t = -0.30, p = 0.7651).

[3] In the first model, the fixed effect of word frequency is not significant ($\beta = 0.03$, $SE = 0.02$, $t = 1.819$, $p = 0.0782$), and the fixed effect of sentence length is also not significant ($\beta = 0.01$, $SE = 0.01$, $t = 1.60$, $p = 0.1145$).

[4] In the second model, the fixed effect of word frequency is also not significant ($\beta = -0.00$, $SE = 0.02$, $t = -0.28$, $p = 0.7835$), and the fixed effect of sentence length is also not significant ($\beta = 0.01$, $SE = 0.01$, $t = 1.19$, $p = 0.2389$).

Table 3
Mean reaction times and error rates in the perceptual identification task per condition of experiment 1.

	Expected	Unexpected	Lure	New
RT in ms	2386(1262)	2279(1155)	2284(1201)	2595(1370)
Error rate in%	4.75(5.19)	4.28(3.92)	4.51(5.30)	5.90(5.75)

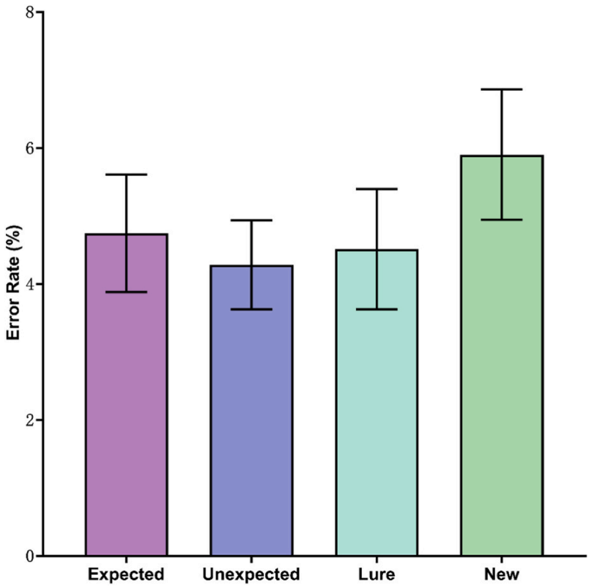

Fig. 3. Error rates in the perceptual identification task across participants and items by condition of experiment 1.

3. Experiment 2

3.1. Methods

3.1.1. Participants

In this study, we recruited 38 college students who were native Chinese speakers. Data analysis led to the removal of 3 subjects' data due to high error rates and false trigger rates, leading to a sample size of 35 participants in the final analyses, consisting of 11 males and 24 females. The sample size calculation for the experiment follows the same process as Experiment 1, which determined that a minimum of 19 participants per group is required. All participants have either normal or corrected-to-normal vision and are exclusively right-handed. None of the participants have a reported history of neurological or mental disorders. Monetary compensation was provided to all participants as remuneration for their participation in the experiment. To ensure active participation from the participants, we made the completion of all tasks a prerequisite for receiving compensation. Furthermore, we informed the participants beforehand that they were required to return promptly after 24 h; failure to do so within the specified time frame would result in not receiving the full compensation. All participants underwent the second experimental session within 24 ± 2 h after the initial reading task to control for the impact of additional experimental variables.

3.1.2. Stimulates and procedure

The procedure of Experiment 2 is similar to that of Experiment 1, with the only difference being that in Experiment 2, participants are required to wait for 24 h after reading the sentences before completing the implicit memory task again. All materials used in Experiment 2 are the same as those used in Experiment 1.

3.1.3. Statistical analyses

The data analysis is the same as experiment 1.

3.2. Experiment 2 results

The principle of data elimination is the same as that of experiment 1. Approximately 4.85 % of the data were excluded, and the remaining 95.15 % of the data were included in the analysis. For the analysis of error rates, responses judged to be incorrect due to false trigger were excluded, with a total of approximately 0.03 % of the total data excluded, ultimately retaining 99.97 % of the data for inclusion in the analysis.

3.2.1. Results of RT

The arithmetic mean of the reaction time data (Table 4)indicates that both expected words and unexpected words still show a tendency towards faster identification speeds compared to new words, but with less significant increases of 184 ms and 151 ms respectively. Additionally, lure words exhibit a significant speed-up effect of 228 ms compared to the new words.

Fig. 4 depicts the data after logarithmic transformation. In the LMM analysis of the logarithmically transformed data, compared to new words, the identification speed of unexpected words was significantly faster (β = -0.08, SE = 0.04,t = -2.13, p = 0.0369),[5,6]. And there was no significant difference in identification speed between previously expected words and new words (β = 0.05, SE = 0.04,t = 1.50, p = 0.1372). The difference between expected words and unexpected words was still not significant (β = 0.02, SE = 0.04,t = 0.62, p = 0.5353). Crucially, lure words still show a significant facilitative effect on identification speed compared to new words (β = -0.08, SE = 0.03,t = -2.85, p = 0.0064).[7]

3.2.2. Results of error rates

According to the analysis result of the error rate (Fig. 5), compared to new words, the error rates of previously expected words were marginally significant. (β = 0.57, SE = 0.32, z = 1.81,p = 0.0703). Previously unexpected words exhibited reduced error rates (β = 0.83, SE = 0.33, z = 2.50,p = 0.0124). Importantly, lure words displayed significantly lower error rates compared to new words (β = 0.77, SE = 0.38, z = 2.04,p = 0.0413).

To further explore the differences in error rates between the two experiments, we conducted a joint analysis of the error rate data from Experiment 1 and Experiment 2 by fitting a generalized linear mixed model. The full model we fitted was *Error ~ Condition * Experiment + (1 | Participant) + (1 | Item)*, simplifying the model step by step to examine the interaction between experimental conditions and experimental versions. The results of the analysis of variance indicated that there was no significant interaction between word type and experimental version (χ^2 = 2.29, p = 0.317), and moreover, the main effect of the experimental version was also not significant (χ^2 = 0.75, p = 0.3854).

