Bilingualism: Language and Cognition

cambridge.org/bil

Inhibitory control facilitates learning new knowledge based on existing knowledge in cross-linguistic word contexts

Zilan Zou and Baoguo Chen

Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education, Faculty of Psychology, Beijing Normal University, Beijing 100875, China

Research Article

Cite this article: Zou, Z. and Chen, B. (2024). Inhibitory control facilitates learning new knowledge based on existing knowledge in cross-linguistic word contexts. *Bilingualism: Language and Cognition*, 1–15
https://doi.org/10.1017/S1366728924000993

Received: 12 January 2024
Revised: 08 October 2024
Accepted: 08 November 2024

Keywords:
inhibitory control; ambiguous word; semantic similarity; word frequency; second language

Corresponding author:
Baoguo Chen;
Email: chenbg@bnu.edu.cn

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

In cross-linguistic word learning, learning new knowledge based on existing knowledge is a common and lifelong process. This study investigated whether inhibitory control would be conducive to this process. We asked Chinese-English bilinguals to learn new meanings for familiar English ambiguous words within two consecutive days, manipulating semantic relatedness and word frequency to create four categories: high-frequency-unrelated, high-frequency-related, low-frequency-unrelated and low-frequency-related ambiguous words. Participants completed translation recognition and production tests immediately after learning and again one week later, with flanker and stop-signal tasks interspersed to measure their interference inhibition and response inhibition. The results indicated that inhibitory control, particularly interference inhibition, significantly aided in learning new meanings when direct knowledge transfer from existing knowledge was unfeasible. This research enhances our comprehension of individual differences in word learning, offering valuable perspectives for broader theories of word learning and targeted educational interventions.

Highlights

- Inhibitory control affects learning new based on existing knowledge;
- Inhibitory control aids in learning new meanings in homonyms and polysemous words;
- Strong inhibition leads to early processing challenges but better L2 word learning;
- It helps to understand how individual factors affect bilingual learning.

1. Introduction

Learning new knowledge based on existing knowledge and experience is a universally important learning process, as numerous studies highlight that existing knowledge can facilitate new knowledge acquisition (Bein et al., 2019; Witherby & Carpenter, 2022), particularly in language learning. For instance, adults can learn new words faster by linking them to existing knowledge (Vitevitch et al., 2014). However, when existing knowledge clashes with new information, it can impede learning (Brod et al., 2013). In such circumstances, individual INHIBITORY CONTROL can be a crucial additional pathway to aid new knowledge learning (Grenell & Carlson, 2021). Concentrating on this general domain control ability can help devise interventions to bridge learning gaps, particularly in language learning, between those facing difficulties and normal learners. Despite its significance, research is lacking in the field of language learning, to confirm the causal link between inhibitory control and this type of learning, and to understand the underlying cognitive mechanisms.

In this study, we investigated the impact of inhibitory control on the new knowledge learning process based on the existing knowledge in second-language (L2) word learning. The reason for choosing this context was the increasing emphasis on bilingual education and the fundamental role of word learning in language acquisition. By examining how inhibitory control interacts with the two word characteristics— SEMANTIC RELATEDNESS and WORD FREQUENCY, our findings can offer insights to deepen our understanding of cognitive processes underlying this learning, contributing to formulating a universally applicable theory of word learning and language learning strategies that engage general cognitive abilities.

1.1. Learning based on existing knowledge in the context of language learning

Hirosh and Degani (2018) posited in their theory that there were two different influences when people learned a new language based on the existing one: direct and indirect. Direct influences of existing language knowledge on new language learning emerge when similarities between the two languages are great, enabling direct knowledge transfer. In cases where this direct transfer is absent, indirect influences prompt learners to draw on existing cognitive or social skills, such as inhibitory control, to support learning. Hiroshi and Degani emphasized that inhibitory control, or the ability to resolve competition, was crucial for multilingual individuals, enabling them to excel in learning new languages. The relationship between the characteristics of existing and new languages assisted learners in adjusting direct and indirect influences, facilitating language learning. This implies that both linguistic and individual cognitive characteristics should be considered when studying language acquisition.

Given that bilingualism is currently the primary environment for upbringing (Serrat et al., 2021), our focus would be on learning new knowledge based on existing knowledge in L2 acquisition. In this context, *learning new meanings for familiar words* (LNM-FW) is particularly suitable for exploring its relationship with inhibitory control. Using a single word to denote multiple meanings is common in natural languages. For example, over 80% of English words have at least two meanings (Rodd et al., 2002). Such words are often termed AMBIGUOUS WORDS (Klepousniotou et al., 2012) and can be classified as either POLYSEMES or HOMONYMS (Rodd et al., 2002). Polysemes, like *walk*, have related meanings such as "going on foot" and "strolling." Homonyms, however, have unrelated meanings that are etymologically distinct and happen to share the same word form (Clark & Clark, 1979). For instance, *bank* can refer to "riverside" or "financial institution." In daily life, the learning of different meanings of ambiguous words typically occurs sequentially. Language learners often update their knowledge by forming new connections between familiar word forms and new meanings (Fang & Perfetti, 2019).

The LNM-FW process can span learners' entire lives, pivotal in building individual lexical knowledge structures (Rodd, 2020). Essentially, LNM-FW involves integrating new lexical knowledge into a word's SEMANTIC SPACE built upon existing knowledge (i.e., familiar word forms and known meanings). Unlike learning entirely new words, LNM-FW focuses cognitive resources more on semantic learning, particularly in allowing new meanings to interact with existing ones (Fang & Perfetti, 2019). However, in L2 acquisition, mastering ambiguous words through this process can be challenging. For example, learning secondary meanings of English ambiguous words can be more difficult than learning primary ones for native Chinese speakers (Mo & Sun, 2004), and this difficulty persists even among advanced English learners (Qu & Zhang, 2005).

1.2. Factors related to direct influence: semantic relatedness and word frequency

Hirosh and Degani (2018) emphasized that in cross-linguistic learning contexts, objects of direct influence primarily included transferable knowledge and representations from early experiences. Moreover, the essence of L2 ambiguous word learning involves creating new semantic representations and associating them with existing word forms (Bracken et al., 2017). Our study therefore examined factors influencing lexical representations: 1) the relatedness between meanings of ambiguous words (i.e., semantic relatedness), which affects the formation of new semantic representations and 2) word frequency, which relates to the bonds between word forms and new meanings.

According to Rodd's Lexical Processing and Acquisition theory with an ambiguity-focused perspective (Rodd, 2020), word meanings reside within a semantic space as distributed representations formed by a combination of semantic features. The meanings of polysemes share semantic features and are thus closely placed in this space, while those of homonyms are unrelated and positioned distantly. Links within the space are twofold: frequently co-occurring, neighboring semantic features form positive and facilitating links, whereas distant, unrelated features exhibit negative and inhibiting links. Therefore, polysemy learning utilizes positive links to transfer semantic features, facilitating the formation of new representations, while homonym learning poses more challenges.

Support for this comes from studies on L2 ambiguous word learning. Studies by Zhang et al. (2018, 2022) found that Chinese-English bilinguals learned related new meanings more quickly compared to unrelated ones. They therefore proposed an interference-facilitation mechanism: for polysemes, existing meanings facilitate the learning of new meanings and semantic integration, whereas for homonyms, the learning process is hindered by existing meanings due to low relatedness that leads to challenges. However, their studies had two limitations: a) they used a binary method to classify ambiguous words, despite researchers like Bracken et al. (2017) recommended considering semantic relatedness as a continuous variable and b) their materials were less ecologically valid by pairing new meanings with familiar monosemic words or assigning two meanings to English pseudowords to simulate the LNM-FW process.

Chen and Chen (2022) addressed these issues by asking Chinese-English bilinguals to learn subdominant meanings for familiar English words with varying relatedness to known meanings. They employed primed lexical decision tasks and analyzed N400 and LPC (Late Positive Component) priming effects to gauge the state of semantic representations of both existing and new meanings. Their study found that higher semantic relatedness could enhance the semantic integration of new meanings by facilitating the update of existing meaning representation in reconsolidation. In essence, semantic relatedness influences how new meanings are acquired: polysemes facilitate the effortless transfer of useful semantic features from related existing meanings to form new semantic representations, while learners need to rely on different ways for homonyms.

Word frequency also affects ambiguous word representation, serving as an effective indicator of the solidity of words' form-meaning bond (Baayen, 2010; Perfetti & Hart, 2002). Fang et al. (2017) found that when English native speakers were asked to learn new meanings of high-frequency and low-frequency words (unrelated to the existing meanings), existing meanings of higher-frequency words tended to interfere with the learning process, leading to greater adjustments in the overall lexical space, but what was surprising was that this was conducive to the lexicalization of new meanings.

Previous research has not explored how word frequency affects polysemous word learning. However, a study on processing suggested that word frequency influenced their lexical representations (Jager et al., 2016). In their study, native English speakers performed semantic categorization and lexical decision tasks. Results showed that high-frequency words were advantageous in more demanding semantic processing tasks (i.e., semantic categorization tasks), while low-frequency words were disadvantageous; in less demanding semantic processing tasks (i.e., lexical decision tasks), this trend reversed. This indicates high-frequency polysemous words, unlike low-frequency ones, have more distinct lexical representations, posing challenges for constructing new form-meaning bonds.

In summary, both semantic relatedness and word frequency affect the construction of new semantic representations and form-

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

meaning bonds in the LNM-FW process. Previous studies indicated that higher semantic relatedness facilitated the direct transfer of similar semantic features from existing meanings, leading to a learning advantage, whereas obstacles occurred in acquiring homonyms (Chen & Chen, 2022; Zhang et al., 2018, 2022). Though the role of word frequency in this area is still underresearched, monolingual studies indicate its significance in lexicalizing new meanings (Fang & Perfetti, 2017). And high-frequency words always have many meanings (Piantadosi et al., 2012), providing more examples to draw upon and aiding in the construction of new representations (Storkel et al., 2013). Thus, we presumed that word frequency also impacted this process in L2 learning. Notably, unlike the impact of semantic relatedness, the form-meaning bonds in high-frequency words, which help learn new meanings, can interfere with each other. All of these hints that learning new meanings requires the involvement of indirect influence.