3.3. Experiment 2 discussion

Experiment 2 investigated whether prediction error and prediction representation influence implicit learning can be maintained beyond 24 h when predictions are incorrect. Participants are presented with a set of strongly constrained sentences that conclude with either an expected or unexpected word. Then, after a 24-h delay, they underwent an implicit perceptual identification task to assess their memory for the target words.

According to the results of reaction time analysis, previously unexpected words still exhibited significantly faster identification speeds compared to new words after a 24-h interval. This suggests that the facilitation of implicit memory by prediction errors can be maintained for a longer duration. On the other hand, previously expected words did not show any advantage in terms of identification speed. This finding indicates that the repetition effect starts to diminish over longer intervals, and also supports our research hypothesis that correct predictions do not enhance implicit memory. The marginally significant phenomenon observed in Experiment 1 is likely attributed solely to the repetition effect. However, it should be noted that a 24-h interval may not completely eliminate the influence of repetition, as evidenced by the lower reaction times observed for both previously seen words (expected and unexpected) numerically.

On the other hand, experiment 2 replicated the impact of lure words on the implicit perceptual identification task. The results show that even after 24 h, lure words still show significantly faster identification speed than new words. This suggests that when the prediction is confirmed, the correct prediction representation remains in the brain and increases the accessibility of the corresponding word for a longer period, even if it is not actually presented. Since this activation state of predicted but not actually presented words does not seem to be able to be sustained for longer periods of time [19], this may imply that these predictive representations have produced long-lasting changes in the implicit memory network, rather than just simple short-term activation. Indeed, this interpretation is consistent with the research findings of Hoeltje and Mecklinger [19]. In their study, participants may have influenced the implicit memory representation of lure words during the initial sentence reading, which led to their increased familiarity with these words on the recognition test. In addition, explicit memory may have influenced participants' responses because they may have consciously recalled lure words based on their prior expectations or predictions. The interaction between implicit and explicit memory processes leads to higher false alarm rate of lure words.

The analysis of error rates revealed a significantly lower error rate for previously unexpected and lure words compared to new words. In addition, the error rates for previously expected words also showed a decreasing trend. Our initial hypothesis was that the shorter the time interval between the subsequent perceptual recognition task and the prior sentence reading task, the lower the error

[5] The results showed that In Experiment 2, the fixed effect of plausibility on reaction time was not statistically significant(β = -0.02, SE = 0.03,t = -0.82, p = 0.4189). This indicates that we can exclude the impact of word plausibility on the experimental outcomes.

[6] In the first model, the fixed effect of word frequency is not significant (β = 0.02, SE = 0.01, t = 1.74, p = 0.0855), and the fixed effect of sentence length is also not significant (β = 0.01, SE = 0.01, t = 0.96, p = 0.3330).

[7] In the second model, the fixed effect of word frequency is not significant (β = −0.02, SE = 0.01, t = −1.26, p = 0.2146), and the fixed effect of sentence length is also not significant (β = −0.00, SE = 0.01, t = −0.16, p = 0.8732).

Table 4
Mean reaction times and error rates in the perceptual identification task per condition of experiment 2.

	Expected	Unexpected	Lure	New
RT in ms	2479(1164)	2446(1155)	2402(1089)	2630(1197)
Error rate in%	3.69(4.26)	2.98(3.29)	3.34(4.26)	6.31(7.52)

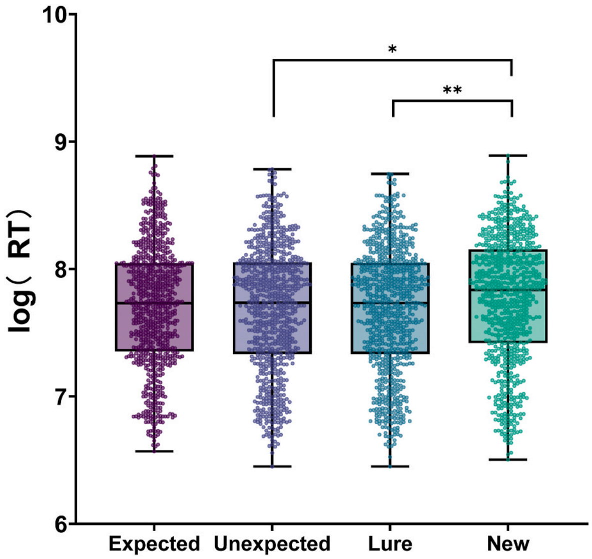

Fig. 4. Reaction times (log-transformed) in the perceptual identification task across participants and items by condition of experiment 2.

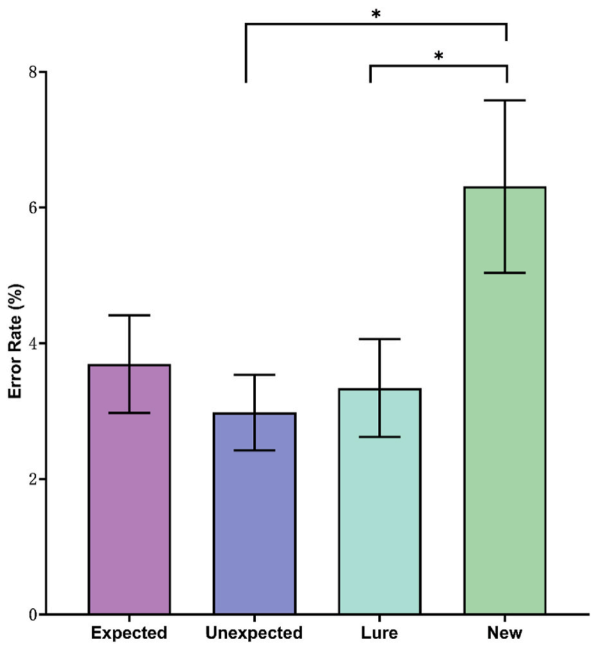

Fig. 5. Error rates in the perceptual identification task across participants and items by condition of experiment 2.

rate of the participants should be. This is because the influence of the previous experimental manipulation should be more pronounced. Over time, for all conditions that have undergone experimental treatment, the error rate under the same condition should increase, while there should be no difference in the error rate for new words. However, the results were the opposite; our study showed that at longer time intervals, the difference between the words treated by the experimental condition and the new words became more significant. A joint analysis of the data from Experiment 1 and Experiment 2 indicated that there was no significant interaction between the experimental condition and the version of the experiment. In one hand, this may be attributed to the longer interval between the sentence reading task and the perceptual identification task in Experiment 2, which resulted in a lower facilitative effect of the experimental conditions on word retrieval. Participants may have exhibited more caution during the perceptual identification task, leading to an overall increase in response time but a lower error rate compared to new words.

On the other hand, this phenomenon might also be credited to the processes of sleep and memory consolidation. Sleep is known to bolster a spectrum of cognitive functions, including decision-making, language, categorization, and memory [64]. Critically, memory consolidation plays an essential role in the enduring retention of both explicit and implicit memories. Prior research has explored the consolidation process in the contexts of implicit learning of phonotactic patterns, context reinstatement, and structural transfer in visual learning [65–67]. In the study by Anderson and Dell, a distinction was made between first-order and second-order patterns in phonotactic learning [65]. It was discovered that first-order patterns significantly influence speech errors within a single experimental session, whereas second-order patterns only affect errors when tested the following day. This suggests that sleep may offer a critical consolidation phase, facilitating the retention and retrieval of memories. Additionally, research by Schechtman and colleagues indicates that memory consolidation during sleep may involve the reinstatement of context, potentially further enhancing the stability and longevity of memories [66]. Notably, sleep, particularly the non-rapid eye movement (NREM) stage, is recognized as the primary period for the consolidation of explicit memories.

Error rates, typically a conscious measure, reflect the explicit aspects of perceptual identification tasks more than they do implicit memory. The benefits derived from a longer delay in memory performance may be more pronounced for explicit memory. In Experiment 2, the extended delay, attributed to sleep, could have permitted a more effective process of memory consolidation. This consolidation is likely particularly advantageous for explicit memory due to its reliance on conscious memory processes. The reactivation of explicit memories during sleep may strengthen memory traces, aiding participants in accurately retrieving words in perceptual identification tasks, even after an extended interval. Conversely, reaction times, which are more closely associated with implicit memory, reflect unconscious cognitive processes that are not typically subject to direct conscious influence. Even if participants exhibited increased caution in explicit tasks, reaction times may not have varied significantly like error rates, as implicit memory retrieval does not necessitate conscious effort.

Thus, the observed results in error rates could be attributed to the 24-h interval between study and test, which may have facilitated the consolidation of representations for unexpected and lure words, an effect not readily evident in reaction time data. It suggests that reaction time may not be a sensitive indicator of the effects of memory consolidation. The consolidation process might require a more extended period to manifest its effects and could be demonstrated through other means, such as a reduction in error rates.

To summarize, the results of Experiment 2 suggest that, with a 24-h time interval, previously unexpected words and lure words still maintain an advantage in identification speed compared to new words. Additionally, we observed differences in error rates among different conditions, with lower error rates for previously expected, previously unexpected, and lure words compared to new words.

4. General discussion

The present study investigated the potential impact of prediction errors during sentence encoding on facilitating implicit learning. We also aim to explore the mechanisms responsible for the maintenance of otherwise accurate prediction representations. To achieve this, a sentence-reading task is used to manipulate the predictability of critical words, followed by a subsequent perceptual identification task to assess implicit memory. Our findings indicate that prediction errors can indeed facilitate implicit learning and can have an enduring effect for at least 24 h. Additionally, our results suggest that the sustenance of prediction representations may be closely related to the process of implicit learning.

4.1. Prediction error and implicit learning

Our research focuses on the potential relationship between prediction errors and implicit learning, a concept initially proposed by Chang et al. and supported by neural network models in the field of cognitive psychology [15–17,48]. The central idea is that when there is a discrepancy between expected and actual inputs, prediction errors can facilitate the process of implicit learning. We hypothesize that words that deviate from predictions can prompt implicit learning, as evidenced by faster identification speed in perceptual identification tasks for previously unexpected words compared to expected words. In line with this hypothesis, the results of Experiment 1 demonstrate that previously unexpected words exhibit faster speeds in perceptual identification tasks, while previously expected words do not display this phenomenon. Surprisingly, we did not find a significant difference in identification speed between these two types of words. Our findings differ from those reported by Hodapp in 2021, whose study indicated that unexpected words were identified more quickly than expected words [18]. We believe this discrepancy may stem from the adjustments we made to the presentation of materials in the sentence reading task. In their research, sentences were presented to participants word-by-word, but in our study, we presented them in a sentence frame plus target word format. This presentation method could lead to a stronger anticipation of the upcoming final word by the participants, and through this enhanced prediction, strengthen the memory encoding of the critical words [19]. In Hoelte's study, final words that were consistent with predictions had a higher recognition accuracy rate compared to those that were not. Since our research focuses more on the fate of predictive representations that were not actually seen, we adopted this sentence presentation method that can enhance prediction [19]. As a result, this may well have prevented a significant differencer between expected reaction times and unexpected reaction times. Additionally, the longer presentation time may reinforce the effects of word repetition, leading to the sustained activation of previously encountered words, as the repetition effect of vocabulary can persist for an extended period [45]. The 10-min N-back task used in our study may not have adequately controlled for possible explicit memory effects.