1.3. Factors related to indirect influence: inhibitory control

From the above, it is evident that high frequency and low semantic relatedness cause interference in learning new meanings for ambiguous words. Thus, drawing on Hirosh and Degani (2018)'s theory, we believed that inhibitory control could influence this learning indirectly.

Inhibitory control, a core component of executive functions, is the capacity to regulate attention, thoughts, behaviors and emotions when faced with strong internal biases or external allurements (Diamond, 2013). It consists of two types: INTERFERENCE INHIBITION, which overcomes conflicts arising from unrelated or incompatible stimuli (Kang et al., 2022), and RESPONSE INHIBITION, which curtails reactive impulses that are inappropriate or no longer needed (Verbruggen & Logan, 2008).

The study by Lev-Ari and Keysar (2014) found that individuals' inhibitory skills impacted their perceived similarity ratings (high ratings mean high similarity) of ambiguous word representations. Specifically, compared to those with stronger inhibitory skills, individuals with weaker inhibitory skills gave higher similarity ratings to homonyms but lower ratings to polysemous words. This suggests that those with strong inhibitory skills are better at discerning the connections between polysemous words' meanings and the distinctions among homonyms' meanings.

Indeed, Lu et al. (2017) were the first to investigate the influence of inhibitory control on L2 ambiguous word learning. They created 30 English pseudowords, 15 of which were paired with a Chinese meaning to form monosemic words and others were paired with two unrelated Chinese meanings to form homonyms. During the three-day learning period, Chinese-English bilinguals began learning the monosemic words and the first meaning of the homonyms starting from the first day, and the second meaning starting from the second day. Before learning, inhibitory control was assessed using a Stroop task. After learning, cross-language semantic relatedness judgment tasks gauged the word learning effect. They found that *reaction times* (RTs) for the second meanings were the longest. Furthermore, inhibitory control significantly predicted RTs for second meanings—stronger inhibitory control resulted in shorter RTs. Thus, they argued that the first meaning hindered the learning of the second meaning. Those with stronger inhibitory control could effectively curb this interference, thereby constructing the representations of the second meaning.

To conclude, inhibitory control could be a pivotal indirect influence in LNM-FW for L2 ambiguous words. Yet, previous studies had their limitations: First, they did not utilize real L2 ambiguous words. The learning materials of the two studies were either words from a graphic-based artificial language or English pseudowords. Second, they all used a single task to measure one type of inhibitory control without examining the different influences arising from interference inhibition and response inhibition simultaneously, both of which were proven to affect L2 word learning in past works (Grant et al., 2015; Kapa & Erikson, 2020). Finally, their learning time was short, which was completely different from real learning situations. In the study of Lu et al. (2017), there was only a one-day gap between the learning times of the two meanings of homonyms, resulting in weak semantic representation stability and huge conflicts between the two meanings, in which case inhibitory control may play a greater role. However, in real life, people are often exposed to the second meaning only after the representation of the first meaning is highly stable, which may diminish the role of inhibitory control.

1.4. The present study

The present study explored how inhibitory control influenced new knowledge learning based on existing knowledge, specifically focusing on the LNM-FW process in L2 ambiguous word learning. Essential for understanding individual learning differences, our research can reveal how individuals leverage inhibition control to learn new knowledge by invoking existing knowledge according to the learning difficulty. These insights contribute to the development of a comprehensive theory of word learning and the design of effective educational interventions.

We asked Chinese-English bilinguals to learn new meanings for familiar English words, with these new meanings either related or unrelated to existing meanings, and words classified as high frequency or low frequency. This resulted in the following categorization: high-frequency semantically unrelated words, low-frequency semantically unrelated words, high-frequency semantically related words and low-frequency semantically related words. Learning tasks were conducted over two days, and participants completed translation recognition and translation production tests on new meanings immediately after learning and again one week later. We assessed their interference inhibition and response inhibition using flanker and stop-signal tasks respectively. Finally, we constructed linear mixed-effects models to examine the role of inhibitory control, using performances in ambiguous word tests as the dependent variables and including scores from inhibitory control tasks as part of the predictors.

Compared with prior studies, our research offered unique contributions: a) We utilized authentic L2 ambiguous words as learning materials, whose meanings were inherent to the words themselves, ensuring higher ecological validity and broader applicability; b) We examined the impact of interference inhibition and response inhibition on LNM-FW. Previous studies showed that stronger interference inhibition led to better performance in L2 words learning (Kapa & Colombo, 2014), and response inhibition could positively predict bilinguals' new word learning (Warmington et al., 2019). Thus, it was necessary to use different tasks to differentiate their impacts; c) We investigated inhibitory control's role in learning both homonyms and polysemes, since polysemes account for a larger portion of ambiguous words (Rodd et al., 2002; Zipf, 1945); d) We assessed how the stability of existing form-meaning bonds influenced new meaning acquisition. Word frequency served as an indicator: high-frequency words exhibited more stable form-meaning bonds, while low-frequency words showed less stability. Learning new meanings can alter the established semantic space,

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

leading to the creation of new semantic representations and bonds. This dynamically adjusts the mapping relationship between form and meaning to a one-to-many structure. Focusing on this indicator helps us better understand the impact of existing knowledge on new knowledge learning.

This study investigated three critical questions: 1) Does inhibitory control play a role in learning semantically unrelated ambiguous words with varying frequencies? 2) Is it also influential in learning semantically related ambiguous words with varying frequencies? 3) Which has a more pronounced impact: interference inhibition or response inhibition?

We hypothesized that inhibitory control would be key to learning semantically unrelated ambiguous words due to possible interference among meanings, and that stronger inhibitory control would facilitate learning, especially for high-frequency words. However, due to the lack of relevant studies, our expectations for polysemes were not definitive. One possibility was that inhibitory control would have little effect on learning polysemes; another was that learning polysemes might require stronger inhibitory control to suppress existing meanings and facilitate learning new ones, as they required greater distinction between meanings. We also expected interference inhibition to play a more vital role in LNM-FW, in contrast to response inhibition which is typically related to curbing habitual physical reactions instead of cognitive behaviors like word learning.

2. Method

2.1. Participants

Forty-six bilinguals (40 females, average age = 21.7, *SD* = 1.88) were recruited from Beijing Normal University. All participants were native Chinese (L1, native language) speakers. The average age at which they started learning English (L2) was 7.72 (*SD* = 1.75), with an average of 14.66 (*SD* = 2.75) years spent learning English. They were all right-handed, with normal or corrected-to-normal vision, and none had lived abroad for over six months or had systematically learned other languages. The formal experiment included three phases: Learning, Immediate Testing and Delayed Testing. Due to COVID-19, only thirty-nine participants (33 females, average age = 21.64, *SD* = 1.97) completed the Delayed Testing on time. All participants signed a written informed consent before the experiment and were paid for participation. This study was approved by the Ethics Committee of Beijing Normal University.

Language history and English proficiency were measured with a Chinese version of the language history questionnaire (LHQ-3; Li et al., 2020) and the Oxford Quick Placement Test (OQPT) (Crosthwaite et al., 2015). The first survey asked participants to subjectively rate their English listening, speaking, reading and writing abilities on a 7-point Likert-type scale (1 = very weak; 7 = very strong). The average scores of these abilities were 3.83 (*SD* = 1.04), 3.41 (*SD* = 1.22), 4.65 (*SD* = 1.02) and 3.89 (*SD* = 1.14), respectively. The OQPT, capped at 60 points and timed at 30 minutes, gauged English proficiency, with higher scores indicating higher proficiency levels. The mean OQPT score was 40.30 (*SD* = 5.14, *N* = 46). According to the above results, all participants' English proficiency was roughly at the intermediate level.

2.2. Materials

2.2.1. Learning materials

The learning materials were 160 English ambiguous words, each with two corresponding Chinese translations. These words were equally divided into four categories based on WORD FREQUENCY and SEMANTIC RELATEDNESS: high-frequency semantically related words, low-frequency semantically related words, high-frequency semantically unrelated words and low-frequency semantically unrelated words (see Table 1 for examples and Supplementary Appendix 1 for the full list).

We collected the LogFreq (Zipf) value from SUBTLEX-UK (van Heuven et al., 2014) as the word frequency for each word, using Zipf = 4 as the criterion to distinguish between high-frequency and low-frequency words. Semantic relatedness was determined by calculating the average score of relatedness between two Chinese meanings, as rated on a 7-point Likert-type scale (1 = completely unrelated, 7 = completely related) by seventeen participants with backgrounds similar to those in the experiment. Additionally, we measured 14 other lexical properties. See Supplementary Appendix 4 for details.

Additionally, the 160 English words and their Chinese meanings were evaluated in three rigorous rounds to meet the LNM-FW criteria (see Supplementary Appendix 4). Feedback from formal experiment participants confirmed the materials' suitability in two ways: a) Only participants who scored over 85% on the Online English Word Test, which evaluated knowledge of existing meanings, were eligible for the experiment. Our participants scored an average of 96% (*SD* = .03), indicating a strong grasp of the words' existing meanings before the experiment. b) At the end of the first day, participants self-reported whether they knew the new meanings beforehand. Those who knew many new meanings were excluded. All participants reported either no prior knowledge or familiarity with only a few, affirming the materials' appropriateness for learning.

Table 1. Examples of different types of L2 ambiguous words and their Chinese translations

	High frequency		Low frequency	
Group	Related	Unrelated	Related	Unrelated
English word	Weight	Plant	Wrinkle	Appendix
Existing meaning	重量	植物	皱纹	附录
Definition	The force of gravity on an object	An organism with stem, leaves, and roots in the soil	Skin crease	End-of-book informational section
New meaning	砝码	卧底	褶皱	阑尾
Definition	Metal piece for specific weight measurement	An undercover person	The raised fold in fabric or paper	Small tissue pouch on the large intestine

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

One-way ANOVA tests showed significant differences in both word frequencies [F (3,156) = 200.46, p < .001] and semantic relatedness [F (3,156) = 208.52, p < .001] among the four word groups. Independent T-tests revealed no significant differences in word frequency (ps > .08) or semantic relatedness (ps > .45) within groups sharing the same level of word frequency or semantic relatedness. This confirmed our effective manipulation of the two word characteristic variables. Four groups showed significant differences in *word familiarity, the age of acquisition for L2 word (L2AOA), word length, Orthographic Levenshtein Distance 20 (OLD20), number of orthographic neighbors, number of syllables, frequency of the existing meaning*, and *familiarity with the existing meaning* (*all Fs* < 5.39, *all ps* < .001). To avoid their influences on the results, these indices were incorporated as covariates in the subsequent linear mixed-effects model.