On the other hand, we are also highly interested in the duration of this facilitation effect. According to previous studies on the N400 repetition effect, this repetition priming effect can last for more than 120 sentences (approximately 45 min) or even longer [45]. Our research predicts that prediction errors can sustain a stimulatory effect on implicit memory over a considerable period, whereas correct predictions do not confer such a benefit. The results of Experiment 2 support our hypothesis, showing that previously unexpected words still demonstrated significantly reduced identification speed compared to new words even after a 24-h interval. By contrast, a similar effect was not observed for previously expected words. It is noteworthy that even though expected words were perceived slightly faster than new words at shorter intervals, this trend significantly dissipated over longer intervals. This may indicate that although word repetition can initially enhance identification speed to a certain extent, this facilitatory effect tends to diminish over time. In contrast, the stimulatory effect of prediction errors on implicit memory appears to endure longer. Moreover, we did not observe a significant difference in identification speed between expected words and unexpected words, suggesting that, even though they had different effects compared to new words, the differences between them were not substantial. One possible reason for this could be that lexical repetition has mitigated the distinction between the two cases. Alternatively, another possible factor contributing to these results could be the high frequency of the target words we used, which are commonly seen in everyday life. The representation probability of these words could be readjusted when participants encounter them again, even multiple times, between the two tasks, potentially affecting the facilitation effect we investigated in our experiment. Therefore, in future experiments, it would be beneficial to shorten the time interval between tasks to further explore the boundary conditions and implement more precise control over the experimental environment.

One alternative theoretical explanation for the facilitation of learning through prediction errors is that prediction errors capture more attention and result in more thorough encoding. In sentence reading tasks, the frequency of unexpected words may be too high, thereby diminishing the unexpectedness or novelty of prediction errors. Reggev et al. investigated word recognition memory in both semantically congruent and incongruent word contexts and conducted experiments by manipulating the proportion of incongruent word pairs [68]. The results revealed that when the proportion of incongruent word pairs decreased and the distinctiveness increased, word memory performance in the incongruent word context promoted. However, memory for congruent words was not affected by the manipulations of proportion. Therefore, the proportion of unpredicted sentences may impact the recognition of previously unexpected words over longer intervals, weakening their facilitatory effect. Taking into consideration the suggestion by Hoeltje, future research should consider controlling the ratio of expected and unexpected words to further investigate its influence on learning outcomes [19].

In Experiment 1, we did not observe any differences in error rates under different conditions. However, in Experiment 2, error rates for previously unexpected words and lure words were significantly reduced compared to new words and previously expected words also exhibited a lower error rate compared to new words. Based on our initial hypothesis, we anticipated that the error rates observed under various conditions would not be expected to diminish further with extended time intervals. One possible reason for the observed phenomenon is that in our experiment, the perceptual identification task was implemented as a word that becomes clearer with flickering. In Experiment 2, the sentence reading task had a longer interval between the perceptual identification task, which resulted in a reduction of the facilitation effect of the experimental conditions on vocabulary. Participants became more cautious when performing the perceptual identification task and were more inclined to wait a little longer before making judgments when faced with blurry words. This ultimately led to an overall increase in response time (slower reaction), but a decrease in error rates (more accurate judgments during response). On the other hand, the observed improvements in error rates may be also attributed to the beneficial effects of sleep on memory consolidation, particularly for explicit memory, which relies on conscious processes. The 24-h interval between study and test sessions could have facilitated the consolidation of representations for unexpected and lure words, an effect that is not as evident in reaction times, which are more closely linked to implicit memory processes. This suggests that while error rates are a sensitive indicator of memory consolidation, reaction times may not be as responsive to the consolidation process, indicating that the latter might require a more extended period or different measures to manifest its full impact.

4.2. The long-term impact of expected but not presented words

A second important aim of the present study was to investigate the representation in memory of words that were predicted but never actually seen in a sentence reading task (i.e., lure words). Our results show that lure words show faster identification speed in a perceptual identification task, both at relatively short intervals and at intervals up to 24 h, suggesting that the increased accessibility of lure words can be sustained for more than one day.

Some researchers have explained this phenomenon as the pre-activation of words that are predicted but not presented. In one explanation for prediction errors, when we encounter a word different from what is predicted, the predicted word form is actively inhibited or suppressed [50]. This idea has been applied to other domains of sentence processing research. For instance, Gernsbacher and Faust proposed that once a possible meaning is selected for integration, irrelevant meanings of homophones are suppressed [69]. According to the strong version of the predictive suppression hypothesis, the processor actively suppresses the activation of highly probable words in all situations. The occurrence of a prediction leads the brain to inhibit input that is inconsistent with the prediction, thereby redirecting attention and resources toward prediction-consistent information [70]. According to this hypothesis, due to the suppression of predicted word activation, participants should not experience facilitation when encountering lure words in the implicit recognition task following the sentence reading task.

In contrast to the inhibition theory, some other theories suggest that the activation of predicted words may persist and thus could influence subsequent processing. Rich and Harris opposed the strong version of the inhibition hypothesis and further differentiated between the active and passive versions of the sustained activation theory [71]. The active version posits that the increased activation of predicted words can be actively maintained in memory, while the passive version suggests that the increased activation of predicted words will gradually decay passively after the prediction is not confirmed. According to this hypothesis, participants should exhibit a significant facilitation effect when encountering lure words in tasks following the sentence reading task. They compared the reading times for words that previously matched predictions after those predictions were negated under different proximity conditions. The experiment included two conditions: a close condition without an additional adverb before the target word and a distant condition with an adverb (e.g., the italicized part in the sentence). The results showed that shorter reading times for previously predicted words were only observed in the close condition, indicating that the increased activation of predictive words can rapidly decay during sentence processing. This finding reveals that the persistence of activation may not stem from a continuously active maintenance, but rather resembles a natural, passive process of gradual decay. The experimental results indicate that when the expectation of predicted words is not fulfilled, their state of activation in memory does not last long, but diminishes rapidly as the process of sentence comprehension unfolds. This scenario is more inclined to support the passive decay model within the sustained activation theory. The results from Hoeltje stand in contrast to these findings [19]. They first had participants engage in a sentence reading task, followed by a recognition test for critical words the next day. The results showed that lure words had a higher false alarm rate compared to new words, indicating that the lingering activation of expected but unpresented words can influence memory decisions even after a retention interval as long as 24 h. This supports the active version of the sustained activation theory. However, it seems unlikely that the high activation of predicted but unpresented words could be actively maintained for several hours. They suggest that this phenomenon might be due to the 1-s delay between the sentence context and the target word, which facilitated the prediction of the target word and the pre-updating of the sentence representation. If a pre-updated sentence representation is formed within the context of the sentence and the actual word presented does not match the expectation, this pre-updated representation may not be thoroughly revised, allowing the predicted but unpresented words to persist in memory without active maintenance. This persistence could explain why these words are still associated with high-frequency false memory decisions the day after the initial sentence processing [19,21].

Our experimental results provide another direction to answer the above question, suggesting that the long-term effects of predicted but unseen words on memory may not solely rely on sustained activation or active maintenance of relevant representations, but also involve persistent changes in the internal representation probability of the associated words. As mentioned earlier, according to N400-related research, larger N400 amplitudes are believed to reflect an implicit learning signal that leads to increased adaptation, which involves higher representation probability and thus better accessibility of the corresponding words in subsequent tasks [17,18]. Specifically, when comprehenders encounter an unexpected word, it induces a larger N400 amplitude. If the unexpectedness of the word is due to its lower internal frequency representation, the estimate will be updated (corrected) to decrease the comprehender's surprise level when encountering the word again. This necessary updating process is thought to be based on the error reflected by the N400. This explains the results of the classic repetition paradigm, where after model adaptation, the prediction error (i.e., N400) decreases when the word is presented for the second time, as the model now has adjusted internal estimates [18]. Meanwhile, a study by Rommers et al. shows that the N400 amplitude for previously expected but unseen words also decreases, indicating a "pseudo-repetition" effect [72]. This suggests that the internal representation probability of previously predicted but unseen words may change before they are actually encountered, resulting in a reduction in N400 amplitude when they are subsequently encountered. Of course, in the short term, we cannot completely rule out the influence of increased word pre-activation. Therefore, in terms of short-term effects, the performance improvement of lure words in related tasks may be attributed to the combined effect of the two mechanisms. On one hand, because the predicted but unseen words remain in a state of increased pre-activation, this may lead to increased fluency in processing the expected lure in testing [19]. On the other hand, it may also be due to the rectification of the representation probability of predicted but unseen words. In terms of long-term effects, the effect of increased word pre-activation diminishes over time, but the correction of its internal representation probability remains intact.

4.3. Limitations of the study

Despite achieving certain results, there are still some limitations in the study that require further exploration and improvement. Firstly, this study failed to control the gradient of interval time, which prevented the investigation of the boundary conditions of how prediction errors facilitate the maintenance of implicit learning over time. Future research can consider exploring the effects of prediction errors on implicit learning by using a more appropriate time range or incorporating other distracting tasks. Secondly, we did not examine the violation of predictions under weak constraint conditions. Future research can further investigate the subsequent impacts of originally correct prediction representations under weaker constraint conditions by considering the degree of sentence constraints. Lastly, while prediction can be a source of learning, we only measured its effects on a specific aspect of memory. Thus, further research is needed to identify which linguistic features can be learned through prediction errors. Additionally, we should explore the relationship between error-driven learning and prediction errors in driving memory structures. Previous studies have suggested that the memory structures driven by prediction errors are unlikely to operate in isolation [73], but the specific mechanisms of this relationship are not yet clear.

5. Conclusion

Our results suggest that a mechanism for the persistence of words that are predicted but are not actually presented may be a long-term change in the probability of the internal representation. And the facilitation of implicit learning by prediction errors can persist for more than 24 h. Incorrect predictions not only promote implicit learning of the observed words but also facilitate implicit learning of the predicted representations that were not actually observed.

CRediT authorship contribution statement

Xuliu Ren: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Xin Xin:** Writing – review & editing, Supervision. **Xiaorong Gao:** Resources. **Guiqin Ren:** Supervision, Funding acquisition.

Ethics statement

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Human Research Ethics Committee of the College of Psychology, Liaoning Normal University (protocol number: LL2023076, date of approval: 30 October 2023). In light of the nature of our research and the specific circumstances, obtaining written informed consent was not feasible or necessary. We have meticulously detailed in our manuscript the reasons for this decision, ensuring that the rationale for using verbal informed consent is clear and justified. Verbal informed consent was obtained from each participant, and this approach was assessed on a case-by-case basis to ensure that participants were fully informed about the study's objectives, procedures, potential risks, benefits, and their rights, including the right to withdraw from the study at any time without repercussions. The study was designed with the utmost respect for the dignity, rights, and welfare of all participants. All data were collected, stored, and analyzed in a manner that safeguarded the privacy and confidentiality of the participants. Our ethical review process ensured compliance with all relevant ethical standards and legal requirements, upholding the principles of transparency and integrity in research.