2.2.2. Materials for the translation recognition test

Translation recognition tests were used to assess the learning effects of new meanings and its stimuli were identical to the learning materials. To avoid practice effects, we created two test versions—A and B, each containing 160 words with new meanings: half corresponding and half false. Specifically, we randomly divided the 160 words with corresponding new meanings into Groups A and B, balancing them for WORD FREQUENCY and SEMANTIC RELATEDNESS. In version A, we mismatched each L2 word in Group B with the L1 translation of another word within the group, yielding choices where participants would press "no." However, Group A remained unchanged, yielding "yes" choices. Version B adopted the same method but with the word-meaning pairs from Group A being mismatched. Each participant completed both versions, and the two versions were counterbalanced across participants. See Supplementary Appendix 1 for full materials.

2.3. Procedures

The experiment comprised three phases: Learning, Immediate Testing and Delayed Testing. The first two phases were on the first two days, and the last phase was on the ninth day. The tests in the two testing phases were identical (i.e., tests in the shaded boxes), aiming to evaluate participants' performance of memorizing and retrieving new meanings over short and long periods.

During the Learning Phase, participants learned the new L1 meanings of the 160 L2 ambiguous words three times per day for two days, followed by a cued-recall test with new meaning feedback each day. In the Immediate Testing phase, participants successively completed a translation production test and a translation recognition test, both focusing on new meanings. On the ninth day, a cued-recall test without feedback preceded the same translation tests. Throughout the experiment, participants also completed Stop-signal and Flanker tasks, an intelligence test, a language history questionnaire, and the OQPT alternately (see Figure 1 for the full procedure).

2.3.1. Learning phase

Word leaning We used the Anki software (https://apps.ankiweb.net/) to create 160 bilingual flashcards (see Figure 2a), a common method for learning L2 words (Elgort & Piasecki, 2014). Each flashcard, displayed in fullscreen on a computer, included an English word's spelling, phonetic symbol, pronunciation video

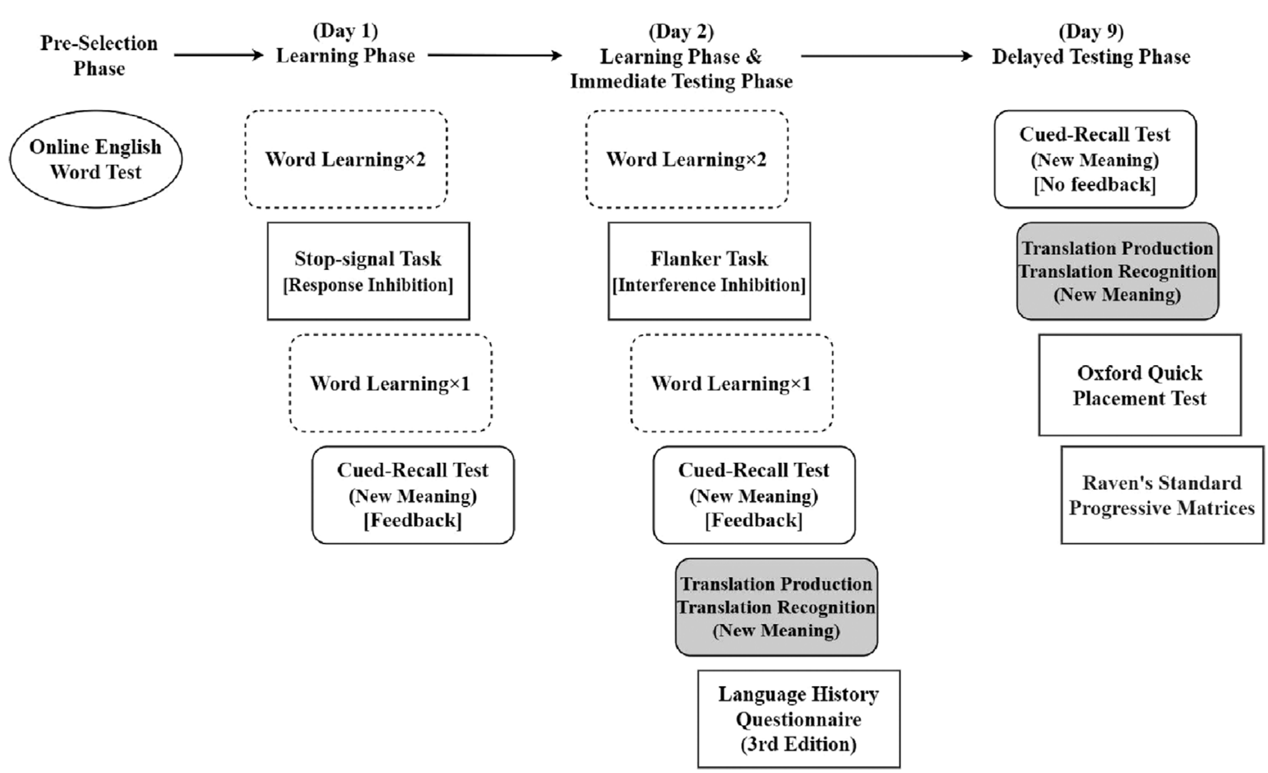

Figure 1. Experimental procedure.
Note: The dashed oval boxes signify word learning, the solid oval boxes signify lexical tests, and the solid square boxes signify tests related to individual ability. The shaded boxes signify immediate and delayed tests for learning effects.

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

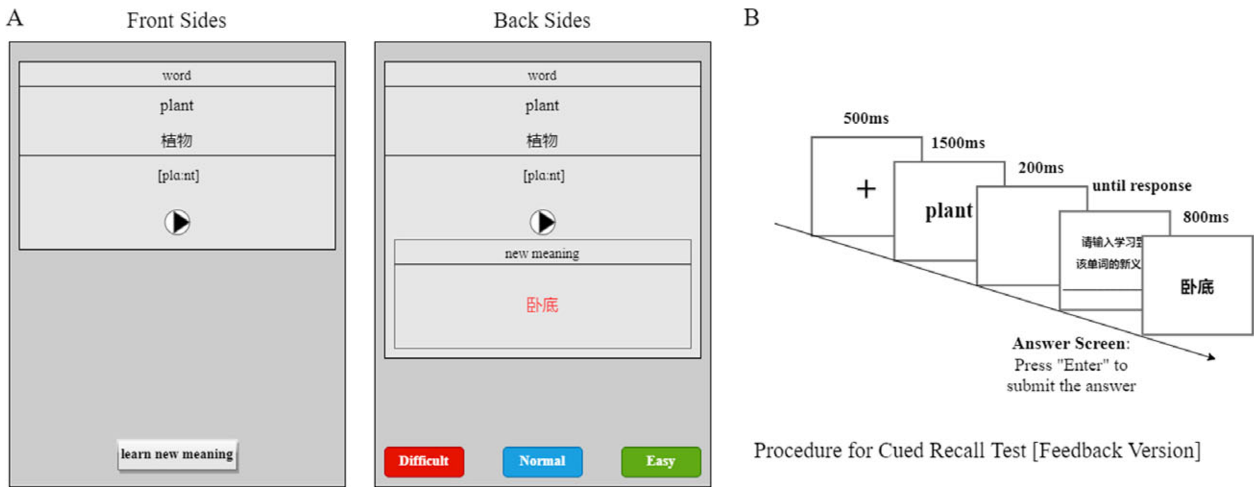

Figure 2. Example of bilingual flashcard (a) and cued-recall test flow chart (b).

and both its existing and new Chinese meanings, each appearing only once per learning round.

Participants wore headphones before learning to listen to each word's pronunciation audio that was automatically played once they saw the front of the card. Then, participants clicked the "learn the new meaning" button to view the back and assessed the learning difficulty before proceeding (to maintain focus and reduce fatigue). During the instruction phase, experimenters highlighted that participants should memorize meanings solely by observing the screen without using paper or pens. They were encouraged to learn at their own pace and take breaks when necessary. A verbal reminder would be given if the learning round exceeded 15 minutes.

Cued-recall test To directly examine the learning effects of new meanings and reinforce retention (Kang et al., 2013), we wrote a test program using E-prime 2.0. It included 160 trials, each beginning with a 500-ms fixation cross, followed by a 1500-ms display of an English word. Subsequently, participants had to quickly recall and type the new meanings on the following screen after a 200-ms blank screen. Upon submission, they either viewed the correct answer for immediate feedback (see Figure 2b) or, an 800-ms blank screen in the no-feedback version. Due to E-prime's restrictions with Chinese character input, we recorded only accuracy. Correct answers must perfectly match the two-character Chinese translations provided during the Learning Phase; any other inputs were marked incorrect.

2.3.2. Testing phase

L2-L1 translation recognition test This test used E-prime 2.0 to present English and Chinese words alternately across 170 trials—10 practice and 160 experimental, and the experimental trials divided into four blocks evenly with short breaks in between. Each trial started with a 500-ms fixation cross, followed by a 250-ms English word and a 2000-ms Chinese word. Participants rapidly and accurately pressed "F" (yes) or "J" (no) to indicate if the Chinese word matched the new meaning of the preceding English word. Button assignments were counterbalanced among participants. The test maintained a 1:1 ratio of "yes" to "no" responses. Trials not answered within 2000 ms were skipped and marked incorrect. A 500-ms interval separated each trial. The test lasted approximately ten minutes.

L2-L1 translation production test The program of this test, written using E-prime 2.0, included seven blocks: the first six each contained 25 trials and the final block comprised 10 trials. Short breaks were allowed between blocks. Each trial featured an 800-ms fixation cross, a 500-ms icon "★" (indicating to answer with new meanings), a 500-ms English word, and a "?" displayed for 3500 ms, signaling participants to vocalize the new meaning into a microphone. RTs were captured via an SR-Box and answers were recorded on a tablet computer. The test took about fifteen minutes.

2.3.3. Tests related to individual abilities

Flanker task To evaluate interference inhibition, participants in this task identified the direction of the central arrow flanked by lines or arrows using the "F" (left) or "J" (right) keys. It consisted of 160 trials: 16 practice and 144 experimental trials which were divided evenly based on three conditions: congruent (e.g., → → → → →), neutral (e.g., – – → – –) and incongruent (e.g., → → ← → →). Each trial began with a 500-ms fixation cross, followed by the stimuli. Following participants' responses or after 1500 ms, a 500-ms interval preceded the next trial. RTs and accuracy were recorded, with the Flanker effect calculated from RT differences between congruent and incongruent trials. The larger the effect, the weaker the interference inhibition ability.