Data availability statement

In adherence to the principles of open science, the data underpinning this study have been deposited in a publicly accessible repository to facilitate independent verification and further research. The data can be accessed via the Open Science Framework (OSF) at the following link: https://osf.io/mtxek/, which was last accessed on December 5, 2023. We believe that sharing research data is essential for building trust in scientific findings and for enabling the broader research community to build upon our work. Our commitment to data sharing is reflected in this practice, and the rationale for making our data publicly available, as well as any potential limitations on access, will be detailed in the published article to ensure transparency and align with best practices in scholarly communication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:G.Ren reports financial support was provided by the National Natural Science Foundation of China (31471075). G. Ren reports financial support was provided by the Scientific Research Program of the Educational Department of the Liaoning Province of China (LJKMZ20221422). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Clark, Whatever next? Predictive brains, situated agents, and the future of cognitive science, Behav. Brain Sci. 36 (3) (2013) 181–204, https://doi.org/10.1017/s0140525x12000477.
[2] K.A. DeLong, T.P. Urbach, M. Kutas, Probabilistic word pre-activation during language comprehension inferred from electrical brain activity, Nat. Neurosci. 8 (8) (2005) 1117–1121, https://doi.org/10.1038/nn1504.
[3] P.J. Schwanenflugel, E.J. Shoben, The influence of sentence constraint on the scope of facilitation for upcoming words, J. Mem. Lang. 24 (2) (1985) 232–252, https://doi.org/10.1016/0749-596x(85)90026-9.
[4] A. Staub, Response time distributional evidence for distinct varieties of number attraction, Cognition 114 (3) (2010) 447–454, https://doi.org/10.1016/j.cognition.2009.11.003.
[5] G.T.M. Altmann, Y. Kamide, Incremental interpretation at verbs: restricting the domain of subsequent reference, Cognition 73 (3) (1999) 247–264, https://doi.org/10.1016/s0010-0277(99)00059-1.
[6] Y. Kamide, G.T.M. Altmann, S.L. Haywood, The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements, J. Mem. Lang. 49 (1) (2003) 133–156, https://doi.org/10.1016/s0749-596x(03)00023-8.
[7] K.D. Federmeier, M. Kutas, A rose by any other name: long-term memory structure and sentence processing, J. Mem. Lang. 41 (4) (1999) 469–495, https://doi.org/10.1006/jmla.1999.2660.
[8] J.J.A. Van Berkum, C.M. Brown, P. Zwitserlood, V. Kooijman, P. Hagoort, Anticipating upcoming words in discourse: evidence from ERPs and reading times, J. Exp. Psychol. Learn. Mem. Cognit. 31 (3) (2005) 443–467, https://doi.org/10.1037/0278-7393.31.3.443.
[9] E.W. Wlotko, K.D. Federmeier, So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning, Neuroimage 62 (1) (2012) 356–366, https://doi.org/10.1016/j.neuroimage.2012.04.054.
[10] A. Greve, E. Cooper, A. Kaula, M.C. Anderson, R. Henson, Does prediction error drive one-shot declarative learning? J. Mem. Lang. 94 (2017) 149–165, https://doi.org/10.1016/j.jml.2016.11.001.
[11] R.N. Henson, P. Gagnepain, Predictive, interactive multiple memory systems, Hippocampus 20 (11) (2010) 1315–1326, https://doi.org/10.1002/hipo.20857.
[12] K.J. Friston, A theory of cortical responses, Phil. Trans. Biol. Sci. 360 (1456) (2005) 815–836, https://doi.org/10.1098/rstb.2005.1622.
[13] M. Ramscar, M. Dye, S.M. McCauley, Error and expectation in language learning: the curious absence of mouses in adult speech, Language 89 (4) (2013) 760–793, https://doi.org/10.1353/lan.2013.0068.