Stop-signal task To assess response inhibition, we used the STOP-IT program (https://osf.io/wuhpv/) to perform a standard stop-signal task (Verbruggen et al., 2013). Trials were split into no-signal and stop-signal categories at a 3:1 ratio. In no-signal trials, participants needed to identify the direction of a central white arrow using the "←" or "→" buttons. In stop-signal trials, they had to refrain from pressing any button when a blue arrow appeared. We utilized R codes from the program to analyze RT and accuracy, calculating the STOP-SIGNAL REACTION TIME (SSRT)—the span from the appearance of the stop signal to the point when participants finished the trial. The higher its value, the weaker the reaction inhibition ability.

Intelligence test To measure individual intelligence, we used Raven's Standard Progressive Matrices (Zhang & Wang, 1986), which required participants to choose the correct pattern from six or eight options to complete a matrix of geometric patterns with a missing bottom-right segment. Given the uniform age cohort of the participants, INTELLIGENCE QUOTIENT (IQ) was directly measured by the raw score (Raven_score), which is the total number of correct responses with a maximum of 60. The Cronbach's alpha reliability for this test was .70.

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

3. Data analysis

We analyzed the data using the package *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) to construct linear mixed-effects models in *R* (Version 4.2.1; R Core Team, 2022). Effect sizes (*Cohen's d*) were calculated using the package *EMA tools* (Kleiman, 2021). All variables were centralized and then incorporated into the model as continuous variables.

We optimized the model structure mainly by focusing on the random slope of random effects, starting from the full model containing the maximum random effect structure (Barr et al., 2013)—random intercepts for both *participants* and *items*, and random slopes for *word frequency, semantic relatedness, inhibition control (flanker effect and SSRT)* and *their interaction.* If nonconvergence occurred, we employed backward stepping until achieving model fit. ANOVA tests were used to compare all converging models, with the best-fitting one selected based on the lowest Akaike Information Criterion (AIC) value. Results only present estimates and statistics from the best-fitting model (see Supplementary Appendix 3). For R codes, see Supplementary Appendix 2.

3.1. Data preprocessing

A researcher coded the translation production test data by listening to participants recorded verbal answers. The criteria were as follows: for accuracy, an answer was marked correct only if it precisely matched the new word meaning; for RT, if the participant used obvious modal particles (e.g., "ah," "er," etc.) or hesitated, the RT of the word was recorded as 0 ms.

We excluded participants with overall accuracies below 70%, retaining 46 for the immediate tests and 39 and 38 for the delayed recognition and production tests, respectively. In subsequent modeling, we only used RTs as the dependent variable. The reasons were twofold: first, the accuracies of both translation recognition tests and the immediate translation production test were close to the ceiling (see Table 2); second, attempts to build a generalized linear mixed-effects model using delayed translation production test accuracy as the dependent variable yielded near-unidentifiable results due to large eigenvalues.

We filtered out erroneous RT data and values beyond *Mean* ± 2.5 *SD*. In translation production tests, RTs under 200 ms or missing were also removed. As a result, 6.90% and 6.73% of data were excluded for immediate and delayed recognition tests, while 6.35% and 13.34% were excluded for immediate and delayed production tests (see Table 2 for details). The distribution of RTs, analyzed using package *moments* (D'Agostino, 1970), showed positive skewness. According to the skewness of the data, we applied reciprocal transformation (1000/RT) to recognition test RTs and log transformation ($\log_{10}$) to production test RTs, respectively.

3.2. Model structure

Before model construction, we performed a co-linearity test on individual cognitive variables—namely, inference inhibition, response inhibition and IQ. Pearson correlations (see Table 3), which revealed no significant correlation among the scores from flanker task (Flanker_effect), stop-signal task (SSRT) and intelligence test (Raven_score) (*all ps* > .25), allowing their direct incorporation into the models.

Four linear mixed-effects models were constructed for both immediate and delayed translation recognition and production tests, respectively. The core predictors for each model in the fixed effects part were *Word Frequency, Semantic Relatedness* (hereafter referred to as *Relatedness*), two inhibitory control component variables (*Flanker_effect* and *SSRT*), and their respective interaction with the first two variables. The covariates contained 8 lexical factors (see Learning Materials for details) and individual *IQ* to exclude their impacts on the fixed effects. To concurrently consider the variance due to participants and the variance related to items, random effects included by-participant and by-item intercepts. Because of our focus on inhibitory control, which was related to individual differences, we did not include the by-participant slope. By-item slope was determined by the backward-stepping procedure described previously.

4. Results

See results of cued-recall test in Supplementary Appendix 4.

4.1. Translation recognition tests

In this test, the dependent variable in all interaction plots was the inverse of RT (i.e., 1000/RT), showing the opposite trend to the actual RT.

4.1.1. Immediate testing phase

In this phase, the main effect of relatedness was significant ($\beta = .06$, $t = 3.08$, $p < .01$), while that of word frequency was marginally significant ($\beta = -.06$, $t = -1.97$, $p = .05$). No other core predictors or interactions in the fixed effects were significant. The model

Table 2. Mean reaction time (ms) and accuracy (%) (SD) among tests and days for the different words' groups

	Immediate testing phase (Day 2)				Delayed testing phase (Day 9)			
	Recognition test		Production test		Recognition test		Production test	
Groups	RT	ACC	RT	ACC	RT	ACC	RT	ACC
HR	590 (141)	96 (19)	830 (313)	98 (15)	637 (171)	97 (18)	880 (337)	93 (25)
HU	609 (145)	95 (22)	874 (311)	96 (19)	656 (174)	95 (23)	940 (353)	87 (34)
LR	606 (148)	97 (17)	912 (314)	97 (17)	645 (169)	97 (18)	956 (334)	91 (28)
LU	611 (144)	96 (21)	875 (306)	97 (17)	658 (167)	96 (19)	928 (323)	88 (32)
Overall	604 (145)	96 (20)	873 (312)	97 (17)	649 (170)	96 (20)	925 (338)	90 (30)

Abbreviations: HR = High-frequency and Semantically Related Ambiguous Words; HU = High-frequency and Semantically Unrelated Ambiguous Words; LR = Low-frequency and Semantically Related Ambiguous Words; LU = Low-frequency and Semantically Unrelated Ambiguous Words.

Table 3. Pearson correlations among individual variables

Measures	1	2	3
1 Flanker_effect	1.00		
2 SSRT	−.03	1.00	
3 Raven_score	.11	−.17	1.00

coefficients indicated that higher relatedness quickened new meaning recognition, while higher word frequency slowed it down.

4.1.2. Delayed testing phase

The main effects of relatedness ($\beta = .07$, $t = 3.31$, $p < .01$) and word frequency ($\beta = -.07$, $t = -2.36$, $p < .05$) were significant. A notable interaction emerged between relatedness and Flanker_effect ($\beta = -.03$, $t = -2.36$, $p < .05$). No other core predictors or interactions in the fixed effects were significant. The model coefficients revealed the same trends as the immediate testing phase.

Using the package *interactions* (Bauer & Curran, 2005), we conducted a simple-slope test via the *Johnson-Neyman* method for the interaction Relatedness × Flanker_effect. We found a significant slope of relatedness on RTs ($p < .05$) when Flanker_effect values resided outside the range of [12.99, 229.80]. Compared with our observed range of [−32.39, 34.41] for participants' Flanker_effect, a lower Flanker_effect yielded a more significant slope. Concisely, stronger interference inhibition amplified the RT differences in recognizing ambiguous words with varying relatedness levels evidenced by longer RTs for unrelated words than those for related ones (see Figure 3).

The performance at the extreme ranges of relatedness is shown in Figure 3. For unrelated words (left end of the plot), stronger interference inhibition (–1*SD*) prolonged RTs. However, for related words (right end of the plot), RTs remained indistinguishable across different levels of interference inhibition (i.e., both –1*SD* and + 1*SD* lines overlap). This indicated that interference inhibition mainly affected RTs for unrelated words, showing an interference inhibitory disadvantage. Specifically, stronger interference inhibition slowed down the recognition of unrelated words without significantly affecting the recognition speed of related ones.

In summary, our translation recognition test results revealed that: a) regarding lexical characters, both word frequency and semantic relatedness impacted postlearning RTs for recognizing new meanings, with high word frequency slowing down and high semantic relatedness accelerating recognition. These effects persisted one week after learning completion; b) regarding individual characteristics, interference inhibition only influenced RTs in the delayed phase, indicating a disadvantage. Stronger interference inhibition was associated with longer RTs; and c) regarding the interactions between the two characteristics, effects were only observed in the delayed phase, indicating an inhibitory disadvantage related to lower relatedness. Specifically, stronger interference inhibition prolonged RTs for recognizing new meanings of semantically unrelated ambiguous words.

4.2. Translation production tests

In this test, log-transformed RT with a base of 10 (i.e., $\log_{10}$RT) was the dependent variable in all interaction plots, showing the consistent trend to actual RT.

4.2.1. Immediate testing phase

In this phase, for main effects, only found that SSRT was significant ($\beta = -.20$, $t = -2.80$, $p < .01$), whereas relatedness and word frequency were not significant (Relatedness: $\beta = -.03$, $t = -1.11$, $p = .27$; Word_Frequency: $\beta = .08$, $t = 1.79$, $p = .08$). No other core predictors showed significant main effects. According to the coefficient, individuals with stronger response inhibition exhibited longer RTs for producing new meanings, indicating a response inhibitory disadvantage. Regarding interactions, a significant three-way interaction emerged among word frequency, relatedness and Flanker_effect ($\beta = -.03$, $t = -3.58$, $p < .001$). No other interactions were significant.

A simple-slope test on the significant interaction revealed the following results: for unrelated words (*Mean* – 1*SD*), the slope for word frequency was significant ($p < .05$) within the range of [−16.99, 189.04] of Flanker_effect values; for related words (*Mean* + 1*SD*), the significance of its slope occurred outside the range of [−24.68, 57.00]. Within our observed participants' Flanker_effect range of [−32.54, 45.07], comparison of these ranges showed diverging trends for related and unrelated words. Specifically, for homonyms, weaker interference inhibition resulted in more significant RT differences when producing new meanings of

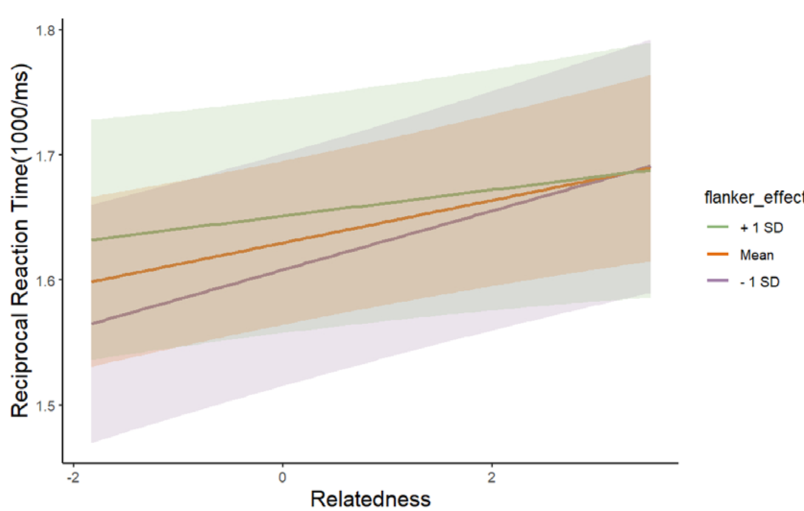

Figure 3. Interaction graph of Day 9 translation recognition test: Relatedness × Flanker_effect.

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

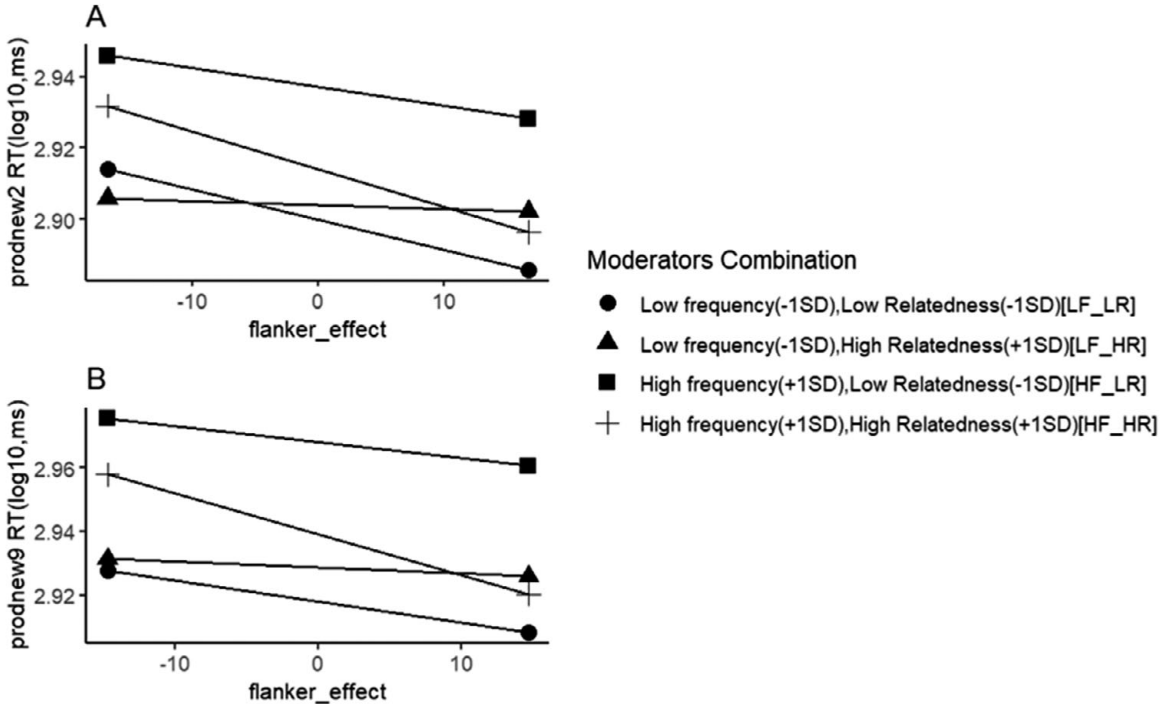

Figure 4. Interaction graph of translation production test: Word_Frequency × Relatedness × Flanker_effect.
Note: Graph A displays results from the second-day test, while Graph B shows the ninth-day test outcomes.

ambiguous words with varying frequencies; for polysemes, stronger interference inhibition resulted in more significant differences. Given the positive regression coefficient for word frequency on RT, this difference meant individuals took more time to produce new meanings for high-frequency words than for low-frequency words.

A simple-slope graph (see Figure 4a) was plotted using the package *pequod* (Mirisola & Seta, 2016). Abbreviations in legends represent different ambiguous words, e.g., LF_LR for low-frequency semantically unrelated ambiguous words with both word frequency and relatedness one standard deviation below the Mean. Specific illustrations based on words with different semantic relatedness are as follows.

First, for unrelated words (HF_LR-■ and LF_LR-●), both low- and high-frequency ones showed that stronger interference inhibition led to longer RTs. High-frequency words consistently had longer RTs than low-frequency words, highlighting the necessity to address interference caused by high-frequency words regardless of individual interference inhibition levels. The slope for low-frequency words (LF_LR) was steeper, suggesting a more prominent role of interference inhibition for low-frequency words compared to high-frequency words.

Second, for related words (LF_HR-▲ and HF_HR - +), only high-frequency ones (HF_HR) showed a production disadvantage —stronger interference inhibition led to longer RTs. The slope for low-frequency words (LF_HR) was relatively flat with no discernible difference in RTs among individuals with varying interference inhibition. Figure 4a depicts specific situations: for those with strong interference inhibition (on the left of Figure 4a), high-frequency words resulted in longer RTs than low-frequency words. Conversely, for those with weak interference inhibition (on the right of Figure 4a), low-frequency words necessitated longer RTs than high-frequency words. Combining this with slope test results, we found that word frequency differences were not significant among those with weak interference inhibition, as their Flanker_effect range of [0, 45.07] fell within the insignificant slope range of [−24.68, 57.00]. In a word, interference inhibition predominantly affected high-frequency words, manifesting an inhibitory disadvantage.

Finally, comparing all four types of ambiguous words, we found the slopes in the graph ranged from steep to flat: HF_HR > LF_LR > HF_LR > LF_HR. This suggested that the influence of interference inhibition decreased in the sequence: high-frequency-related words > low-frequency-unrelated words > high-frequency-unrelated words > low-frequency-related words. However, as mentioned earlier, interference inhibition had an insignificant impact on LF_HR. Therefore, interference inhibition mainly affected unrelated words and high-frequency related words, with the most prominent influence on the latter.

4.2.2. Delayed testing phase

Here, both word frequency (β = .10, t = 2.11, p < .05) and SSRT (β = −.14, t = −2.23, p < .05) exhibited significant main effects, whereas relatedness did not (β = −.03, t = −1.14, p = .26). Same model coefficients directions revealed the same trends as the immediate testing phase. No other core predictors yielded significant main effects. Notably, three significant interactions emerged: a) the interaction among word frequency, relatedness and Flanker_effect (β = −.03, t = −2.07, p < .05); b) the interaction between word frequency and relatedness (β = −.07, t = −2.34, p < .05); c) the interaction between relatedness and SSRT (β = −.03, t = −2.09, p < .05).

First, regarding the Word frequency× Relatedness × Flanker_effect interaction, the simple-slope test showed that for unrelated words (*Mean* – 1*SD*), the slope of word frequency was significant within the Flanker_effect range of [−48.55, 57.09]; and for related

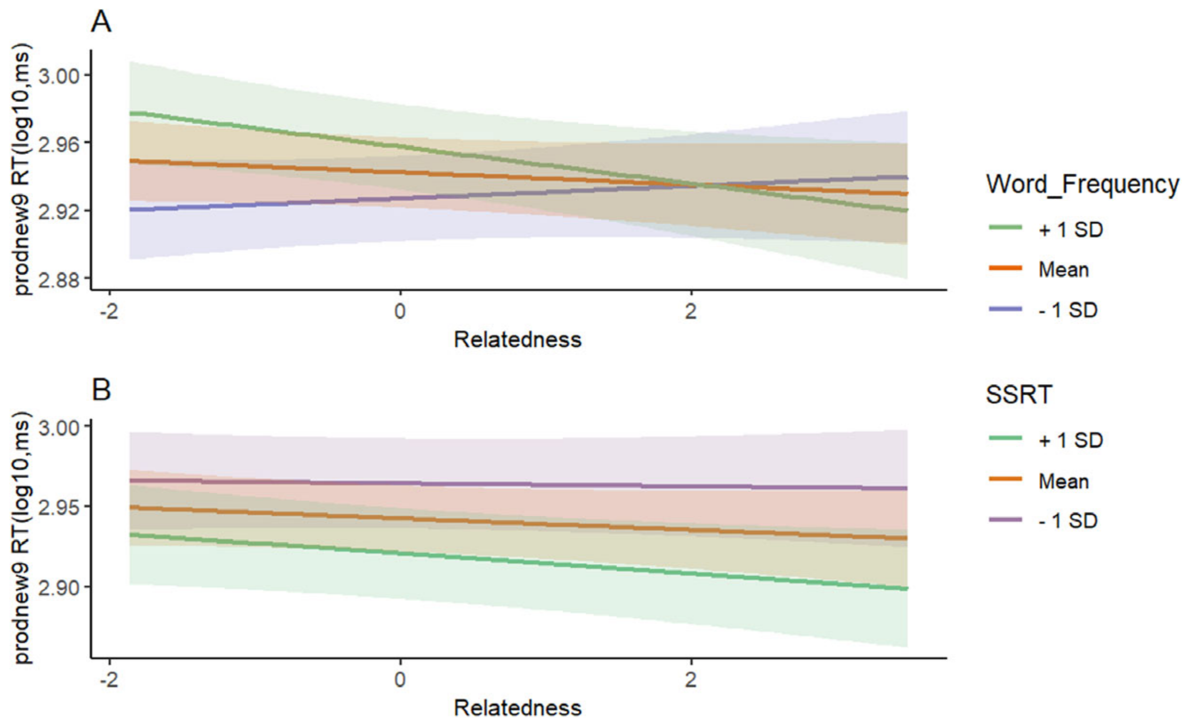

Figure 5. Interaction graph of Day 9 translation production test: Word Frequency × Relatedness (a) and Relatedness × SSRT (b).

words (*Mean* + 1*SD*), significant slope of word frequency emerged outside the Flanker_effect range of [−28.20, 61.29]. Within our observed participants' Flanker_effect range of [−32.81, 34.00], comparison of these ranges showed the following trends: for unrelated words, the slope of word frequency was significant regardless of Flanker_effect values; for related words, a smaller Flanker_effect led to a more significant slope. In other words, for homonyms, word frequency consistently influenced RTs, regardless of interference inhibition levels; for polysemes, stronger interference inhibition resulted in more significant RT differences across varying word frequencies. Similarly, due to the positive regression coefficient for word frequency, the influence of or the differences in word frequency implied that higher-frequency words yielded longer RTs compared to lower-frequency ones.