[14] R.P.N. Rao, D.H. Ballard, Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects, Nat. Neurosci. 2 (1) (1999) 79–87, https://doi.org/10.1038/4580.
[15] F. Chang, G.S. Dell, K. Bock, Becoming syntactic, Psychol. Rev. 113 (2) (2006) 234–272, https://doi.org/10.1037/0033-295x.113.2.234.
[16] G.S. Dell, F. Chang, The P-chain: relating sentence production and its disorders to comprehension and acquisition, Phil. Trans. Biol. Sci. 369 (1634) (2014) 20120394, https://doi.org/10.1098/rstb.2012.0394.
[17] M. Rabovsky, S.S. Hansen, J.L. McClelland, Modelling the N400 brain potential as change in a probabilistic representation of meaning, Nat. Human Behav. 2 (9) (2018) 693–705, https://doi.org/10.1038/s41562-018-0406-4.
[18] A. Hodapp, M. Rabovsky, The N400 ERP component reflects an error-based implicit learning signal during language comprehension, European Journal of Neroscience 54 (9) (2021) 7125–7140, https://doi.org/10.1111/ejn.15462.
[19] G. Hoeltje, A. Mecklinger, Benefits and costs of predictive processing: how sentential constraint and word expectedness affect memory formation, Brain Res. 1788 (2022) 147942, https://doi.org/10.1016/j.brainres.2022.147942.
[20] R.J. Hubbard, J. Rommers, C.L. Jacobs, K.D. Federmeier, Downstream behavioral and electrophysiological consequences of word prediction on recognition memory, Front. Hum. Neurosci. 13 (291) (2019), https://doi.org/10.3389/fnhum.2019.00291.
[21] J. Rommers, K.D. Federmeier, Lingering expectations: a pseudo-repetition effect for words previously expected but not presented, Neuroimage 183 (2018) 263–272, https://doi.org/10.1016/j.neuroimage.2018.08.023.
[22] F. Huettig, Four central questions about prediction in language processing, Brain Res. 1626 (2015) 118–135, https://doi.org/10.1016/j.brainres.2015.02.014.
[23] M.J. Pickering, C. Gambi, Predicting while comprehending language: a theory and review, Psychol. Bull. 144 (10) (2018) 1002–1044, https://doi.org/10.1037/bul0000158.
[24] T.P. Mcnamara, Semantic priming: perspectives from memory and word recognition, Conscious. Cognit. (2005).
[25] E. Drake, M. Corley, Articulatory imaging implicates prediction during spoken language comprehension, Mem. Cognit. 43 (8) (2015) 1136–1147, https://doi.org/10.3758/s13421-015-0530-6.
[26] C.D. Martin, F.M. Branzi, M. Bar, Prediction is Production: the missing link between language production and com-prehension, Sci. Rep. 8 (2018) 1079, https://doi.org/10.1038/s41598-018-19499-4.
[27] K.D. Federmeier, Thinking ahead: the role and roots of prediction in language comprehension, Psychophysiology 44 (4) (2007) 491–505, https://doi.org/10.1111/j.1469-8986.2007.00531.x.
[28] M. Corley, L.J. MacGregor, D.I. Donaldson, It's the way that you, er, say it: hesitations in speech affect language comprehension, Cognition 105 (3) (2007) 658–668, https://doi.org/10.1016/j.cognition.2006.10.010.
[29] K.I. Haeuser, J. Kray, Effects of prediction error on episodic memory retrieval: evidence from sentence reading and word recognition, Language Cognition and Neuroscience 38 (4) (2023) 558–574, https://doi.org/10.1080/23273798.2021.1924387.
[30] G. Brod, M. Werkle-Bergner, Y.L. Shing, The influence of prior knowledge on memory: a developmental cognitive neuroscience perspective, Frontiers in Behavioral Neuroscience 7 (2013) 139, https://doi.org/10.3389/fnbeh.2013.00139.
[31] P. Graf, D.L. Schacter, Implicit and explicit memory for new associations in normal and amnesic subjects, J. Exp. Psychol. Learn. Mem. Cognit. 11 (3) (1985) 501–518, https://doi.org/10.1037/0278-7393.11.3.501.
[32] D.L. Schacter, P. Graf, Effects of elaborative processing on implicit and explicit memory for new associations, J. Exp. Psychol. Learn. Mem. Cognit. 12 (1) (1986) 21–29, https://doi.org/10.1037/0278-7393.12.1.21.
[33] D.L. Schacter, P. Graf, Preserved implicit memory in amnesic patients, J. Clin. Exp. Neuropsychol. 8 (3) (1986) 291–314, https://doi.org/10.1080/01688638608401069.
[34] R.P.G. VAN Gompel, M. Arai, Structural priming in bilinguals, Biling. Lang. Cognit. 21 (3) (2018) 448–455. https://doi:10.1017/S1366728917000542.
[35] H.P. Branigan, K. Messenger, Consistent and cumulative effects of syntactic experience in children's sentence production: evidence for error-based implicit learning, Cognition 157 (2016) 250–256, https://doi.org/10.1016/j.cognition.2016.09.004.
[36] C.F. Rowland, F. Chang, B. Ambridge, J.M. Pine, E.V.M. Lieven, The development of abstract syntax: evidence from structural priming and the lexical boost, Cognition 125 (1) (2012) 49–63, https://doi.org/10.1016/j.cognition.2012.06.008.
[37] M.P. Kaschak, T.J. Kutta, J.M. Coyle, Long and short term cumulative structural priming effects, Language, Cognition and Neuroscience 29 (6) (2012) 728–743, https://doi.org/10.1080/01690965.2011.641387.
[38] L.O.H. Kroczek, T.C. Gunter, Communicative predictions can overrule linguistic priors, Sci. Rep. 7 (2017) 17581, https://doi.org/10.1038/s41598-017-17907-9.
[39] K. Bock, G.S. Dell, F. Chang, K.H. Onishi, Persistent structural priming from language comprehension to language production, Cognition 104 (3) (2007) 437–458, https://doi.org/10.1016/j.cognition.2006.07.003.
[40] F. Chang, E. Kidd, C.F. Rowland, Prediction in processing is a by-product of language learning, Behav. Brain Sci. 36 (4) (2013) 350–351, https://doi.org/10.1017/s0140525x12002518.
[41] K. Bock, Z.M. Griffin, The persistence of structural priming: transient activation or implicit learning? J. Exp. Psychol. Gen. 129 (2) (2000) 177–192, https://doi.org/10.1037/0096-3445.129.2.177.
[42] J.M. Olichney, C. Van Petten, K.A. Paller, D.P. Salmon, V.J. Iragui, M. Kutas, Word repetition in amnesia - electro-physiological measures of impaired and spared memory, Brain 123 (2000) 1948–1963, https://doi.org/10.1093/brain/123.9.1948.
[43] M. Kutas, K.D. Federmeier, Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP), Annu. Rev. Psychol. 62 (2011) 621–647, https://doi.org/10.1146/annurev.psych.093008.131123.
[44] K.A. Paller, M. Kutas, Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming, J. Cognit. Neurosci. 4 (4) (1992) 375–391, https://doi.org/10.1162/jocn.1992.4.4.375.
[45] M. Besson, M. Kutas, C. Vanpetten, An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences, J. Cognit. Neurosci. 4 (2) (1992) 132–149, https://doi.org/10.1162/jocn.1992.4.2.132.
[46] C.N. Jackson, H. Hopp, Prediction error and implicit learning in L1 and L2 syntactic priming, Int. J. BiLing. 24 (5–6) (2020) 895–911, https://doi.org/10.1177/1367006920902851, 1367006920902851.
[47] M. Rabovsky, K. McRae, Simulating the N400 ERP component as semantic network error: insights from a feature-based connectionist attractor model of word meaning, Cognition 132 (1) (2014) 68–89, https://doi.org/10.1098/rstb.2019.0313.
[48] H. Fitz, F. Chang, Language ERPs reflect learning through prediction error propagation, Cognit. Psychol. 111 (2019) 15–52, https://doi.org/10.1016/j.cogpsych.2019.03.002.
[49] A.B. Fine, T.F. Jaeger, The role of verb repetition in cumulative structural priming in comprehension, J. Exp. Psychol. Learn. Mem. Cognit. 42 (9) (2016) 1362–1376, https://doi.org/10.1037/xlm0000236.
[50] M. Kutas, In the company of other words: electrophysiological evidence for single-word and sentence context effects, Lang. Cognit. Process. 8 (4) (1993) 533–572, https://doi.org/10.1080/01690969308407587.
[51] T. Ness, A. Meltzer-Asscher, Predictive pre-updating and working memory capacity: evidence from event-related potentials, J. Cognit. Neurosci. 30 (12) (2018) 1916–1938, https://doi.org/10.1162/jocn_a_01322.
[52] C.V. Petten, M. Kutas, R. Kluender, M. Mitchiner, H. McIsaac, Fractionating the word repetition effect with event-related potentials, J. Cognit. Neurosci. 3 (2) (1991) 131–150, https://doi.org/10.1162/jocn.1991.3.2.131.
[53] M.D. Rugg, The effects of semantic priming and word repetition on event-related potentials, Psychophysiology 22 (6) (1985) 642–647, https://doi.org/10.1016/j.neuropsychologia.2004.11.020.
[54] E. Gibson, Linguistic complexity: locality of syntactic dependencies, Cognition 68 (1) (1998) 1–76, https://doi.org/10.1016/s0010-0277(98)00034-1.
[55] E.F. Lau, P.J. Holcomb, G.R. Kuperberg, Dissociating N400 effects of prediction from association in single-word contexts, J. Cognit. Neurosci. 25 (3) (2013) 484–502, https://doi.org/10.1162/jocn_a_00328.