Compared with RTs in the immediate phase, RTs in this phase were generally longer (See Figure 4b). Trends for different ambiguous words were similar, with some distinctions. For unrelated words (HF_LR-■ and LF_LR-●), interference inhibition effects on different word frequencies remained consistent, showing similar steepness in both lines, which was unlike the immediate testing phase where the slope of low-frequency words was steeper than that of high-frequency words. Further combined with the simple test results of the two phases, the effect of word frequency was only seen in individuals with weak interference inhibition initially, but later in individuals with varying interference inhibition. In other words, as individuals produced new meanings for unrelated words with different frequencies, the role of interference inhibition diminished over time. For related words (LF_HR-▲ and HF_HR - +) and the overall steepness of four lines, their trends aligned with the immediate testing phase. Consequently, in the delayed phase, interference inhibition also influenced unrelated words and high-frequency related words, with the most prominent influence on the latter.

Second, regarding the Word frequency × Relatedness interaction, the simple-slope test revealed a significant relatedness slope (p < .05) when word frequency fell outside the range of [−3.16, 0.41]. Given the word frequency range of [−2.95, 1.63], the slope was more significant for higher-frequency words. This suggested that in high-frequency words (+1*SD*), unrelated words resulted in longer RTs (See Figure 5a). In Figure 5a, the trend for unrelated words (left end of the plot) was that high frequency led to longer RTs, whereas the trend for related words (right end of the plot) was the opposite. In short, word frequency and relatedness influenced RTs through distinct processes.

Third, regarding the Relatedness × SSRT interaction, the simple-slope test revealed a significant relatedness slope (p < .05) beyond the SSRT range of [−1055.27, 54.34]. Given the participants' SSRT range of [−96.77, 137.20], the relatedness slope was more significant in individuals with larger SSRTs. This suggested that weaker response inhibition (+1*SD*) was linked to greater RTs disparity between related and unrelated words (See Figure 5b). In Figure 5b, RT differences among individuals with varying response inhibition were larger in related words and relatively smaller in unrelated words, with differences indicating response inhibitory disadvantages—stronger inhibition led to longer RTs.

In summary, our translation production test results revealed that: a) regarding lexical characteristics, higher word frequency resulted in longer RTs. The interaction between word frequency and relatedness revealed different influence processes; b) regarding individual characteristics, both inhibitory components exhibited significant effects, showing a disadvantage, with stronger inhibitory control leading to longer RTs for producing new meanings; and c) regarding the interaction between the two characteristics, both inhibitory components exhibited inhibitory disadvantages in words with different relatedness, which were more prominent in related words. Furthermore, interference inhibition only affected high-frequency semantically related words.

5. Discussion

This study aims to investigate the impact of inhibitory control on LNM-FW in the context of L2 ambiguous word learning, focusing on Chinese-English bilinguals. Uniquely, our findings indicated an inhibitory disadvantage in this learning process, a novel observation contrasting with prior research (Lev-Ari & Keysar, 2014; Lu et al., 2017). We will delve into these results in the following.

5.1. Inhibitory control's effects on learning homonyms of varied word frequency

For semantically unrelated ambiguous words (i.e., homonyms), previous studies have mostly emphasized the interference among meanings (Chen & Chen, 2022; Zhang et al., 2018), which is especially prominent in high-frequency words (Fang et al., 2017). Thus, given the effects of low semantic relatedness and high word frequency, we initially hypothesized a significant role for inhibitory control in learning these words. Since learners cannot directly benefit from existing meanings, inhibitory control is needed to reduce interference, which is essential for establishing and connecting new semantic representations. Studies involving English pseudowords (Lu et al., 2017) support this view, showing that stronger inhibitory control facilitates second meaning learning. However, contrary to our expectations and past findings, we observed an inhibitory disadvantage in both translation recognition and production tests.

We believe that this is due to differences in the solidity of the existing meaning representations between the two learning processes. In Lu et al. (2017)'s study, learners encountered two unrelated meanings of English pseudowords within three days, with only a one-day gap between the learning of the first and second meanings. Although lexical tests showed a significant processing speed difference for the two meanings, this likely stemmed from varying learning times rather than representation solidity. In this scenario, the short learning period for both meanings did not allow for solid representations, leading to easier interference and competition. Here, inhibitory control influenced the learning of the second meaning, as it was necessary to overcome competing meanings for successful representation. In our study, participants learned new meanings based on a good command of existing meanings, leading to a stark contrast between representations and a low likelihood of fierce competition, a case particularly evident for high-frequency words. Under these circumstances, inhibitory control exerted a disadvantageous effect. We will further discuss the reasons for this inhibitory disadvantage later.

5.2. Inhibitory control's effects on learning polysemes of varied word frequency

For semantically related ambiguous words (i.e., polysemes), we hypothesized two potential roles of inhibitory control in LNM-FW: weak or strong. Previous studies (Chen & Chen, 2022; Rodd, 2020) have suggested that bilinguals effectively utilize positive links in polysemes, directly transferring similar representations from existing meanings to new ones. This implies that learners may not heavily rely on inhibitory control as an indirect aid. However, given that this process inherently involves modifying form-meaning mappings, learners might need to use inhibitory control to suppress existing meanings for the smooth establishment of similar new semantic representations and form-meaning bonds. Our results suggest that inhibitory control plays a relatively weak role in learning new meanings for polysemes, as this effect is primarily observed in high-frequency semantically related words, rather than uniformly across words of different frequencies, which is likely due to interference stemming from the greater stability of their existing form-meaning bonds.

Our results found that the influence of inhibitory control only emerged in translations production tests, especially for high-frequency related words, showing a similar disadvantage as observed in learning homonyms. This influence was not found in translation recognition tests, likely due to differences in semantic processing levels between the two tests. They represent two distinct activation processes of new meanings: BOTTOM-UP and TOP-DOWN (Kroll et al., 2010), each demanding different levels of semantic processing. In the mental lexicon, polysemes, due to overlapping representations, are stored as unified larger entries, while homonyms, with less overlap, are stored as distinct, smaller entries (Jager et al., 2016). Therefore, regardless of semantic processing levels required by tasks, specific meanings need to be extracted for homonyms, while for polysemes, it remains uncertain.

In translation recognition tests that require lower-level semantic processing, individuals simply activate the general meaning entry to assess whether the target word matches its new meaning. For example, *queen* can signify both "a female king" and "the lord of bees," with both meanings converging on the broader concept of "group leader" and stored under the entry *queen*. Deep semantic processing may not be necessary for a "yes" response because recognizing the general concept of "group leader" is sufficient. For a "no" response, individuals simply exclude any presented meaning unrelated to "group leader." Thus, in these tests, activating any meaning, either existing or new, triggers the broader meaning entry, thereby fulfilling task requirements. This process does not involve detailed extraction of new meaning representations, precluding the assessment of inhibitory control's influence.

In translation production tests that require higher-level semantic processing, individuals treat polysemes similarly to homonyms, extracting precise meaning representations to meet task requirements. For instance, processing *queen* involves reaching the deep concept of "the leader of the bee colony" to extract the specific meaning of "the lord of bees" instead of "female king." These tests demand finer discrimination between new and existing meanings within the larger entry to determine which representation to extract precisely, especially in high-frequency words with low overlapping of representations. In such cases, inhibitory control is needed, as existing meanings can interfere with new ones. However, this also displays an inhibitory disadvantage, where stronger inhibitory control leads to longer RTs. The reasons for this inhibitory disadvantage are discussed later.

5.3. Inhibitory disadvantage in tests reflects inhibition advantage in learning

From the preceding discussion, inhibitory control exhibits an inhibitory disadvantage in lexical tests for both homonyms and polysemes, especially in translation production tests. This indicates that stronger inhibitory control leads to slower recognition or production of new meanings. However, this disadvantage does not necessarily imply that those with stronger inhibitory control are less effective at learning new meanings. The cued-recall test conducted on Day 2 revealed that participants, regardless of their inhibitory control level, achieved over 98% accuracy in recalling new meanings for different L2 ambiguous words. We propose two possible explanations for this INHIBITORY DISADVANTAGE EFFECT:

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

One explanation is from Rodd (2020)'s perspective of lexical-semantic space reshaping degree. We assume that inhibitory disadvantage arises from diverse reshaping degrees in the lexical-semantic space among individuals with different levels of inhibition. Specifically, Rodd emphasized that when individuals learn new meanings, the initial stable state of lexical-semantic space for quick assessment (McLeod et al., 2000) was broken to integrate new meanings, forming a new stable state. Therefore, whenever new meanings are learned, the lexical-semantic space undergoes reshaping, transitioning from an existing stable state to a new one. Previous studies reveal that individuals with stronger inhibitory control excel in L2 word learning by efficiently engaging brain regions associated with inhibitory control like the anterior cingulate cortex and the right ventrolateral prefrontal cortex, especially in the early stages of word learning (Rueschemeyer & Gaskell, 2018). They tend to expedite learning through greater early activation of the frontal inhibitory network. In contrast, individuals with weaker inhibitory control prolong this process and rely on the inhibitory control-related brain network to gradually adjust their vocabulary knowledge in later stages of word learning (Grant et al., 2015).