[56] M. Rabovsky, J.L. McClelland, Quasi-compositional mapping from form to meaning: a neural network-based approach to capturing neural responses during human language comprehension, Phil. Trans. Biol. Sci. 375 (1791) (2020) 20190313, https://doi.org/10.1098/rstb.2019.0313.
[57] C.E.L. Stark, J.L. McClelland, Repetition priming of words, pseudowords, and nonwords, J. Exp. Psychol. Learn. Mem. Cognit. 26 (4) (2000) 945–972, https://doi.org/10.1037//0278-7393.26.4.945.
[58] A. Buchner, W. Wippich, On the reliability of implicit and explicit memory measures, Cognit. Psychol. 40 (3) (2000) 227–259, https://doi.org/10.1006/cogp.1999.0731.
[59] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2023.
[60] S. Champely, Pwr: basic functions for power analysis, R package version 1 (2023), 3-0.
[61] C.C. Sun, P. Hendrix, J.Q. Ma, R.H. Baayen, Chinese lexical database (CLD): a large-scale lexical database for simplified Mandarin Chinese, Behav. Res. Methods 50 (6) (2018) 2606–2629, https://doi.org/10.3758/s13428-018-1038-3.
[62] A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen, lmerTest package: tests in linear mixed effects models, J. Stat. Software 82 (13) (2017) 1–26.
[63] L.M. Zhu, J. Li, emmeans: estimated marginal means, aka least-squares means, R package version 1.8.8 (2022). Retrieved from, https://CRAN.R-project.org/package=emmeans.
[64] A.M. Chambers, The role of sleep in cognitive processing: focusing on memory consolidation, Wiley Interdisciplinary Reviews: Cognit. Sci. 8 (3) (2017) e1433.
[65] N.D. Anderson, G.S. Dell, The role of consolidation in learning context-dependent phonotactic patterns in speech and digital sequence production, Proc. Natl. Acad. Sci. USA 115 (14) (2018) 3617–3622.
[66] E. Schechtman, J. Heilberg, K.A. Paller, Memory consolidation during sleep involves context reinstatement in humans, Cell Rep. 42 (4) (2023), https://doi.org/10.1016/j.celrep.2023.112331.
[67] D. Garber, J. Fiser, Structure transfer and consolidation in visual implicit learning, bioRxiv 2024–06 (2024).
[68] N. Reggev, R. Sharoni, A. Maril, Distinctiveness benefits novelty (and not familiarity), but only up to a limit: the prior knowledge perspective, Cognit. Sci. 42 (1) (2018) 103–128, https://doi.org/10.1111/cogs.12498.
[69] M.A. Gernsbacher, M.E. Faust, The mechanism of suppression: a component of general comprehension skill, J. Exp. Psychol. Learn. Mem. Cognit. 17 (2) (1991) 245–262, https://doi.org/10.1037/0278-7393.17.2.245.
[70] M. Bar, Predictions: a universal principle in the operation of the human brain, Phil. Trans. Biol. Sci. 364 (1521) (2009) 1181–1182, https://doi.org/10.1098/rstb.2008.0321.
[71] S. Rich, J.A. Harris, Unexpected guests: when disconfirmed predictions linger. Proceedings of the Annual Meeting of the Cognitive Science Society, 2021. https://escholarship.org/uc/item/7q0000c8.
[72] J. Rommers, K.D. Federmeier, Lingering expectations: a pseudo-repetition effect for words previously expected but not presented, Neuroimage 183 (2018) 263–272, https://doi.org/10.1016/j.neuroimage.2018.08.023.
[73] G. Bovolenta, E. Marsden, Prediction and error-based learning in L2 processing and acquisition a conceptual review, Stud. Sec. Lang. Acquis. 44 (5) (2022) 1384–1409, https://doi.org/10.1017/s0272263121000723.