Therefore, we speculate that stronger inhibitory control facilitates early reshaping of lexical-semantic space in L2 ambiguous word learning, leading to quicker formation of stable lexical access. However, complete reshaping in the strong inhibition group can greatly damage the lexical-semantic space, making it challenging to extract new meanings quickly due to the unstable state before reshaping finishes. Therefore, in the early, incomplete stages of reshaping, stronger inhibitory control may result in a disadvantage in different lexical tests. The reshaping of our participants was not complete when they took the immediate and delayed tests because they had only learned the new meanings six times without any review or consolidation after each time's study. Over time, processing speed and accuracy for these new meanings gradually declined, as seen in the shorter RTs on the second day compared to the ninth day (see Table 2). Once space reshaping is completed and new meanings reach a stable state, without deliberate practice, individuals should process these meanings at a relatively consistent speed and accuracy at any time. Future research aiming to validate this explanation should consider a longer-term process, focusing on how different levels of inhibitory control influence learning after the lexical-semantic space has been fully reshaped.

A similar phenomenon has been observed in research on WORKING MEMORY (WM, another component of executive function). For instance, Tokowicz et al., (2004) found that high working memory capacity (WMC) individuals with study-abroad experience made more semantic errors during word translation compared to their low-capacity counterparts. They suggested that this might be because the immersive foreign language environment encouraged approximate translation, which requires holding multiple items in memory simultaneously—a task more manageable for high-capacity individuals. This additional cognitive load may reduce accuracy, or it may reflect high WMC individuals' greater ability to restructure their lexical networks to accommodate frequent L2 use.

Another possible explanation centers on competitive activation. Maciejewski and Klepousniotou (2020) emphasized that the disadvantage in processing ambiguous words arose from semantic activation competition rather than difficulties in response selection. This competition becomes apparent when different meanings are activated to a similar degree. Therefore, the inhibitory disadvantage could arise from fierce competition between existing and new meanings in L2 ambiguous word learning for individuals with strong inhibitory control. Stronger inhibitory control allows for better indirect influence on new meaning learning, compensating for the lack of direct transfer. Hence, if the learning of new meanings is effective, it causes a significant activation of both the new and existing meanings, resulting in competition and requiring more time for resolution, thus exhibiting the disadvantage effect. In contrast, people with weaker inhibitory control may not build new meanings that are strong enough to compete with existing ones, thus avoiding competition and processing words faster.

Interestingly, Kroll et al. (2002) observed a similar pattern in high WMC bilinguals, who performed worse on cognate word translations compared to low WMC individuals, yet better on noncognate translations. Typically, cognates are processed more easily, which supports the competitive activation explanation: high WMC bilinguals can activate multiple meanings simultaneously, leading to greater competition and slower processing speeds.

Whether viewed from the perspective of restructuring the lexical-semantic space or competition activation, both explanations highlight a common phenomenon: learners with stronger inhibitory control face greater learning costs during the early stages of L2 word learning compared to those with weaker inhibitory control. This aligns with the concept of DESIRABLE DIFFICULTIES proposed by Bogulski et al. (2019) during the early stage of new language acquisition for adult bilinguals, where these initial challenges—increased learning costs caused by induced errors or conceptual elaboration—ultimately benefit long-term learning and memory retention.

More importantly, Bogulski et al. (2019) denied the cognitive control advantage hypothesis, suggesting that superior cognitive control scores were not necessarily linked to better word learning outcomes. Instead, they attributed these desirable difficulties to bilinguals' enhanced ability to regulate their L1. Our research extends this idea by showing that inhibitory control might be a contributing factor to the mechanism behind these desirable difficulties. Taken together with similar findings from WM research (Kroll et al., 2002; Tokowicz et al., 2004), it suggests that future research should further explore the link between executive functions and desirable difficulties in language acquisition.

Additionally, we would like to re-emphasize the role of word frequency. As previously mentioned, word frequency serves as an effective indicator of a word's form-meaning bond solidity, which our findings have confirmed. High-frequency L2 ambiguous words have more stable existing meanings compared to low-frequency words, making it harder to adjust the form-meaning mappings both for homonyms and polysemes. As we all know, changing something already established is always difficult, and the same applies to word learning.

Although both types of lexical tests in our study involve semantic processing at different levels, evaluating the lexicalization effect of new meanings across individuals with varying inhibition remains challenging. Fang et al. (2017) found that high-frequency words presented a disadvantage in semantic judgment tasks requiring extensive semantic processing. However, these words showed an advantage in lexical decision tasks, which require minimal semantic processing. They thus concluded that while high-frequency words interfered more with new meanings, they also enhanced co-activation of new and existing meanings, aiding the learning process. Future research could employ tasks like lexical decision, which require less semantic processing, to compare overall learning outcomes of new meanings across varying levels of inhibitory control.

https://doi.org/10.1017/S1366728924000993 Published online by Cambridge University Press

5.4. Which one is more important: interference inhibition or response inhibition?

The previous study emphasized interference inhibition's crucial role in L2 ambiguous words learning, but the role of response inhibition remained unexplored (Lu et al., 2017). Employing flanker and stop-signal tasks to assess both types of inhibitions in LNM-FW for L2 ambiguous words, we hypothesized interference inhibition played a greater role. According to their definitions (Diamond, 2013), interference inhibition is influenced by the lexical features of ambiguous words, while response inhibition is linked to prior knowledge (i.e., the control of impulsive behaviors developed in earlier learning).

Indeed, our findings affirmed this. Our results highlighted the dominance of interference inhibition across various tests. For instance, significant interactions between Flanker_effect, and word frequency and relatedness were observed in both immediate and delayed phases of the translation production tests. Response inhibition had a lesser yet significant role, such as the two-way interaction during the delayed phase of the translation production tests. Despite this, both types exhibited inhibitory disadvantages: the disadvantage of interference inhibition related to the state of lexical-semantic space or the extent of new meaning learning, while that of response inhibition was only found in production tests that involved speaking and mouth muscle movements. Lev-Ari and Keysar (2014) discovered that individuals with weaker inhibitory skills struggled to distinguish meanings of polysemes and homonyms, facing challenges in suppressing impulsive reactions of existing meanings compared to new ones. However, this limitation is helpful when people initially learning new meanings. Our tests, which do not require detailed distinctions between meanings, show that extracting existing meanings can accelerate their response to tasks. But those with strong response inhibition focus on accurately extracting new meanings, as they are more adept at detecting differences in meanings of ambiguous words, which is beneficial for long-term semantic access. This may explain the response inhibition disadvantage observed in our research.

Overall, both interference and response inhibition contribute to the overall dynamic adjustment of the lexical-semantic space in L2 ambiguous word learning, with interference inhibition being more pronounced. Although both exhibit an inhibitory disadvantage, this disadvantage aids in learning new meanings. Future research might benefit from longer-term studies or lexical tests evaluating the lexicalization of new meanings. Also, considering the design of our study, which did not involve revisiting new meanings, the reshaping process might have remained incomplete. Future studies could investigate if immediate postlearning review expedites this reshaping. And from a broader perspective, our study reaffirms the importance of executive function elements in language acquisition and underscores the importance of considering both learner characteristics and lexical features in language learning research.

6. Conclusion

The present study is the first to comprehensively investigate the role of inhibitory control in learning new knowledge based on existing knowledge within the context of L2 ambiguous word learning, focusing on the process of learning new meanings for familiar words. We found that inhibitory control could be an effective supplementary pathway in this learning process. Specifically, in our study, learning new meanings for both high-frequency and low-frequency homonyms necessitated inhibitory control, whereas for polysemes, only high-frequency words required it. Inhibitory control plays an important role in learning new meanings of L2 ambiguous words, especially when lexical characteristics like high frequency or low semantic relatedness hinder direct knowledge transfer. However, this can initially result in intensified competition between new and existing meanings or substantial reshaping of the lexical-semantic space, temporarily creating a disadvantage in lexical access.

Supplementary material. To view supplementary material for this article, please visit https://osf.io/dz2aj/.

Data availability statement. The data supporting the findings of this study are available at OSF (https://osf.io/dz2aj/).

Acknowledgments. This work was supported by funding from the National Natural Science Foundation of China (31970976) for Baoguo Chen. We would like to thank Geyang Xu for her linguistic support in preparing the final version.

References

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, **5**(3), 436–461. https://doi.org/10.1075/ml.5.3.10baa.

Barr, D. J., **Levy, R.**, **Scheepers, C.**, & **Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**(3), 255=278. https://doi.org/10.1016/j.jml.2012.11.001.

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

Bauer, D. J., & **Curran, P. J.** (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, **40**(3), 373–400. https://doi.org/10.1207/s15327906mbr4003_5.

Bein, O., **Trzewik, M.**, & **Maril, A.** (2019). The role of prior knowledge in incremental associative learning: An empirical and computational approach. *Journal of Memory and Language*, **107**, 1–24. https://doi.org/10.1016/j.jml.2019.03.006.

Bogulski, C. A., **Bice, K.**, & **Kroll, J. F.** (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, **22**(5), 1052–1067. https://doi.org/10.1017/S1366728918000858.

Bracken, J., **Degani, T.**, **Eddington, C.**, & **Tokowicz, N.** (2017). Translation semantic variability: How semantic relatedness affects learning of translation-ambiguous words. *Bilingualism: Language and Cognition*, **20**(4), 783–794. https://doi.org/10.1017/S1366728916000274.

Brod, G., **Werkle-Bergner, M.**, & **Shing, Y. L.** (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience*, **7**. https://doi.org/10.3389/fnbeh.2013.00139.

Chen, J., & **Chen, B.** (2022). Learning new meanings for known L2 words: Long-term semantic representation is updated to integrate new information after consolidation. *Psychophysiology*. https://doi.org/10.1111/psyp.14228.

Clark, E. V., & **Clark, H. H.** (1979). When nouns surface as verbs. *Language*, **55**(4), 767. https://doi.org/10.2307/412745.

Crosthwaite, P., **Bailey, D.**, & **Meeker, A.** (2015). Assessing in-class participation for EFL: Considerations of effectiveness and fairness for different learning styles. *Language Testing in Asia*, **5**. https://doi.org/10.1186/s40468-015-0017-1.

D'Agostino, R. B. (1970). Transformation to normality of the null distribution of g1. *Biometrika*, **57**(3), 679–681. https://doi.org/10.2307/2334794.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, **64**(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750.

Elgort, I., & **Piasecki, A. E.** (2014). The effect of a bilingual learning mode on the establishment of lexical semantic representations in the L2. *Bilingualism: Language and Cognition*, **17**(3), 572–588. https://doi.org/10.1017/S1366728913000588.

Fang, X., & **Perfetti, C. A.** (2017). Perturbation of old knowledge precedes integration of new knowledge. *Neuropsychologia*, **99**, 270–278. https://doi.org/10.1016/j.neuropsychologia.2017.03.015.

Fang, X., & **Perfetti, C. A.** (2019). Learning new meanings for known words: Perturbation of original meanings and retention of new meanings. *Memory & Cognition*, **47**(1), 130–144. https://doi.org/10.3758/s13421-018-0855-z.

Fang, X., **Perfetti, C.**, & **Stafura, J.** (2017). Learning new meanings for known words: Biphasic effects of prior knowledge. *Language, Cognition and Neuroscience*, **32**(5), 637–649. https://doi.org/10.1080/23273798.2016.1252050.

Grant, A. M., **Fang, S.-Y.**, & **Li, P.** (2015). Second language lexical development and cognitive control: A longitudinal fMRI study. *Brain and Language*, **144**, 35–47. https://doi.org/10.1016/j.bandl.2015.03.010.

Grenell, A., & **Carlson, S. M.** (2021). Individual differences in executive function and learning: The role of knowledge type and conflict with prior knowledge. *Journal of Experimental Child Psychology*, **206**, 105079. https://doi.org/10.1016/j.jecp.2020.105079.

Hirosh, Z., & **Degani, T.** (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, **25**(3), 892–916. https://doi.org/10.3758/s13423-017-1315-7.

Jager, B., **Green, M. J.**, & **Cleland, A. A.** (2016). Polysemy in the mental lexicon: Relatedness and frequency affect representational overlap. *Language, Cognition and Neuroscience*, **31**(3), 425–429. https://doi.org/10.1080/23273798.2015.1105986.

Kang, S. H. K., **Gollan, T. H.**, & **Pashler, H.** (2013). Don't just repeat after me: Retrieval practice is better than imitation for foreign vocabulary learning. *Psychonomic Bulletin & Review*, **20**(6), 1259–1265. https://doi.org/10.3758/s13423-013-0450-z.

Kang, W., **Hernández, S. P.**, **Rahman, M. S.**, **Voigt, K.**, & **Malvaso, A.** (2022). Inhibitory control development: A network neuroscience perspective. *Frontiers in Psychology*, **13**, 651547. https://doi.org/10.3389/fpsyg.2022.651547.

Kapa, L. L., & **Colombo, J.** (2014). Executive function predicts artificial language learning. *Journal of Memory and Language*, **76**, 237–252. https://doi.org/10.1016/j.jml.2014.07.004.

Kapa, L. L., & **Erikson, J. A.** (2020). The relationship between word learning and executive function in preschoolers with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research*, **63**(7), 2293–2307. https://doi.org/10.1044/2020_JSLHR-19-00342.

Kleiman, E. (2021). *Package 'EMAtools'* (CRAN; Version 0.1.4) [Computer software]. https://cran.r-project.org/web/packages/EMAtools/index.html.

Klepousniotou, E., **Pike, G. B.**, **Steinhauer, K.**, & **Gracco, V.** (2012). Not all ambiguous words are created equal: An EEG investigation of homonymy and polysemy. *Brain and Language*, **123**(1), 11–21. https://doi.org/10.1016/j.bandl.2012.06.007.

Kroll, J. F., **Hell, J. G. V.**, **Tokowicz, N.**, & **Green, D. W.** (2010). The revised hierarchical model: A critical review and assessment*. *Bilingualism: Language and Cognition*, **13**(3), 373–381. https://doi.org/10.1017/S136672891000009X.

Kroll, J. F., **Michael, E.**, **Tokowicz, N.**, & **Dufour, R.** (2002). The development of lexical fluency in a second language. *Second Language Research*, **18**(2), 137–171. https://doi.org/10.1191/0267658302sr201oa.

Kuznetsova, A., **Brockhoff, P. B.**, & **Christensen, R. H. B.** (2017). **lmerTest** package: Tests in linear mixed effects models. *Journal of Statistical Software*, **82**(13). https://doi.org/10.18637/jss.v082.i13.

Lev-Ari, S., & **Keysar, B.** (2014). Executive control influences linguistic representations. *Memory & Cognition*, **42**(2), 247–263. https://doi.org/10.3758/s13421-013-0352-3.

Li, P., **Zhang, F.**, **Yu, A.**, & **Zhao, X.** (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, **23**(5), 938–944. https://doi.org/10.1017/S1366728918001153.

Lu, Y., **Wu, J.**, **Dunlap, S.**, & **Chen, B.** (2017). The inhibitory mechanism in learning ambiguous words in a second language. *Frontiers in Psychology*, **8**, 636. https://doi.org/10.3389/fpsyg.2017.00636.

Maciejewski, G., & **Klepousniotou, E.** (2020). Disambiguating the ambiguity disadvantage effect: Behavioral and electrophysiological evidence for semantic competition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **46**(9), 1682–1700. https://doi.org/10.1037/xlm0000842.

McLeod, P., **Shallice, T.**, & **Plaut, D. C.** (2000). Attractor dynamics in word recognition: Converging evidence from errors by normal subjects, dyslexic patients and a connectionist model. *Cognition*, **74**(1), 91–114. https://doi.org/10.1016/S0010-0277(99)00067-0.

Mirisola, A., & **Seta, L.** (2016). *Package 'pequod'* (CRAN; Version 0.0–5) [Computer software]. https://cran.r-project.org/web/packages/pequod/index.html

Mo, Q., & **Sun, L.** (2004). On the semantic development of english high-frequency verbs among Chinese students. *Journal of PLA University of Foreign Languages*, **27**(1), 65–69. https://doi.org/10.3969/j.issn.1002-722X.2004.01.014.

Perfetti, C., & **Hart, L.** (2002). *The lexical quality hypothesis*, **11**, 189–213. https://doi.org/10.1075/swll.11.14per.

Piantadosi, S. T., **Tily, H.**, & **Gibson, E.** (2012). The communicative function of ambiguity in language. *Cognition*, **122**(3), 280–291. https://doi.org/10.1016/j.cognition.2011.10.004.

Qu, Y., & **Zhang, J.** (2005). An empirical study of English polysemy netwoks acquisition. *Foreign Languages Research*, **2**, 43–48. https://doi.org/10.3969/j.issn.1005-7242.2005.02.010.

R Core Team. (2022). *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* (Version 4.2.1) [Computer software]. https://www.R-project.org/.

Rodd, J. (2020). Settling into semantic space: An ambiguity-focused account of word-meaning access. *Perspectives on Psychological Science*, **15**, 174569161988586. https://doi.org/10.1177/1745691619885860.

Rodd, J., **Gaskell, G.**, & **Marslen-Wilson, W.** (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, **46**(2), 245–266. https://doi.org/10.1006/jmla.2001.2810.

Rueschemeyer, S.-A., & **Gaskell, M. G.** (2018). *The Oxford handbook of psycholinguistics* (2nd ed.). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.001.0001.

Serrat, E., **Amadó, A.**, **Bonet, A.**, **Feijóo, S.**, & **Aguilar-Mediavilla, E.** (2021). Growing up in a socially bilingual environment: Simultaneous acquisition of a typologically close language (Crecer en un entorno social bilingüe: adquisición de una lengua de forma simultánea a otra lengua tipológicamente similar). *Journal for the Study of Education and Development*, **44**(2), 336–369. https://doi.org/10.1080/02103702.2021.1888490.

Storkel, H. L., **Maekawa, J.**, & **Aschenbrenner, A. J.** (2013). The effect of homonymy on learning correctly articulated versus misarticulated words. *Journal of Speech, Language and Hearing Research (Online)*, **56**(2), 694–707. https://doi.org/10.1044/1092-4388(2012/12-0122.

Tokowicz, N., **Michael, E. B.**, & **Kroll, J. F.** (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition*, **7**(3), 255–272. https://doi.org/10.1017/S1366728904001634.

van Heuven, W. J. B., **Mandera, P.**, **Keuleers, E.**, & **Brysbaert, M.** (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, **67**(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521.

Verbruggen, F., **Chambers, C. D.**, & **Logan, G. D.** (2013). Fictitious inhibitory differences: How skewness and slowing distort the estimation of stopping latencies. *Psychological Science*, **24**(3), 352–362. https://doi.org/10.1177/0956797612457390.

Verbruggen, F., & **Logan, G. D.** (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences*, **12**(11), 418–424. https://doi.org/10.1016/j.tics.2008.07.005.

Vitevitch, M. S., **Storkel, H. L.**, **Francisco, A. C.**, **Evans, K. J.**, & **Goldstein, R.** (2014). The influence of known-word frequency on the acquisition of new neighbours in adults: Evidence for exemplar representations in word learning. *Language, Cognition and Neuroscience*, **29**(10), 1311–1316. https://doi.org/10.1080/23273798.2014.912342.

Warmington, M. A., **Kandru-Pothineni, S.**, & **Hitch, G. J.** (2019). Novel-word learning, executive control and working memory: A bilingual advantage. *Bilingualism: Language and Cognition*, **22**(04), 763–782. https://doi.org/10.1017/S136672891800041X.

Witherby, A. E., & **Carpenter, S. K.** (2022). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **48**(4), 483–498. https://doi.org/10.1037/xlm0000996.

Zhang, H., & **Wang, X.** (1986). *Raven's Standard Reasoning Test manual* (Chinese Urban Revision). Beijing Normal University.

Zhang, Y., **Chen, B.**, **Tang, Y.**, **Yao, P.**, & **Lu, Y.** (2018). Semantic similarity to known second language words impacts learning of new meanings. *Frontiers in Psychology*, **9**, 2048. https://doi.org/10.3389/fpsyg.2018.02048.

Zhang, Y., **Yang, X.**, **Liang, L.**, & **Chen, B.** (2022). The impact of learning new meaning on the previously learned meaning of L2 ambiguous words: The role of semantic similarity. *Psychophysiology*, **59**(4). https://doi.org/10.1111/psyp.13992.

Zipf, G. K. (1945). The meaning-frequency relationship of words. *The Journal of General Psychology*, **33**(2), 251–256. https://doi.org/10.1080/00221309.1945.10544509.